sustainability

MDPI

Article

Factors Associated with Spatial Heterogeneity in Regional Agricultural Output: A GeoDetector-Based Analysis for Sustainable Agricultural Recovery in Ukraine

Olena Pavlova [1,2,*], Kostiantyn Pavlov [2,3,*], Sławomir Ziółkowski [1], Antoni Machowski [1], Viktor Diachenko [4], Olena Smihunova [5], Oleksandr Yefymenko [6], Kateryna Bogomolova [7] and Andrii Yavorskyi [8]

1 Faculty of Management, AGH University of Kraków, Gramatyka 10, 30-059 Kraków, Poland; slawomir@agh.edu.pl (S.Z.); amachowski@agh.edu.pl (A.M.)
2 Faculty of Economics and Management, Lesya Ukrainka Volyn National University, Voli Ave., 13, 43025 Lutsk, Ukraine
3 Faculty of Administration and Social Sciences, WSEI University in Lublin, Projektowa 4, 20-209 Lublin, Poland
4 FLUGART Limited Liability Company, Heroes of Maidan 10, 29005 Khmelnytskyi, Ukraine; dyachenko2023@gmail.com
5 Livestock Farming Institute, National Academy of Agrarian Sciences of Ukraine, Tvarynnykiv 1A, 61026 Kharkiv, Ukraine; elenasmigunova@gmail.com
6 Mechanical Faculty, Kharkiv National Automobile and Highway University, Yaroslava Mudroho 25, 61002 Kharkiv, Ukraine; khadi.alef@gmail.com
7 Department of Management, Business and Administration, State Biotechnological University, Alchevskih 44, 61002 Kharkiv, Ukraine; katbogomolova80@gmail.com
8 Faculty of Philology and Journalism, Lesya Ukrainka Volyn National University, Voli Ave., 13, 43025 Lutsk, Ukraine; javorskyi.andrii@vnu.edu.ua
* Correspondence: opavlova@agh.edu.pl (O.P.); pavlovkostya@gmail.com (K.P.)

Abstract

Spatially balanced agricultural development underpins food security and sustainable recovery in war-affected regions. This study examines factors associated with spatial differences in agricultural output across all 24 districts of Chernihiv, Sumy, Kharkiv and Dnipropetrovsk oblasts using GIS and GeoDetector. The database combines published district statistics for 2025 with open satellite, reanalysis, infrastructure and conflict-event data, so that every value is traceable to a public source. Eight factors were assessed with the four GeoDetector modules, selection-adjusted permutation tests with multiplicity adjustment and a regression benchmark. Crop structure—the maize share of the grain area -showed the strongest association with output ($q = 0.430$; $q = 0.610$ for crop output, Holm-adjusted $p = 0.012$), followed by distance to the front line ($q = 0.324$) and to the Russian border ($q = 0.272$). Vegetation condition, drought, road density and conflict intensity were weakly associated, and no joint stratification added significantly to its stronger component. Output was spatially clustered (Moran's $I = 0.104$, $p = 0.016$). The results describe cross-sectional associations; they suggest that recovery policies should consider crop specialization and proximity to hostilities when targeting support to districts.

Keywords: sustainable agricultural development; sustainable recovery; food security; spatial heterogeneity; GeoDetector; GIS; crop structure; Ukraine; spatial stratified heterogeneity; frontline regions

Check for updates

Academic Editor: Jianming Cai

Received: 23 October 2026
Revised: 14 November 2026
Accepted: 16 November 2026
Published: 18 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Agricultural development exhibits marked spatial heterogeneity, driven by variations in regional resource endowments, infrastructure, and economic efficiency [1]. This uneven

distribution stems from the interplay of environmental conditions, institutional capacity, and, increasingly, security-related disruptions [2]. Analyzing these mechanisms requires a spatial approach, as growth determinants rarely operate independently and often show nonlinear geographic variations [3]; nearby locations also tend to be more strongly related than distant ones, reinforcing the need for explicitly spatial analysis [4]. The theoretical basis for this lies in competitive advantage [5], New Economic Geography [6], and spatial agglomeration, which suggest that regional competitiveness relies on spatial accessibility and cumulative effects as much as on natural endowments [7]. Recent regional evidence likewise demonstrates substantial spatial heterogeneity across agricultural and livestock production systems, reinforcing the need for integrated regional analysis [8].

Advances in GIS and remote sensing now enable more precise investigations of these regional systems [9]. Spatially explicit models demonstrate that agricultural performance analysis must account for geographic dependence and multidimensional variable interactions [10]. However, empirical agricultural economics still relies heavily on linear regression models that assume spatial stationarity and variable independence [11]. These assumptions often fail in real-world systems, preventing conventional econometrics from capturing nonlinear effects or stratified geographic variations [12].

Variance-based spatial inference methods, notably GeoDetector, address these limitations by measuring spatial stratified heterogeneity without assuming normality or linearity [11,13]. While GeoDetector is widely used in ecological and public health studies [11,13], its application in agricultural economics—particularly when integrating environmental, economic, infrastructural and security-related drivers simultaneously—remains limited.

Ukraine is a critical case: its regions differ in natural resources, infrastructure and socioeconomic profiles, and military disruptions and logistical bottlenecks have further transformed them. Strengthening food security and agricultural resilience requires tools that identify these concurrent spatial drivers [14], yet comprehensive spatial assessments remain scarce [13]. Persistent disparities in agricultural output translate into uneven rural incomes, unequal access to food and inefficient land use, so sustainable recovery requires evidence on where and why output differs.

This study addresses these research gaps by answering the following research questions:

RQ1. Which environmental, economic, infrastructural and security-related factors are most strongly associated with the spatial heterogeneity of agricultural development across districts of the study area?

RQ2. How are joint stratifications of these factors associated with spatial differentiation, and do they strengthen or weaken the explanatory power of individual factors?

RQ3. To what extent do the GeoDetector results agree with, or add to, those of a conventional correlation and regression benchmark estimated on the same district-level data?

The approach combines regional economic theory, which emphasizes the cumulative effects of productive specialization, accessibility and local capacity, with GIS and spatial stratified heterogeneity analysis: GeoDetector quantifies how spatial strata explain outcome variance and how factors interact without imposing a linear functional form [5–7,11,13].

The study contributes in three ways. First, it applies spatial stratified heterogeneity analysis to agricultural output in frontline Ukrainian regions using a district database in which every value can be traced to a published or open source. Second, it addresses the small-sample limitations of GeoDetector with selection-adjusted permutation tests, multiplicity adjustment, conditional tests of joint stratifications and alternative outcome definitions. Third, it compares the GeoDetector results with a regression benchmark, clarifying what the stratified approach adds and where its inferential limits lie.

2. Materials and Methods

2.1. Research Framework and Analytical Design

GeoDetector (implemented in Python, version 3.13.18; see Section 2.5.1) was used to quantify spatial stratified heterogeneity and interactions among environmental, agricultural-structure, infrastructure and security factors. The workflow comprised database compilation, normalization, discretization, the four GeoDetector modules and robustness checks (alternative class numbers, Global Moran's I, leave-one-district-out resampling, selection-adjusted permutation tests, secondary outcomes and a regression benchmark) [11–13].

2.2. Study Area and Agricultural Economic Database

The analysis covers all 24 districts (raions) of Chernihiv (5), Sumy (5), Kharkiv (7) and Dnipropetrovsk (7) oblasts under the 2020 administrative reform; i.e., the complete macro-region rather than a sample (Figure 1; Table S1). The database refers to 2025, the most recent year with complete district tables in all four regional statistics offices; district production series for 2022–2024 are not published for all four oblasts. Agricultural indicators were taken from the district tables of the regional statistics offices [15–18] and district boundaries, areas and codes from OCHA COD-AB [19]. Environmental indicators were derived from MODIS and ERA5 [20–23], proximity indicators from country boundaries and occupied-territory maps [24,25], infrastructure indicators from OpenStreetMap [26,27] and conflict intensity from a georeferenced event dataset. Only indicators available for every district from public sources were retained, so that each value can be traced to its source (Table S13).

