*Original Research*

# Dynamic Changes of the Mu Us Sandy Land and Relevant Causes

**Xianghui Wang, Yan Du, Zhiruo Wang, Yang Zhang\***

School of Economics and Management, Beijing Forestry University, Beijing 100083, China

## Abstract

In this research, panel data and variance analysis were used to study the dynamic changes of the Mu Us sandy land from 2000 to 2015 along with relevant causes. It was found that the forest coverage in the relevant areas of the Mu Us sandy land increased continuously, and the corresponding growth was evident in some social and economic factors. Through a fixed-effects model, it was concluded that there was a remarkable positive correlation between GDP and forest coverage in this region; that is, for every 1% increase in GDP, the forest coverage in this region would increase by 1.876%. In addition, ANOVA showed that policies, including The National Plan for Combating Desertification (2005-2010) and The National Plan for Combating Desertification (2011-2020) also played a significant part in increasing forest coverage.

## Introduction

As one of the most severe ecological problems China faces, desertification is also a key yet difficult point of ecological protection. It deteriorates the ecological environment, aggravates poverty, and overwhelms the space needed for the existence and development of human beings, which eventually leads to a vicious cycle between economic growth and environmental degradation. However, the principle that "lucid waters and lush mountains are invaluable assets" has become China's strategy to promote ecological construction in today's society. This principle posits that to protect the environment is to protect the productive forces. We require both economic growth and environmental protection, which is conducive to resolving the contradiction between them and forming a virtuous cycle of economy and environment with the improvement of production. In this context, the achievements in the Mu Us sandy land have attracted considerable attention.

The Mu Us sandy land is located north of the Great Wall of Yulin City, Shaanxi Province, China, with an area of approximately 42,200 square kilometers. As a transition zone between temperate desert and steppe in the natural zone, the Mu Us sandy land's ecological environment is fragile and sensitive to climate change. However, the Mu Us sandy land is suitable for vegetation growth and human survival with 400-440 mm annual rainfall in its southeast and abundant surface water and underground water [1]. The government has implemented a number of ecological governance programs since 2000. Furthermore, the Mu Us sandy land, which is rich in underground freshwater

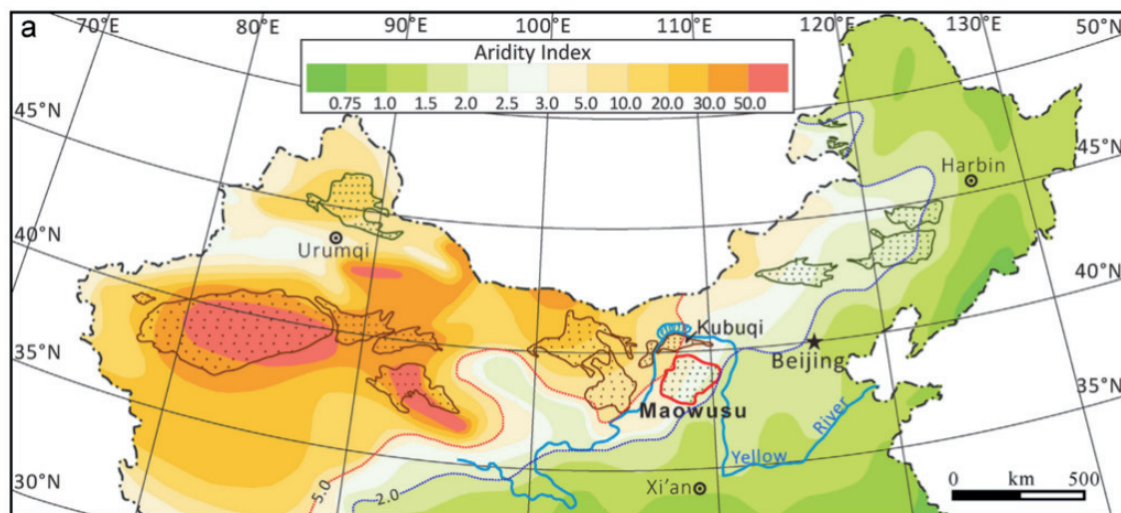---
*\*e-mail: yangzhang@bjfu.edu.cn*

Fig. 1. Aridity index in region researched.

resources, has become a major location for agricultural development with the development of agricultural technology [2]. Researchers introduced the location and aridity index of the Mu Us sandy land when studying its spatial landscape pattern and its influencing factors [3], as shown in Fig. 1, where "Maowusu" is the Mu Us sandy land.

In recent years, the dynamic changes in the Mu Us sandy land have attracted much attention. With the support of remote sensing technology and natural geographic information systems, researchers have made considerable advances in research from the initial static sandy area measurement to decades of research on the status of desertification, most of which supports the conclusion that the Mu Us sandy land shows a reversal trend. Some scholars believe that the Mu Us sandy land underwent rapid development from the 1950s to the 1970s with an overall reversal process after the 1980s due to policy factors [4, 5]. Other follow-up studies have more rigorously proven that the Mu Us Sandy Land enjoyed sound development for 30 years from 1990 to 2010 [6]. Further research showed the actual natural landscape changes in the Mu Us sandy land and presented the spatial-temporal changes in the Mu Us sandy land in a visual form through satelite technology [7, 8].

The discussion on the causes of the dynamic changes in the Mu Us sandy land mainly lies in the following two aspects: social and natural factors. In early climate research, scholars unanimously believed that the region was growing warmer and dryer [9-11], which was not consistent with desertification reversal. Therefore, some scholars began to shift their focus from natural factors to social ones. However, only a few studies simply compared the corresponding changes of vegetation coverage with land use patterns, industrial structure, and policy as influencing factors [5]. Furthermore, selecting the Mu Us Sandy Land as a case and dividing it into eastern, central, and western regions, some

scholars emphasized the influence of policy on the Mu Us sandy land through disparities in different regions [12].

The climate has undergone remarkable changes since the beginning of 21st century. Under these circumstances, studies suggested that more rainfall and rising temperatures played decisive roles in the reversal of desertification in Mu Us and still used vegetation coverage or forest coverage as the response variable to represent disparity of desertification [13, 14]. In terms of actual landscape measurement, most studies indicated that the vegetation coverage of the Mu Us sandy land showed overall improvement and a partial deterioration according to analysis through satellite remote sensing [15], from which it can be further verified that natural factors had little effect on the reversal of the deterioration [16]. Still, other studies showed that vegetation coverage was mainly influenced by wind erosion, considering no remarkable changes in temperature and rainfall in the Mu Us sandy land [17].

Due to the unremarkable effect of natural factors on the desertification reversal in the Mu Us sandy land, many scholars have shifted their attention back to social factors. Some researchers found through analysis of changes of land use and landscape patterns that compared with the climate, social factors, including per capita net income, population, GDP, and gross output value of agriculture, forestry, and animal husbandry in social activities had significant impacts on the reversal [18]. Based on the policy of returning farmland to forest, some studies found that human arable land activities strongly interfered with the ecosystem of the Mu Us Sandy Land using satellite monitoring data, and the protection policies implemented by managers could effectively limit the source of this interference [19]. Subsequently, some studies based on the work mentioned above indicated that a social factor, namely the three-carbon economy industry of biomass power plants, was taken advantage of to alter the Mu Us sandy

land, thus posing a win-win strategy to promote the reversal of desertification and economic growth [20].

