

Article

Statistically Derived Marginal Contribution Thresholds and Key Drivers of Sustainable Agricultural Development in Yunnan, China, Under Multidimensional Constraints

Zhenli Wang ^{1,*} and Longfei Ren ²¹ Business School, Yunnan University of Finance and Economics, Kunming 650221, China² Fuzhou International Joint Institute, Tianjin University, Tianjin 300072, China

* Correspondence: zz2228@ynufe.edu.cn

Abstract

Sustainable agricultural development requires regional agricultural systems to balance output growth, resource efficiency, ecological protection, and long-term resilience. In mountainous and ecologically sensitive regions, identifying the development constraints and statistically derived marginal contribution thresholds of agriculture is essential for promoting green transformation and sustainable land use. Taking Yunnan Province, China, as a representative plateau mountainous agricultural region, this study uses provincial annual data from 1990 to 2023 to quantitatively identify the key drivers and threshold characteristics of agricultural development under multidimensional constraints. A multidimensional indicator system was constructed covering fiscal and investment support, agricultural production inputs, rural infrastructure, and labor and population conditions. Ridge regression was employed to address multicollinearity among explanatory variables, Bootstrap approximate inference was used to improve the robustness of coefficient estimation, and the SHAP interpretation framework was introduced to rank key driving factors and identify marginal contribution thresholds. By integrating ridge regression, Bootstrap approximate inference, SHAP-based contribution ranking, and threshold identification, the proposed framework advances prior agricultural sustainability studies by linking coefficient-based factor analysis with interpretable marginal contribution thresholds under conditions of high multicollinearity and multidimensional resource constraints. The results show that agricultural development in Yunnan is characterized by multidimensional resource and infrastructure constraints. Rural electricity consumption, total reservoir storage capacity, fixed asset investment in agriculture, forestry, animal husbandry and fisheries, local public fiscal budget expenditure, and agricultural population generally act as positive supporting factors. Rural electricity consumption is the most stable and core driver across the aggregate and three sectoral models. In contrast, pesticide and fertilizer inputs show significant negative associations in most models, suggesting that future agricultural development in Yunnan is unlikely to be sustainably supported by continued expansion of high-intensity chemical inputs. Sectoral heterogeneity is also evident: agriculture and animal husbandry are more dependent on energy, water resources, and mechanization, whereas forestry shows a more distinct operational structure. The SHAP dependence analysis identifies several statistically derived marginal contribution thresholds, including rural electricity consumption of approximately 6.055 billion kWh, total reservoir storage capacity of approximately 10.395 billion m³, total agricultural machinery power of approximately 19.8324 million kW, pesticide use of approximately 37,500 tons, and fertilizer application of approximately 1.5238 million tons. These values should be interpreted as empirical transition points in the modeled marginal contributions rather than definitive biophysical



Academic Editor: Seung-Hoon Yoo

Received: 30 May 2026

Revised: 27 June 2026

Accepted: 1 July 2026

Published: 4 July 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

ecological limits. They indicate that the sustainability-related constraint structure of agricultural development in Yunnan is not a single output ceiling but a composite interval shaped by infrastructure support capacity, factor allocation conditions, and the declining marginal contribution of high-intensity chemical inputs. The findings provide directional quantitative evidence for sustainable agricultural governance, agricultural green transformation, and differentiated policy discussion in mountainous agricultural regions and offer reference implications for advancing SDG 2 and SDG 15 through the coordination of food-related production, resource use efficiency, and ecosystem conservation. The identified thresholds should be interpreted as model-derived marginal contribution transition points rather than operational policy cutoffs or directly enforceable ecological standards.

Keywords: sustainable agricultural development; statistically derived marginal contribution thresholds; agricultural green transformation; resource–environment constraints; ridge regression; SHAP

1. Introduction

Sustainable agricultural development is a central issue in sustainability science because it requires agricultural systems to simultaneously support food-related production, resource efficiency, ecological protection, and rural socioeconomic development [1]. In the context of climate change, land degradation, water scarcity, biodiversity loss, and increasing pressure on rural transformation, agricultural growth can no longer be understood only as the expansion of production factors such as land, labor, capital, and technology [2]. Instead, regional agricultural systems must operate within resource–environment carrying capacities and statistically derived marginal contribution thresholds while maintaining long-term resilience. This issue is closely related to the United Nations Sustainable Development Goals, particularly SDG 2, which emphasizes food security and sustainable agriculture, and SDG 15, which focuses on terrestrial ecosystem protection, sustainable land use, and the prevention of land degradation. However, applying sustainability boundary concepts to regional agriculture requires distinguishing between ecological limits in a strict biophysical sense and empirical thresholds estimated from socioeconomic and production data. This distinction is important because statistical thresholds can provide management reference points, but they should not be interpreted as direct measurements of ecological tipping points.

Traditional studies of agricultural growth have often explained output expansion through factor inputs, technological progress, capital accumulation, and institutional improvement [3]. These perspectives are important for understanding the sources of agricultural productivity, but they are not sufficient to explain the structural constraints and sustainability boundaries faced by contemporary agricultural systems. With the intensification of resource and environmental pressures, the key question of agricultural development is shifting from “how to increase output” to “how to achieve stable growth within ecological and resource limits.” [4] In particular, excessive dependence on chemical inputs, inefficient use of water and energy, inadequate infrastructure, and mismatched factor allocation may reduce the long-term sustainability of agriculture, even if they contribute to short-term output growth [5]. Therefore, identifying the limits and thresholds of agricultural development is not only a technical issue in regional agricultural planning but also an important scientific issue in the study of sustainable agricultural transformation.

The concept of sustainability-related development constraints has strong theoretical relevance to sustainability research. Rockström et al. proposed the idea of a “safe operating

space,” emphasizing that human development should remain within the boundaries that Earth systems can support [6]. Steffen et al. further developed the planetary boundaries framework and highlighted the systemic risks associated with exceeding environmental limits [7]. In the agricultural field, Tilman et al. and Foley et al. showed that although agricultural intensification can increase production, it may also lead to unsustainable outcomes if resource consumption and ecological costs are ignored [8,9]. These studies suggest that agricultural development should be examined not only from the perspective of productivity improvement, but also from the perspective of statistically derived marginal contribution thresholds, resource constraints, and long-term system resilience [10]. In this study, the “limit” of agricultural development does not refer to a fixed output ceiling. Rather, it refers to a sustainability-related boundary formed by infrastructure carrying capacity, factor allocation thresholds, and the potential upper limits of chemical input-related marginal contributions. However, the application of planetary boundaries and ecological threshold concepts to regional agricultural systems requires careful operationalization. Global boundary frameworks identify broad Earth-system processes, whereas regional agricultural studies usually observe proxy indicators such as fertilizer application, pesticide use, water storage capacity, energy consumption, mechanization, and land use intensity. Therefore, this study does not directly measure planetary boundaries at the provincial scale. Instead, it translates the boundary perspective into an empirical framework that examines whether key agricultural inputs and supporting conditions exhibit marginal contribution turning points within a long-term regional dataset. In this sense, the identified thresholds are statistical and management-relevant reference points, not direct estimates of global ecological boundaries.

Yunnan Province provides a representative case for examining these issues. Yunnan is especially relevant because its agricultural development is shaped by the interaction between mountainous agricultural conditions and strong ecological conservation responsibilities [11]. Recent provincial forestry statistics show that Yunnan’s forest coverage rate has reached 57.87%, with a forest area of approximately 333 million mu, forest stock volume of approximately 2.374 billion m³, and standing tree stock volume of approximately 2.681 billion m³ [12]. These indicators demonstrate that Yunnan is not only an important agricultural province in southwestern China but also an ecologically significant region where agricultural development, forest conservation, biodiversity protection, and land use sustainability are closely intertwined [13]. At the same time, mountainous and plateau areas account for the dominant share of the province’s landform structure, while cultivated land is relatively fragmented and spatially constrained. Compared with many plains-based agricultural provinces, Yunnan therefore faces higher costs for irrigation, mechanization, transport connectivity, and rural energy supply [14]. These characteristics make it a suitable case for examining how agricultural development proceeds under infrastructure, resource, and ecological constraints [15]. As a typical plateau mountainous agricultural region in southwestern China, Yunnan also has complex topography, fragmented farmland, uneven spatial and temporal distribution of water resources, and relatively high costs of infrastructure construction [11]. Against this background, this study aims to quantitatively identify the key drivers, constraint structures, and marginal contribution thresholds of sustainable agricultural development in Yunnan Province from 1990 to 2023. A multidimensional indicator system is constructed from four aspects: fiscal and investment support, agricultural production inputs, rural infrastructure, and labor and population conditions. Ridge regression is used to address multicollinearity among explanatory variables, Bootstrap approximate inference is employed to improve the robustness of coefficient estimation, and the SHAP interpretation framework is introduced to rank key driving factors and identify threshold characteristics. By combining stable coefficient estimation, contribution ranking,

threshold identification, and sectoral comparison, this study attempts to move beyond conventional regional agricultural factor analysis and provide a quantitative framework for identifying sustainability-related development constraints in agricultural systems.

The potential contributions of this study are threefold. First, the novelty of this study lies not in examining agricultural inputs individually but in combining four analytical tasks within one empirical framework: stabilizing coefficient estimation under severe multicollinearity, approximating coefficient uncertainty through Bootstrap resampling, ranking variable-level predictive contributions using SHAP values, and identifying marginal contribution thresholds for both aggregate and sectoral agricultural models. This integrated design allows the study to distinguish common constraints from sector-specific constraints and to interpret sustainability-related development constraints as empirical transition intervals rather than fixed ecological ceilings. Second, it provides empirical evidence on how infrastructure, water resources, energy, capital, labor, fertilizers, and pesticides jointly shape the development boundaries of agriculture, forestry, and animal husbandry in a mountainous agricultural province. Third, it offers policy-relevant insights for sustainable agricultural governance, including the optimization of input structures, the control of excessive chemical inputs, and the design of differentiated policies for different agricultural sectors. The findings may also provide methodological and practical references for other mountainous, ecologically sensitive, or resource-constrained agricultural regions facing similar sustainability challenges.

2. Literature Review

2.1. Sustainable Agricultural Development and Resource–Environment Constraints

Sustainable agricultural development has become a significant subject in sustainability science, since it is connected with food production, resource consumption, ecological protection, rural living standards and regional progress. Different from the traditional study of agricultural growth, which mainly stresses increased output and enhanced productivity, sustainable agricultural development aims at the harmonious coexistence of economic, environmental and social objectives [5]. Velten et al. carried out a comprehensive review on the concept of sustainable agriculture and found that though different disciplines have different definitions, most investigations highlight the necessity to achieve a balance between productivity, environmental preservation, social welfare and the long-term stability of the system [16]. Moreover, Bathaei et al. indicated that the evaluation of agricultural sustainability needs multidimensional frameworks and indices rather than single-output assessment [17]. These researchers offer crucial theoretical support for investigating agricultural development from the aspect of sustainability limits.

Resource–environment constraints have also become central to agricultural sustainability research. Building on the planetary boundaries perspective, previous studies have shown that agriculture is closely related to land use change, freshwater use, biogeochemical cycles, biodiversity loss, and climate change [6,7]. Empirical agricultural studies further indicate that intensification may increase production but can weaken long-term sustainability when resource consumption and ecological costs are ignored [8,9]. Therefore, the literature increasingly emphasizes not only productivity growth but also the identification of input ranges, carrying conditions, and marginal turning points associated with sustainable agricultural transformation.

In the Chinese situation, the sustainable development of agriculture is closely connected with the national plans for ecological civilization, green agricultural change, rural reconstruction and high-quality development. Some research has investigated the sustainability of agriculture in China from the viewpoints of environmental laws, fertilizer application, rural progress and land use arrangements. For instance, Wang et al. studied

the relation between environmental laws and the performance of sustainable agricultural development in China, which indicates that policy interference is significant for enhancing the sustainability of agriculture [18]. Wu et al. examined the fertilizer usage in China and pointed out the importance of agricultural support policies, suggesting that a reduction in chemical inputs and the optimization of policies are the main problems in the sustainable transformation of agriculture [19]. Therefore, it is clear that the sustainability of Chinese agriculture not only relies on increasing input factors but also on improving input efficiency, decreasing environmental stress and reinforcing policy coordination. Existing empirical studies have increasingly attempted to quantify agricultural sustainability limits through carrying capacity assessment, input–output efficiency analysis, environmental pressure indices, and threshold regression models. These studies show that agricultural sustainability is often affected not only by the absolute scale of inputs but also by the point at which additional inputs generate diminishing returns or environmental pressure. Nevertheless, many regional studies still focus on evaluating sustainability performance or estimating average factor effects while giving less attention to the marginal turning points of specific inputs under multidimensional constraints.

2.2. Factor Inputs, Infrastructure, and Chemical Input Limits in Agricultural Systems

Many previous researchers have investigated the influence of factor inputs on agricultural progress. Land, labor, capital, machinery, technology, irrigation, electricity, fertilizers and pesticides are generally considered significant factors affecting agricultural yield. From the viewpoint of agricultural modernization, investments, mechanization, water-conserving facilities and rural electricity supply can enhance production capacity, decrease production risks and assist in changing traditional agriculture towards more efficient production methods [20]. In mountainous agricultural areas, infrastructure factors are especially important since the complicated topography, the divided fields and the uneven resource distribution raise the cost of agricultural production and restrict the spatial extension of agricultural operations [21].

Moreover, the connection between factor input and sustainable agricultural development is not always direct. Although the construction of infrastructure and productive investments can improve agricultural capability, the overuse of chemicals may result in negative environmental side effects and decreasing marginal benefits [22]. Excessive application of fertilizers and pesticides can lead to soil deterioration, non-point source pollution, water contamination, reduced biodiversity and increased ecological risks [23]. It should also be noted that chemical inputs do not always have negative effects. Improved nutrient management, soil testing and formula fertilization, integrated pest management, slow-release fertilizers, organic substitution, and precision application technologies may increase productivity while reducing environmental pressure. Therefore, the sustainability issue is not the simple presence of fertilizers or pesticides but whether their application intensity, timing, and management efficiency remain within a range where productivity gains exceed ecological costs. Wu et al. pointed out that the fertilizer usage in China is closely related to agricultural support policies, indicating that policy design may affect both the intensity of inputs and the sustainability outcomes [19]. Wang et al. also found that environmental policies are relevant to the sustainable development performance of agriculture in China, which implies that the sustainability of agriculture depends on the combination of production inputs, environmental controls and policy implementations [18]. Hence, from a sustainable point of view, the main problem is not merely whether an input increases agricultural output but also whether its additional effect remains positive within an acceptable ecological limit. This is especially significant for areas where ecological protection and agricultural production have overlapping spatial distributions.

The literature concerning sustainable land use and rural sustainability indicates that agriculture has other functions besides producing output. Lovell pointed out that agriculture can supply food, ecosystem services, landscape functions and social advantages and thus should be included in sustainable land use planning [24]. Similarly, Gullino et al. stressed that rural sustainability evaluations should take into account the relationship among agricultural actions, landscape preservation and environmental management [25]. These findings are applicable to Yunnan Province because it integrates agricultural production, biological diversity protection, mountainous scenery and rural living functions. Therefore, the agricultural development in Yunnan should be assessed not only in terms of output growth but also in regard to the compatibility of important inputs and infrastructures with ecological protection and long-term land use sustainability.

2.3. Quantitative Identification of Agricultural Development Constraints and Research Gaps

In regard to the methods, previous research has applied production functions, regression models, efficiency evaluations, index systems, spatial analyses and system dynamics to investigate agricultural growth and agricultural sustainability [26]. These techniques offer valuable means to determine influential factors and evaluate sustainability performance. Nevertheless, some limitations still exist. Firstly, most studies concentrate on whether certain factors significantly influence agricultural yield but ignore the critical ranges in which the contributions of the factors change. Secondly, ordinary regression analysis frequently encounters multicollinearity problems when a large number of agricultural inputs and infrastructural variables are considered together [27]. Thirdly, although sustainability studies stress multidimensional restrictions, few studies integrate the estimation of coefficients, robustness inference, contribution ranking and threshold identification into a consistent empirical model. Fourthly, regional studies generally do not have a definite relation between empirical agricultural results and broader sustainability ideas, such as statistically derived marginal contribution thresholds, resource–environmental carrying capacity and green transformation. Compared with studies that separately examine agricultural productivity, environmental regulation, fertilizer use, rural infrastructure, or sustainability indicators, fewer studies jointly address severe multicollinearity among long-term agricultural indicators, uncertainty in coefficient estimation, contribution-based ranking, and threshold identification in one framework. This gap is particularly evident for mountainous provinces, where infrastructure, water storage, energy supply, labor conditions, and chemical input intensity interact strongly over time.

Ridge regression is an effective approach to handling multicollinearity among explanatory variables by adding a penalty term and obtaining more reliable coefficient estimates [28]. Bootstrap approximate inference can also enhance the stability of parameter estimation, particularly when the sample size is small. Moreover, the SHAP interpretation framework can break down model predictions into component contributions at the variable level, which enables us to rank important influencing factors and detect nonlinear marginal effects [29]. Compared with the traditional significance-based interpretation, SHAP analysis gives a clearer way to comprehend how various factors affect agricultural development under multiple constraints.