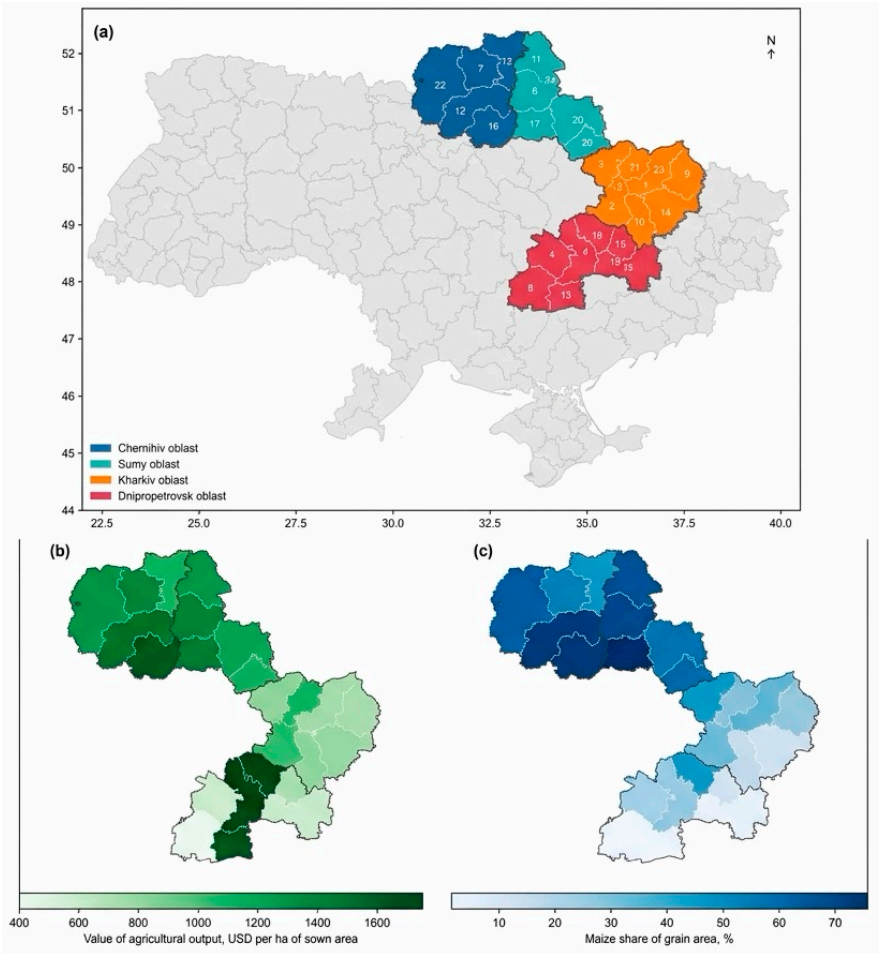


Figure 1. Study area and spatial distribution of the main variables: (**a**) location of the 24 districts of Chernihiv, Sumy, Kharkiv and Dnipropetrovsk oblasts (district numbers as in Table S1); (**b**) value of

agricultural output per hectare of sown area, USD/ha; (**c**) maize share of the grain area, %. Source: authors' elaboration based on OCHA COD-AB administrative boundaries (version 05, 2025) [19] and district tables of the regional statistics offices [15–18].

District statistics cover enterprises that meet the criteria of the statistical methodology and exclude temporarily occupied territories and areas of hostilities; values withheld for confidentiality (symbol "к") were not used to construct agricultural output. For four districts (Bohodukhivskyi, Kupianskyi, Kamianskyi and Synelnykovskyi), suppressed sown areas of roots, tubers and vegetables were treated as zero when constructing total sown area (Table S13). Spatial processing used the ETRS89-LAEA (EPSG:3035) projection. Output values are presented in USD at the 2025 NBU average rate (UAH 41.69 per USD); the conversion does not affect any q-statistic or test.

2.3. Indicator Selection and Variable Construction

The dependent variable (Y) is the value of the agricultural output of enterprises in 2025. Because district tables report production in physical units, Y was computed as the sum of grain, sunflower and livestock (live weight) production valued at common prices—the mean of the four oblasts' average sale prices of enterprises in January–November 2025 (UAH 8560.4, 22,004.5 and 68,943.2 per tonne, respectively) [15–18]—so that Y is a volume measure comparable across districts. These are the only products with published values for all 24 districts (rapeseed, milk and eggs are partly withheld for confidentiality; soybean is not published by district), and for each product the sum over districts equals the published oblast total. The eight explanatory variables (X1–X8) are defined in Table 1; descriptive statistics and correlations are given in Tables S2 and S3.

Table 1. Variables included in the GeoDetector analysis.

Variable	Indicator	Unit	Period/ Aggregation	Hypothesized Direction	Data Source
Y	Value of agricultural output	USD million	2025	–	Regional statistics offices [15–18]
X1	NDVI	index	23 April–11 July 2025, mean	+	MODIS MOD13Q1 [20]
X2	Sown-area share	%	2025	+	Regional statistics offices [15–18]; OCHA [19]
X3	SPEI	index	April–August 2025	±	ERA5 [21,22]; method [23]
X4	Maize share of grain area	%	2025	+	Regional statistics offices [15–18]
X5	Distance to the Russian border	km	static	+	Natural Earth [24]
X6	Distance to the front line	km	1 July 2025	+	DeepState map [25]
X7	Main-road density	km/km^2	November 2026 snapshot	+	OpenStreetMap [26,27]
X8	Conflict intensity	events per 1000 km^2	2022–2025, cumulative	–	UCDP GED v26.1 [29]

Note: NDVI—Normalized Difference Vegetation Index; SPEI—Standardized Precipitation-Evapotranspiration Index; + = benefit-type indicator; – = cost-type indicator; ± = indeterminate direction. Monetary values are converted at UAH 41.69 per USD. Hypothesized directions are tested with Spearman correlations in Section 3.6. Source: own composition (see Data Source column).

NDVI (X1) is the district mean of good-quality MODIS MOD13Q1 v6.1 pixels [20], averaged over the 16-day composites of 23 April–11 July 2025 (later composites were not available in the Microsoft Planetary Computer archive). Sown-area share (X2) is the sown area of enterprises as a percentage of district area, and maize share (X4) the maize area as a percentage of the grain area. SPEI (X3) was computed from ERA5 [21] daily precipitation and FAO-56 reference evapotranspiration (Open-Meteo [22]): the April–August 2025 water

balance was standardized with a log-logistic distribution fitted to 1991–2020 [23] at 62 grid points (0.5°) and averaged within districts. X5 and X6 are the distances from the district centroid to the Russian Federation [24] and to the nearest occupied territory on the DeepState map of 1 July 2025 [25]. Main-road density (X7) is the length of motorway to tertiary roads per km^2 [26,27]. All geometric quantities were computed using GeoPandas (version 1.2.0) in Python (version 3.13.18) [28]. Conflict intensity (X8) is the number of UCDP GED v26.1 events with location precision of 25 km or better per 1000 km^2 in 2022–2025 [29].

Because Y increases with district size, output per hectare of sown area (Y_{ha}) and crop output only (grain and sunflower) were analyzed as secondary outcomes; total output (Y) is the primary outcome (Section 3.5). Table S13 documents the source of every published value and the construction of derived variables.

2.4. Indicator Processing and Spatial Stratification

2.4.1. Indicator Array and Normalization

The observations were organized in a third-order indicator structure to maintain consistent indexing of districts, indicators and the observation period:

$$D = \{d_{ijp}\},\ i = 1,\ldots, m;\ j = 1,\ldots, n;\ p = 1,\ldots, P \tag{1}$$

where D is the complete indicator array; d_{ijp} is the observed value of indicator j for district i in observation period p; $i = 1,\ldots$, m indexes districts; $j = 1, \ldots, n$ indexes indicators; $p = 1, \ldots, P$ indexes observation periods; m is the number of districts; n is the number of indicators; and P is the number of observation periods. In the empirical application, $P = 1$ because the database is cross-sectional. The period index p is retained so that the same notation applies to a future panel extension; in the present application, the array reduces to an $m \times n$ district-by-indicator matrix.

Indicators were transformed to dimensionless values using logistic normalization and then oriented so that larger values consistently represent more favorable conditions:

$$z_{ijp} = \frac{1}{1 + exp[-(\ \frac{d_{ijp} - \mu_j}{\gamma \sigma_j}\ )]}, \tag{2}$$

where z_{ijp} is the logistic-normalized value of d_{ijp}; d_{ijp} is the original value of indicator j for district i in period p; μ_j is the sample mean of indicator j; σ_j is its sample standard deviation; $\gamma > 0$ is the smoothness parameter of the logistic transformation; and $\exp(\cdot)$ denotes the exponential function. For benefit-type indicators, larger original values are favorable. Cost-type indicators are reverse-oriented after normalization according to Equation (3).

$$\tilde{z_{ijp}} = \begin{cases} z_{ijp}\ benefit\ indicator \\ 1 - z_{ijp}\ cost\ indicator. \end{cases} \tag{3}$$

In Equation (3), $\tilde{z_{ijp}}$ denotes the orientation-consistent normalized value used in subsequent analysis: $\tilde{z_{ijp}} = z_{ijp}$ for a benefit-type indicator and $\tilde{z_{ijp}} = 1 - z_{ijp}$ for a cost-type indicator. Thus, higher $\tilde{z_{ijp}}$ values consistently represent more favorable conditions. This orientation provides a common analytical scale before spatial stratification.

Because q depends only on the partition of districts into strata, the orientation of cost-type indicators leaves all q-statistics unchanged, and the monotonic logistic transformation leaves quantile classes unchanged; equal-interval classes were constructed on the original values, so γ does not affect any reported statistic. Normalization therefore serves mainly to order strata from unfavorable to favorable for the Risk Detector; SPEI was treated as a benefit-type indicator.