Above all, the influence of social and natural factors on the dynamic changes of Mu Us sandy land has been discussed. The analysis methods have been changed from the qualitative analysis to descriptive statistics based on satellite remote sensing technology in some studies [21]. It can be concluded that many studies have put forward subjective judgments and corresponding suggestions only through literature reviews without using qualified factors [22, 23]. Considering the aspects mentioned above, the current paper analyzes the reasons for dynamic changes in the Mu Us sandy land and provides specific and reliable suggestions while considering social factors (economy, policies) and natural factors (rainfall, temperature) as the influencing sources. Methods introduced in this paper, including panel data and variance analysis, were used to qualify the statistics and analysis.

## Materials and Methods

### Data Acquisition

Two issues needed to be addressed before the selection of the database.

(1) Changes in statistical caliber. After 2015, changes to the yearbook statistical caliber made it difficult to find available panel data. In addition, the data were only updated up to 2018 after the change of statistical caliber, contributing to an incomplete stage. Therefore, the time period selected for this research ended in 2015.

(2) Interference with existing policies. Some major national ecological projects, such as the Natural Forest Protection Project and the Project of Returning Farmland to Forests and Grasslands, had been implemented prior to 2000. In order to avoid additional policy effects of such ecological projects, 2000 was selected in this study as the starting point of the research.

In 2005, the State Council approved The National Plan for Desertification Control (2005-2010), and then The National Plan for Desertification Control (2011-2020) [24]. The years 2005 and 2011 were important time points for the national fundamental layout of sand prevention and control, thus dividing three different planning levels. Therefore, relevant data of some counties and cities around the Mu Us sandy land were selected for analysis. The data sources were from China's economic and social big data platforms, particularly the Shaanxi Economic and Social Development Statistical Database, the Inner Mongolia Autonomous Region Economic and Social Development Statistical Database, and the Ningxia Hui Autonomous Region Economic and Social Development Statistical Database from 2000 to 2015. A total of 160 samples were selected for the research. The relevant states were all within the Mu Us sandy land, including Etuoke Front Banner, Etuoke Banner,

Wushen Banner, Ejin Horo Banner, Yanchi, Yuyang, Shenmu, Hengshan, Jingbian, and Dingbian.

### Descriptions of the Variables

On the basis of early research, the selected variables were as follows:

Forest coverage rate (f_ratio) refers to the percentage of forest area in the total land area of a certain region. It is an important index that reflects the richness of forest resources and ecological balance. When calculating the forest coverage rate, forest areas include arbor woodland areas; bamboo woodland areas with a canopy density of 0.2 or more; shrub woodland areas specially stipulated by the nation; farmland forest networks; and coverage areas of forest around villages, roads, water areas, and houses.

Annual rainfall (rain) refers to the depth of the accumulated rain that falls to the horizontal surface without leakage, loss, or evaporation for one year. It is measured in millimeters (mm).

Average annual temperature (temp) refers to the temperature of the air, measured by a blind box 1.5 meters above the ground, and is measured in degrees Celsius. The average annual temperature is obtained by dividing the sum of 12 months of average temperature by 12, and the monthly average temperature is calculated

Table 1. Statistics for variables.

| Variable | Transformation | Formula | p-value |
|---|---|---|---|
| gdp | Square | gdp^2 | 0.000*** |
| | Identity | gdp | 0.000*** |
| | Square root | sqrt(gdp) | 0.001*** |
| | Log | log(gdp) | 0.000*** |
| | Inverse | 1/gdp | 0.000*** |
| prod | Square | prod^2 | 0.000*** |
| | Identity | prod | 0.000*** |
| | Square root | sqrt(prod) | 0.000*** |
| | Log | log(prod) | 0.000*** |
| | Inverse | 1/prod | 0.000*** |
| inc | Square | inc^2 | 0.000*** |
| | Identity | inc | 0.000*** |
| | Square root | sqrt(inc) | 0.000*** |
| | Log | log(inc) | 0.000*** |
| | Inverse | 1/inc | 0.000*** |

[1] Note: *** $p<0.01$ ** $p<0.05$ * $p<0.1$.

[2] Note: gdp: regional GDP. prod: the total output of regional agriculture, forestry, animal husbandry, and fisheries. inc: per capita disposable income of permanent residents in rural pastoral areas.

by adding the average temperature of each day of the entire month and then dividing by the number of days in the month.

As shown in Table 1, gdp, prod, and inc refer to regional Gross Domestic Products, the total output value of agriculture, forestry, animal husbandry and fishery in pastoral region, the per capita disposable income of rural residents in pastoral region respectively. The various statistical transformations to gdp, prod, and inc were consistent with the normal distribution in the Skewness-Kurtosis test (p<0.01), but for convenient explanation, the logarithmic value (hereinafter referred to as "log value") of the regional GDP (lngdp); the log value of the total output value of agriculture, forestry, animal husbandry, and fisheries (lnprod); and the log value of the per capita disposable income of rural residents in pastoral areas (lninc) are regarded as the explanatory variables. This was also beneficial for use in the semi-elastic model, as described below. The specific explanation for these log values is as follows [25-29].

Logarithmic value of regional GDP (lngdp): Regional GDP refers to achievements made through a region's production activities in a certain period of time. There are three manifestations, namely value form, income form, and product form. In actual accounting, three methods, namely production method, income method, and expenditure method, can be used to reflect different aspects of the regional gross product and its composition.

Logarithmic value of the total output of regional agriculture, forestry, animal husbandry, and fisheries (lnprod): The total output of regional agriculture, forestry, animal husbandry, and fisheries consists of the output in agriculture, forestry, animal husbandry, and fisheries. Among them, agricultural output includes the planting industry and the part of the industry operated by people in rural areas. The planting industry refers to food crops, cash crops, vegetables, and fruits. Forestry output includes forest management, forest products, and timber harvesting in villages and smaller settlements. Animal husbandry output includes livestock and poultry breeding and hunting. Fishery output was calculated uniformly.

Logarithmic value of per capita disposable income of permanent residents in rural pastoral areas (lninc): Per capita disposable income of residents refers to the sum of residents' final expenditure for consumption and savings; that is, income that residents can dispose of freely includes income both in cash and in kind. In terms of the sources of income, disposable income includes four kinds, namely wage income, net operating income, net property income, and net transfer income. In this paper, it is referred to as the per capita disposable income of permanent residents.