According to the previous documents, this research indicates that the “limit” of agricultural development cannot be considered a single maximum output value. It should be regarded as a sustainability-related limit influenced by infrastructure capacity, the allocation boundaries of resources and the upper limits of ecological destruction caused by excessive chemical inputs. Some previous studies have given valuable information about agricultural sustainability, the effects of factor inputs and environmental restrictions, but there is still a shortage of empirical research that quantitatively determines

the sustainability-related development constraints and the critical influencing factors in mountainous agricultural areas. Thus, taking Yunnan Province as an example of a plateau mountainous area, this study combines ridge regression, Bootstrap approximate inference, SHAP-based factor ranking, threshold detection and sectoral heterogeneity analysis to show the multidimensional constraint system of sustainable agricultural development. This method can supplement regional agricultural sustainability research and offer support for the formulation of different agricultural green transformation policies.

3. Data Source and Variable Selection

3.1. Data Source

This paper forms a time-series sample with the provincial annual data of Yunnan Province from 1990 to 2023. The time length is chosen according to the following points: Firstly, since 1990, Yunnan's agriculture has experienced a gradual change from traditional production techniques to a modern agricultural system. The sampling period is adequate to include important phases like building agricultural facilities, increasing financial assistance, rearranging production resources and adjusting the industrial structure. Secondly, a longer time series is beneficial in detecting medium- and long-term restrictions on agricultural growth, instead of just showing short-term variations. Furthermore, the ridge regression and SHAP explanation methods used in our study are more appropriate for examining the stable correlations and slight differences in the influence of variables based on relatively complete time-series data.

The data mainly originates from the Yunnan Statistical Yearbook, the annual public information of agricultural and rural authorities and the relevant national agricultural science data. Considering possible changes in the calculation methods of statistical indicators among different years, we standardized the indicator names during the data cleaning process and used rule-based correction and interpolation techniques to handle the missing historical values, thus ensuring the continuity and comparability of the sample series. The final analysis sample was based on "Yunnan Province—Annual" as the basic unit, and it consisted of a total of 34 annual samples.

To investigate both the overall growth restrictions in Yunnan's agriculture and its sectoral heterogeneity characteristics, this research uses a macro-level model together with three sector-specific models, which include four dependent variables: the total output value of agriculture, forestry, animal husbandry and fisheries; the total agricultural output value; the total forestry output value; and the total livestock output value. The total quantity model can accurately find the main supportive elements and constraint limits of the agricultural system in Yunnan. The industry-specific model can also show the differences in sensitivity of different industries within agriculture to the same factor, thus preventing the loss of structural information when only the total output value is used as the study object.

According to the present agricultural progress in Yunnan, some suitable materials and statistical information are chosen to determine 14 explanatory variables and classify them into four main groups. In Table 1, it is observed that (1) financial assistance and investment, including the local government expenditure (100 million yuan) and the total social fixed asset investment in agriculture, forestry, animal husbandry and fishing (10,000 yuan), reflect the amount of public funds used and the possibility of social capital. Previous researchers have indicated that the government's fiscal expenditure and fixed asset investment are of great significance in improving the agricultural production environment, building agricultural infrastructure and promoting industrial transformation. In areas like Yunnan with complex geographical conditions, the impact of finance and investment on agricultural development is not only manifested in increased direct

output but also in enhanced public services and the indirect effects caused by infrastructure. (2) The agricultural production inputs consist of rural electricity consumption (100 million kilowatt-hours), the total power of agricultural machinery (10,000 kilowatts), the use of pesticides (10,000 tons), the application of fertilizers (10,000 tons) and the total cultivated land (10,000 hectares). These variables show agricultural production capability from the points of view of energy supply, mechanization level, chemical input and production scale. Rural electricity consumption can be regarded as an important index for evaluating agricultural production status and also demonstrating the degree of agricultural modernization. The total power of agricultural machinery reflects the ability to substitute manual work with mechanization. Pesticides and fertilizers are typical direct input factors, although their marginal effects are not always linear. The total cultivated land indicates the scale of agricultural production. (3) Rural infrastructure includes the number of villages supplied with tap water, the number of villages accessible by car, the number of villages with telephone facilities, the number of rural power stations, the installed capacity of 100 million kilowatts and the total storage capacity of reservoirs. These variables reflect rural production, life environment and external supporting conditions for agriculture, such as village water supply, transportation, communication, power facilities and water conservancy storage capacity. An improvement in infrastructure generally promotes agricultural growth by decreasing transaction costs, increasing the availability of resources, enhancing the ability to deal with risks and optimizing the industrial structure. (4) Regarding the categories of workforce and population, only the agricultural population (in thousands of people) is considered. The agricultural population variable indicates the supply of labor force for the agricultural sector in Yunnan. With the stable development of agricultural mechanization and scale, the importance of labor in agricultural progress has not been completely eliminated. Especially in mountain agriculture, special agriculture and labor-intensive industries, the agricultural population still has a considerable effect on agricultural growth.

Given that this study uses 34 annual observations and 14 explanatory variables, the predictor-to-sample-size ratio is relatively high. This modeling choice was made because this study aims to represent the multidimensional constraint structure of agricultural development rather than to estimate a parsimonious causal model. The 14 predictors were selected according to theoretical relevance, data continuity, and correspondence to four dimensions of agricultural sustainability-related constraints: fiscal and investment support, agricultural production inputs, rural infrastructure, and labor population conditions. To reduce the risk of overinterpretation caused by the limited sample size and high predictor dimensionality, this study uses ridge regression rather than ordinary least squares, standardizes all explanatory variables, applies Bootstrap approximate inference, and interprets the coefficients as conditional associations rather than isolated causal effects.

Table 1. Variable definitions, units and classifications.

Variable Name	Variable Symbol	Unit	Type of Variable	Classification
Total output value of agriculture, forestry, animal husbandry, and fisheries	Y_total	Billion yuan	dependent variable	Total Output
Total value of farm output	Y_agri	Billion yuan	dependent variable	Output by Industry
Total forestry value	Y_forestry	Billion yuan	dependent variable	Output by Industry
Total output value of animal husbandry	Y_husbandry	Billion yuan	dependent variable	Output by Industry
Local public financial budget expenditure	X1	Billion yuan	independent variable	Financial and Investment Support Category
The total invested fixed assets in society are divided among agriculture, forestry, animal husbandry and fisheries	X2	Ten thousand yuan	independent variable	Financial and Investment Support Category
Rural electricity consumption	X3	Gigawatt-hour	independent variable	Agricultural Production Input Category
Total power of agricultural machinery	X4	Thousand kilowatts	independent variable	Agricultural Production Input Category
Pesticide usage amount	X5	Ten thousand tons	independent variable	Agricultural Production Input Category
Fertilizer application rate	X6	Ten thousand tons	independent variable	Agricultural Production Input Category
Total sown area	X7	Ten thousand hectares	independent variable	Agricultural Production Input Category
Number of villages benefiting from tap water supply	X8	Unit	independent variable	Rural Infrastructure Category
Total number of connected villages	X9	Unit	independent variable	Rural Infrastructure Category
Number of villages contacted by phone	X10	Unit	independent variable	Rural Infrastructure Category
Building a power station in the rural area	X11	Unit	independent variable	Rural Infrastructure Category
Installed capacity	X12	Billion kilowatts	independent variable	Rural Infrastructure Category
Total storage capacity of the reservoir	X13	Billion cubic meters	independent variable	Rural Infrastructure Category
Agricultural population	X14	Ten thousand people	independent variable	Labor and Population Category

Table note: The dependent variables are employed to establish models for the total of agriculture, forestry, animal husbandry and fisheries, as well as agricultural, forestry and animal husbandry models. The independent variables are divided into four categories according to the research content: fiscal and investment support, agricultural production investment, rural infrastructure and labor and population. All the explanatory variables were standardized before being entered into the ridge regression model to avoid the influence of the differences in dimensions on the parameter estimation.

3.2. Data Preprocessing

Because there are some variations, such as changes in indicator names, line breaks, spaces, additional explanations, etc., in the long-term sequence of the annual statistical data, this paper first carries out a unified and standardized processing of the original indicator names. We deleted unnecessary spaces, combined line breaks, standardized the bracket format and modified some differences in the column names to achieve an accurate match of the same indicator in different years during the data transformation. Moreover, we removed the auxiliary explanatory rows that are obviously not included in the observed data, for example, the descriptions of the data sources and other non-analytical information.

Long-term statistical data may contain some missing years and indicators. If these missing years are deleted, the sample size will be reduced greatly and the continuity of the sequence will be affected. To retain as much original information as possible, the missing value treatment followed a hierarchical procedure. First, rule-based filling was used only when missing values could be inferred from official statistical definitions, unit conversions, or adjacent-year reporting continuities. Second, linear interpolation was applied to short internal gaps in variables with smooth temporal evolution, such as infrastructure coverage and population-related indicators. Third, K-nearest neighbor imputation was used for the remaining missing values by borrowing information from variables with similar temporal and structural patterns. This combined approach was adopted to preserve the continuity of the 1990–2023 time series while reducing the loss of observations. Because different imputation methods may affect estimates, the interpolated results were checked against variable trends and official statistical ranges to avoid implausible values. Because the units of each explanatory variable are different, for example, fiscal expenditure in billions of yuan, investment in ten thousand yuan, electricity in billion kilowatt-hours, reservoir capacity in billion cubic meters and population in ten thousand people, the unit differences may affect the comparison of the coefficients and decrease the penalty term. Therefore, before performing the ridge regression estimation in this paper, the explanatory variables were standardized.

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}$$

where x_{ij} denotes the original value of the i -th sample on the j -th variable, $\bar{x}_j$ represents the mean value of variable j , s_j represents the standard deviation of variable j , and Z_{ij} is the standardized variable value. After standardization, the mean of each variable is 0 and the standard deviation is 1, which is advantageous for implementing comparable parameter shrinkage for different variables in ridge regression.

Because the dataset is a provincial annual time series covering 1990–2023, additional time-series diagnostics were conducted before interpreting the level-form regression results. For level-form variables, ADF and KPSS tests were specified with a constant and linear trend. For first-differenced variables, both tests were specified with a constant only. The lag length in the ADF test was selected using an AIC-based automatic lag selection procedure, while the KPSS bandwidth was selected using automatic bandwidth selection. These diagnostics were used to assess whether the main results should be interpreted as long-run conditional associations rather than short-run causal estimates.

3.3. Descriptive Statistics

This paper computes the mean, standard deviation, minimum value, maximum value, skewness, kurtosis and number of missing values for both the dependent variable and the explanatory variables and calculates descriptive statistics based on the common sample, as presented in Table 2. From the viewpoint of variable types, the explanatory variables not only include macro fiscal investment and capital formation variables but also infrastructure

coverage indicators, resource carrying indicators and agricultural direct input indicators. The number of variables is moderate, and the dimensions are relatively rich, which can well reflect the multi-level characteristics of the agricultural development constraints in Yunnan. Because this research uses long-term annual provincial data, the variables show distinct temporal changes. Therefore, the descriptive statistics not only describe the sample distribution but also serve as a basis for studying the relationships between variables and possible multicollinearity.

Table 2. Descriptive statistical results of major variables.

Variable	Mean	Standard Deviation	Minimum Value	Maximum Value	Skewness	Kurtosis
Y_total	2175.226928	2039.618169	211.7233	6834.5	1.075117026	0.023592024
Y_agri	1068.760588	988.7128105	119.63	3629.9	1.239647158	0.596144397
Y_forestry	175.6955882	153.837087	18.27	497.33	0.732658084	−0.806659918
Y_husbandry	651.6894118	650.5610792	54.01	2315.41	1.28594915	0.815042105
X1	2458.159844	2543.767822	131.99	6974.0176	0.707616715	−1.174764169
X2	677,103.4118	487,951.6878	23,383	1,503,490	0.171912571	−1.397747649
X3	60.55215882	40.02567761	12.19	145.5339	0.716782495	−0.709002358
X4	1983.235294	922.8459624	649	3535	0.131444111	−1.421429276
X5	3.747042353	1.33506408	1.4414	5.8648	0.037186604	−1.08789102
X6	152.3769	54.9034	57.4273	235.5831	−0.0787	−1.2793
X7	605.9458026	89.77124731	451.4527	718.9187	−0.182282899	−1.33192662
X8	10,644.02941	2629.094135	3936	13,190	−1.279597024	0.511746011
X9	12,185.52941	924.276793	9804	12,914	−1.483018173	0.878484353
X10	12,421.76471	166.5087237	12,065	12,704	−0.109111326	−0.577763279
X11	1285.235294	1310.655308	235	4334	1.09702917	−0.166062245
X12	29.35885	4.817440722	23.3	44.29	1.591004028	2.576889217
X13	103.9535353	28.38814344	57.8094	154.54	0.043898158	−1.120853045
X14	3739.419956	255.7750725	3271.7	4144.667	−0.152451086	−1.098185513

Table Note: The observation period ranges from 1990 to 2023, including 34 years. Descriptive statistics were computed based on the original variables without using the standardized values. Skewness indicates the symmetry of the variable distribution, whereas kurtosis shows the sharpness of the variable distribution. Variables were considered in the statistics after interpolation or substitution of missing values, as described in the main data processing part.

The descriptive statistics also indicate clear long-term structural changes. Output variables increased substantially from their early-period minimum values to their later-period maximum values, reflecting the expansion and modernization of Yunnan's agricultural economy. Infrastructure and support variables, including fiscal expenditure, fixed asset investment, rural electricity consumption, machinery power, and reservoir storage capacity, also show large ranges, suggesting a long-term transition from traditional factor-dependent production toward infrastructure-supported agricultural modernization. In contrast, pesticide and fertilizer variables display patterns consistent with input intensification followed by potential marginal weakening, which provides an empirical basis for subsequent threshold analysis.

4. Research Methodology

4.1. Correlation Matrix Analysis

This paper calculates the correlation matrices between the explanatory variables and the dependent variable, as well as among the explanatory variables themselves by using the Pearson correlation coefficient, and then constructs a correlation heat map. The correlation matrix can clearly show the existence of a high correlation between variables and also give some basis for understanding the direction of model coefficients and SHAP ranking results in the future.

As illustrated in Figure 1, the main variables have a significant overall correlation. Particularly, the total output value of agriculture, forestry, animal husbandry and fisheries, as well as the output values of different industries, have a notable positive correlation with most input and basic condition variables. Among them, the correlation coefficients between X3, X1, X2, X13, X14, X7 and X4 and various output indicators are generally high. For example, Y_total has a highly positive correlation with X3 (0.989), X1 (0.972), X2 (0.939) and X13 (0.929), indicating that rural electricity consumption, local public financial budget expenditure, fixed asset investment in agriculture, forestry, animal husbandry and fisheries, and total reservoir capacity are closely related to the increase in total agricultural output. Sectorally, Y_forestry shows significant positive correlations with X1 (0.983), X3 (0.991), X2 (0.969) and X13 (0.951). These findings suggest that fiscal investment, energy use and water resource conditions play an important supportive role in promoting the output values of forestry and fisheries. On the contrary, X11 exhibits a strong negative correlation with all the output variables. For instance, its correlation coefficients with Y_total and Y_forestry are -0.648 and -0.699 , showing that the factors represented by this variable are in the opposite direction of the changes in agricultural output. Moreover, the correlation coefficients between X12 and X10 and the dependent variables are generally low, indicating a relatively weak linear correlation. Additionally, there exists a high degree of correlation among the explanatory variables. The correlation coefficients between pairs of variables such as X2, X14, X13, X4, X6 and X7 are generally high, suggesting significant multicollinearity in the sample. There is a high positive correlation between the dependent variables. For example, the correlation coefficients between Y1 total and Y1 agricultural and Y1 livestock are 0.993 and 0.989, respectively, reflecting the strong synchronism of the output values of different sub-industries in agriculture, forestry, and animal husbandry in terms of time change. The results of the correlation analysis provide a preliminary basis for the subsequent Variance Inflation Factor test, ridge regression estimation and SHAP interpretation analysis.

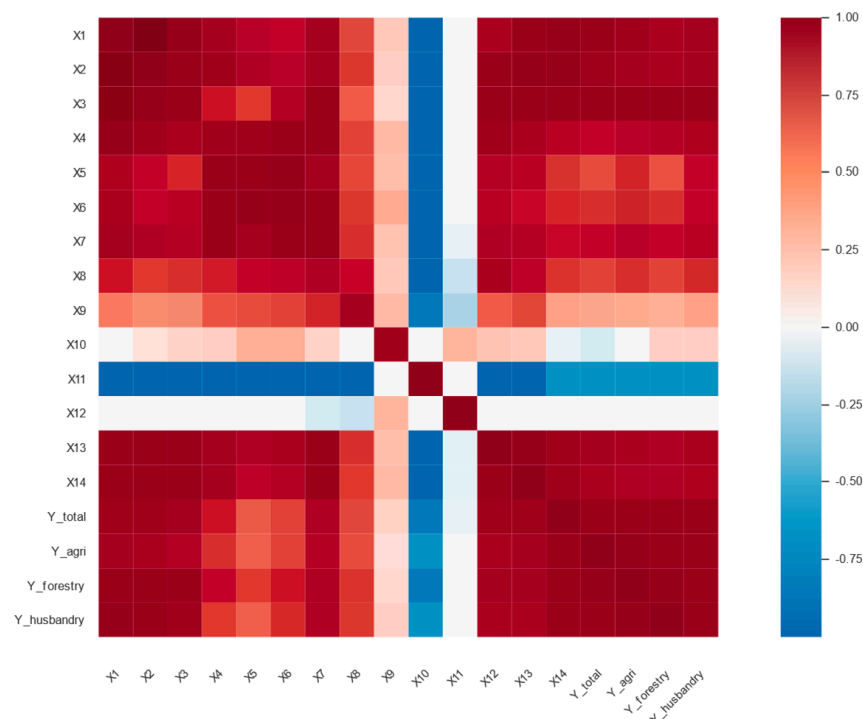


Figure 1. Heatmap of the correlation matrix among main variables.