2.4.2. Spatial Discretization

Following the optimal-parameter approach of Song et al. [12], quantile and equal-interval classifications with four classes were compared, and the one with the higher q was retained for each variable. Because this outcome-guided selection can inflate q in small samples, (i) significance was assessed with permutation tests that repeat the selection (Section 2.5.1); (ii) all analyses were repeated with quantile classes (k = 4; about six districts per class); and (iii) k = 3–6 was examined. Ties were assigned to the lower class; class boundaries are reported in Table S4.

$$X_j = \{s_1, s_2, \ldots, s_L\}, \quad (4)$$

In Equation (4), X_j denotes the discretized form of explanatory variable j; s_1, s_2, ..., s_L are the L mutually exclusive spatial strata (classes) assigned to that variable; and L = k is the selected number of discretization classes. The resulting strata constitute the categorical input to GeoDetector.

2.4.3. Analytical Feature Space

After normalization and spatial discretization, the complete dataset was transformed into an analytical feature space:

$$X = \{X_1, X_2, \ldots, X_8\}, \quad (5)$$

In Equation (5), X denotes the complete explanatory feature space and X_1, ..., X_8 are the eight discretized explanatory variables defined in Table 1. The feature space combines the environmental, agricultural-structure, infrastructure and security factors used in the detector analyses.

2.5. GeoDetector Analytical Framework

Four complementary GeoDetector modules were applied: factor, interaction, ecological and risk detectors [11,13].

2.5.1. Factor Detector

The Factor Detector quantifies the explanatory contribution of each individual factor to the observed spatial heterogeneity of agricultural economic development [11,13]. The explanatory power is measured using the q-statistic:

$$q = 1 - \frac{\sum_{h=1}^{L} N_h \sigma_h^2}{N\sigma^2}, \; 0 \le q \le 1, \quad (6)$$

where q is the GeoDetector explanatory-power statistic ($0 \le q \le 1$); $h = 1, \ldots, L$ indexes the spatial strata of explanatory factor X; L is the number of strata; N is the total number of districts; N_h is the number of districts in stratum h; σ^2 is the variance of the dependent variable Y across all districts; and ${\sigma_h}^2$ is the variance of Y within stratum h. Larger q values indicate stronger correspondence between the spatial stratification of X and the spatial heterogeneity of agricultural output Y.

Significance was assessed with the non-central F-test [13] and with a selection-adjusted permutation test: Y was permuted 999 times, and the full selection procedure was repeated for each permutation; the same procedure was applied to joint q-statistics. The eight factor tests form the confirmatory family (RQ1); interaction, ecological and risk comparisons are exploratory. For the factor tests, permutation p-values were recomputed with 9999 permutations and adjusted with the Holm procedure, with Benjamini–Hochberg and Westfall–Young (minP) adjustments as sensitivity analyses (Table S8). Analyses were run

in Python 3.13.18 (NumPy 2.5.3, pandas 3.0.5, SciPy 1.18.1, statsmodels 0.15.0, GeoPandas 1.2.0) with fixed seeds (20261103 and 20261111); the script (File S1) reproduces all reported statistics, and file checksums are provided in Supplementary Materials.

2.5.2. Interaction Detector

The Interaction Detector compares $q(X_a)$, $q(X_b)$ and $q(X_a \cap X_b)$ to determine whether two factors weaken, act independently, or enhance each other. The interaction categories used in this study are reported in Table 2.

Table 2. Interaction types identified by the Interaction Detector.

Interaction Type	Criterion
Nonlinear weakening	$q(X_a \cap X_b) < \min[q(X_a), q(X_b)]$
Uni-factor nonlinear weakening	$\min[q(X_a), q(X_b)] < q(X_a \cap X_b) < \max[q(X_a), q(X_b)]$
Bi-factor enhancement	$\max[q(X_a), q(X_b)] < q(X_a \cap X_b) < q(X_a) + q(X_b)$
Independence	$q(X_a \cap X_b) = q(X_a) + q(X_b)$
Nonlinear enhancement	$q(X_a \cap X_b) > q(X_a) + q(X_b)$

Note: X_a and X_b denote two explanatory factors; $q(X_a)$ and $q(X_b)$ are their individual explanatory powers; $q(X_a \cap X_b)$ is the explanatory power after overlaying their spatial strata. The criteria classify whether the joint effect weakens, is independent, or enhances explanatory power. Source: adapted from the GeoDetector interaction classification framework of Wang et al. [11] and Wang et al. [13]. The criteria are mutually exclusive.

For each pair, the increment over the stronger factor, $\Delta q = q(X_a \cap X_b) - \max[q(X_a), q(X_b)]$, the numbers of non-empty and single-district joint strata and the permutation p-value are reported. Because two four-class stratifications can create up to 16 joint strata for 24 districts, joint q-statistics tend towards one by construction; the interaction types of Table 2 are therefore used only as descriptive labels, and "enhancement" does not imply substantive synergy or a causal interaction.

Because the permutation test of the joint q only shows that the joint stratification is associated with Y, Δq was tested with a conditional permutation test: Y was permuted within the strata of the stronger factor (leaving its q unchanged), the discretization of the added factor was re-selected in each of 9999 permutations, and p-values were Holm-adjusted across the 28 pairs. As a sparse-stratum check, Δq was recomputed with the added factor split at its median (Table S9).

2.5.3. Ecological Detector

The Ecological Detector evaluates whether two explanatory factors differ significantly in their ability to explain the spatial heterogeneity of the dependent variable [11,13]. The comparison is performed using the F-statistic:

$$F = \frac{N_a(N_b - 1)\sum N_{a,h}\sigma^2_{a,h}}{N_b(N_a - 1)\sum N_{b,h}\sigma^2_{b,h}}, \quad (7)$$

where F compares the explanatory capacities of factors X_a and X_b; N_a and N_b denote the total numbers of observations used for the two factor stratifications; $N_{a,h}$ and $N_{b,h}$ are the numbers of observations in stratum h of X_a and X_b, respectively; $\sigma^2_{a,h}$ and $\sigma^2_{b,h}$ are the within-stratum variances of the dependent variable Y under the corresponding factor stratifications; and the summation is taken over all strata h. A statistically significant F-test indicates that the explanatory capacities of the two factors differ beyond random variation. Under the null hypothesis, F follows an F distribution with $(N_a - 1, N_b - 1) = (23, 23)$ degrees of freedom; a two-sided test was used, and F statistics and exact p-values are reported in Section 3.3. The test assumes independent within-stratum sums of squares and approximately normal deviations; because both stratifications use the same districts, the p-values are approximate and were checked with a district bootstrap of the difference in q (2000 resamples; Table S11).

2.5.4. Risk Detector

The Risk Detector identifies statistically significant differences in agricultural development among the spatial strata of individual explanatory variables [11,13]. Differences between strata are evaluated using Welch's *t*-test, which does not assume equal variances:

$$t = \frac{\overline{y}_a - \overline{y}_b}{\sqrt{\frac{\sigma_a^2}{N_a} + \frac{\sigma_b^2}{N_b}}}, \quad (8)$$

where t is the Welch t-statistic for comparing two strata a and b; $\overline{y}_a$ and $\overline{y}_b$ are the mean values of the dependent variable Y in strata a and b; σ_a^2 and σ_b^2 are the corresponding within-stratum variances of Y; and N_a and N_b are the numbers of districts in the two strata. A statistically significant test indicates that mean agricultural output differs between the compared strata. Degrees of freedom were approximated with the Welch–Satterthwaite equation, p-values were adjusted for multiple comparisons with the Holm procedure, and Hedges' g is reported as the effect size. Because several strata contain only one to six districts, these tests have low power and their results are interpreted as indicative.

2.6. Integrated Interpretation Framework

The four modules were interpreted jointly: the Factor Detector ranked individual factors, the Interaction Detector assessed combined effects, the Ecological Detector tested differences in explanatory capacity, and the Risk Detector identified statistically distinct outcome levels across factor strata.

Global Moran's I of Y was computed with row-standardized inverse-distance weights between district centroids (no cut-off) and, alternatively, with first-order queen contiguity weights built from the OCHA COD-AB v05 boundaries joined by P-codes (Table S12), using 999 permutations [3,30]. The test is two-sided (H1: $I \neq E(I) = -1/(N-1)$), with the pseudo p-value $p = \min[1, 2(1+m)/1000]$, where m is the number of permuted values in the smaller tail. Leave-one-district-out robustness was assessed by re-estimating q 24 times, each time omitting one district and re-deriving the classes, for the three highest-ranked factors and the strongest interaction.