In order to show the variables more intuitively, the ggplot2 package in R language and the matplotlib package in Python were selected for visual presentation. It can be seen from Fig. 2 that there was a generally growing trend in terms of the forest coverage rate with
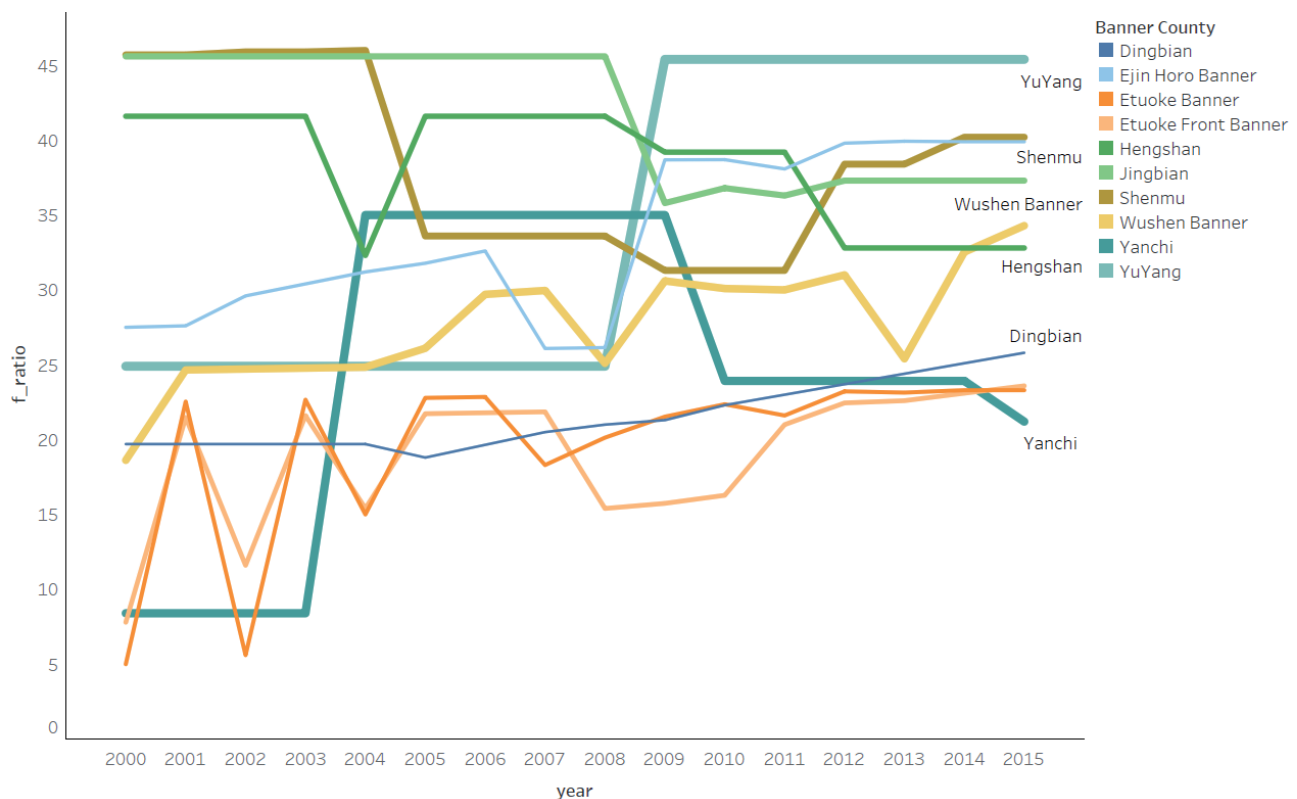


Fig. 2. Annual changes of f_ratio in select banners and counties, the Mu Us sandy land.

large increases in 2004 and 2009, which were most likely influenced by government policies. On average, Yuyang was the region with the highest forest coverage rate in all regions, while Dingbian had the lowest long-term forest coverage rate.

The changes of the two natural factors, rain and temp, are shown in Fig. 3. There was little annual change in various regions in terms of rain, generally maintaining below 500 mm. Only Etuoke Banner experienced a sharp annual precipitation increase in
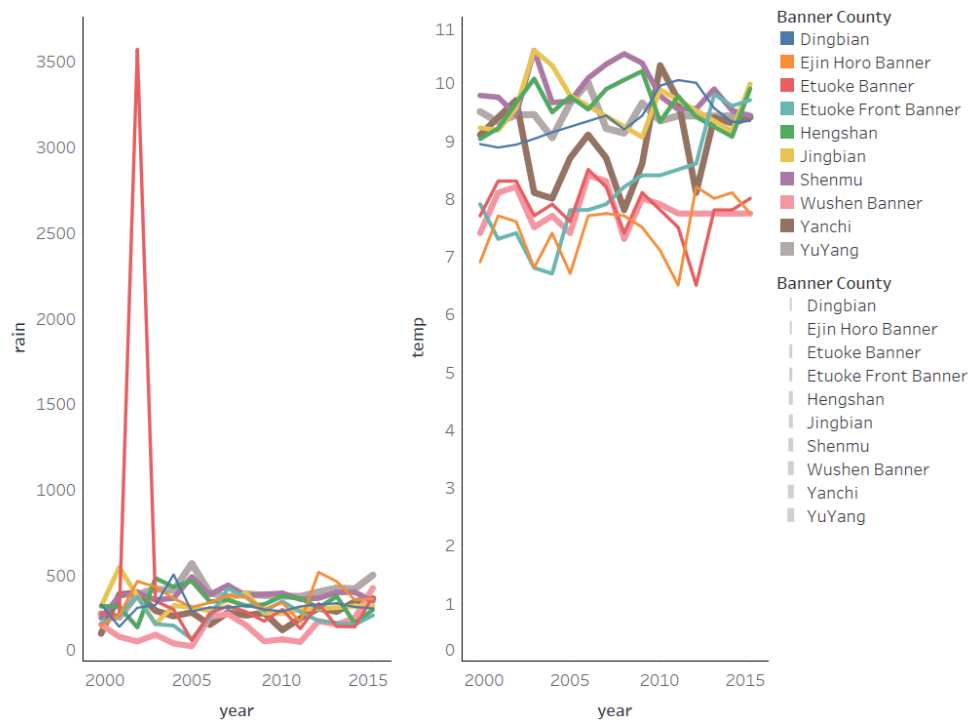


Fig. 3. Annual changes of rain and temp in select banners and counties, the Mu Us sandy land.
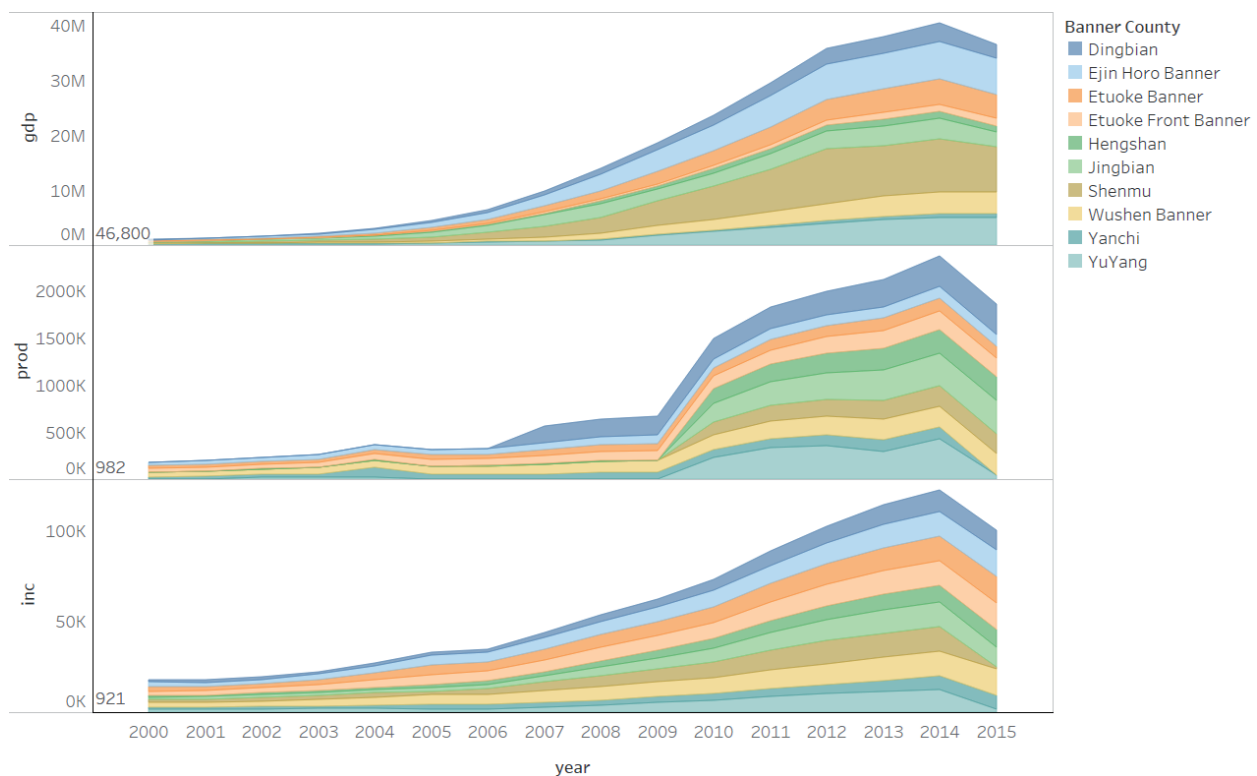