4.2. Variance Inflation Factor (VIF) Test

To investigate multicollinearity among the explanatory variables, this research also computed the Variance Inflation Factor (VIF). For the j -th explanatory variable, its VIF is defined as:

$$VIF_j = \frac{1}{1 - R_j^2}$$

where R_j^2 denotes the coefficient of determination in an auxiliary regression with the j -th variable as the dependent variable and the other explanatory variables as independent variables. When a variable has a closer relationship with other explanatory variables, R_j^2 increases and so does the corresponding VIF_j .

Usually, when VIF is larger than 5, it suggests that the variables may have strong multicollinearity. If the VIF value is more than 10, the collinearity problem is generally very serious. Since the variables investigated in this research, such as finance, investment, energy, water resources, machinery and population, are also interrelated elements in the agricultural system, the removal of high-VIF variables may lead to the loss of important information, which is not beneficial for understanding the overall constraint structure of agricultural growth. Therefore, after detecting collinearity, this paper did not use the stepwise variable elimination method but applied ridge regression to get reliable parameter estimates.

As depicted in Figure 2, the VIF test results demonstrate that a multicollinearity problem exists among the model's explanatory variables, and the degree is rather high. Only X12, with a VIF value of 5.263, is within an acceptable but significant range, while the VIF values of the other variables are all greater than the empirical threshold of 10, suggesting that the linear relationship between variables is largely evident. The VIF values for X14, X8, X13, X3 and X6 are 3528.389, 2410.762, 1482.618, 1306.664 and 1266.841, respectively, indicating strong multicollinearity among the agricultural population, rural infrastructure, water resources conditions, energy input and fertilizer application. If ordinary least squares regression is used directly, it may result in unstable parameter estimates, increased standard errors and false significant tests. Hence, applying the ridge regression method will provide a more reliable basis for parameter estimation and interpretative analysis in subsequent steps. The extremely high VIF values indicate that several explanatory variables share strong common time trends and overlapping information. This is expected in a long-term provincial agricultural dataset because fiscal expenditure, investment, infrastructure, machinery, energy use, water conservancy capacity, and population conditions often evolve together during modernization. Therefore, individual coefficients should not be interpreted as isolated causal effects. Rather, they should be interpreted as conditional associations after shrinkage, while SHAP values are used as complementary evidence on predictive contribution. Ridge regression was selected because deleting highly correlated variables would remove theoretically important dimensions of the agricultural constraint system.

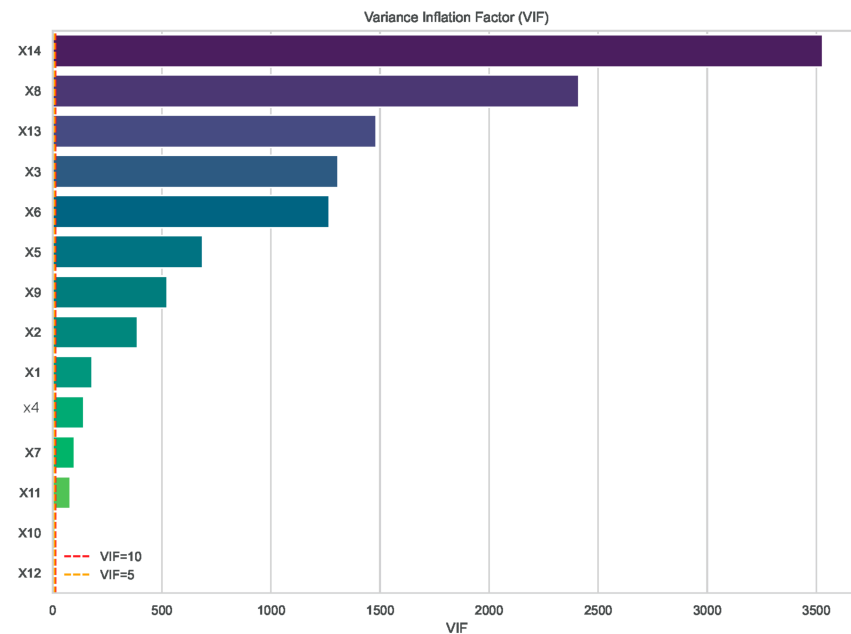


Figure 2. Interpretation of the VIF test results for the variables.

4.3. Ridge Regression Model Construction

4.3.1. Basic Model Specification

Within the traditional linear regression model, we denote y_i as the agricultural output value of the i -th year sample and x_{ij} as the value of the i -th sample for the j -th explanatory variable. The general linear model can then be written as:

$$y_i = \beta_0 + \sum_{j=1}^p \beta_j x_{ij} + \varepsilon_i, \quad i = 1, 2, \dots, n$$

where β_0 signifies the constant term, β_j stands for the coefficient of the explanatory variable, and ε_i indicates the random error term.

Ridge regression determines the coefficients by incorporating a penalty term of the square of the parameter into the sum of squared residuals, and its objective function is:

$$\min_{\beta_0, \beta} \left\{ \sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^p \beta_j^2 \right\}$$

where λ is the ridge parameter, also named the penalty coefficient. When $\lambda = 0$, the model turns into OLS; as λ increases, the degree of coefficient contraction increases and the estimation fluctuations caused by multicollinearity are decreased.

In matrix form, the ridge regression coefficient estimator can be expressed as:

$$\hat{\beta}_{ridge} = (X'X + \lambda I)^{-1} X'Y$$

where X is the standardized explanatory variable matrix, Y represents the dependent variable vector, and I is the identity matrix.

4.3.2. Model Grouping and Ridge Parameter Selection

To illustrate the differences in agricultural systems and industries, this paper develops four models: Model 1 represents the total agricultural, forestry and animal husbandry model, Model 2 represents the agricultural model, Model 3 represents the forestry model, and Model 4 represents the animal husbandry model. All these models adopt the same

14 explanatory variables. By doing so, the variations in responses of different industries to fiscal support, infrastructure, production inputs and demographic conditions can be compared under the same variable system.

The main point of ridge regression is to choose the penalty parameter λ . If λ is too small, it cannot effectively reduce multicollinearity; if λ is too large, it may cause excessive shrinkage and result in a decrease in the fitting ability of the model. In this paper, the optimal ridge parameters are searched by using the cross-validation method in the preset parameter range. Finally, the optimal λ values obtained by each model are: total agriculture, forestry, animal husbandry and fisheries model—0.2500, agricultural model—0.1055, forestry model—0.0084, and animal husbandry model—0.4630. The ridge penalty parameter was selected through cross-validation over a predefined grid of candidate lambda values. For each candidate value, model performance was evaluated using prediction error criteria, including RMSE and MAE. The lambda with the lowest cross-validated prediction error was selected as the optimal penalty parameter. Sensitivity checks showed that the main signs and relative importance of core variables, especially rural electricity consumption, reservoir storage capacity, pesticide use, and fertilizer application, remained broadly stable around the selected lambda values.

4.3.3. Coefficient Inference: Bootstrap Approximation Method

This paper uses the Bootstrap repeated sampling technique to approximately estimate the ridge regression coefficients. The detailed procedures are: draw samples with replacement from the original data, estimate the ridge regression model repeatedly, save the coefficient vector for each sample, and finally compute the Bootstrap mean, standard error, approximate z-value and p-value for each variable's coefficient. In this study, the Bootstrap procedure was repeated 1000 times for each model. Each replication drew annual observations with replacement from the original sample, re-estimated the ridge regression model using the selected lambda, and stored the coefficient vector. The empirical distribution of the 1000 replicated coefficients was then used to calculate standard errors, z-values, and approximate p-values.

The ridge regression coefficient of variable j in the b -th Bootstrap sampling is denoted as $\hat{\beta}_j^{(b)}$, and this process is repeated B times. Then, the Bootstrap mean and standard error are as follows:

$$\bar{\beta}_j = \frac{1}{B} \sum_{b=1}^B \hat{\beta}_j^{(b)}$$

$$SE(\hat{\beta}_j) = \sqrt{\frac{1}{B-1} \sum_{b=1}^B (\hat{\beta}_j^{(b)} - \bar{\beta}_j)^2}$$

Then, further approximate statistics are obtained:

$$z_j = \frac{\bar{\beta}_j}{SE(\hat{\beta}_j)}$$

The significant markers employed in our research adhere to the following criteria: $p < 0.01$ is indicated by ***, $p < 0.05$ is also indicated by **, and $p < 0.1$ is marked by *. The significance here does not indicate exact and closed-form inference in the OLS method, but rather approximate statistical results obtained from the Bootstrap distribution.

4.4. SHAP Interpretation Framework and Ranking of Key Driving Factors

Despite the fact that ridge regression can provide more stable estimation of coefficients under high correlation conditions, the coefficients themselves mainly indicate the direction

and relative impact of the standardized variations in the independent variables on the predicted values of the dependent variable, and it is still not easy to directly answer questions concerning “which variables are most significant in model prediction” and “how much each variable contributes to different samples”. Therefore, this paper introduces the SHAP (Shapley Additive Explanations) method. Although SHAP is widely used in machine learning models, it is also useful for linear ridge regression because it decomposes model predictions at the sample level and provides a contribution-based ranking that is not identical to coefficient magnitude. Ridge coefficients indicate the direction and average conditional association of standardized variables, whereas SHAP values show how much each variable contributes to fitted output values across observations. In a highly collinear time-series setting, using both coefficients and SHAP values helps distinguish coefficient direction from predictive importance.

SHAP comes from the Shapley value in cooperative game theory. The basic idea is that the model’s prediction results are considered the outcome when several features “cooperate”, and then the total predicted value is allocated fairly among the features to evaluate the contribution of each feature. For a sample x , the predicted value of the model can be expressed as:

$$f(x) = \phi_0 + \sum_{j=1}^p \phi_j$$

where ϕ_0 denotes the base predicted value and ϕ_j indicates the marginal effect of the j -th feature on the predicted result of the sample.

The general expression for the Shapley value is:

$$\phi_j = \sum_{S \subseteq F \setminus \{j\}} \frac{|S|!(|F|-|S|-1)!}{|F|!} [f_{S \cup \{j\}}(x_{S \cup \{j\}}) - f_S(x_S)]$$

where F denotes the whole set of features and S represents any subset that does not contain feature j . For the linear ridge regression model in this paper, the SHAP values can decompose the impact of variables on the prediction results in a relatively stable manner and supplement the interpretation with the regression coefficients. The former focuses more on the predictive contribution and rank, whereas the latter focuses more on the parameter direction and statistical inference.

To evaluate the overall importance of various variables in both the population and sector-specific models, this research uses the average absolute SHAP value as a global importance measure, defined as:

$$Importance_j = \frac{1}{n} \sum_{i=1}^n |\phi_{ij}|$$

where ϕ_{ij} denotes the SHAP value of variable j in the i -th sample. The larger the value of this index, the greater the influence of variable j on the model’s prediction outcomes. In this paper, the average absolute SHAP values of 14 explanatory variables in the four models are calculated, and these results are used to rank the main influencing factors and compare the industry heterogeneity.

4.5. Factor Threshold Identification Method

Only knowing that a particular variable is important is not sufficient for understanding its marginal contribution pattern. A further issue is whether the fitted contribution of a factor changes direction within the observed sample range. For example, rural electricity consumption may show a positive contribution after the agricultural system reaches a certain infrastructure support condition, whereas pesticide or fertilizer inputs may show

weakening marginal contributions after long-term intensification. This study therefore regards threshold identification as an interpretive procedure for detecting the point at which the fitted SHAP contribution of an explanatory variable crosses zero.

The thresholds identified in this study should be interpreted as statistical marginal contribution transition points. They indicate the observed point in the sample at which the fitted SHAP contribution of a variable changes sign relative to the model baseline. Such thresholds can provide directional reference information for monitoring and policy discussion, but they are not equivalent to strict ecological thresholds, causal optima, regulatory limits, or biophysical carrying capacity boundaries. Ecological thresholds would require direct ecological process indicators, experimental evidence, or spatially explicit environmental response data. Therefore, this study uses the term threshold in an empirical and model-based sense.

The SHAP-based threshold approach was used instead of conventional threshold regression for three reasons. First, regular threshold regression usually requires a pre-specified threshold variable and estimates regime-dependent coefficients, whereas this study aims to explore marginal contribution turning points for multiple explanatory variables across aggregate and sectoral models. Second, conventional threshold regression is more suitable for larger samples or panel datasets; in contrast, this study uses 34 annual provincial observations, which limits the reliability of estimating multiple regime-switching models. Third, the purpose of the threshold analysis in this study is interpretive rather than causal. The SHAP-based approach is consistent with the ridge regression framework because it identifies the point at which the fitted contribution of a variable changes sign in model prediction. Therefore, the identified thresholds are empirical marginal contribution transition points, not causal regime thresholds or definitive ecological limits.

For any variable X_j in any industry, this paper specifies a linear connection between the SHAP value and the original variable value.

$$SHAP_j = a_j + b_j X_j$$

where a_j denotes the intercept and b_j denotes the slope. If $b_j \neq 0$, we set $SHAP_j = 0$ to obtain the linear threshold of the variable:

$$X_j^* = -\frac{a_j}{b_j}$$

Its economic significance is that when the variable X_j reaches X_j^* , the influence of this variable on the output value of the target industry becomes zero. If the variable value increases more, its effect may change from negative to positive or from positive to negative according to the sign of the slope.

When identifying the experience threshold for LOWESS, since there may exist non-linear correlations between some variables and SHAP values, applying only linear fitting to determine the threshold may neglect the local curved characteristics. In this paper, LOWESS is used for non-parametric fitting of the relationship between variable values and SHAP values, and the intersection point of the smoothing curve with the zero axis is obtained to find the empirical threshold. If the linear threshold is close to the LOWESS threshold, it suggests that the relation between the variable and its SHAP contribution is relatively stable within the sample range.

After getting the SHAP importance and threshold of each variable in each industry, this paper also builds a “factor–industry” constraint heatmap. This heatmap expresses the constraint intensity through the average absolute SHAP value, thus distinguishing the variables that are common constraints among different industries from those that are specific to certain industries.

4.6. Time-Series Diagnostic and Robustness Strategy

To address the time-series nature of the data, this study implemented three supplementary diagnostic and robustness procedures. First, ADF and KPSS tests were used to examine the unit root and stationarity properties of each dependent and explanatory variable. The ADF test treats the presence of a unit root as the null hypothesis, whereas the KPSS test treats stationarity as the null hypothesis. Using both tests helps avoid relying on a single diagnostic criterion. For level-form variables, both tests were specified with a constant and a linear trend. For first-differenced variables, both tests were specified with a constant only. The ADF lag length was selected using an AIC-based automatic lag selection procedure, and the KPSS bandwidth was selected using automatic bandwidth selection.

Second, Johansen cointegration tests were conducted for selected core variable groups to examine whether the level-form relationships reflected possible long-run equilibrium associations rather than unrelated common trends. Because the sample contains only 34 annual observations, Johansen tests were applied to theoretically relevant core variable groups rather than to the full system with all explanatory variables.

Third, first-difference ridge regression was used as a robustness check to examine whether the signs of the main variables remained broadly consistent after removing common time trends. These procedures were not used to claim short-run causality. Instead, they were used to clarify the appropriate interpretation of the ridge regression and SHAP results.

5. Results

This section presents the empirical results of the ridge regression estimation, Bootstrap approximate inference, SHAP importance ranking, model fitting evaluation, threshold identification, and cross-sector comparison.

5.1. Ridge Regression Estimation Results

Because the explanatory variables show substantial multicollinearity, ridge regression was used to obtain more stable coefficient estimates. Bootstrap resampling was further applied to approximate the standard errors, z-values, and significance levels of the ridge coefficients. The results are reported for the total output value of agriculture, forestry, animal husbandry and fisheries and for the three sectoral models of agriculture, forestry, and animal husbandry.

5.1.1. Results of the Total Output Model of Agriculture, Forestry, Animal Husbandry and Fisheries

As shown in Table 3, in the total output model, rural electricity consumption has the largest positive standardized coefficient, reaching 755.760, and is significant at the 1% level. Total reservoir storage capacity and local public fiscal budget expenditure also show significant positive coefficients of 498.734 and 471.521, respectively, both significant at the 1% level. Agricultural population and total fixed asset investment in agriculture, forestry, animal husbandry and fisheries also show significant positive associations, with coefficients of 346.088 and 312.265, respectively.

By contrast, pesticide use and fertilizer application have significant negative coefficients of -558.095 and -313.924 , respectively. Both are significant at the 1% level. The number of villages accessible by road also shows a significant negative coefficient of -257.988 . These results indicate that the total output of agriculture, forestry, animal husbandry and fisheries in Yunnan is positively associated with electricity use, reservoir storage capacity, public fiscal expenditure, capital investment, and agricultural population, while pesticide use and fertilizer application are negatively associated with total output in the

ridge regression estimation. The negative coefficients of pesticide and fertilizer inputs should be interpreted cautiously. They may reflect the declining marginal productivity of chemical inputs after long-term intensification, but they may also be related to reverse causality, historical overuse, changes in cropping structure, improvements in nonchemical technologies, or broader structural transformation of the agricultural system. Therefore, the results should not be read as evidence that all chemical inputs are harmful. Rather, they suggest that further output growth in Yunnan is unlikely to be achieved sustainably through continued expansion of high-intensity chemical input use.