2.7. Benchmark Correlation and Regression Analysis

For RQ3, the GeoDetector results were compared with a benchmark estimated on the same data. Spearman correlations show the direction of association, which q does not convey, and test the hypothesized directions in Table 1. Because eight predictors cannot be estimated reliably with N = 24, a parsimonious OLS model with the three leading factors and conflict intensity was estimated with standardized continuous predictors: $Y_i = \beta_0 + \beta_1 X4_i + \beta_2 X6_i + \beta_3 X5_i + \beta_4 X8_i + \varepsilon_i$. Diagnostics comprise variance inflation factors (VIF), the Breusch–Pagan and Shapiro–Wilk tests and Moran's I of residuals, with HC3 standard errors. Because the predictors were selected on the same data, stability was assessed with HC3 and bootstrap intervals, Cook's distance and DFBETAS, leave-one-district-out coefficients and prediction errors (RMSE, MAE and Q^2) and a selection-aware cross-validation (Table S10). Coefficients are interpreted as conditional associations, not causal effects. The model was also estimated for Y_{ha}.

3. Results

3.1. Principal Factors Associated with Spatial Heterogeneity

Crop structure—the maize share of the grain area (X4)—has the highest explanatory power (q = 0.4304), followed by distance to the front line (X6, q = 0.3237) and to the Russian

border (X5, q = 0.2723), whereas vegetation condition, drought, road density and conflict intensity have low q-statistics (Figure 2; Table 3).

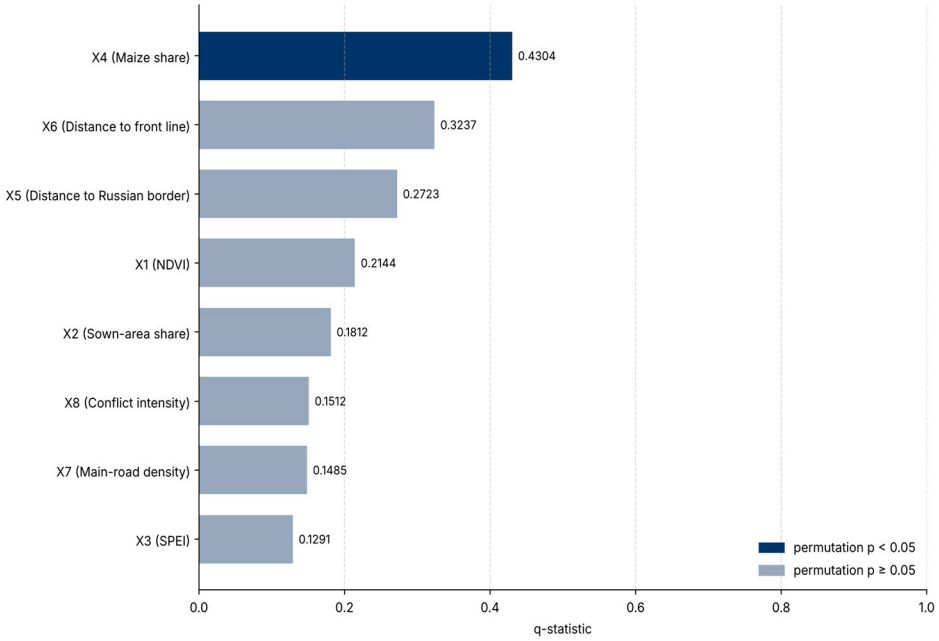


Figure 2. Explanatory power (q-statistics) of individual factors (Factor Detector, Equation (6)); dark bars: selection-adjusted permutation $p < 0.05$. X1–X8 are defined in Table 1. Source: authors' calculations.

Table 3. Factor Detector results for agricultural economic development.

Rank	Domain	Factor	Indicator	q-Statistic	*p* (F-Test)	*p* (Perm.)	q, Quantile k = 4
1	Agricultural structure	X4	Maize share of grain area	0.4304	0.040	0.020	0.4304
2	Security	X6	Distance to the front line	0.3237	0.514	0.096	0.2867
3	Security	X5	Distance to the Russian border	0.2723	0.310	0.151	0.2316
4	Environmental	X1	NDVI	0.2144	0.269	0.270	0.2144
5	Agricultural structure	X2	Sown-area share	0.1812	0.342	0.380	0.1812
6	Security	X8	Conflict intensity	0.1512	0.393	0.519	0.0998
7	Infrastructure	X7	Main-road density	0.1485	0.429	0.518	0.1485
8	Environmental	X3	SPEI	0.1291	0.489	0.605	0.1291

Note: q is the GeoDetector Factor Detector statistic defined in Equation (6). Factors X1–X8 are defined in Table 1. Rank 1 denotes the largest explanatory-power statistic. Source: authors' calculations. *p* (F-test): significance of q based on the non-central F-test [13]; *p* (perm.): selection-adjusted permutation *p*-value (999 permutations, full discretization selection repeated); q, quantile k = 4: q-statistic under the single pre-specified classification rule.

Equal-interval classes were retained for X5, X6 and X8 and quantile classes for the other factors. Under the fixed quantile rule, maize share remains first (q = 0.4304), whereas the q-statistics of X6 and X5 decrease to 0.2867 and 0.2316. Maize share is the only factor significant under both the F-test ($p = 0.040$) and the permutation test ($p = 0.020$). However, after Holm adjustment for the eight tests (9999 permutations), no factor is significant at the 5% level (maize share: Holm $p = 0.113$; Benjamini–Hochberg 0.113; Westfall–Young 0.101; Table S8). Consequently, in line with our conservative approach to multiple comparisons in a small sample, the associations with total output are strictly interpreted as exploratory.

The eight factors produce eight distinct stratifications, and none is a near-duplicate (largest $|\rho| = 0.75$, between NDVI and distance to the Russian border; Table S3); the two distance measures are only weakly correlated ($\rho = 0.26$) and capture exposure in the south-east and north-east, respectively.

3.2. Interactions Among Explanatory Factors

All ten leading pairs are classified as nonlinear enhancement and none as weakening (Figure 3; Table 4); five of the six strongest pairs include distance to the front line.

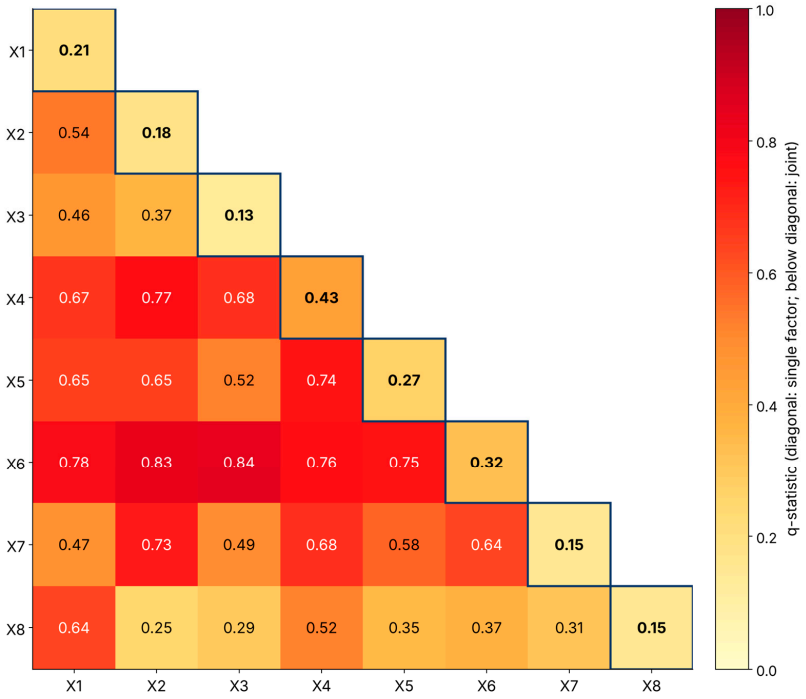


Figure 3. Interaction Detector: joint q-statistics $q(X_a \cap X_b)$ below the diagonal and single-factor q-statistics on the diagonal (outlined). X1–X8 are defined in Table 1. Source: authors' calculations.

Table 4. Top 10 interacting variables and their combined explanatory power.

Rank	Interacting Pair	q(a)	q(b)	q(a ∩ b)	Δq	Joint Strata (Non-Empty/ Single)	*p* (Perm.)	Interaction Type
1	X3 (SPEI) ∩ X6 (Front line)	0.1291	0.3237	0.8390	0.5153	12/5	0.012	Nonlinear enhancement
2	X2 (Sown share) ∩ X6 (Front line)	0.1812	0.3237	0.8350	0.5112	12/6	0.013	Nonlinear enhancement
3	X1 (NDVI) ∩ X6 (Front line)	0.2144	0.3237	0.7751	0.4514	12/3	0.092	Nonlinear enhancement
4	X2 (Sown share) ∩ X4 (Maize share)	0.1812	0.4304	0.7713	0.3409	13/5	0.122	Nonlinear enhancement
5	X4 (Maize share) ∩ X6 (Front line)	0.4304	0.3237	0.7561	0.3258	9/1	0.024	Nonlinear enhancement
6	X5 (Russian border) ∩ X6 (Front line)	0.2723	0.3237	0.7515	0.4277	11/4	0.044	Nonlinear enhancement
7	X4 (Maize share) ∩ X5 (Russian border)	0.4304	0.2723	0.7421	0.3118	9/2	0.029	Nonlinear enhancement
8	X2 (Sown share) ∩ X7 (Roads)	0.1812	0.1485	0.7276	0.5465	14/5	0.101	Nonlinear enhancement
9	X4 (Maize share) ∩ X7 (Roads)	0.4304	0.1485	0.6845	0.2542	14/7	0.189	Nonlinear enhancement
10	X3 (SPEI) ∩ X4 (Maize share)	0.1291	0.4304	0.6792	0.2488	13/6	0.162	Nonlinear enhancement

Note: $\Delta q = q(a \cap b) - \max[q(a), q(b)]$. Joint strata: non-empty strata (maximum 16)/single-district strata. *p* (perm.): selection-adjusted permutation *p*-value of the joint q (conditional tests of Δq in Table S9). Interaction type: descriptive category (Table 2). Source: authors' calculations.