Fig. 4. Annual Changes of gdp, prod, and inc in select banners and counties, the Mu Us sandy land.

2002, and there was no significant change in annual precipitation in other years. There still was little annual change in various regions in terms of temp. Among them, temp were relatively high in Hengshan, Jingbian, and Shenmu, maintaining an average of about 10 degrees Celsius, while the temp were relatively low in Etuoke Banner, Etuoke Qianqi, and Wushen Banner, maintaining an average of around 8 degrees Celsius.

The changes in the three economic factors of gdp, prod, and inc are shown in Fig. 4. They all showed an upward trend as a whole with the most obvious rises in 2006 and 2009. Specifically, Shenmu's gdp began to grow rapidly in 2004, making it the region with the highest gdp in 2012, while the changes in prod and inc were relatively balanced among regions, maintaining a long-term upward trend.

It can be seen that economic characteristic factors showed obvious changes as time went by, while natural environmental factors remained constant throughout the year. Based on the results mentioned above, it could be concluded that changes in forest coverage were more likely to be induced by economic characteristics before more rigorous statistical verification of the data was conducted.

Due to two important years included in the selected time period of the basic layout of national sand prevention and control, this work introduced two variables, policy_1 and policy_2; the former represents "The National Sand Control Plan (2005-2010)", the latter "The National Plan for Desertification Control (2011-2020)". In terms of the data, the sample data from 2000 to 2004 were classified as policy 0, and the time variables were assigned as policy_1 = 0 and policy_2 = 0. The sample data from 2005 to 2010 were classified as policy 1, and the time variables were assigned as policy_1 = 1 and policy_2 = 0. The sample data of 2011-2015 were classified as policy 2, and the time variables were assigned as policy_1 = 0 and policy_2 = 1.

## Methods

### *Introduction to Panel Data Analysis*

Panel data analysis has unique advantages in solving missing variables and limitations on sample size. Among them, the individual effect model in panel data analysis assumes that the regression equation of each individual maintains the same slope but that its intercepts are different (as shown in Formula 1).

$$y_{it} = x'_{it}\beta + z'_i\delta + u_i + \varepsilon_{it} (i = 1, \cdots, n; t = 1, \cdots, T)$$
$$(1)$$

The individual effects model is divided into a fixed-effects model and random effect model according to whether the null hypothesis is rejected or not (that is, $u_i$ is irrelevant to $x'_{it}$, $z'_i$). The general method is the Hausman test, whose statistics are shown in Formula 2:

$$\left(\widehat{\beta}_{FE} - \widehat{\beta}_{RE}\right)' \left[Var(\widehat{\beta}_{FE})\right) - Var(\widehat{\beta}_{RE})\right]^{-1}$$
$$\left(\widehat{\beta}_{FE} - \widehat{\beta}_{RE}\right) \xrightarrow{d} \chi^2(K) \qquad (2)$$

These statistics are based on the following principles. Regardless of whether the null hypothesis is true, the mean value of the estimator in the fixed-effects model will be equal to the true value in a large sample. If the null hypothesis is true, the variance of the estimated value in the random-effects model will be smaller than that in the fixed-effects model. If the null hypothesis is not true, the mean value of the estimator of the random-effects model will not be equal to the true value in a large sample. Therefore, if the null hypothesis is true, there is little difference between estimators in the fixed-effects model and the random-effects model. On the contrary, if the gap between the two is too large, the null hypothesis should be rejected.

In terms of the fixed-effects model, if $u_i$ is related to a certain explanatory variable, it is called a fixed-effects model. In this case, the estimated value of Ordinary Least Squares regression (OLS) is inconsistent, so $u_i$ needs to be eliminated to obtain a consistent estimate. In a fixed-effects model, averaging both sides of Formula 1 at the same time can result in Formula 3 for an individual:

$$\overline{y}_i = \overline{x}'\beta + z'_i\delta + u_i + \overline{\varepsilon}_i \qquad (3)$$

Formula 3 can be subtracted from Formula 1 to obtain the dispersion form of the original model, as shown in Formula 4:

$$y_{it} - \overline{y}_i = (x_{it} - \overline{x}_i)'\beta + (\varepsilon_{it} - \overline{\varepsilon}_i) \qquad (4)$$

Given that $\tilde{y}_{it} \equiv y_{it} - \overline{y}_i, \tilde{x}_{it} \equiv x_{it} - \overline{x}_i, \tilde{\varepsilon}_{it} \equiv \varepsilon_{it} - \overline{\varepsilon}_i$, Formula 4 can be simplified as Formula 5.

$$\tilde{y}_{it} = \tilde{x}'_{it}\beta + \tilde{\varepsilon}_{it} \qquad (5)$$

At this time, according to the general assumptions of OLS that $\tilde{\varepsilon}_{it}$ is not related to $\tilde{x}_{it}$, the estimated value of OLS is consistent, and this estimated value is also called the fixed effects estimator.

### *Introduction to Theory of Dimensionality Reduction*

Dimensionality reduction is a process of selecting a low-dimensional feature set from a high-dimensional feature set. As a preprocessing step of machine learning, it can effectively eliminate redundant features, making the computation more effective, the estimation more accurate, and estimation results more comprehensive. Generally speaking, dimensionality reduction can be divided into feature extraction and feature selection. Feature extraction transforms the original feature space to regenerate a feature space with a smaller dimension

and more independence among each dimension. Feature extraction can be divided into linear methods and nonlinear methods. The former mainly include principal component analysis (PCA), independent component analysis (ICA), linear discriminant analysis (LDA), and factor analysis (FA).

Among the above methods, factor analysis can decompose the original variables and classify the more correlated variables into one category with the premise that there is low correlation among different types. Accordingly, fewer variables can be used to represent the information of most of the original variables, and the problem of multicollinearity among variables can be effectively solved.