Table 3. Ridge regression estimation results for the total output of agriculture, forestry, animal husbandry and fisheries.

Variable	Coefficient	Standard Error	Z-Value	p-Value	Significance	Absolute Value Sort
X3	755.76	77.95	9.69	0.00	***	1
X5	−558.09	117.94	−4.73	0.00	***	2
X13	498.73	87.24	5.72	0.00	***	3
X1	471.52	136.85	3.45	0.00	***	4
X14	346.09	50.85	6.81	0.00	***	5
X6	−313.92	91.59	−3.43	0.00	***	6
X7	313.51	87.54	3.58	0.00	***	7
X2	312.27	67.11	4.65	0.00	***	8
X9	−257.99	58.23	−4.43	0.00	***	9
X4	176.52	165.36	1.07	0.29		10
X10	54.18	29.05	1.86	0.06	*	11
X12	−30.28	40.97	−0.74	0.46		12
X11	28.43	43.52	0.65	0.51		13
X8	7.48	55.48	0.13	0.89		14

Table notes: Standardized ridge regression coefficients are given; standard errors and significance levels are estimated by Bootstrap resampling; *** and * indicate significance at the 1% and 10% levels, respectively.

5.1.2. Results of the Agricultural Model

Table 4 reports the ridge regression results of the agricultural output model. Rural electricity consumption remains the largest positive factor, with a coefficient of 570.785, significant at the 1% level. The total power of agricultural machinery has a coefficient of 352.467 and is also significant at the 1% level. Total reservoir storage capacity and total fixed asset investment in agriculture, forestry, animal husbandry and fisheries have positive coefficients of 245.214 and 163.394, respectively, both significant at the 5% level. Pesticide use and fertilizer application have significant negative coefficients of −276.568 and −290.284, respectively, both significant at the 1% level. Agricultural population has a positive coefficient of 118.637 and is significant at the 1% level. Local public fiscal budget expenditure does not reach conventional significance in the agricultural model. These results show that agricultural output in Yunnan is strongly associated with electricity, mechanization, water conservancy conditions, investment, and labor support, while pesticide and fertilizer inputs show negative estimated effects.

5.1.3. Results of the Forestry Model

Table 5 shows the results of the forestry output model. Compared with the total output and agricultural models, the forestry model has fewer statistically significant variables. Rural electricity consumption has a positive coefficient of 105.778 and is significant at the 5% level. Pesticide use has a positive coefficient of 85.268 and is significant at the 10% level. Fertilizer application has a negative coefficient of −96.001 and is significant at the 10% level. The number of villages accessible by road also has a negative coefficient of −39.421 and is significant at the 10% level. Other variables, including fixed asset investment, agricultural

population, local public fiscal budget expenditure, and total reservoir storage capacity, do not reach conventional levels of significance in the forestry model. These results indicate that the factor structure of forestry differs from that of agriculture and animal husbandry. The forestry model differs from the agricultural and animal husbandry models because forestry production has longer biological cycles, stronger ecological functions, and weaker dependence on annual mechanized field operations and reservoir-based irrigation. The positive coefficient of pesticide use in forestry may reflect pest and disease control needs in economic forests and forest product management, rather than a general endorsement of chemical intensification. This result supports sector-specific input management: crop farming and animal husbandry should prioritize chemical input reduction and substitution, whereas forestry should emphasize precise, ecological, and risk-based pest control.

Table 4. Agricultural ridge regression estimation results.

Variable	Coefficient	Standard Error	Z-Value	p-Value	Significance	Absolute Value Sort
X3	570.79	90.75	6.29	0.00	***	1
X4	352.47	132.82	2.65	0.01	***	2
X6	−290.28	111.03	−2.61	0.01	***	3
X5	−276.57	91.22	−3.03	0.00	***	4
X13	245.21	103.52	2.37	0.02	**	5
X2	163.39	70.60	2.31	0.02	**	6
X7	130.98	100.73	1.30	0.19		7
X14	118.64	42.07	2.82	0.00	***	8
X11	94.81	49.11	1.93	0.05	*	9
X9	−78.55	58.79	−1.34	0.18		10
X8	−46.49	65.37	−0.71	0.48		11
X12	−21.25	42.62	−0.50	0.62		12
X1	16.86	123.66	0.14	0.89		13
X10	1.74	25.39	0.07	0.95		14

Table notes: Coefficients are the results of standardized ridge regression; standard errors and significance are estimated by approximate inference from Bootstrap resampling; ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table 5. Forest ridge regression estimation results.

Variable	Coefficient	Standard Error	Z-Value	p-Value	Significance	Absolute Value Sort
X3	105.78	45.91	2.30	0.02	**	1
X6	−96.00	50.57	−1.90	0.06	*	2
X5	85.27	47.70	1.79	0.07	*	3
X14	47.45	39.59	1.20	0.23		4
X9	−39.42	20.70	−1.90	0.06	*	5
X2	22.77	33.76	0.67	0.50		6
X8	19.57	34.53	0.57	0.57		7
X7	−15.56	21.96	−0.71	0.48		8
X4	−9.89	32.91	−0.30	0.76		9
X13	8.52	48.43	0.18	0.86		10
X1	6.84	36.99	0.18	0.85		11
X10	4.73	6.89	0.69	0.49		12
X11	−4.59	13.68	−0.34	0.74		13
X12	1.18	7.16	0.16	0.87		14

Table note: The coefficients are standardized ridge regression coefficients; the standard errors and significances are based on approximate inference from Bootstrap resampling; ** and * indicate significance at the 5% and 10% levels, respectively.

5.1.4. Results of the Animal Husbandry Model

Table 6 presents the ridge regression results of the animal husbandry model. Rural electricity consumption has the largest positive coefficient of 233.914 and is significant at the 1% level. Total reservoir storage capacity has a coefficient of 198.356 and is significant at the 1% level. Total fixed asset investment in agriculture, forestry, animal husbandry and fisheries has a coefficient of 129.000 and is significant at the 1% level. Agricultural population also shows a significant positive coefficient of 114.618. The total power of agricultural machinery has a positive coefficient of 142.561 and is significant at the 10% level. Pesticide use and fertilizer application have negative coefficients of −224.572 and −143.130, respectively, both significant at the 1% level. These results show that animal husbandry output is positively associated with electricity, water storage capacity, investment, agricultural machinery, and labor conditions, while pesticide and fertilizer inputs are negatively associated with output.

Table 6. Ridge regression estimation results for animal husbandry.

Variable	Coefficient	Standard Error	Z-Value	p-Value	Significance	Absolute Value Sort
X3	233.91	38.30	6.11	0.00	***	1
X5	−224.57	69.27	−3.24	0.00	***	2
X13	198.36	58.79	3.37	0.00	***	3
X6	−143.13	46.55	−3.07	0.00	***	4
X4	142.56	85.05	1.68	0.09	*	5
X2	129.00	22.83	5.65	0.00	***	6
X14	114.62	19.66	5.83	0.00	***	7
X9	−95.03	34.31	−2.77	0.01	***	8
X7	84.99	41.75	2.04	0.04	**	9
X1	82.14	90.93	0.90	0.37		10
X11	24.08	25.26	0.95	0.34		11
X8	23.55	29.49	0.80	0.42		12
X10	12.44	12.42	1.00	0.32		13
X12	−2.90	28.56	−0.10	0.92		14

Table note: Coefficients are the results of standardized ridge regression; standard errors and significances are estimated by Bootstrap resampling; ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

In summary, the ridge regression results show that rural electricity consumption is the most stable positive factor across the total and sectoral models. Total reservoir storage capacity, fixed asset investment, fiscal expenditure, agricultural machinery, and agricultural population also show positive effects in several models. Pesticide and fertilizer inputs are significant negative factors in the total, agricultural, and animal husbandry models, while forestry presents a more differentiated coefficient structure.

5.2. SHAP Value Ranking and Identification of Key Driving Factors

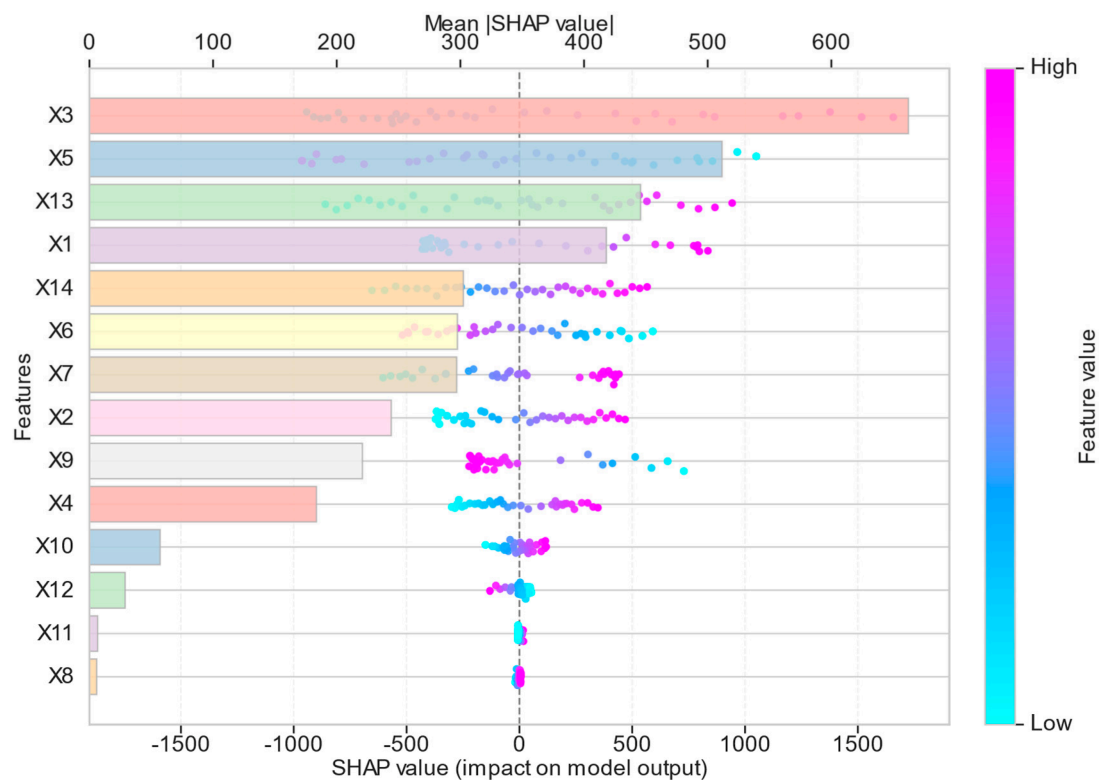
To further evaluate the contribution of each explanatory variable to model prediction, this study calculates the mean absolute SHAP values for all variables in each model. A larger mean absolute SHAP value indicates a greater contribution of the variable to the model output. The results are shown in Table 7.

5.2.1. SHAP Ranking for the Total Output Model

In the total output model, the five variables with the largest mean absolute SHAP values are rural electricity consumption, pesticide use, total reservoir storage capacity, local public fiscal budget expenditure, and agricultural population. Their SHAP values are 662.290, 511.582, 445.682, 417.987, and 302.362, respectively (see Figure 3).

Table 7. Ranking of SHAP importance of explanatory variables under different models.

Variable	Y_total	Y_agri	Y_forestry	Y_husbandry
X1	417.987	14.469	4.585	61.907
X2	244.051	131.180	17.411	110.439
X3	662.290	542.790	103.305	206.971
X4	183.564	347.410	11.991	157.828
X5	511.582	228.717	94.231	199.363
X6	297.575	298.479	110.870	143.184
X7	296.984	101.753	12.785	70.183
X8	5.824	44.125	23.795	15.721
X9	220.768	71.395	37.244	83.258
X10	57.149	6.965	4.813	13.994
X11	6.621	66.470	7.598	13.151
X12	28.842	23.995	0.186	9.734
X13	445.682	223.617	3.219	180.129
X14	302.362	93.514	44.598	96.511

**Figure 3.** SHAP analysis results for the total volume of agriculture, forestry, animal husbandry, and fisheries.

This ranking is broadly consistent with the ridge regression results. Rural electricity consumption has both the largest positive coefficient and the highest SHAP importance. Pesticide use ranks second in SHAP importance, although its ridge coefficient is negative. This indicates that pesticide use is a highly influential constraint variable in the total output model.

5.2.2. SHAP Ranking for the Agricultural Model

In the agricultural model, the five variables with the highest mean absolute SHAP values are rural electricity consumption, total power of agricultural machinery, fertilizer

application, pesticide use, and total reservoir storage capacity. Their SHAP values are 542.790, 347.410, 298.479, 228.717, and 223.617, respectively (see Figure 4).

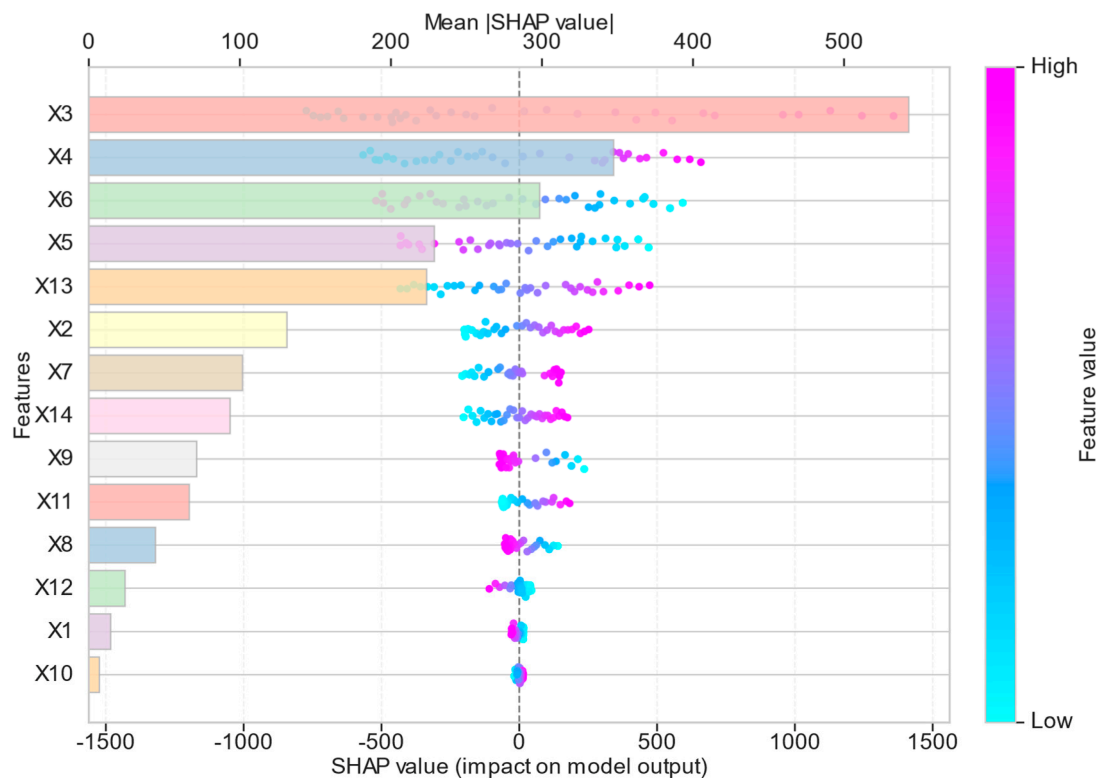


Figure 4. Results of agricultural SHAP analysis.

These results show that crop output is mainly influenced by electricity, mechanization, fertilizer use, pesticide use, and water storage conditions. Compared with the total output model, agricultural machinery has a more prominent position in the agricultural model.

5.2.3. SHAP Ranking for the Forestry Model

In the forestry model, the five most important variables are fertilizer application, rural electricity consumption, pesticide use, agricultural population, and the number of villages accessible by road. Their SHAP values are 110.870, 103.305, 94.231, 44.598, and 37.244, respectively (see Figure 5).

Forestry differs from agriculture and animal husbandry in that pesticide use has a positive ridge coefficient, while fertilizer application has a negative coefficient. The SHAP ranking also indicates that forestry output is influenced by chemical input variables, electricity, labor conditions, and spatial infrastructure.

5.2.4. SHAP Ranking for the Animal Husbandry Model

In the animal husbandry model, the five variables with the highest mean absolute SHAP values are rural electricity consumption, pesticide use, total reservoir storage capacity, total power of agricultural machinery, and fertilizer application. Their SHAP values are 206.971, 199.363, 180.129, 157.828, and 143.184, respectively (see Figure 6).

The SHAP results show that animal husbandry output is associated with electricity, water storage capacity, machinery, and chemical inputs. Similar to the agricultural model, pesticide and fertilizer inputs have high importance but negative ridge coefficients.