The strongest joint stratification combines SPEI with distance to the front line (q = 0.8390), followed by sown-area share with distance to the front line (q = 0.8350). Because the individual q-statistics are low, the increments over the stronger factor are large (Δq = 0.25–0.55), but they

arise in sparse partitions with 9–14 non-empty strata, 1–7 of them single-district. The joint permutation *p*-values are below 0.05 for five pairs, but they test the joint q, not Δq.

Under the conditional null, the expected Δq of the ten leading pairs is already 0.18–0.35. The observed Δq exceeds it nominally only for SPEI, sown-area share and maize share combined with distance to the front line ($p = 0.014$, 0.048 and 0.033). Crucially, no pair is significant after Holm adjustment across the 28 pairs (smallest adjusted $p = 0.395$; Table S9). With the added factor split at its median, only the increment of distance to the Russian border over maize share is nominally significant ($\Delta q = 0.309$; $p = 0.002$). Therefore, the interactions are interpreted as descriptive joint partitions rather than statistically confirmed synergy.

3.3. Significance of Explanatory Capacities

The Ecological Detector tested whether selected factors differed significantly in explanatory capacity (Table 5).

Table 5. Ecological Detector results: significance, F statistics and exact *p*-values (df = 23, 23).

vs.	X4 (Maize Share)	X6 (Front Line)	X5 (Russian Border)	X1 (NDVI)	X2 (Sown Share)
X4 (Maize share)	—	N; F = 0.84; $p = 0.684$	N; F = 0.78; $p = 0.562$	N; F = 0.73; $p = 0.447$	N; F = 0.70; $p = 0.391$
X6 (Front line)	N; F = 1.19; $p = 0.684$	—	N; F = 0.93; $p = 0.862$	N; F = 0.86; $p = 0.722$	N; F = 0.83; $p = 0.650$
X5 (Russian border)	N; F = 1.28; $p = 0.562$	N; F = 1.08; $p = 0.862$	—	N; F = 0.93; $p = 0.856$	N; F = 0.89; $p = 0.780$
X1 (NDVI)	N; F = 1.38; $p = 0.447$	N; F = 1.16; $p = 0.722$	N; F = 1.08; $p = 0.856$	—	N; F = 0.96; $p = 0.922$
X2 (Sown share)	N; F = 1.44; $p = 0.391$	N; F = 1.21; $p = 0.650$	N; F = 1.13; $p = 0.780$	N; F = 1.04; $p = 0.922$	—

Note: Y indicates a statistically significant difference between the explanatory capacities of two factors at the 95% confidence level; N denotes the absence of a statistically significant difference. Source: authors' calculations. F—Ecological Detector statistic (Equation (7)); *p*—exact two-sided *p*-value.

No pair of the five leading factors differs significantly in explanatory capacity (F-test $p = 0.39$–0.92), and all district-bootstrap intervals of the differences in q include zero (Table S11).

3.4. Differences in Mean Output Across Factor Strata

The Risk Detector compares the mean value of agricultural output across strata of the leading factors and identifies statistically significant differences between adjacent strata (Figure 4; Table 6).

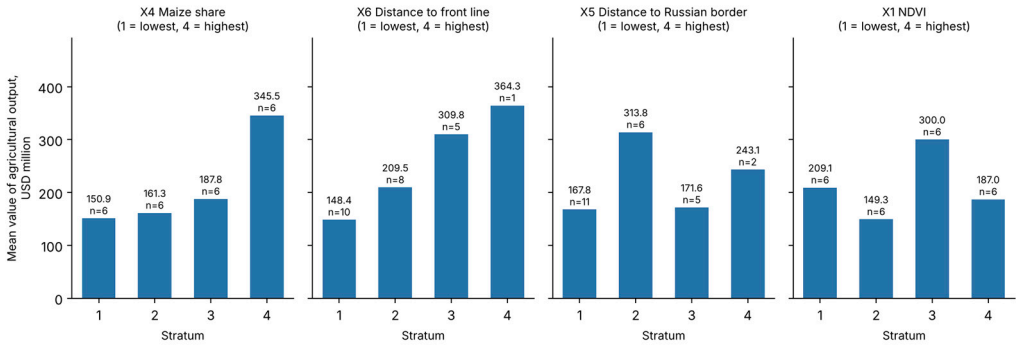


Figure 4. Risk Detector: mean value of agricultural output (USD million) across the k = 4 strata of the four leading factors (1 = lowest, 4 = highest value of the factor). No adjacent difference is significant after Holm adjustment (Table 6). Source: authors' calculations.

Table 6. Risk Detector analysis of dominant explanatory variables across spatial strata.

Variable	Description	Stratum	Mean Output, USD Million	N	Significant Difference	Welch t (df); p_{Holm}; g
X4	Maize share of grain area	1	150.93	6	—	—
		2	161.25	6	No (1 → 2)	t = 0.20 (7.5); p = 1.000; g = 0.11
		3	187.75	6	No (2 → 3)	t = 0.51 (7.8); p = 1.000; g = 0.27
		4	345.49	6	No (3 → 4)	t = 2.53 (6.9); p = 0.119; g = 1.35
X6	Distance to the front line	1	148.37	10	—	—
		2	209.46	8	No (1 → 2)	t = 1.57 (12.0); p = 0.283; g = 0.74
		3	309.79	5	No (2 → 3)	t = 1.13 (5.3); p = 0.305; g = 0.70
		4	364.28	1	n/a (n = 1)	—
X5	Distance to the Russian border	1	167.77	11	—	—
		2	313.81	6	No (1 → 2)	t = 2.06 (6.2); p = 0.252; g = 1.22
		3	171.60	5	No (2 → 3)	t = −1.85 (7.6); p = 0.252; g = −0.97
		4	243.12	2	No (3 → 4)	t = 0.56 (1.2); p = 0.664; g = 0.56
X1	NDVI	1	209.10	6	—	—
		2	149.28	6	No (1 → 2)	t = −1.30 (7.0); p = 0.441; g = −0.69
		3	300.01	6	No (2 → 3)	t = 1.87 (5.6); p = 0.341; g = 1.00
		4	187.03	6	No (3 → 4)	t = −1.36 (6.3); p = 0.441; g = −0.72

Note: Yes/No—adjacent-stratum difference significant/not significant (Welch *t*-test, Holm-adjusted $p < 0.05$); n/a—not applicable (Welch's t-test cannot be computed for a single-district stratum, n = 1); g—Hedges' g. Strata are ordered by the value of the factor (1 = lowest). USD converted at UAH 41.69 per USD. Source: authors' calculations.

Mean output rises with maize share (from USD 150.9 to 345.5 million) and with distance to the front line (from USD 148.4 million in the ten districts closest to it to USD 309.8 million at 144–188 km; the most distant stratum contains one district), whereas NDVI and distance to the Russian border show non-monotonic patterns. Some adjacent differences are large (Hedges' g = 1.35 for maize-share strata 3–4 and 1.22 for border-distance strata 1–2). However, none is significant after Holm adjustment ($p \geq 0.12$). Because class boundaries result from discretization, they are not economic thresholds.

3.5. Robustness and Validity Checks

The stability of the findings was assessed with alternative class numbers, Global Moran's I, leave-one-district-out resampling and secondary outcomes. First, q was recomputed for k = 3–6 under both schemes, retaining the higher value for each k (Table 7, Figure 5).

Maize share ranks first for k = 4–6 (q = 0.4284–0.4887), whereas for k = 3 distance to the Russian border has a slightly higher q (0.3456 versus 0.3066), so the ranking of the leading factors depends partly on the classification. With k = 6, classes contain about four

districts, and the general increase in q (e.g., for sown-area share from 0.1812 to 0.4099) reflects overfitting rather than stronger associations.

Table 7. Sensitivity of selected q-statistics to the number of discretization classes.