Assuming that there are $p$ variables, namely $x_1, x_2,...,$ $x_p$, the original data matrix in $n$ samples is formed as shown in Formula 6:

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix} \tag{6}$$

The basic assumption of factor analysis is that the $p$ variables can be represented by the new $p$ variables in a linear way, as shown in Formula 7:

$$x_j = a_{j1}F_1 + a_{j2}F_2 + \cdots + a_{jp}F_p (j = 1,2,\cdots,p) \tag{7}$$

Among them, $a_{ij}(i,j - 1,2,..., p)$ constitutes the factor loading matrix. If the $p$ factors in Formula 7 are arranged in descending order of variance contribution rate, then the research usually selects factors with eigenvalues greater than 1. In this circumstance, Formula 7 can be rewritten as Formula 8 with $e_j$ as the error term:

$$x_j = a_{j1}F_1 + a_{j2}F_2 + \cdots + a_{jm}F_m + e_j \tag{8}$$

### *Introduction to Analysis of Variance and Covariance*

In terms of one-way ANOVA, the following hypotheses were generally made: (1) the sample values of the same factor level should come from the same normal population, and (2) the random impact has the same influence on different factor levels; that is, in terms of factors, if the $i$ level contains $n$ observations, the model for one-way ANOVA is as shown in Formula 9.

$$\begin{cases} y_{ij} = \mu_i + \varepsilon_{ij} (i = 1,\dots,m, j = 1,\dots,n) \\ \varepsilon_{ij} \sim N(0,\delta^2) \end{cases} \tag{9}$$

Among them, $\mu_j$ stands for the $j$-th observation value under the $i$-th level, $\mu_j$ is the mean value under the $i$-th level, and $\varepsilon_{ij}$ is the random disturbance term of the $j$-th observation value of the $i$-th level. To determine whether

the corresponding influencing factors are obvious, the essence is to determine whether they are true. To test the above equations, three concepts were introduced, namely total dispersion square sum ($S_T$), sum of squares between groups ($S_A$), and sum of squares within groups ($S_E$), as shown in Formula 10.

$$\begin{cases} S_T &= \sum_{i=1}^{m}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}\right)^2 \\ S_A &= \sum_{i=1}^{m}\sum_{j=1}^{n}\left(\bar{y_i} - \bar{y}\right)^2 \\ S_E &= \sum_{i=1}^{m}\sum_{j=1}^{n}\left(y_{ij} - \bar{y_i}\right) \end{cases} \tag{10}$$

When testing $H_0$: $\mu_1 = \mu_2 = \dots = \mu$, it could be concluded that $\frac{S_A/(m-1)}{S_E/(n-m)} \sim F(m-1, n-m)$.

One-way analysis of covariance (one-way ANCOVA) is the extension of one-way analysis of variance (one-way ANOVA). It involves one or more quantitative covariates, which are usually called confounding factors. They can also explain the differences among groups. Therefore, these confounding factors are introduced into the analysis of variance as covariates so as to avoid their interference. The model of single-factor covariance analysis is shown in Formula 11.

$$y_{ij} = \mu + \tau_i + \beta_i\left(x_{ij} - \bar{x}\right) + \varepsilon_{ij}, i \\ = 1,2,3,\cdots,r; j = 1,2,3,\cdots,n_i \tag{11}$$

$y_{ij}$ represents the dependent variable of the $j$ observation sample in the $i$ group, $x_{ij}$ the covariate of the sample, $\bar{x}, \bar{y}$ the mean of all their samples, and $\tau_i$ the $i$ type of grouping effect. The model has the following assumptions: $c_{ij} \sim NID(0,\sigma^2)$, where each of the variables are independent of each other. $x_{ij}$ is not affected by $\tau_i$, there is a linear relationship between $y$ and $x$ and $\beta_i$, and the linear regression coefficients in each group are equal.

## Results and Discussion

### Analysis of the Panel Data

First, mixed regression was performed as a basic reference to compare the characteristics of the panel data. Normal standard errors and robust standard errors were used in the ordinary regression process; the latter was expressed by "vce (cluster state)".

After the comparison of the normal standard error and the robust standard error shown in Table 2, it was found that the confounding factors under the same state among different years are generally self-relevant. Therefore, the normal standard error underestimated

Table 2. Normal standard error regression.

| f_ratio | Coef. | Std. Err. | t | P>t | [95% Conf. | Interval] |
|---------|-------|-----------|------|-------|-----------|-----------|
| rain | -0.0048 | 0.0024 | -1.91 | 0.052 | -0.0098 | 0.0001 |
| temp | 1.9731 | 0.7732 | 2.55 | 0.012 | 0.4455 | 3.5007 |
| lngdp | 4.5158 | 0.7913 | 5.71 | 0.000 | 2.9526 | 6.0791 |
| lnprod | -1.3952 | 0.5270 | -2.65 | 0.009 | -2.4364 | -0.3541 |
| lninc | -3.4836 | 1.7052 | -2.04 | 0.043 | -6.8524 | -0.1149 |
| _cons | -2.8815 | 10.4580 | -0.28 | 0.783 | -23.5413 | 17.7783 |

Table 3. Robust standard error regression.

| f_ratio | Coef. | Robust SE | t | P>t | [95% Conf. | Interval] |
|---------|-------|-----------|------|-------|-----------|-----------|
| rain | -0.0048 | 0.0009 | -5.04 | 0.001 | -0.0070 | -0.0026 |
| temp | 1.9731 | 1.8153 | 1.09 | 0.305 | -2.1333 | 6.0796 |
| lngdp | 4.5158 | 1.1591 | 3.90 | 0.004 | 1.8936 | 7.1381 |
| lnprod | -1.3952 | 0.7543 | -1.85 | 0.097 | -3.1017 | 0.3111 |
| lninc | -3.4836 | 1.9619 | -1.78 | 0.11 | -7.9219 | 0.9545 |
| _cons | -2.8815 | 15.2840 | -0.19 | 0.855 | -37.4565 | 31.6934 |

the true standard error, as its result showed that temperature and GDP had a significant positive effect on forest coverage, while other factors such as rainfall did not. This conclusion is not desirable. According to the robust standard error, only rainfall and GDP were notable at the 95% confidence level, but the coefficient of rainfall was very small.

As shown in Table 3, some missing variables might exist without changes during a period of time due to the different situations of the state. Therefore, a fixed effect model (FE) was applied in this research.

As shown in Table 4, the constant term (_cons) in the output represented the average value of effects of the single factor. Rho = 0.6298 in the last row of the above table indicates that the variance of the compound

confounding factors largely resulting from the change of individual effects. In the individual test, the p-value of the F test was 0.0000, strongly rejecting the null hypothesis that the fixed-effects model was significantly better than mixed regression. Individual factors in the model should be allowed to have their own intercepts but can still exist in random effects (RE) form. The Hausman test is shown in Table 5. Regardless of whether the null hypothesis is true, the fixed-effects model is unbiased, while the random-effects model may be biased ($p = 0.0803 < 0.1$). Therefore, the fixed-effects model was chosen for this study and the Hausman test was conducted.