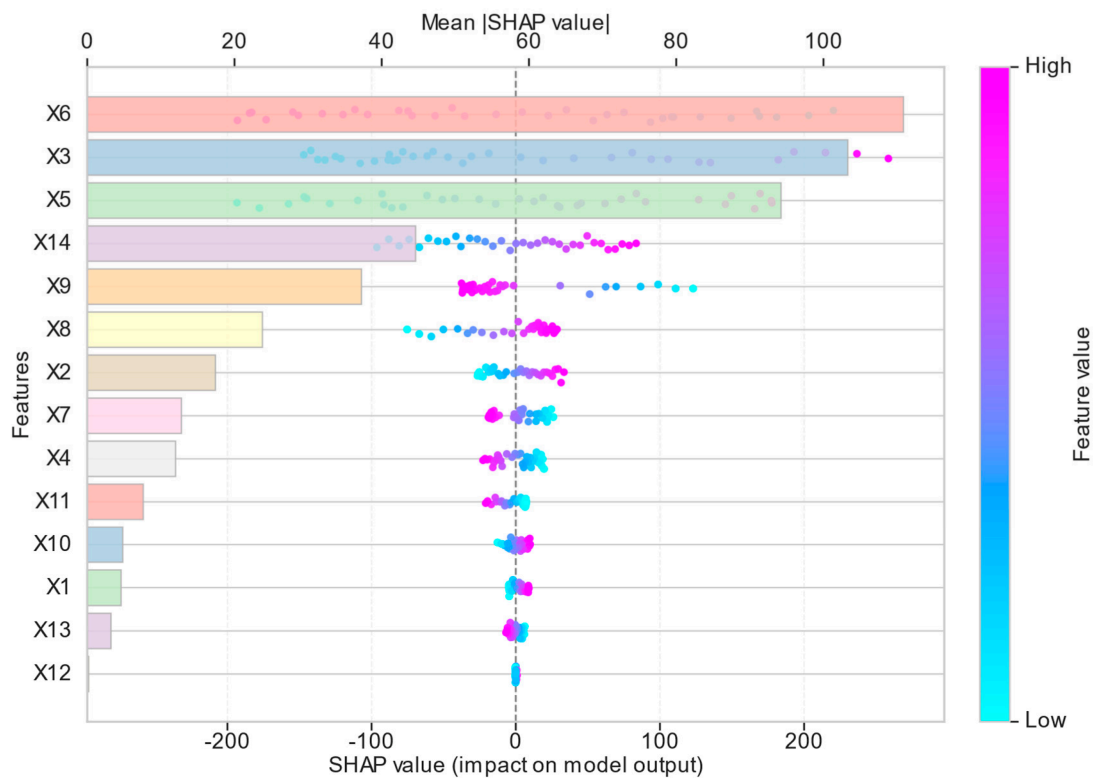


Figure 5. Results of forestry SHAP analysis.

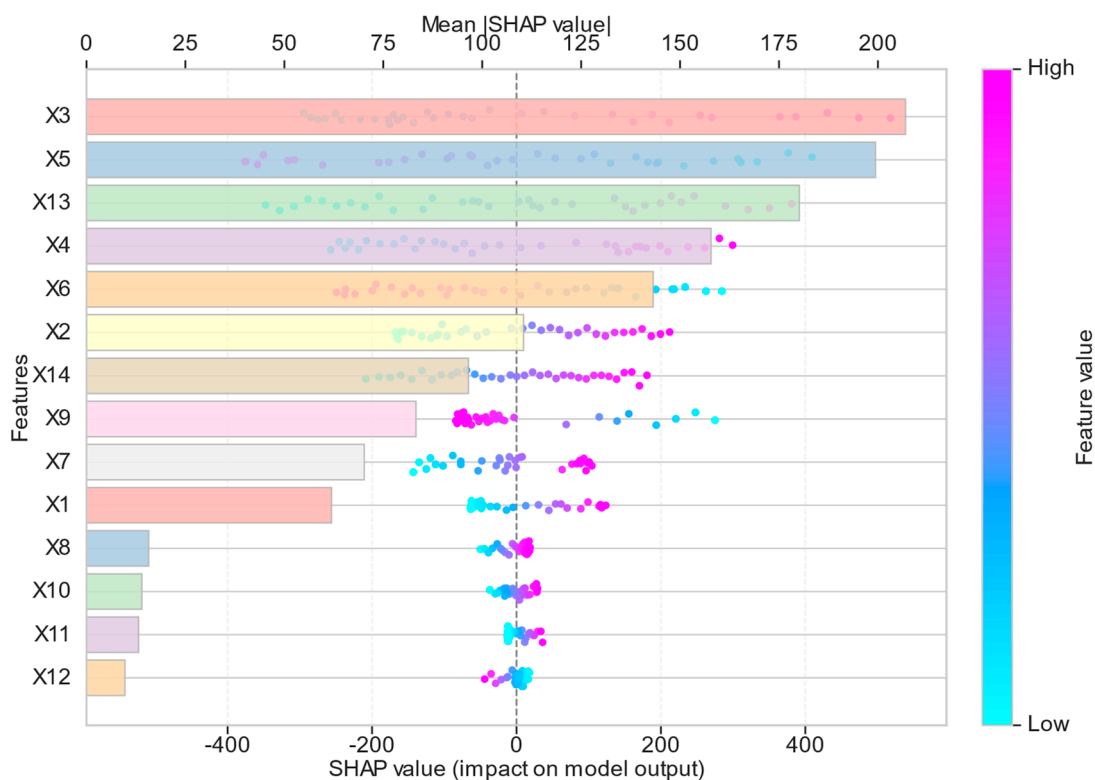


Figure 6. Results of the livestock SHAP analysis.

Overall, the SHAP analysis confirms that rural electricity consumption is the most stable key factor across models. Chemical inputs, particularly pesticides and fertilizers, are also highly influential variables, but their estimated effects are mainly negative in the total, agricultural, and animal husbandry models. Agriculture and animal husbandry are

more strongly associated with electricity, machinery, and water conditions, while forestry shows a more specific and differentiated factor structure. The comparison between ridge coefficients and SHAP rankings shows broad consistency for the core variables. Rural electricity consumption has both positive coefficients and the highest SHAP importance in most models, confirming its role as a stable supporting factor. Pesticide and fertilizer inputs also rank highly in SHAP importance, but their ridge coefficients are negative in the total, agricultural, and animal husbandry models, indicating that they are influential constraint variables rather than positive growth drivers. Thus, SHAP analysis complements ridge regression by showing the predictive importance of variables whose coefficient directions may indicate sustainability-related constraints.

5.3. Heterogeneity in Restriction Structures Across Different Sectors

Based on the ridge regression coefficients and SHAP rankings, this study further compares the factor constraint structures of different sectors. Figure 7 presents the factor–industry heatmap based on mean absolute SHAP values.

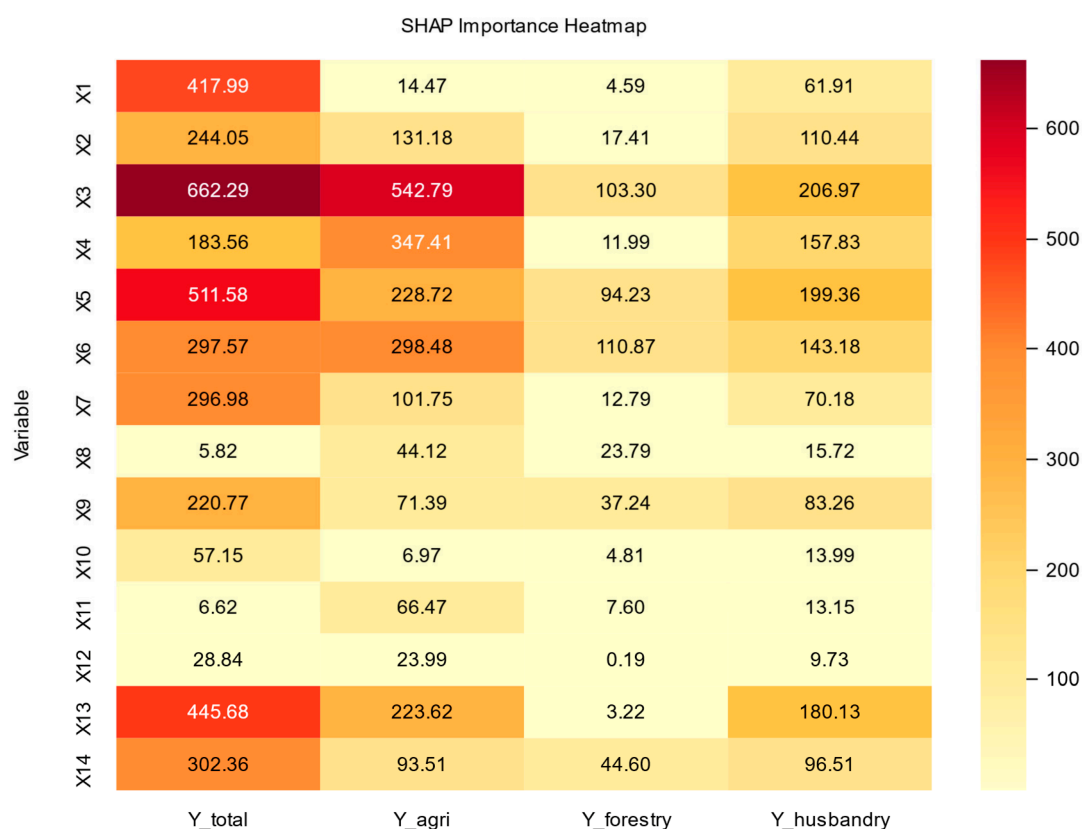


Figure 7. Heatmap of “factor–industry” constraints. Figure note: The thermal intensity is indicated by the mean absolute SHAP value. The greater the value, the greater the influence of the variable on the variation in the output value of the corresponding industry.

The heatmap shows that rural electricity consumption is a common influential factor across the total output model and all sectoral models. Total reservoir storage capacity is also highly important in the total output, agricultural, and animal husbandry models. These results indicate that energy and water conservancy conditions are common supporting factors for Yunnan’s agricultural system.

Agriculture and animal husbandry show similar factor structures. Both sectors are influenced by electricity, machinery, water conservancy conditions, pesticides, and fertilizers.

In both models, electricity, machinery, and reservoir capacity are mainly positive factors, while pesticides and fertilizers are negative factors.

Forestry has a more distinct factor structure. Its most important variables include fertilizer application, rural electricity consumption, pesticide use, agricultural population, and road accessibility. Compared with agriculture and animal husbandry, forestry is less dependent on machinery and reservoir storage capacity, while chemical input variables and labor-related conditions play a more visible role.

These results indicate that the factor constraints of Yunnan's agricultural development are not uniform across sectors. Common constraints exist, especially in energy and water conditions, but sector-specific differences are also evident.

5.4. Model Fitting Performance and Predictive Power

To evaluate model performance, this study uses R^2 , RMSE, and MAE. The results are presented in Table 8 and Figure 8.

Table 8. Evaluation results of the total agricultural output and sector-specific models in Yunnan.

Metric	Y_total	Y_agri	Y_forestry	Y_husbandry
Best_lambda	0.250	0.106	0.008	0.463
R^2	0.997	0.994	0.995	0.978
RMSE	103.763	73.862	10.325	93.979
MAE	75.323	56.436	7.272	51.568

Table note: Drawn by the author based on the model calculation results.

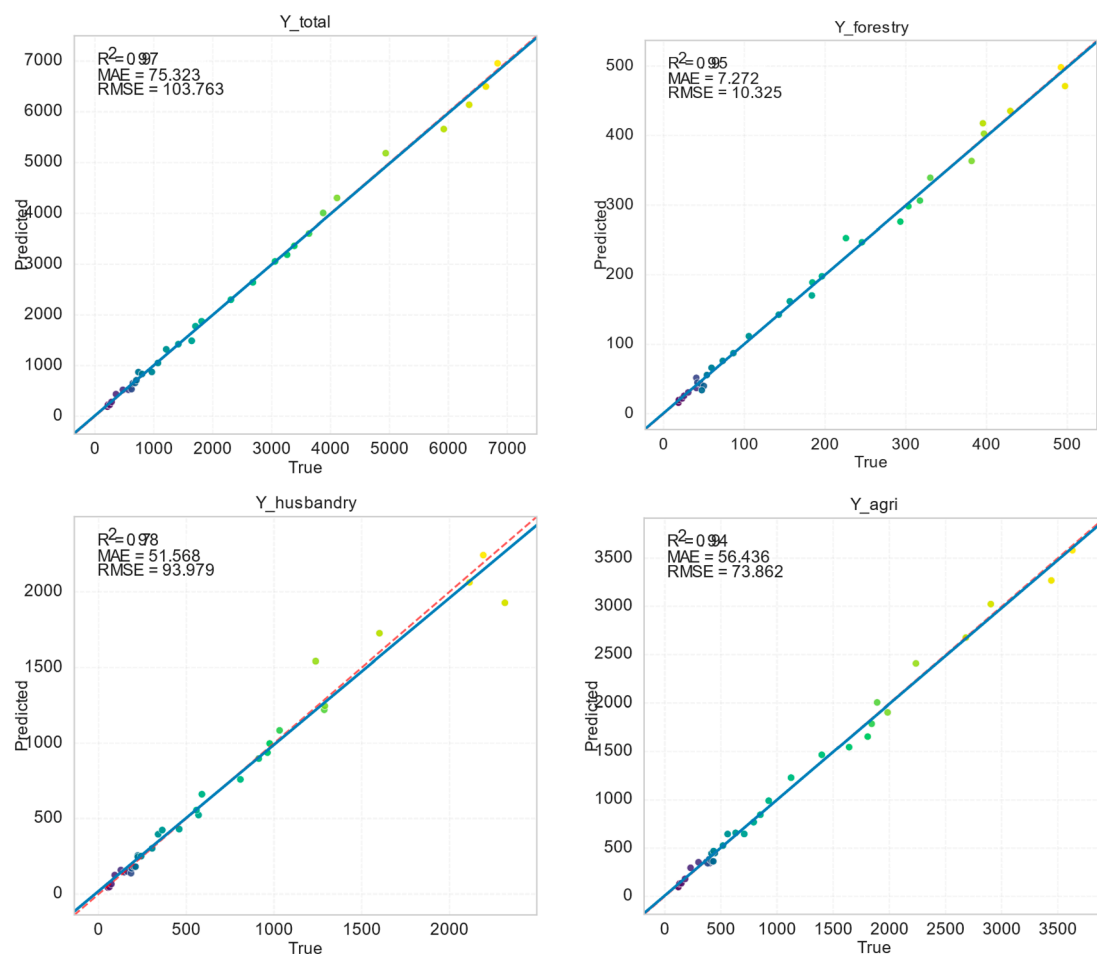


Figure 8. Comparison of actual and fitted values across models. Figure note: Drawn by the author based on the model calculation results.

According to Table 8 and Figure 8, the results show that all models have high in-sample explanatory power. The optimal lambda value of the total output model is 0.2500, with $R^2 = 0.997$, RMSE = 103.763, and MAE = 75.323. The agricultural model has an optimal lambda value of 0.1055, with $R^2 = 0.994$, RMSE = 73.862, and MAE = 56.436. The forestry model has an optimal lambda value of 0.0084, with $R^2 = 0.995$, RMSE = 10.325, and MAE = 7.272. The animal husbandry model has an optimal lambda value of 0.4630, with $R^2 = 0.978$, RMSE = 93.979, and MAE = 51.568.

The R^2 values of all models are higher than 0.97, indicating that the selected explanatory variables can effectively explain the long-term variation in the total and sectoral agricultural output values of Yunnan. Differences in RMSE and MAE among models mainly reflect differences in output value scale across sectors.

5.5. Time-Series Diagnostics and Robustness Checks

Because the study uses annual time-series data from 1990 to 2023, potential non-stationarity, common trends, and long-run equilibrium relationships may affect the interpretation of the level-form ridge regression results. To address this issue, detailed ADF and KPSS diagnostics were conducted for all dependent and explanatory variables. In the level-form tests, both ADF and KPSS were specified with a constant and a linear trend. In the first-difference tests, both tests were specified with a constant only. The ADF lag length was selected using an AIC-based automatic lag selection procedure, and the KPSS bandwidth was selected using automatic bandwidth selection. Table 9a reports the detailed results for the core variables.

The ADF and KPSS diagnostics indicate mixed time-series properties across variables. The four output variables show inconclusive or mixed evidence, reflecting strong long-term trends and structural changes in Yunnan's agricultural development. Several explanatory variables, including X1, X2, X5, X7, X10, X12, and X13, are more consistent with first-order integration, while X6 and X11 are closer to stationary series. Other variables, including X3, X4, X8, X9, and X14, show mixed or inconclusive evidence. This result is not unexpected for a 34-year provincial macroeconomic and agricultural dataset, in which many variables contain deterministic trends, modernization effects, and structural transitions. Therefore, the level-form ridge regression results should not be interpreted as short-run causal estimates. Instead, they are treated as long-run conditional associations within Yunnan's agricultural development system.

To further examine whether the level-form relationships are purely driven by unrelated common trends, Johansen cointegration tests were conducted for selected theoretically relevant core variable groups. Because the sample size is limited, the tests were not applied to the full system with all explanatory variables. The results are reported in Table 9b.

The Johansen results provide supplementary evidence that the selected core variable groups contain cointegrating relationships. The total output model shows three cointegrating relationships according to the trace test and two according to the max-eigen test. The agricultural model shows two cointegrating relationships under both tests. The forestry model provides weaker but observable evidence of cointegration, with one cointegrating relationship according to the trace test and two according to the max-eigen test. The animal husbandry model shows three cointegrating relationships under both tests. These findings suggest that the level-form associations are not merely arbitrary correlations among unrelated trending variables. However, given the small sample size and the exploratory purpose of this study, the Johansen results are used only as supportive diagnostics rather than as a basis for causal inference.

First-difference ridge regression was then used as a robustness check. After first differencing, the results remain broadly consistent with the main conclusions for the most important variables. The summary results are shown in Table 9c.

Table 9. (a) Detailed ADF and KPSS diagnostics for core variables. (b) Johansen cointegration diagnostics for selected core variable groups. (c) First-difference ridge robustness summary.