Variable	Quantile k = 4	Equal k = 4	k = 3 (Best)	k = 5 (Best)	k = 6 (Best)
X4 Maize share	0.4304	0.2889	0.3066	0.4284	0.4887
X6 Front line	0.2867	0.3237	0.2903	0.3204	0.4328
X5 Russian border	0.2316	0.2723	0.3456	0.3416	0.4055
X1 NDVI	0.2144	0.0294	0.0325	0.1911	0.2515
X8 Conflict	0.0998	0.1512	0.1480	0.2378	0.2481
X3 SPEI	0.1291	0.0449	0.0471	0.3181	0.1591

Note: k—number of discretization classes. The Quantile k = 4 and Equal k = 4 columns show both candidate specifications. For k = 3, 5 and 6, the table reports the higher q-statistic from the quantile and equal-interval alternatives. q is defined in Equation (6). Source: authors' calculations.

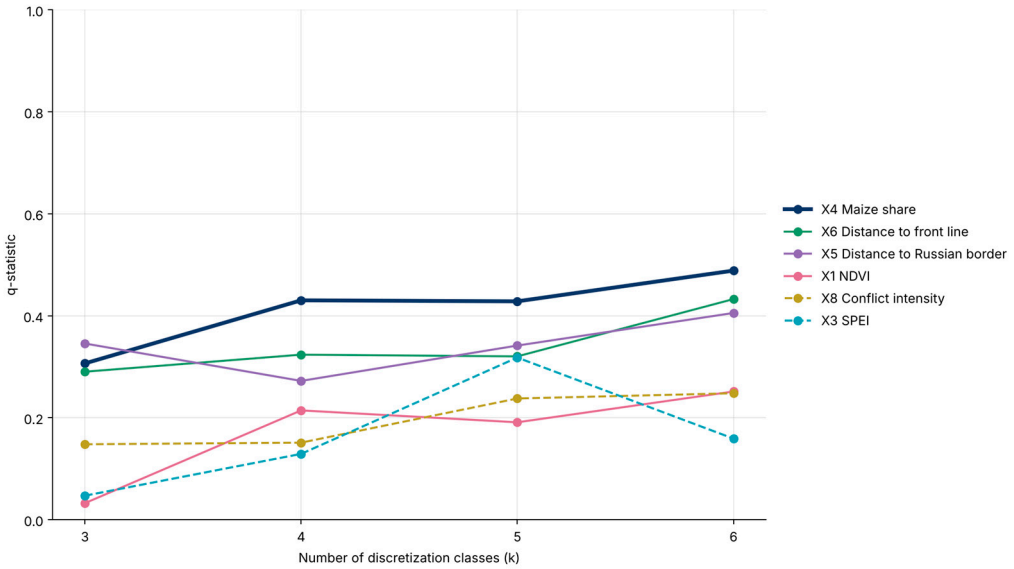


Figure 5. Sensitivity of q-statistics to the number of discretization classes (k = 3–6); for each k, the higher q of the quantile and equal-interval schemes is shown. X4—maize share; X6—distance to the front line; X5—distance to the Russian border; X1—NDVI; X8—conflict intensity; X3—SPEI. Source: authors' calculations.

Global Moran's I of output is 0.104 with row-standardized inverse-distance weights (two-sided $p = 0.016$), 0.098 without standardization ($p = 0.016$) and 0.322 with queen contiguity ($p = 0.024$), compared with $E(I) = -0.043$. Output is thus positively spatially autocorrelated: high-output districts are concentrated in Chernihiv and Sumy oblasts and lower-output districts in Kharkiv oblast. Stratified heterogeneity therefore coexists with spatial dependence, under which the permutation tests are approximate; the benchmark residuals show no autocorrelation (Section 3.6).

Leave-one-district-out q-statistics range from 0.260 to 0.535 for maize share (mean 0.406), 0.242–0.451 for distance to the front line, 0.183–0.376 for distance to the Russian border and 0.632–0.864 for X3 ∩ X6 (Table S6); maize share thus depends noticeably on individual districts but never falls below 0.26.

Per hectare of sown area, maize share again ranks first (q = 0.3262; $p = 0.060$), followed by distance to the front line (0.2592) and NDVI (0.1919), with no factor significant after Holm adjustment. For crop output only, the association with maize share is stronger (q = 0.6101; $p = 0.003$; Holm $p = 0.012$), followed by distance to the Russian border (0.3877) and to the front line (0.3341) (Table S7). Crop structure thus ranks first under all three definitions but

is robust to multiplicity adjustment only for crop output, where the association is partly mechanical because maize yields more per hectare than other grains.

3.6. Comparison with a Conventional Regression Benchmark

Table 8 compares the GeoDetector results with Spearman rank correlations and the parsimonious OLS benchmark described in Section 2.7.

Table 8. GeoDetector q-statistics compared with Spearman correlations and the OLS benchmark.

Rank	Factor	Indicator	Hyp. Dir.	q (Geo Detector)	Spearman ρ (p)	OLS std. β (HC3 p)
1	X4	Maize share of grain area	+	0.4304	0.548 (0.006)	0.957 (<0.001)
2	X6	Distance to the front line	+	0.3237	0.451 (0.027)	0.087 (0.636)
3	X5	Distance to the Russian border	+	0.2723	0.220 (0.302)	0.739 (0.008)
4	X1	NDVI	+	0.2144	0.035 (0.872)	—
5	X2	Sown-area share	+	0.1812	0.507 (0.011)	—
6	X8	Conflict intensity	−	0.1512	−0.377 (0.069)	0.156 (0.283)
7	X7	Main-road density	+	0.1485	0.050 (0.815)	—
8	X3	SPEI	±	0.1291	−0.002 (0.994)	—
	OLS model fit	R^2/adjusted R^2				0.650/0.576
		Maximum VIF				2.88
		Breusch–Pagan test, p				0.206
		Shapiro–Wilk test of residuals, p				0.831
		Moran's I of residuals (p)				−0.085 (0.290)

Note: Hyp. dir.—hypothesized direction from Table 1; q—Factor Detector statistic (Table 3); Spearman ρ—rank correlation with Y (two-sided p in parentheses); OLS std. β—standardized coefficient of the benchmark model $Y = \beta_0 + \beta_1 X4 + \beta_2 X6 + \beta_3 X5 + \beta_4 X8 + \varepsilon$, with HC3-robust p-values in parentheses; — = not included in the OLS model. Source: authors' calculations.

Both approaches identify maize share as most strongly associated with output ($\rho = 0.548$; standardized $\beta = 0.957$, $p < 0.001$). The OLS model explains $R^2 = 0.650$ (adjusted 0.576), with well-behaved residuals (Table 8). Sown-area share ($\rho = 0.507$) and distance to the front line ($\rho = 0.451$) are also positively rank-correlated with output, conflict intensity negatively but not significantly ($\rho = -0.377$, $p = 0.069$), and NDVI, SPEI and road density not at all ($|\rho| \leq 0.05$); the hypothesized directions are confirmed for X2, X4 and X6. The approaches differ in three respects. First, distance to the Russian border has no monotonic association ($\rho = 0.220$) but a significant conditional coefficient ($\beta = 0.739$, $p = 0.008$), because districts near the border have higher maize shares ($\rho = -0.47$). Second, distance to the front line is not significant conditional on the other predictors ($\beta = 0.087$, $p = 0.636$), as its information overlaps with that of conflict intensity and crop structure. Third, GeoDetector describes joint stratifications for all 28 pairs (which a regression with N = 24 cannot estimate) and stratum-specific differences that a linear slope averages out. Per hectare, the model ($R^2 = 0.431$) again identifies maize share as dominant ($\beta = 0.898$, $p < 0.001$); its residuals show no spatial autocorrelation ($I = -0.106$, $p = 0.060$) but some heteroskedasticity and non-normality ($p = 0.030$ and 0.046), partly addressed by the HC3 errors.

The coefficients of maize share and distance to the Russian border keep their sign in all 24 leave-one-district-out fits, and their HC3 and bootstrap intervals exclude zero, whereas those of distance to the front line and conflict intensity include zero (Table S10). Two districts exceed Cook's distance 4/N (maximum 0.33). The out-of-sample Q^2 is 0.452 (RMSE USD 88.7 million), compared with 0.172 for maize share alone and 0.369 when predictors are re-selected within each fold, indicating moderate predictive power; the coefficients describe conditional associations, not causal policy effects.

4. Discussion

The results indicate a moderate hierarchy: crop structure has the strongest individual association with output, followed by proximity to the front line and to the Russian border,

whereas vegetation condition, drought, road density and conflict intensity are weakly associated. This is consistent with regional competitiveness and New Economic Geography, which emphasize specialization, accessibility and local capacity [5–7].

The leading role of maize share points to production specialization as a key differentiating condition. High-maize districts, mainly in Chernihiv and Sumy oblasts, combine high-yielding grain with large sown areas, whereas Dnipropetrovsk districts specialize in lower-yielding wheat and sunflower. The association is partly mechanical, sensitive to individual districts and robust to multiplicity adjustment only for crop output. Crop structure itself reflects soils and climate, the capital and drying and storage capacity that maize requires and, since 2022, operating risks near the front line; the cross-sectional design cannot separate these channels or exclude reverse causality. The finding complements research linking agricultural competitiveness to specialization, infrastructure and innovation capacity [5].