Table 4. Fixed effect model analysis.

| f_ratio | Coef. | Std. Err. | t | P>t | [95% Conf. | Interval] |
|---------|-------|-----------|------|-------|-----------|-----------|
| rain | -0.0034 | 0.0019 | -1.83 | 0.070 | -0.0071 | 0.0003 |
| temp | -0.5566 | 1.0271 | -0.54 | 0.589 | -2.5868 | 1.4734 |
| lngdp | 1.8764 | 0.8382 | 2.24 | 0.027 | 0.2197 | 3.5332 |
| lnprod | -0.4027 | 0.4316 | -0.93 | 0.352 | -1.2558 | 0.4503 |
| lninc | -0.3975 | 1.4633 | -0.27 | 0.786 | -3.2898 | 2.4946 |
| _cons | 18.1574 | 9.4826 | 1.91 | 0.057 | -0.5845 | 36.8994 |
| sigma_u | 7.9699 | | | | | |
| sigma_e | 6.1230 | | | | | |
| rho | 0 .6288 | | | | | |

Table 5. Hausmann test.

| f_ratio | FE | RE | Difference | S.E. |
|---|---|---|---|---|
| rain | -0.0034 | -0.0035 | 0.0001 | 0.0002 |
| temp | -0.5566 | 0.3687 | -0.9254 | 0.4742 |
| lngdp | 1.8764 | 2.4841 | -0.6076 | 0.2783 |
| lnprod | -0.4027 | -0.5696 | 0.1668 | 0.0838 |
| lninc | -0.3975 | -0.5696 | 0.8849 | 0.3692 |
| _cons | 18.1574 | 11.0588 | 7.0985 | 2.7203 |
| chi2(5) | 9.8300 | | | |
| Prob>chi2 | 0.0803 | | | |

[1] Note: FE: fixed-effects model. RE: random-effects model.

Table 6 shows the coefficient estimates and standard errors of the above methods for comparison.

As shown in Table 6, in the robust standard error estimation results of the fixed effect model, only rainfall and GDP have significant effects on forest coverage in terms of p-value, but further observation of the coefficients shows that the coefficient of rainfall is very small (-0.003), indicating little effect on the forest coverage. By contrast, the logarithmic impact coefficient of GDP is 1.876, which can ensure that the p-value is less than 0.1 even under the requirement of robust standard errors. Therefore, this model fully proved the positive effect of GDP growth on the increase of forest coverage. The particularity of this semi-elastic model is explained as follows. For the general semi-elastic model:

$$f\_ratio = \alpha + \beta lngdp + \varepsilon \qquad (12)$$

$$\Delta f\_ratio = (\beta/100)\% \Delta lngdp \qquad (13)$$

In the horizontal-logarithmic model shown in Formula 12, with only the logarithmic value as the independent variable, its coefficient could be interpreted as that with every 1% increase in the independent variable, the dependent variable increases by $\beta/100$, as shown in Formula 13. Specifically, this coefficient means that for every 1% increase in GDP, the forest coverage rate in the region will increase by 1.876%.

Dimension Reduction Processing

First, the corrplot package in R language was used to visualize the correlation of variables, as shown in Fig. 5. The largest circles were observed for lngdp, lnprod, and lninc. It can be seen from the colors that their correlation coefficients are all more than 0.5, which represents a significant positive correlation among the three factors. There is a positive correlation between rain and temperature with a small correlation coefficient. At the same time, they still have low correlation with lngdp, lnprod, and lninc. In addition, there is a relatively significant negative correlation between policy_1 and policy_2; policy_1 has a lower correlation with other factors, while policy_2 has a higher correlation with lngdp, lnprod, and lninc.

Exploratory factor analysis is mainly used to discover latent variables in measurable data. These latent variables are virtual and not directly observable, so the corresponding latent variables and their structure can only be obtained indirectly. Therefore, this study utilized R language to perform exploratory factor analysis (EFA) to reduce dimensions and analyze the internal structure in a reasonable way.

Table 6. Summary of model.

| | OLS | OLS_robust | FE |
|---|---|---|---|
| rain | -0.005* | -0.005*** | -0.003* |
| | (0.002) | (0.001) | (0.002) |
| temp | 1.973** | 1.973 | -0.557 |
| | (0.773) | (1.815) | (1.027) |
| lngdp | 4.516*** | 4.516** | 1.876** |
| | (0.791) | (1.159) | (0.838) |
| lnprod | -1.395*** | -1.395 | -0.403 |
| | (0.727) | (0.754) | (0.432) |
| lninc | -3.484** | -3.484 | -0.398 |
| | (1.705) | (1.962) | (1.463) |
| _cons | -2.882 | -2.882 | 18.160 |
| | (10.458) | (15.280) | (9.483) |

[1] Note: *** p<0.01 ** p<0.05 * p<0.1.
[2] Note: OLS: Ordinary Least Squares regression.
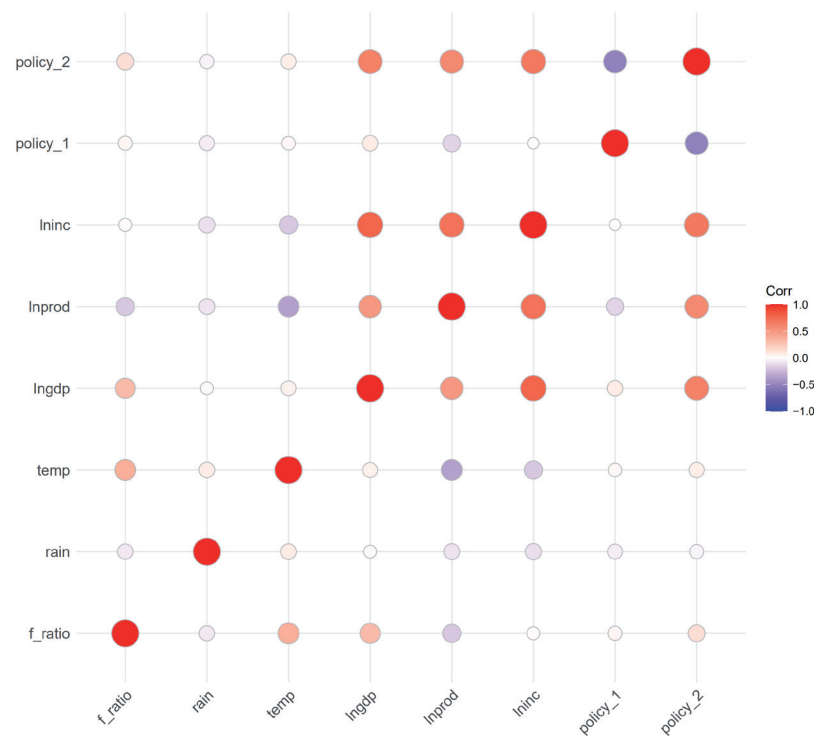FE: fixed-effects model.

Fig. 5. Correlation analysis diagram.

*Economic Factors*

Factor analysis was conducted for the three variables lngdp, lnprod, and lninc, as shown in Fig. 6. First, common factors with eigenvalues higher than 1 were retained based on the Kaiser–Harris criterion, and the first common factor was selected corresponding to Fig. 6. Second, common factors above the maximum inflection point of the broken line were retained. As seen in Fig. 6, the second common factor appeared in the corner, so the first common factor was selected. The variance contribution rate was further considered. As shown in Table 7, the variance contribution rate when extracting the first common factor reached 73%,
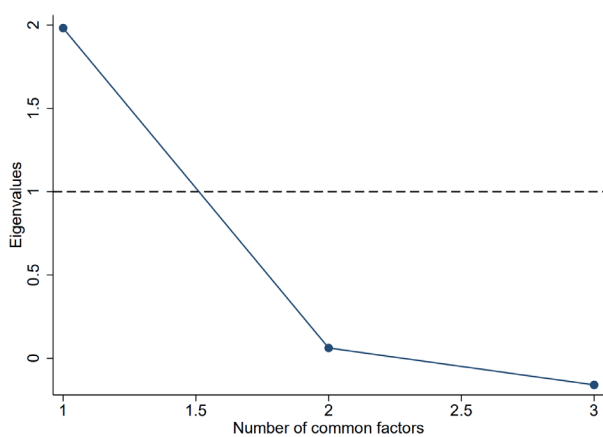
whereas the cumulative variance contribution rate when extracting the first two common factors was only 74%. Based on the results mentioned above, the three variables of lngdp, lnprod, and lninc can be explained by a common factor (fa_economy).