(a)												
Variable	Transformation	ADF Specification	ADF Lag	ADF Statistic	ADF <i>p</i> -Value	ADF 5% CV	KPSS Specification	KPSS Lag	KPSS Statistic	KPSS <i>p</i> -Value	KPSS 5% CV	Interpretation
Y_total	Level	constant + trend	10	1.640	1.000	−3.622	constant + trend	3	0.236	0.010	0.146	Mixed
Y_total	First-difference	constant only	10	0.829	0.992	−3.005	constant only	3	0.640	0.019	0.463	Mixed
Y_agri	Level	constant + trend	4	2.321	1.000	−3.574	constant + trend	3	0.237	0.010	0.146	Mixed
Y_agri	First-difference	constant only	6	1.162	0.996	−2.981	constant only	3	0.661	0.017	0.463	Mixed
Y_forestry	Level	constant + trend	3	−0.770	0.968	−3.569	constant + trend	3	0.245	0.010	0.146	Mixed
Y_forestry	First-difference	constant only	0	−7.053	<0.001	−2.957	constant only	1	0.799	0.010	0.463	Mixed
Y_husbandry	Level	constant + trend	10	3.826	1.000	−3.622	constant + trend	3	0.230	0.010	0.146	Mixed
Y_husbandry	First-difference	constant only	10	2.179	0.999	−3.005	constant only	5	0.455	0.053	0.463	Mixed
X3	Level	constant + trend	1	0.406	0.997	−3.558	constant + trend	3	0.244	0.010	0.146	Mixed
X3	First-difference	constant only	1	−1.836	0.363	−2.961	constant only	3	0.736	0.010	0.463	Mixed
X4	Level	constant + trend	8	−0.605	0.979	−3.603	constant + trend	3	0.100	0.100	0.146	Mixed
X4	First-difference	constant only	8	−1.247	0.653	−2.992	constant only	1	0.130	0.100	0.463	Mixed
X5	Level	constant + trend	5	−4.649	0.001	−3.581	constant + trend	3	0.152	0.045	0.146	More consistent with I(1)
X5	First-difference	constant only	6	−3.788	0.003	−2.981	constant only	3	0.389	0.082	0.463	More consistent with I(1)
X6	Level	constant + trend	6	−4.235	0.004	−3.588	constant + trend	3	0.145	0.053	0.146	Closer to I(0)
X6	First-difference	constant only	0	−1.596	0.486	−2.957	constant only	3	0.394	0.080	0.463	Closer to I(0)
X13	Level	constant + trend	0	−2.393	0.383	−3.553	constant + trend	2	0.066	0.100	0.146	More consistent with I(1)
X13	First-difference	constant only	0	−5.387	<0.001	−2.957	constant only	2	0.119	0.100	0.463	More consistent with I(1)
X14	Level	constant + trend	6	−3.827	0.015	−3.588	constant + trend	3	0.229	0.010	0.146	Mixed
X14	First-difference	constant only	6	−3.887	0.002	−2.981	constant only	3	0.720	0.012	0.463	Mixed
(b)												
Model group	Variable set	Deterministic order		Lagged differences		Trace rank at 5%	Max-eigen rank at 5%		Interpretation			
Total output	Y_total, X3, X4, X5, X6, X13	det_order = 1		k_ar_diff = 1		3	2		Cointegration evidence observed			
Agriculture	Y_agri, X3, X4, X5, X6, X7, X13	det_order = 1		k_ar_diff = 1		2	2		Cointegration evidence observed			
Forestry	Y_forestry, X1, X2, X5, X13, X14	det_order = 1		k_ar_diff = 1		1	2		Weaker but observable cointegration evidence			
Animal husbandry	Y_husbandry, X3, X4, X5, X6, X13	det_order = 1		k_ar_diff = 1		3	3		Cointegration evidence observed			
(c)												
Model	Selected alpha	RMSE	MAE	X3	X13	X5	X6	Consistency				
Y_total	17.629	133.848	85.300	37.026	38.011	−46.225	−34.682	Broadly consistent Broadly consistent Differentiated forestry pattern remains Broadly consistent				
Y_agri	35.297	95.793	60.604	18.952	13.481	−10.074	−16.974					
Y_forestry	0.022	13.016	9.170	1.260	−3.540	29.021	−22.146					
Y_husbandry	100.000	137.490	78.941	2.531	8.915	−7.548	−8.027					

(a) note: ADF denotes the Augmented Dickey–Fuller test, and KPSS denotes the Kwiatkowski–Phillips–Schmidt–Shin test. The ADF null hypothesis is the presence of a unit root, whereas the KPSS null hypothesis is stationarity. Level–form tests include a constant and a linear trend, while first-difference tests include a constant only. ADF lag length was selected using AIC-based automatic lag selection, and KPSS bandwidth was selected using automatic bandwidth selection. KPSS *p*-values reported as 0.010 or 0.100 may represent boundary values returned by the statistical software. (b) note: Johansen ranks are determined sequentially according to the 5% critical values. Because the sample contains only 34 annual observations, the test was conducted for selected theoretically relevant core variable groups rather than the full system with all predictors. The results are used as supportive diagnostics rather than as evidence of causal relationships. (c) note: Coefficients are estimated using first-differenced variables and ridge regression. X3 = rural electricity consumption; X13 = total reservoir storage capacity; X5 = pesticide use; X6 = fertilizer application. The first-difference ridge model is used as a robustness check rather than as the main estimation model.

The first-difference ridge results show that rural electricity consumption remains positive in all four models, and reservoir storage capacity remains positive in the total output,

agricultural, and animal husbandry models. Pesticide use and fertilizer application remain negative in the total output, agricultural, and animal husbandry models. The forestry model continues to show a differentiated structure, with pesticide use positive, reservoir storage capacity negative, and fertilizer application negative. These patterns support the main finding that energy and water resource conditions are important supporting factors in most models, whereas high-intensity chemical inputs are associated with weakening marginal contributions in the total, agricultural, and animal husbandry models.

Overall, the time-series diagnostics provide additional support for the reliability of the main empirical patterns while clarifying their appropriate scope of interpretation. The level-form results are best understood as model-based long-run conditional associations, not as short-run causal estimates. The SHAP-derived thresholds should accordingly be interpreted as statistically derived marginal contribution transition points rather than definitive ecological limits or operational policy cutoffs.

5.6. Threshold Identification and Cross-Sector Comparison

After estimating coefficients and ranking variable importance, this study further identifies the point at which the marginal contributions of key factors change. Thresholds are identified based on the relationship between SHAP values and the original values of explanatory variables. Linear thresholds and LOWESS empirical thresholds are calculated for each factor. The results for the total output model are shown in Table 10 and Figure 9.

Table 10. Threshold identification results of key factors for the total output of agriculture, forestry, animal husbandry, and fisheries in Yunnan.

Variable	Mean_abs_SHAP	Linear_a	Linear_b	Linear_R2	Linear_Threshold	LOWESS_Threshold	Direction
X1	417.9865168	−454.7698825	0.185004195	1	2458.159844	2458.159844	Positive
X2	244.0507737	−384.0372437	0.000567177	1	677,103.4118	677,103.4118	Positive
X3	662.2899825	−1180.642023	19.49793444	1	60.55215882	60.55215882	Positive
X4	183.56379	−444.9456098	0.224353414	1	1983.235294	1983.235294	Positive
X5	511.5821634	1706.22372	−455.352131	1	3.747042353	3.747042353	Negative
X6	297.5748818	948.9280955	−6.227506781	1	152.3768868	152.3768868	Negative
X7	296.9843562	−2369.889274	3.911058157	1	605.9458026	605.9458026	Positive
X8	5.824323759	−29.27155809	0.002750045	1	10,644.02941	10,644.02941	Positive
X9	220.7681682	3730.494441	−0.306141351	1	12,185.52941	12,185.52941	Negative
X10	57.14912111	−5235.118487	0.421447243	1	12,421.76471	12,421.76471	Positive
X11	6.620619908	−7.67550514	0.005972062	1	1285.235294	1285.235294	Positive
X12	28.8421988	256.8469525	−8.748535875	1	29.35885	29.35885	Negative
X13	445.6820892	−1937.814761	18.64116267	1	103.9535353	103.9535353	Positive
X14	302.3615046	−5224.548625	1.397154823	1	3739.419956	3739.419956	Positive

Table note: Compiled by the author based on the model calculation results.

The results show that the linear thresholds and LOWESS thresholds of most variables are highly consistent, suggesting a relatively stable first-order relationship between these variables and their SHAP contributions within the sample range. It should be emphasized that the thresholds identified in this study refer to marginal contribution turning points in model interpretation. They are statistically derived empirical transition points rather than strict causal optima, definitive sustainability boundaries, operational policy cutoffs, or absolute biophysical limits of the agricultural ecosystem. Therefore, these thresholds should be used as directional reference information for monitoring, diagnosis, and policy discussion and should be interpreted together with ecological monitoring, field-level evidence, and policy judgment before any operational management standard is considered.

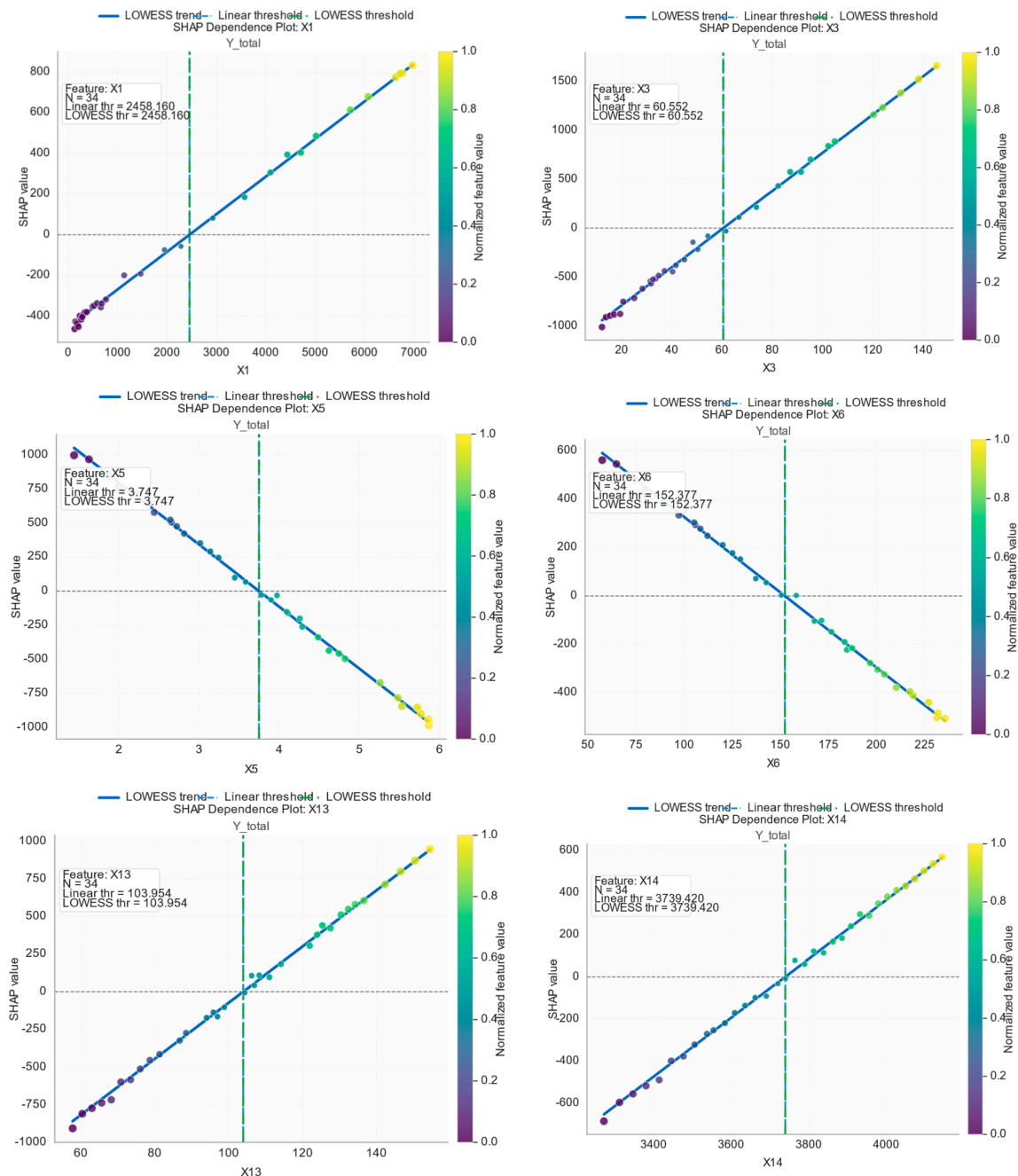


Figure 9. SHAP dependence plots of key variables for the total output of agriculture, forestry, animal husbandry, and fisheries in Yunnan. Figure note: The threshold refers to the turning point in the relationship between the SHAP value and the variable's original value at which the marginal contribution changes from positive to negative or from negative to positive.

5.6.1. Threshold Structure of Agriculture and Animal Husbandry

The agricultural model shows a dual-threshold structure. The detailed results are shown in Figure 10. On the one hand, rural electricity consumption, total agricultural machinery power, and total reservoir storage capacity show positive threshold effects, with thresholds of approximately 6.055 billion kWh, 19.8324 million kW, and 10.395 billion m³, respectively. Fixed asset investment, total sown area, rural power stations, and agricultural population also show positive effects.

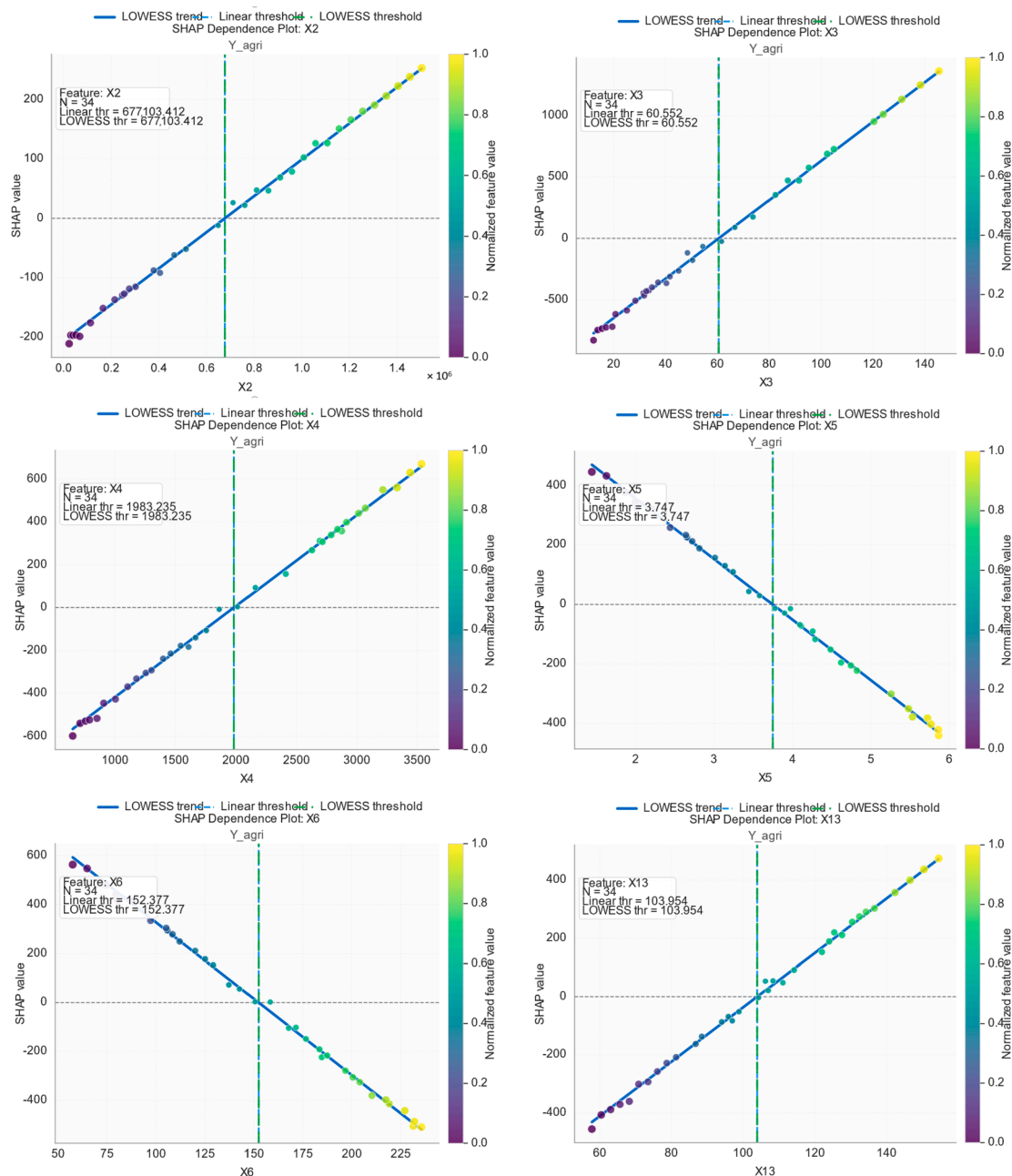


Figure 10. SHAP dependence plots of key variables for agriculture in Yunnan. Figure note: The threshold indicates the boundary in the association between the SHAP value and the original value of a variable, at which the marginal contribution transforms from positive to negative or from negative to positive.