NDVI ($q = 0.2144$) and SPEI ($q = 0.1291$) show no monotonic association with output. In 2025, vegetation was greenest in the wetter north and weakest in the drought-affected south of Dnipropetrovsk oblast (district SPEI −1.37 to 0.48), but these gradients are only weakly aligned with output value, which depends more on crop structure and district size; favorable biophysical conditions alone do not explain output differences [2,31].

The interaction results are descriptive. GeoDetector can describe joint partitions without a linear interaction specification [11–13], but joint stratifications raise q by construction, and with 24 districts and up to 16 joint strata none of the increments remains significant after adjustment, so they indicate joint partitioning rather than synergy.

Districts farther from the front line and the Russian border tend to have higher output, whereas the number of recorded conflict events is only weakly associated with it. Proximity may act through land withdrawn from use, operating risks and lower investment, but the distance measures also capture the long-standing north–south gradient in specialization, and conflict events are concentrated in a few districts (Kupianskyi, Kharkivskyi and Chuhuivskyi). This is compatible with resilience research on the interaction between external shocks and pre-existing territorial capacities [32].

Combining official statistics, remote sensing and conflict-event data within GIS allows these relationships to be represented in one spatial framework; GeoDetector and regression are complementary, the former describing stratified and joint differences without a functional form and the latter providing direction and effect size [10,13,33].

The leading role of specialization is consistent with the competitiveness literature [5–7] and with regional evidence of pronounced heterogeneity in agricultural and livestock production systems [8]. Satellite studies document war-induced losses of crop area and yields in Ukraine, larger on small farms [34], effects on wheat biomass [35] and conflict-related losses affecting nearly half of productive cropland after 2022 [36]. In the present district cross-section, proximity to hostilities is more closely associated with output than event counts, suggesting that war effects operate through exposure and risk at finer scales than district counts capture. As in other GeoDetector applications, a few factors dominate and interactions appear enhancing [11,13], but q-statistics are sensitive to discretization in small samples, which the optimal-parameter approach addresses only partly [12]. The robust geographical detector [37] could reduce this sensitivity in future work.

These findings must be interpreted in light of specific methodological and data constraints. With a sample of 24 districts, statistical power is inherently limited. Fine and joint stratifications can inflate q-statistics, and because multiple tests were performed on the same data, conservative adjustments render the associations with total output and interaction increments exploratory. Furthermore, GeoDetector identifies spatial associations rather than causal mechanisms. The cross-sectional design based on a single, dry year (2025)

precludes establishing the direction of the observed relationships or capturing longitudinal dynamics, particularly since complete multi-year series for 2022–2024 remain unpublished. Additionally, the focus on four frontline oblasts means the results may not be representative of the agricultural situation in western or central Ukraine. The analysis is also subject to the modifiable areal unit problem (MAUP), relies solely on enterprise statistics excluding occupied territories, and is restricted by confidentiality to grain, sunflower, and livestock output. Finally, output is spatially autocorrelated, so the permutation tests are approximate; spatial lag, spatial error or geographically weighted regression models would require more spatial units [38].

5. Conclusions

This study applied a GIS-GeoDetector framework to 24 districts in four Ukrainian oblasts using eight environmental, agricultural-structure, infrastructure and security variables compiled from published district statistics and open geospatial data.

For RQ1, the maize share of the grain area is most strongly associated with agricultural output ($q = 0.4304$; 0.6101 for crop output), followed by distance to the front line (0.3237) and to the Russian border (0.2723). Importantly, after multiplicity adjustment, only the association with crop output remains statistically significant. For RQ2, joint stratifications produce higher q-statistics (up to 0.8390 for SPEI and distance to the front line); however, no interaction increment is significant after adjustment, indicating that they describe joint spatial partitions rather than confirmed synergy. The ecological and risk detectors show no significant differences among the leading factors or between adjacent strata.

For RQ3, the regression benchmark confirms the leading role of maize share and reveals a conditional association with distance to the Russian border, whereas GeoDetector adds stratum-specific and joint information that a parsimonious linear model cannot capture with $N = 24$.

The leading position of maize share holds for $k = 4$–6 and for all three outcome definitions, and output shows positive spatial autocorrelation.

Subject to longitudinal confirmation, recovery support should take into account the crop specialization of districts and their exposure to hostilities rather than being uniform. Because the framework uses only published and open data, it can be updated as district statistics are released and supports spatially targeted recovery that addresses the economic, social and environmental dimensions of sustainable agricultural development [39–41].

The results describe a single-year cross-section of 24 districts, not causal dynamics for all of Ukraine. Future work can use multi-year district series as publication resumes, community- or farm-level data and indicators of investment and war damage.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/su182211753/s1, Table S1: Names, P-codes and areas of the 24 districts; Table S2: Descriptive statistics; Table S3: Spearman correlation matrix of the explanatory variables; Table S4: Class boundaries and stratum counts; Table S5: Interaction Detector results for all 28 factor pairs; Table S6: Leave-one-district-out estimates; Table S7: Factor Detector results for secondary outcomes; Table S8: Multiple-comparison sensitivity of the Factor Detector; Table S9: Conditional permutation tests of incremental joint explanatory power (Δq) and sparse-stratum sensitivity; Table S10: Stability of the OLS benchmark; Table S11: Ecological Detector F-tests and district-bootstrap intervals; Table S12: District centroids and queen-contiguity neighbors; Table S13: Data provenance of every published value used; File S1: Analysis script and district-level dataset; File S2: District boundary files (OCHA COD-AB v05).

Author Contributions: Conceptualization, O.P., K.P., S.Z., A.M., V.D., O.S., O.Y., K.B. and A.Y.; methodology, O.P., K.P., S.Z., A.M., V.D., O.S., O.Y., K.B. and A.Y.; software, O.P., K.P., S.Z., A.M., V.D., O.S., O.Y., K.B. and A.Y.; validation, O.P., K.P., S.Z., A.M., V.D., O.S., O.Y., K.B. and A.Y.; formal analysis, O.P., K.P., S.Z., A.M., V.D., O.S., O.Y., K.B. and A.Y.; investigation, O.P., K.P., S.Z., A.M., V.D., O.S., O.Y., K.B. and A.Y.; resources, O.P., K.P., S.Z., A.M., V.D., O.S., O.Y., K.B. and A.Y.; data curation, O.P., K.P., S.Z., A.M., V.D., O.S., O.Y., K.B. and A.Y.; writing—original draft preparation, O.P., K.P., S.Z., A.M., V.D., O.S., O.Y., K.B. and A.Y.; writing—review and editing, O.P., K.P., S.Z., A.M., V.D., O.S., O.Y. and A.Y.; visualization, O.P., K.P., S.Z., A.M., V.D., O.S., O.Y., K.B. and A.Y.; supervision, K.P. and O.P.; project administration, K.P. and O.P.; funding acquisition, O.P., S.Z. and A.M. All authors have read and agreed to the published version of the manuscript.

Funding: Research is partly supported by the program "Excellence Initiative—Research University" for the AGH University of Krakow, Poland, as well as through contributions from the authors.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The analysis dataset (Y and X1–X8 with P-codes), the source of every published value (Table S13), the district boundary files (OCHA COD-AB v05, File S2), the analysis script, software versions, seeds and checksums are provided as Supplementary Materials. All underlying data are openly available from the regional statistics offices of the four oblasts, OCHA COD-AB (Humanitarian Data Exchange), MODIS MOD13Q1 (Microsoft Planetary Computer), ERA5 (Open-Meteo), OpenStreetMap (HOT export on the Humanitarian Data Exchange), the DeepState map repository and UCDP GED v26.1.