*Natural Environmental Factors*

Second, factor analysis of the two variables of rain and temp was conducted, as shown in the scree plot in Fig. 7. First, if the common factors with eigenvalues higher than 1 were retained based on the Kaiser–Harris criterion, no common factor could be selected corresponding to Fig. 7. Secondly, if the common factors were above the maximum inflection point of the broken line, no common factors could be selected considering no inflection point in Fig. 7. Furthermore, considering the variance contribution rate factor, as shown in Table 7, the variance contribution rate when extracting a common factor was only 27%, much lower than 50%, suggesting that the common factor could not represent the two original variables, and thus the two variables rain and temp were still used.

*Homology Deviation Test*

Considering the problem of sample homology deviation, this study also involved a homology deviation test. According to Harman's one-way test standard, if the homology deviation exists, the measured item will form a factor that explains that the variance is dominant. The items f_ratio, rain, temp, lngdp, lnprod, and lninc



Fig. 6. Scree plot of economic factors.

Table 7. Summary of model.

| Type | Factor | Eigen | Difference | Proportion | Cumulative |
|---|---|---|---|---|---|
| Economic factors | factor1 | 2.339 | 1.838 | 0.730 | 0.730 |
| | factor2 | 0.501 | 0.341 | 0.011 | 0.741 |
| | factor3 | 0.160 | 0.000 | 0.259 | 1.000 |
| Natural environmental factors | factor1 | 0.087 | 0.160 | 0.273 | 0.273 |
| | factor2 | -0.074 | 0.000 | 0.727 | 1.000 |

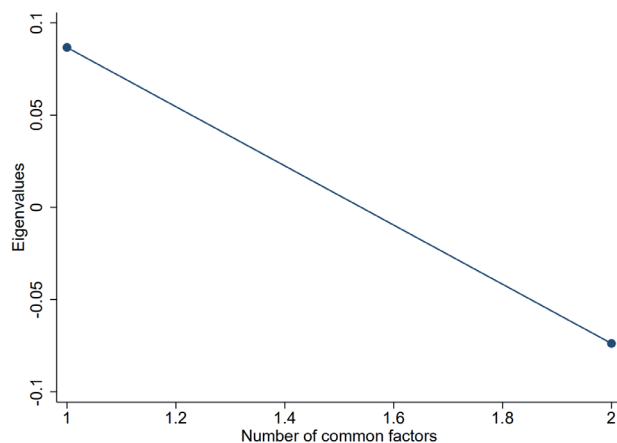[1] Note: *** $p<0.01$ ** $p<0.05$ * $p<0.1$



Fig. 7. Scree plot of natural environmental factors.

were grouped together for principal component analysis and exploratory factor analysis. The results show that the variance contribution rates of the first two principal components were 31% and 17%, respectively, while the first two common factors explained the variance of 38% and 23%, respectively. There was no dominant factor, and neither factor exceeded the recommended limit. The standard value of 50% [30-33] showed that the problem of homology bias did not exist in this study [34], so the reliability of the research conclusion would not be affected accordingly.

## Variance Analysis

### *One-Way Variance Analysis*

According to the above theory, on the basis of factor dimensionality reduction, the modeling policy for f_

ratio requires one-way ANOVA, and any sample in this study would not be affected by other policy levels at the same time. Therefore, the individual factors were inter-grouped, and this section could also be called inter-group one-way ANOVA. The preliminary analysis of variance is shown in Table 8:

It was seen from the output that one-way ANOVA had obvious effects on the F test of the policy ($p<0.01$), indicating that significant differences in forest coverage rate were found for the three different policy states. However, the above results only demonstrated the significant impact of factor variables as a whole. Considering the three categories of policy, early studies recommended the use of multiple comparison tests instead of t-tests [35-38], in order to clarify the difference between the two categories. Tukey's Honest Significant Difference (HSD) test was used in this study for multiple comparison tests, and the results are shown in Fig. 8:

At the top of Fig. 8, groups with the same letter represent no obvious differences in means. Therefore, it could be found that the difference between policy 0 and policy 1 was not obvious (the same letter, a); neither was the difference between policy 1 and policy 2 (the same letter, b); but the difference between policy 0 and policy 2 was obvious (no matching letter), indicating that policy 2, compared with policy 0, can significantly improve the f_ratio.

### *One-Way Covariance Analysis (One-Way ANCOVA)*

In this study, f_ratio was used as the dependent variable, with policy as the individual factor, and three covariates of rain, temp, and fa_economy were added. The results were as follows.

Table 8. Summary of one-way ANOVA.

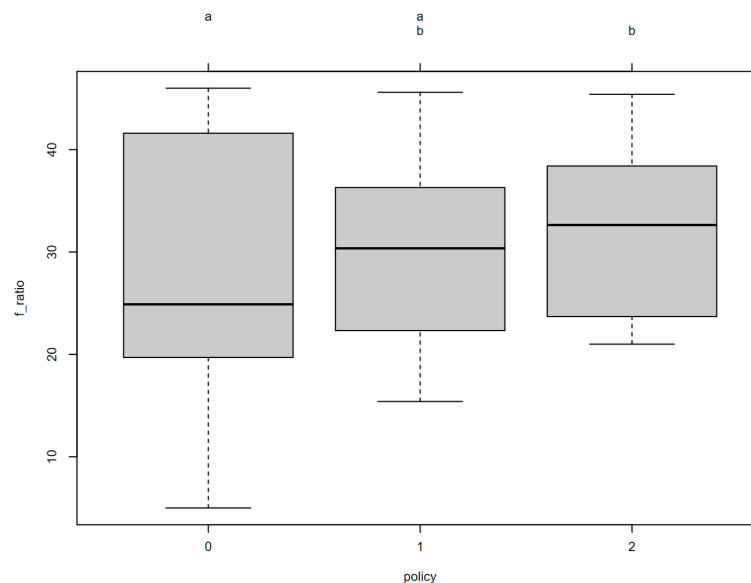| Variance | Degrees of freedom | Sum of squares | Mean sum of squares | F value | p-value |
|---|---|---|---|---|---|
| Policy | 1.000 | 0.076 | 0.076 | 8.195 | 0.005*** |
| Error | 158.000 | 1.474 | 0.009 | | |

[1] Note: *** $p<0.01$ ** $p<0.05$ * $p<0.1$

Fig. 8. Diagram of multiple comparison test.