On the other hand, fertilizer application and pesticide use have negative effects. The threshold for fertilizer application is approximately 1.5238 million tons, and the threshold for pesticide use is approximately 37,500 tons. These results indicate that agricultural output is associated with both positive infrastructure-related thresholds and negative chemical input constraints. The animal husbandry model shows a factor structure similar to that of agriculture.

5.6.2. Threshold Structure of Forestry

The forestry model presents a more differentiated threshold structure. The detailed results are shown in Figure 11. Fertilizer application has the largest mean absolute SHAP value in the forestry model, but its direction is negative. Rural electricity consumption and

pesticide use are also important variables. Pesticide use shows a positive threshold effect, with a threshold of approximately 3.75, while rural electricity consumption and agricultural population also show positive effects.

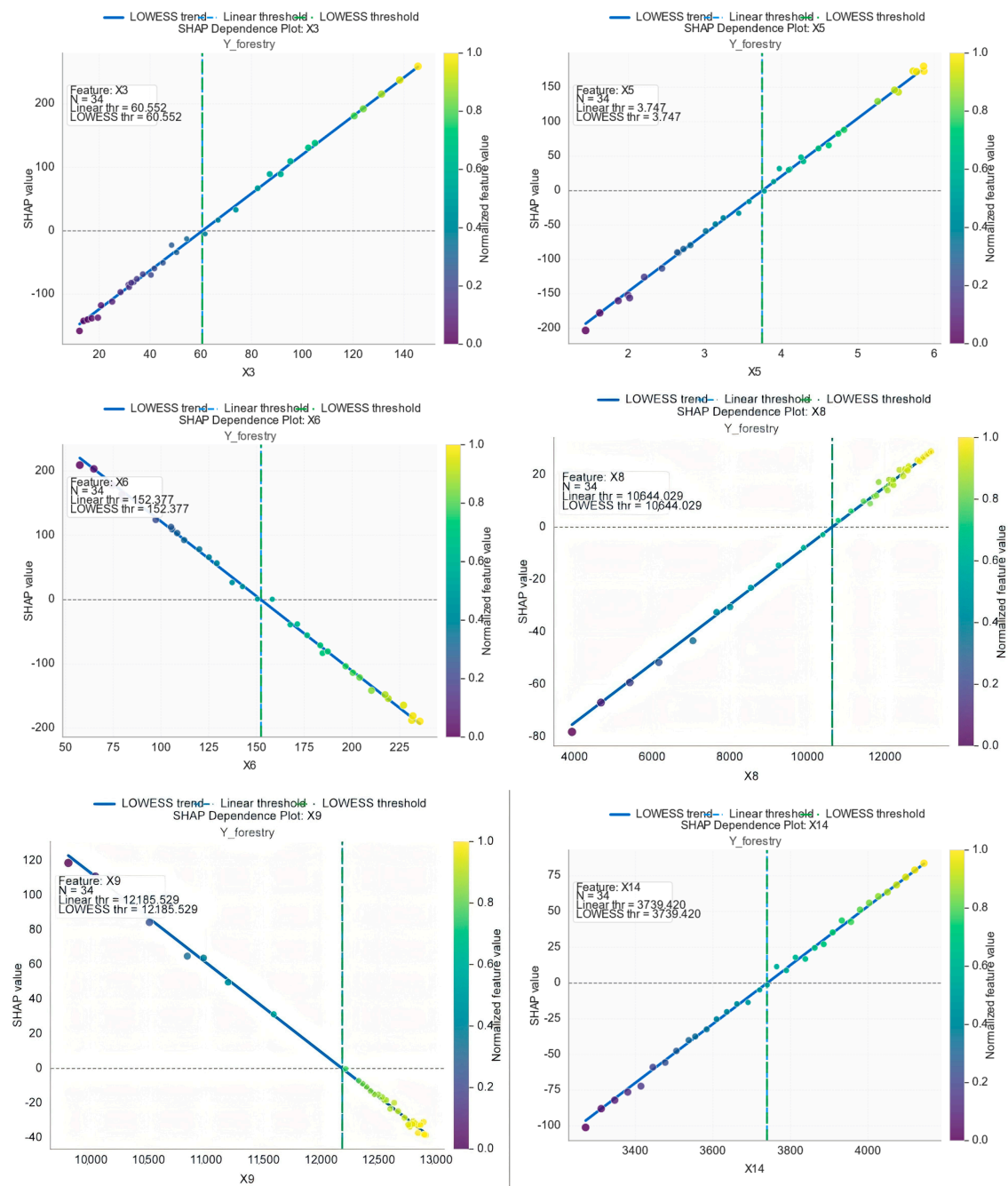


Figure 11. SHAP dependence plots of key variables for forestry in Yunnan. Figure note: The threshold indicates the boundary in the association between the SHAP value and the original value of a variable, at which the marginal contribution transforms from positive to negative or from negative to positive.

In contrast, fertilizer application, road accessibility, total sown area, total agricultural machinery power, and total reservoir storage capacity show negative effects in the forestry model. These results indicate that forestry does not follow the same factor structure as agriculture and animal husbandry.

5.7. Structural Boundaries of Yunnan's Agricultural Development

Combining the ridge regression results, SHAP importance ranking, and threshold identification, the empirical results show that the “limit” of Yunnan's agricultural development is not a single output ceiling. Instead, it is reflected in a composite structural boundary formed by three types of constraints.

First, there is an infrastructure carrying boundary. Rural electricity consumption and total reservoir storage capacity show stable positive thresholds across several models. Second, there is a factor allocation threshold. Agricultural machinery power, fixed asset investment, fiscal expenditure, and agricultural population influence output only when they reach certain levels of contribution. Third, there is an upper limit of chemical input contribution. Pesticide and fertilizer inputs show negative marginal contribution directions in the total, agricultural, and animal husbandry models.

Figure 12 further shows that different sectors have different sensitivities to these boundaries. Agriculture and animal husbandry show a dual structure of positive infrastructure thresholds and negative chemical input constraints. Forestry has a more specific structure, in which pest control, labor organization, and management-related factors play a more distinctive role.

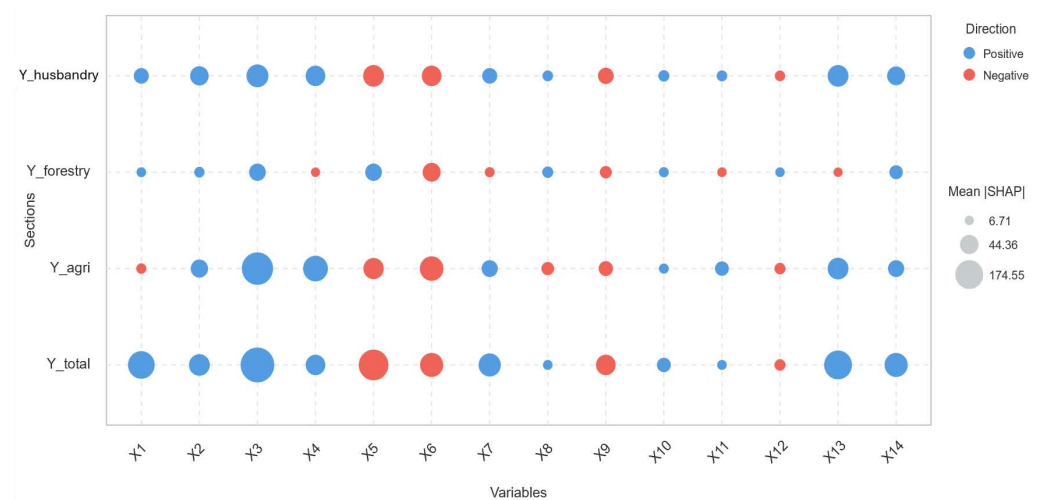


Figure 12. Threshold direction bubble diagram.

Therefore, the empirical results indicate that Yunnan's agricultural development limit should be understood as a composite constraint interval formed by infrastructure thresholds, factor allocation thresholds, and the marginal upper limits of chemical inputs.

6. Discussion

The study presents primary empirical results concerning sustainability science, ecological threshold theory, green agricultural change and policy systems like the Sustainable Development Goals and China's ecological civilization strategy. It aims to contribute to the literature on quantitatively determining the limitations of agricultural development under multiple constraints by going beyond traditional regional agricultural factor analysis.

6.1. From Agricultural Growth Factors to Sustainability Boundaries

An important result of this research is that agricultural development in Yunnan is influenced not only by the existence of production factors but also by the upper limits of their marginal effects. This finding is significant as it changes the research emphasis from “which factors accelerate agricultural growth” to “in what circumstances do these factors facilitate or restrict sustainable agricultural development.” Accordingly, the threshold

values reported in this study are better understood as empirical transition points in marginal contribution rather than definitive sustainability boundaries. They provide quantitative reference values for identifying potential bottlenecks and risk zones, but they should be combined with ecological monitoring, field-level evidence, and policy judgment before being used as regulatory limits.

Traditional studies on agricultural advancement focus chiefly on increasing the amounts of capital, labor, machinery, land and chemical products. But our research indicates that an increase in these factors is not always beneficial. The infrastructural parts, such as the utilization of rural electricity and the storage capacity of reservoirs, have positive threshold properties, i.e., they function well only when a certain quantity is reached. Conversely, the chemical inputs like pesticides and fertilizers often show a negative effect in most models, suggesting that further increases may lead to decreasing marginal benefits or affect sustainability.

This distinction is connected with the limit concept in sustainability science. The “limit” mentioned in this research is not an unchangeable physical upper limit, but a complicated boundary considering infrastructure capacity, factor allocation efficiency and ecological input constraints. Therefore, the study identifies agricultural development constraints by discovering empirical thresholds. This meets the requirement to link agricultural economic analysis with the whole sustainability issue of ensuring agricultural production within feasible ecological and resource limits.

6.2. Sustainability Implications of Infrastructure Thresholds

The results indicate that rural electricity consumption is the most stable and significant positive factor in both the total model and the three sectoral models. Its importance was verified by ridge regression and SHAP analysis. This suggests that electricity is not only a common production factor but also an essential precondition for agricultural advancement in hilly regions.

For Yunnan, with its complicated topography, scattered villages and unequal resource allocation, rural electricity is necessary for various aspects of agricultural work. It helps in irrigation, agricultural processing, refrigerated transportation, mechanical operations, livestock breeding houses and digital agricultural services. Hence, the value of 6.055 billion kWh should not be regarded as a planning quota or operational minimum requirement. Instead, it indicates a model-derived transition point at which rural electricity consumption begins to show a stable positive fitted contribution within the observed sample. It can therefore serve as a directional reference for discussing agricultural infrastructure support, but not as a direct standard for energy allocation.

Similarly, the total reservoir storage capacity is an important aspect in the total, agricultural and animal husbandry systems. Water conservancy is particularly crucial for Yunnan because of its seasonal droughts, unequal rainfall distribution and mountainous terrain, which result in serious threats to agricultural production. The reservoir storage capacity value of about 10.395 billion m³ indicates a model-derived reference point for water resource support capacity, rather than a fixed hydrological planning standard. Otherwise, agricultural production will be more affected by climate variations and resource instability.

These results are closely related to the objective of enhancing agricultural resistance to change. In the situation of climate change, infrastructures such as electricity supply and water storage are not only important for economic development but also reflect adaptability capacities. They indicate whether mountainous agricultural systems can maintain stable production and decrease their sensitivity to environmental disturbances.

6.3. Chemical Input Constraints and Green Agricultural Transformation

Another significant result in both the overall model and the three sectoral models is that the use of pesticides and the application of fertilizers are important factors, but their influences are mainly negative in the total, agricultural and animal husbandry models. The SHAP values indicate that these variables have high importance, whereas the ridge regression coefficients show negative impacts in most models. The threshold analysis also indicates that pesticide use of about 37,500 tons and fertilizer application of about 1.5238 million tons represent statistically derived marginal contribution turning points within the sample period.

This result has important effects on the green transformation of agriculture. During the early stages of agricultural progress, the application of fertilizers and pesticides can enhance crop yields by supplying nutrients and preventing pest and disease infestations. Nevertheless, when the quantity of chemical inputs exceeds a particular level, the additional advantages may decrease and the associated environmental costs may rise. The overuse of fertilizers and pesticides may lead to soil deterioration, water contamination, reduction in biodiversity, increase in greenhouse gas emissions and food safety hazards. Consequently, the negative impact of chemical inputs found in our study indicates that agricultural development in Yunnan cannot depend on the expansion of high-intensity chemical input usage.

This outcome agrees with the shift from intensive agriculture requiring large amounts of inputs to efficient and environmentally friendly agriculture. The policy implication is therefore not to simply reduce all chemical inputs but to promote precision fertilization, integrated pest management, soil testing and formula fertilization, organic substitution, biological control techniques, and circular agricultural practices. In this respect, the detected thresholds provide directional evidence for identifying potential risk zones of excessive dependence on conventional chemical inputs, rather than numerical standards for setting maximum allowable input levels.

The forestry model requires particular consideration as pesticide use has a positive coefficient in forestry, different from that in agriculture and animal husbandry. It suggests that pesticide use in forestry may imply the necessity of pest and disease control under certain operational conditions, particularly for economic forests and forest product management. Consequently, a uniform input-reduction policy should not be applied to all sectors. Instead, differentiated chemical input management is necessary. Therefore, chemical inputs in crop farming and animal husbandry should be reduced and substituted, while in forestry, precise and ecological pest control should be emphasized instead of merely reducing them.

6.4. Sectoral Heterogeneity and Differentiated Sustainability Strategies

The empirical results indicate that there is heterogeneity among different agricultural sectors. Both agriculture and animal husbandry have similar structures, depending on electricity, machinery, water conservancy, investment and labor conditions, but are limited by pesticide and fertilizer use. In contrast, forestry has a more scattered factor structure and is less influenced by the use of machinery and reservoir storage capacity.

This multi-disciplinary field is crucial for sustainable policies. If agriculture is considered as a whole uniform system, the policies might overlook the particular difficulties of each branch. For crop cultivation, it is essential to improve electric power, irrigation, and mechanization and substitute the application of green resources. In animal husbandry, it is still important to ensure a continuous supply of energy and water, but the policies should also take into consideration fodder production, breeding sheds, waste disposal, disease control and the ability of livestock systems. For forestry, the development plans should consider long-term production, ecological functions, pest control requirements,

labor organization and the equilibrium between economic forest development and the preservation of biological diversity.

Thus, the results indicate the necessity of an individualized administrative method. The progressive development of agriculture in Yunnan cannot rely on an increase in inputs; rather, it must rely on the adjustment of thresholds according to different fields. Each field has its own set of advantageous thresholds and restrictive upper limits. Detecting these distinctions may enhance the accuracy of agricultural policies and prevent the improper or uneconomic distribution of resources.

6.5. Connection with SDGs and Broader Sustainability Frameworks

Initially, this research is connected with SDG 2: Zero Hunger, particularly the objective of promoting sustainable agriculture. By determining the infrastructure and factor limits in both the total model and the three sector models, this study offers evidence for enhancing agricultural productivity and stability in mountainous regions. Electricity, water conservancy, machinery and investment are not only production factors but also the basis for ensuring a stable food and agricultural supply.

Furthermore, the results are associated with SDG 12: Responsible Consumption and Production. The adverse impacts of fertilizer and pesticide applications indicate that the agricultural system in Yunnan should decrease its reliance on intensive chemical input strategies. It is necessary to adopt more efficient and environmentally friendly production techniques to enhance the utilization of resources and lessen ecological expenses.

Furthermore, this research is related to SDG 13: Climate Action. Water storage capacity, rural electricity facilities and mechanization affect the stability of agricultural systems in the face of climate change. In mountainous regions that are susceptible to drought and severe weather, strengthening the infrastructure level can improve adaptive ability.

Fourth, this study is closely related to SDG 15: Life on Land. The use of excessive chemical inputs may endanger soil health, biodiversity and ecosystem services. Determining the upper limits of chemical inputs provides empirical evidence for the protection of land ecosystems along with agricultural development.

In the Chinese governmental framework, the results are consistent with the ecological civilization policy and the national aim of accelerating green agricultural progress. The agricultural policy in China stresses a decrease in fertilizer and pesticide application, resource conservation, ecological preservation and an improvement in rural areas. The present study offers statistical support for these policy directions at the provincial level.

6.6. Theoretical Contributions

This study makes three main theoretical contributions.

First, it expands the investigation of agricultural growth from the determination of factors to the determination of boundaries. Most previous research has considered whether fiscal funding, investment, manpower, machines, water supply or chemical inputs influence agricultural production. Our study also investigates where the additional effect of these factors is significant. Thus, it introduces the concept of “development limitations” in the empirical analysis of agricultural growth.

Furthermore, this research combines ridge regression, Bootstrap approximate inference and SHAP interpretation in one analytical system. Ridge regression deals with the multicollinearity problem among agricultural development indicators, Bootstrap inference enhances the reliability of statistical interpretation, and SHAP analysis determines the significance and contribution levels of the variables. This system offers a practical methodological way to investigate agricultural networks with multidimensional and highly correlated macro-indicators.

Furthermore, the investigation compares the development limitations of each agricultural sector. Through the analysis of the total model and the three sectoral models for agriculture, forestry and animal husbandry, it is found that the agricultural sectors have different restriction patterns. This helps to obtain a clearer picture of the sustainable transformation of agriculture, particularly in mountainous and ecologically sensitive areas.