Conflicts of Interest: Author Viktor Diachenko was employed by FLUGART Limited Liability Company. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

GIS	Geographic Information System
NDVI	Normalized Difference Vegetation Index
SPEI	Standardized Precipitation-Evapotranspiration Index
USD	United States dollar
UCDP	Uppsala Conflict Data Program
GED	Georeferenced Event Dataset
OCHA	United Nations Office for the Coordination of Humanitarian Affairs
HDX	Humanitarian Data Exchange
MAUP	Modifiable Areal Unit Problem
OLS	Ordinary Least Squares
VIF	Variance Inflation Factor
HC3	Heteroskedasticity-consistent covariance estimator (type 3)
MODIS	Moderate Resolution Imaging Spectroradiometer
ERA5	Fifth-generation ECMWF atmospheric reanalysis
ECMWF	European Centre for Medium-Range Weather Forecasts
OSM	OpenStreetMap

References

1. Fischer, G.; Shah, M.; van Velthuizen, H. *Climate Change and Agricultural Vulnerability*; International Institute for Applied Systems Analysis (IIASA): Laxenburg, Austria, 2002. Available online: https://pure.iiasa.ac.at/id/eprint/6670/ (accessed on 1 January 2026).
2. Tittonell, P. Ecological intensification of agriculture—Sustainable by nature. *Curr. Opin. Environ. Sustain.* **2014**, *8*, 53–61. [CrossRef]
3. Anselin, L. *Spatial Econometrics: Methods and Models*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1988. [CrossRef]
4. Tobler, W.R. A computer movie simulating urban growth in the Detroit region. *Econ. Geogr.* **1970**, *46*, 234–240. [CrossRef] [PubMed]

5. Porter, M.E. *The Competitive Advantage of Nations*; Free Press: New York, NY, USA, 1990.
6. Krugman, P. *Geography and Trade*; MIT Press: Cambridge, MA, USA, 1991.
7. Fujita, M.; Krugman, P.; Venables, A.J. *The Spatial Economy: Cities, Regions and International Trade*; MIT Press: Cambridge, MA, USA, 1999. [CrossRef]
8. Sala, D.; Pavlova, O.; Pavlov, K.; Korniietskyi, O.; Orel, V.; Orel, A.; Jammal, M.; Pyzalski, M. Bayesian Tree–Spline Modeling of Circular Bioeconomy Potential in Regional Agricultural and Livestock Production Systems. *Sustainability* **2026**, *18*, 7747. [CrossRef]
9. Longley, P.A.; Goodchild, M.F.; Maguire, D.J.; Rhind, D.W. *Geographic Information Science and Systems*, 4th ed.; Wiley: Hoboken, NJ, USA, 2015.
10. Goodchild, M.F. Geographical information science. *Int. J. Geogr. Inf. Syst.* **1992**, *6*, 31–45. [CrossRef]
11. Wang, J.-F.; Li, X.-H.; Christakos, G.; Liao, Y.-L.; Zhang, T.; Gu, X.; Zheng, X.-Y. Geographical detectors-based health risk assessment and its application in the neural tube defects study of the Heshun region, China. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 107–127. [CrossRef]
12. Song, Y.; Wang, J.; Ge, Y.; Xu, C. An optimal parameters-based geographical detector model enhances geographic characteristics of explanatory variables for spatial heterogeneity analysis. *GISci. Remote Sens.* **2020**, *57*, 593–610. [CrossRef]
13. Wang, J.-F.; Zhang, T.-L.; Fu, B.-J. A measure of spatial stratified heterogeneity. *Ecol. Indic.* **2016**, *67*, 250–256. [CrossRef]
14. Goodchild, M.F. Twenty years of progress: GIScience in 2010. *J. Spat. Inf. Sci.* **2010**, *1*, 3–20. [CrossRef]
15. Chernihiv Statistics Office (Main Department of Statistics in Chernihiv Oblast). Statistical Information on Agriculture by Districts of Chernihiv Oblast. Available online: https://www.chernigivstat.gov.ua/statdani/S_hos/ (accessed on 11 November 2026).
16. Sumy Statistics Office (Main Department of Statistics in Sumy Oblast). Statistical Information on Agriculture by Districts of Sumy Oblast. Available online: https://sumy.ukrstat.gov.ua/?menu=629&level=3 (accessed on 11 November 2026).
17. Kharkiv Statistics Office (Main Department of Statistics in Kharkiv Oblast). Statistical Information on Agriculture by Districts of Kharkiv Oblast. Available online: https://kh.ukrstat.gov.ua/silske-lisove-ta-rybne-hospodarstvo-stat (accessed on 11 November 2026).
18. Dnipropetrovsk Statistics Office (Main Department of Statistics in Dnipropetrovsk Oblast). Statistical Information on Agriculture by Districts of Dnipropetrovsk Oblast. Available online: http://www.dneprstat.gov.ua/statinfo/sg/ (accessed on 11 November 2026).
19. OCHA (United Nations Office for the Coordination of Humanitarian Affairs). Ukraine—Subnational Administrative Boundaries (COD-AB), Version 05. Humanitarian Data Exchange. Available online: https://data.humdata.org/dataset/cod-ab-ukr (accessed on 11 November 2026).
20. Didan, K. *MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V061*; NASA EOSDIS Land Processes Distributed Active Archive Center: Sioux Falls, SD, USA, 2021. [CrossRef]
21. Hersbach, H.; Bell, B.; Berrisford, P.; Hirahara, S.; Horányi, A.; Muñoz-Sabater, J.; Nicolas, J.; Peubey, C.; Radu, R.; Schepers, D.; et al. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **2020**, *146*, 1999–2049. [CrossRef]
22. Zippenfenig, P. *Open-Meteo.com Weather API*, version 1.4.0; Zenodo: Geneva, Switzerland, 2023. [CrossRef]
23. Vicente-Serrano, S.M.; Beguería, S.; López-Moreno, J.I. A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index. *J. Clim.* **2010**, *23*, 1696–1718. [CrossRef]
24. Natural Earth. Admin 0—Countries, 1:50m Cultural Vectors. Available online: https://www.naturalearthdata.com/downloads/50m-cultural-vectors/ (accessed on 11 November 2026).
25. DeepStateMap. Daily Polygons of the Occupied Territory of Ukraine (Deepstate-Map-Data Repository). Available online: https://github.com/cyterat/deepstate-map-data (accessed on 11 November 2026).
26. OpenStreetMap Contributors. OpenStreetMap Data. OpenStreetMap Foundation. Available online: https://www.openstreetmap.org (accessed on 15 September 2026).
27. Humanitarian OpenStreetMap Team. Ukraine Roads (OpenStreetMap Export). Humanitarian Data Exchange. Available online: https://data.humdata.org/dataset/hotosm_ukr_roads (accessed on 11 November 2026).
28. Burrough, P.A.; McDonnell, R.A.; Lloyd, C.D. *Principles of Geographical Information Systems*, 3rd ed.; Oxford University Press: Oxford, UK, 2015.
29. Sundberg, R.; Melander, E. Introducing the UCDP Georeferenced Event Dataset. *J. Peace Res.* **2013**, *50*, 523–532. [CrossRef]
30. Anselin, L. Local indicators of spatial association—LISA. *Geogr. Anal.* **1995**, *27*, 93–115. [CrossRef]
31. Ashworth, A.J.; Kharel, T.; Sauer, T.J. Spatial monitoring technologies for coupling the soil–plant–water–animal nexus. *Sci. Rep.* **2022**, *12*, 3508. [CrossRef] [PubMed]
32. Wilson, G.A. Community resilience, globalization, and transitional pathways of decision-making. *J. Rural Stud.* **2012**, *28*, 458–468. [CrossRef]
33. Wang, J.-F.; Hu, Y. Environmental health risk detection with GeogDetector. *Environ. Model. Softw.* **2012**, *33*, 114–115. [CrossRef]
34. Deininger, K.; Ali, D.A.; Kussul, N.; Shelestov, A.; Lemoine, G.; Yailimova, H. Quantifying war-induced crop losses in Ukraine in near real time to strengthen local and global food security. *Food Policy* **2023**, *115*, 102418. [CrossRef]

35. Antonenko, V.; Al Bitar, A.; Danylenko, I.; Wijmer, T.; Colin, J.; Dejoux, J.-F.; Lefebvre, A.; Knibbe, M.; Ceschia, E.; Gascoin, S. Impact of the Russian invasion on wheat biomass in Ukraine. *Environ. Res. Lett.* **2024**, *19*, 124027. [CrossRef]
36. Li, Y.; Yao, K.; Meng, Q.; Wang, Y.; Xiao, R.; Liu, Y.; Wu, S.; Li, Y. Dynamic patterns and driving factors of productive cropland in Ukraine before and after Russia–Ukraine conflict. *Geogr. Sustain.* **2026**, *7*, 100401. [CrossRef]
37. Zhang, Z.; Song, Y.; Wu, P. Robust geographical detector. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *109*, 102782. [CrossRef]
38. Elhorst, J.P. *Spatial Econometrics: From Cross-Sectional Data to Spatial Panels*; SpringerBriefs in Regional Science; Springer: Berlin/Heidelberg, Germany, 2014. [CrossRef]
39. Wang, J.; Tong, J.; Fang, Z. Assessing the Drivers of Sustained Agricultural Economic Development in China: Agricultural Productivity and Poverty Reduction Efficiency. *Sustainability* **2024**, *16*, 2073. [CrossRef]
40. Lozano Paredes, D.L.; Okada Saavedra, H.; Moscoso Cuaresma, J.R.; Azabache Moran, C.A.; Yesquén Delgado, K.N.d.P.; Diaz Cruz, M.E.; Salazar Seminario, V.L.; Pastor Pinto, J.; Amer Layseca, T. Fairtrade in Peru: Challenges and Opportunities for Sustainable and Equitable Agricultural Development. *Sustainability* **2025**, *17*, 486. [CrossRef]
41. Eliseu, E.E.; Lima, T.M.; Gaspar, P.D. Sustainable Development Strategies and Good Agricultural Practices for Enhancing Agricultural Productivity: Insights and Applicability in Developing Contexts—The Case of Angola. *Sustainability* **2024**, *16*, 9878. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.