6.7. Practical Implications

The following practical implications should be interpreted as directional policy guidance rather than operational cutoff rules. The thresholds identified in this study are model-derived marginal contribution transition points. They should not be directly converted into administrative standards, regulatory red lines, fixed input quotas, or resource allocation targets without further ecological monitoring, field validation, and local policy assessment.

First, rural energy infrastructure can be regarded as an important supporting condition for sustainable agricultural development. The stable importance of rural electricity consumption in the aggregate and sectoral models suggests that improving electricity supply reliability, agricultural electrification, electric agricultural machinery, and energy accessibility may support agricultural modernization in Yunnan. However, the electricity-related threshold should be understood as a directional reference for identifying infrastructure support stages rather than as a fixed planning requirement.

Second, water conservancy capacity remains an important reference dimension for improving agricultural resilience. Reservoir storage capacity is positively associated with total output, agriculture, and animal husbandry. This suggests that water storage, irrigation efficiency, drought-resistance capacity, and water-saving technologies deserve continued policy attention. Nevertheless, the reservoir storage threshold should be interpreted as a model-derived reference point for discussing water resource support capacity, not as a direct standard for water allocation or infrastructure construction.

Third, chemical input management should remain an important component of green agricultural transformation. The estimated effects of pesticides and fertilizers in most models suggest that Yunnan's future agricultural growth is unlikely to be sustainably supported by continued expansion of high-intensity chemical inputs. The corresponding thresholds may help identify potential warning zones of weakening marginal contribution, but they should not be interpreted as direct regulatory limits. Any operational standard for pesticide or fertilizer management should be based on additional evidence from soil testing, crop structure, pest control demand, ecological monitoring, and local agronomic conditions.

Fourth, policy discussion should be differentiated across agricultural sectors. Agriculture and animal husbandry show stronger dependence on electricity, machinery, water conservancy, and green input management, whereas forestry shows a more distinct pattern related to pest control, labor organization, ecological protection, and long-term forest resource management. Therefore, the findings support sector-specific monitoring and adaptive management rather than uniform threshold-based control rules.

Fifth, threshold information can be used as a preliminary screening reference to identify possible bottlenecks, risk zones, and priority areas for further investigation. Instead of increasing all inputs equally or treating the estimated values as fixed cutoffs, policymakers may use the threshold results to support monitoring prioritization, policy discussion, and adaptive management. The practical value of the thresholds lies in helping identify where marginal contributions may change, not in defining enforceable operational boundaries.

6.8. Generalizability and Transferability

Though the present research concentrates on Yunnan Province, the analytical method is also applicable to other areas. Various mountainous, plateau and ecologically sensitive

farming districts encounter similar problems such as divided fields, unbalanced water supply, high construction expenses, shortage of labor force and the necessity to decrease the use of chemical fertilizers. Consequently, the procedure of “stable estimation–importance ranking–threshold determination–comparative analysis” can be used in other regions or agricultural systems.

Nevertheless, the particular threshold values determined in this research cannot be applied directly to other provinces or countries. These values are highly dependent on the resource endowments, industrial structure, historical development process and statistical sample period of Yunnan. What can be extended is not the numerical threshold itself but the evaluation method: sustainable agricultural development should be assessed by determining infrastructure thresholds, factor allocation thresholds and maximum limits of ecological inputs.

Our future research might use the same method to investigate other mountainous areas in China, like Guizhou, Sichuan, Guangxi or Tibet, or in other developing areas facing similar environmental issues. Further comparisons can determine if the structural limitations found in Yunnan are local characteristics or indicate a general trend in agricultural development under multiple restrictions.

6.9. Limitations and Future Research

The analysis is based on provincial annual data from 1990 to 2023. Although this long-term series reflects important changes in Yunnan’s agricultural development, the number of observations remains limited. Future studies may employ prefecture-level, county-level, or farm-level data to enhance spatial resolution and increase sample size. Second, provincial aggregate data may conceal differences across agroecological zones, production systems, and local infrastructure conditions. Third, omitted variables such as climate shocks, market prices, technological adoption, land quality, environmental regulation, and policy changes may affect the estimated relationships. Fourth, the identified thresholds are model-based empirical transition points rather than causal ecological limits. Future research should combine socioeconomic indicators with spatial ecological indicators and direct environmental outcome variables to test whether similar threshold structures exist in other mountainous agricultural regions. For policymakers, the results suggest that agricultural modernization in Yunnan should focus on strengthening energy and water infrastructure, improving factor allocation efficiency, and reducing reliance on extensive chemical input expansion through precision and ecological management.

7. Conclusions

Based on provincial annual data from Yunnan Province from 1990 to 2023, this study constructed a multidimensional analytical framework to identify the limits of agricultural development under factor constraints. Ridge regression was used to address multicollinearity among explanatory variables, Bootstrap resampling was applied for approximate coefficient inference, and SHAP analysis was introduced to rank key driving factors and identify marginal contribution thresholds.

The study examined one aggregate model for the total output value of agriculture, forestry, animal husbandry and fisheries and three sectoral models for agriculture, forestry, and animal husbandry.

The empirical results lead to the following main conclusions.

First, Yunnan’s agricultural development is constrained by multiple factors, including fiscal and investment support, agricultural production inputs, rural infrastructure, water conservancy conditions, and labor population factors. The models show strong explanatory

power, with R^2 values of 0.997 for the total output model, 0.994 for the agricultural model, 0.995 for the forestry model, and 0.978 for the animal husbandry model.

Second, rural electricity consumption is the most stable and core positive factor across the aggregate and three sectoral models. Both the ridge regression and SHAP results confirm its dominant role. The identified threshold of rural electricity consumption is approximately 6.055 billion kWh, indicating that rural energy infrastructure is a fundamental condition for agricultural modernization in Yunnan.

Third, reservoir storage capacity, fixed asset investment, fiscal expenditure, agricultural machinery, and agricultural population form an important support system for agricultural development. The results identify several key positive thresholds, including total reservoir storage capacity of approximately 10.395 billion m^3 and total agricultural machinery power of approximately 19.8324 million kW.

Fourth, across the aggregate and sectoral models, pesticide use and fertilizer application are important constraint variables, with mainly negative effects in the total, agricultural, and animal husbandry models. The identified statistically derived marginal contribution thresholds are approximately 37,500 tons for pesticide use and 1.5238 million tons for fertilizer application. These findings should not be interpreted as evidence that all chemical inputs are harmful, but they suggest that Yunnan's future agricultural growth is unlikely to be achieved sustainably through continued expansion of high-intensity chemical input use.

Fifth, significant heterogeneity exists among the three sectoral models. Agriculture and animal husbandry are more dependent on electricity, machinery, water conservancy, investment, and labor conditions while also being constrained by pesticides and fertilizers. Forestry has a more differentiated factor structure, in which pesticide use shows a positive effect, whereas fertilizer application shows a negative effect.

Overall, the sustainability-related development constraints of Yunnan's agriculture should not be understood as a single output ceiling but as a composite structural boundary. It consists of infrastructure support thresholds, factor allocation thresholds, and upper limits of chemical-input-related marginal contribution. This finding expands the study of agricultural growth from the identification of driving factors to the quantitative identification of empirical transition points under multidimensional constraints.

The study contributes to sustainability science by linking agricultural development constraints with statistically derived marginal contribution thresholds, green transformation, and sector-specific policy design. It also provides practical evidence for supporting SDG 2, SDG 12, SDG 13, and SDG 15, as well as China's ecological civilization and green agricultural development strategies. For mountainous and ecologically fragile regions, sustainable agricultural development should focus not on unlimited input expansion but on identifying key bottlenecks, strengthening infrastructure thresholds, improving factor allocation efficiency, and controlling ecological input pressures. Several limitations should be noted. First, the analysis is based on 34 annual provincial observations, and the sample size remains limited despite the long time span. Second, provincial aggregate data may conceal differences across prefectures, counties, farms, and agroecological zones. Third, omitted variables such as climate shocks, market prices, technological adoption, land quality, and policy changes may affect the estimated relationships. Fourth, the identified thresholds are model-based empirical transition points rather than causal ecological limits, operational policy cutoffs, or directly enforceable ecological standards. Future research should use prefecture-level or county-level panel data, spatial ecological indicators, direct environmental outcome variables, and field validation to test whether similar threshold structures exist in other mountainous agricultural regions and whether they can support more specific management applications.

Author Contributions: Concept and design: Z.W.; data collection and analysis: Z.W.; drafting of the article: Z.W.; critical revision of the article for important intellectual content: L.R.; study supervision: L.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Yunnan Fundamental Research Projects (grant NO. 202301AU070089), which is financially aided by the Yunnan Provincial Department of Science and Technology.

Institutional Review Board Statement: Not applicable. This study did not involve human participants, animal subjects, or individually identifiable personal data. The analysis was based entirely on publicly available, non-identifiable aggregate statistical data from official public databases.

Informed Consent Statement: Not applicable. This study did not involve human participants, and no individually identifiable information was collected or analyzed.

Data Availability Statement: The macro-level agricultural and economic data used in this research are accessible in the public databases of the National Bureau of Statistics of China (<https://www.stats.gov.cn/>), the Yunnan Provincial Bureau of Statistics (<https://stats.yn.gov.cn/>), and the Ministry of Agriculture and Rural Affairs of China (<http://www.moa.gov.cn.>).

Conflicts of Interest: The authors report there are no competing interests to declare.

References

1. Lampridi, M.G.; Sørensen, C.G.; Bochtis, D. Agricultural Sustainability: A Review of Concepts and Methods. *Sustainability* **2019**, *11*, 5120. [CrossRef]
2. Arfini, F.; Antonioli, F.; Cozzi, E.; Donati, M.; Guareschi, M.; Mancini, M.C.; Veneziani, M. Sustainability, Innovation and Rural Development: The Case of Parmigiano-Reggiano PDO. *Sustainability* **2019**, *11*, 4978. [CrossRef]
3. Stellmacher, T.; Kelboro, G. Family Farms, Agricultural Productivity, and the Terrain of Food (In)security in Ethiopia. *Sustainability* **2019**, *11*, 4981. [CrossRef]
4. Jiang, C.; Liao, M.; Hu, Y.; Wang, Y.; Guan, Y.; Wang, B.; Xiao, Y. Policy-Driven Industrial Transformation and Associated Agriculture–Ecology–Economy Nexus Shifts within Social–Ecological System: A Case of the Greenest Coastal Province of China. *Front. Environ. Sci.* **2026**, *14*, 1799981. [CrossRef]
5. Carswell, A.M.; Willcock, S.; Blackwell, M.S.A.; Upadhayay, H.R.; Harris, P.; McAuliffe, G.; Neal, A.L.; Rivero, M.J.; Cardenas, L.M.; Haefele, S.M.; et al. Agricultural Practices Can Threaten Soil Resilience through Changing Feedback Loops. *npj Sustain. Agric.* **2025**, *3*, 56. [CrossRef] [PubMed]
6. Rockström, J.; Steffen, W.; Noone, K.; Persson, Å.; Chapin, F.S., III; Lambin, E.F.; Lenton, T.M.; Scheffer, M.; Folke, C.; Schellnhuber, H.J.; et al. A Safe Operating Space for Humanity. *Nature* **2009**, *461*, 472–475. [CrossRef] [PubMed]
7. Steffen, W.; Richardson, K.; Rockström, J.; Cornell, S.E.; Fetzer, I.; Bennett, E.M.; Biggs, R.; Carpenter, S.R.; De Vries, W.; De Wit, C.A.; et al. Planetary Boundaries: Guiding Human Development on a Changing Planet. *Science* **2015**, *347*, 1259855. [CrossRef] [PubMed]
8. Tilman, D.; Cassman, K.G.; Matson, P.A.; Naylor, R.; Polasky, S. Agricultural Sustainability and Intensive Production Practices. *Nature* **2002**, *418*, 671–678. [CrossRef] [PubMed]
9. Foley, J.A.; Ramankutty, N.; Brauman, K.A.; Cassidy, E.S.; Gerber, J.S.; Johnston, M.; Mueller, N.D.; O’Connell, C.; Ray, D.K.; West, P.C.; et al. Solutions for a Cultivated Planet. *Nature* **2011**, *478*, 337–342. [CrossRef] [PubMed]
10. Duan, Y.; Peng, H.; Chen, D.; Zhang, R.; Wu, X.; He, Z. Whether Balance between Grain Production Resilience and Ecosystem Services? Evidence from Shandong Section of the Yellow River Irrigation District in China. *Environ. Dev. Sustain.* **2026**. [CrossRef]
11. Ma, Z.; Li, W.; Li, Y.; Li, J.; Wang, J. Research on the Current Situation and Sustainable Development Measures for Urban–Rural Water Supply Integration in Yunnan Province, China. *Water* **2024**, *16*, 3232. [CrossRef]
12. Song, J.; Wu, J.; Li, X.; Ning, Q.; Liu, W.; Liu, Z.; Yuan, G.; Huang, H.; Guo, Y.; Zhuang, Y. Tobacco Cultivation in Yunnan: An Integrated Analysis of Economic Trends and Diversification. *Front. Sustain. Food Syst.* **2026**, *9*, 1622916. [CrossRef]
13. Solís, D.; Alvarez, S.; Nguyen, L. Agriculture Under Pressure: The Economic, Environmental, and Development Drivers Transforming Florida Agriculture. *Agriculture* **2026**, *16*, 661. [CrossRef]
14. Huang, Q.; Guo, W.; Wang, Y. A Study of the Impact of New Quality Productive Forces on Agricultural Modernization: Empirical Evidence from China. *Agriculture* **2024**, *14*, 1935. [CrossRef]
15. Martinez, D.A.; Loening, U.E.; Graham, M.C.; Gathorne-Hardy, A. When the Medicine Feeds the Problem; Do Nitrogen Fertilisers and Pesticides Enhance the Nutritional Quality of Crops for Their Pests and Pathogens? *Front. Sustain. Food Syst.* **2021**, *5*, 701310. [CrossRef]

16. Velten, S.; Leventon, J.; Jager, N.; Newig, J. What Is Sustainable Agriculture? A Systematic Review. *Sustainability* **2015**, *7*, 7833–7865. [\[CrossRef\]](#)
17. Bathaei, A.; Štreimikienė, D. A Systematic Review of Agricultural Sustainability Indicators. *Agriculture* **2023**, *13*, 241. [\[CrossRef\]](#)
18. Wang, G.; Qian, Z.; Deng, X. Analysis of Environmental Policy and the Performance of Sustainable Agricultural Development in China. *Sustainability* **2020**, *12*, 10453. [\[CrossRef\]](#)
19. Wu, Y.; Wang, E.; Miao, C. Fertilizer Use in China: The Role of Agricultural Support Policies. *Sustainability* **2019**, *11*, 4391. [\[CrossRef\]](#)
20. Ren, J. Developing New-Quality Productive Forces for China's Farmland: Connotation, Challenges, and Strategies. *Sustainability* **2025**, *17*, 11220. [\[CrossRef\]](#)
21. Liu, F.; Hu, W.; Min, M.; Liu, T.; Li, S.; Rong, Q. How Does Road Networks Impact Cultivated Land Multifunctionality? The Role of Land Fragmentation: Evidence from Hubei Province, China. *Front. Sustain. Food Syst.* **2025**, *9*, 1686176. [\[CrossRef\]](#)
22. Sapkota, T.B.; Singh, B. India's Fertilizer Policies: Implications for Food Security, Environmental Sustainability, and Climate Change. *Reg. Environ. Change* **2025**, *25*, 63. [\[CrossRef\]](#)
23. Tang, F.H.M.; Wyckhuys, K.A.G.; Li, Z.; Maggi, F.; Silva, V. Transboundary Impacts of Pesticide Use in Food Production. *Nat. Rev. Earth Environ.* **2025**, *6*, 383–400. [\[CrossRef\]](#)
24. Lovell, S.T. Multifunctional Urban Agriculture for Sustainable Land Use Planning in the United States. *Sustainability* **2010**, *2*, 2499–2522. [\[CrossRef\]](#)
25. Gullino, P.; Beccaro, G.L.; Larcher, F. Assessing and Monitoring the Sustainability in Rural World Heritage Sites. *Sustainability* **2015**, *7*, 14186–14210. [\[CrossRef\]](#)
26. Keshyagol, K.; Vanarotti, M.; Madiwal, S.; Kulkarni, P.; Bongale, A.M. Sustainable Farming Systems Integrating Soil Health Water Use Efficiency and Crop Productivity for Climate Resilience. *Discov. Sustain.* **2026**, *7*, 481. [\[CrossRef\]](#)
27. Onyenma, G.C.; Selejio, O.; Ekunwe, P.A.; Alavaisha, E. Estimating the Effects of Climate Variables on Agricultural Total Factor Productivity Growth in Sub-Saharan Africa. *Environ. Resour. Econ.* **2026**, *89*, 28. [\[CrossRef\]](#)
28. Augustine, P.; John, N. Detection and Dealing of Multicollinearity Using a New Ridge Estimator. In *Proceedings of International Conference on Computing Systems and Intelligent Applications, ComSIA 2025*; Jaiswal, A., Anand, S., Hassanien, A.E., Azar, A.T., Eds.; Lecture Notes in Networks and Systems; Springer: Singapore, 2026; Volume 1501. [\[CrossRef\]](#)
29. Yang, C.; Guan, X.; Xu, Q.; Xing, W.; Chen, X.; Chen, J.; Jia, P. How Can SHAP (SHapley Additive exPlanations) Interpretations Improve Deep Learning Based Urban Cellular Automata Model? *Comput. Environ. Urban Syst.* **2024**, *111*, 102133. [\[CrossRef\]](#)

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.