Table 9. Summary of one-way ANCOVA.

| Variance | Degrees of freedom | Sum of squares | Mean sum of squares | F value | p-value |
|---|---|---|---|---|---|
| Rain | 1.000 | 0.139 | 0.139 | 17.295 | 0.000*** |
| Temp | 1.000 | 0.101 | 0.101 | 12.537 | 0.001*** |
| fa_economy | 1.000 | 0.017 | 0.017 | 2.055 | 0.154 |
| Policy | 1.000 | 0.046 | 0.046 | 5.721 | 0.018** |
| Error | 155.000 | 1.248 | 0.008 | | |

[1] Note: *** p<0.01 ** p<0.05 * p<0.1

The F test of one-way ANCOVA showed that rain and temp were both related to f_ratio, while fa_economy was not. Policy was related to f_ratio with the control of rain, temp, and fa_economy. There were indeed obvious differences in forest coverage with control variables under different policy states. Further multiple comparison tests are shown in Table 9:

Similarly, at the top of Fig. 9, groups with the same letter represent no clear differences in means. It was found that the difference between policy 0 and policy 1 was not obvious (the same letter, a); neither was the difference between policy 1 and policy 2 (the same letter, b); but the difference between policy 0 and policy 2 was obvious (no matching letter). Therefore, the test results of ANCOVA and one-way variance analysis were consistent.

## Discussion

Human activities and climate change are very important factors that shape landscapes. The socioeconomic factors closely related to human activities have an even greater impact on sandy lands or deserts. In terms of the Mu Us sandy land, many researchers have long focused on the factors mentioned above [5,12,18,19], but they only used simple statements and descriptions in most studies as a method without rigorous statistical proof. As a result, the conclusions drawn from their work are more likely to be a guess rather than a causal judgment. In addition, socioeconomic factors are just used as a supplement to natural environmental factors [39-42], which places more emphasis on the natural environmental factors. Given that natural environmental factors remain almost constant throughout the year, they have little effect on sandy or desert landforms.

However, more scientific analysis methods have been applied to the research on the factors influencing sand or desert, especially socioeconomic factors. For example, a study found that diverse policies and measures have caused significant changes in the entire ecosystem including vegetation coverage in Northeast China from 2000 to 2015 through remote sensing technology [43]. Similarly, the important role of policy measures was also explored in this study. Two policies, policy 1 and policy 2, had a positive effect on the
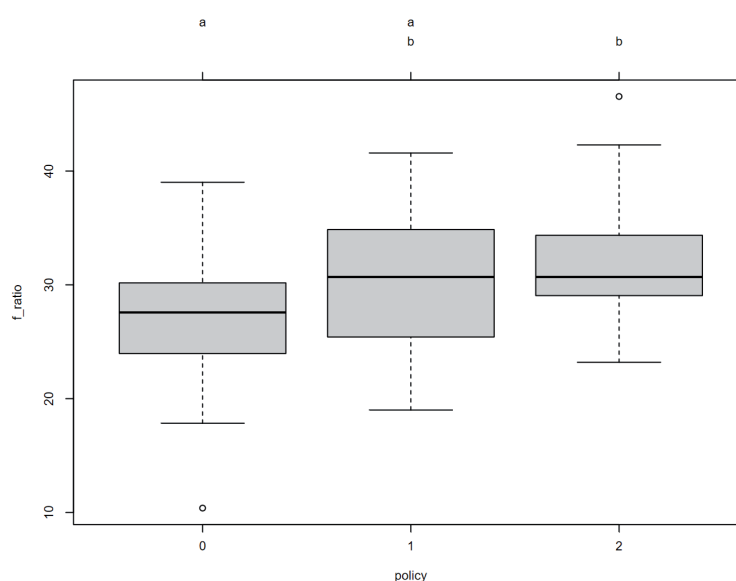
Fig. 9. Diagram of multiple comparison test.

reversal of desertification in the Mu Us sandy land. Specifically, if the variables were not controlled, the F value of policy in the one-way analysis of variance was 8.195, indicating that policy factors have an important impact on forest coverage. Specifically, the most obvious difference was found for the forest coverage rate after implementing policy 2, indicating that policy 2 was more effective than policy 1. After introducing these two control variables and the natural environmental and social and economic factors, the same conclusion was also drawn from the one-way ANOVA analysis. It can be seen that ANOVA could better prove the causality compared with remote sensing technology, despite the difference in methods and research fields between the above study and the present study.

In addition, some scholars have concluded that the value of ecological services has a negative correlation with GDP, but they still believed that economy and ecology should maintain a harmonious relationship in order to achieve sustainable development [44]. This result is different from the result obtained in this research. The conclusions of this study are that for every 1% increase in GDP, the forest coverage in the Mu Us sandy land will increase by 1.876% based on panel data analysis. This can also prove that economy and ecology can coexist harmoniously to some extent, which is the suggestion made by the scholars mentioned above.

## Conclusions

The conclusions were made as follows through multiple methods of analysis, based on the statistical data of the Mu US sandy land:

(1) The forest coverage in the Mu Us sandy land maintained a steady upward trend from 2000 to 2015.

(2) Analysis of panel data showed that GDP had a significant positive effect on the forest coverage of the Mu Us sandy land.

(3) Analyses of variance and covariance showed that policies and measures had a significant effect on the environmental improvement of the Mu Us sandy land.

(4) Environmental protection policies should focus on guiding rather than on compulsory means, and a more efficient economic development model should be adapted to promote environmental protection, rather than protection at the expense of economic development. Specifically, regional economic development should be carried out in a circular economy model. The government and enterprises should coordinate with the environment, namely by giving full play to the government's role in policy guidance and stimulating corporate environmental technology innovation, eventually contributing to the sustainable use of the environment and sustainable economic development.

## Acknowledgements

## Conflict of Interest

The authors declare no conflict of interest.

## References

1. DENG J. Fitting the revised Gash analytical model of rainfall interception to Mongolian Scots pines in Mu Us

Sandy Land, China. Trees, Forests and People, **1**, 100007, **2020**.

2. WANG N., XIE J., HAN J. A sand control and development model in sandy land based on mixed experiments of arsenic sandstone and sand: A case study in Mu Us Sandy Land in China. Chinese Geographical Science, **23**, 700, **2013**.

3. HUI Y., CAO J., HOU X. Characteristics of Aeolian Dune, Wind Regime and Sand Transport in Hobq Desert, China. Applied Sciences, **9**, 5543, **2019**.

4. SIQUAN Y., WEI W. Evaluation of desertification in Mu Us Sandy Land. Arid Land Geography, **33**, 258, **2010**.

5. LIANG P., YANG X. Landscape spatial patterns in the Maowusu, Mu Us Sandy Land, northern China and their impact factors. Catena, **145**, 321, **2016**.

6. LIU J. Dynamics and driving forces of landscape patterns in Mu Us Sandy Land, from 1990 to 2014. Pratacultural Science, **2**, 7, **2017**.

7. MIAO Y., JIN H., CUI J. Human activity accelerating the rapid desertification of the Mu Us Sandy Lands, North China. Scientific Reports, **6**, 23003, **2016**.

8. LIU Q., YANG X. Geochemical composition and provenance of aeolian sands in the Ordos Deserts, northern China. Geomorphology, **318**, 354, **2018**.

9. WANG X. Key driving forces of desertification in the Mu Us Desert, China. Scientific Reports, **7**, 3933, **2017**.

10. LI H.J., GAO J.E., ZHANG H.C., ZHANG Y.X., ZHANG, Y.Y. Response of extreme precipitation to solar activity and El Nino Events in typical regions of the Loess Plateau. Advances in Meteorology, **2017**, 1, **2017**.

11. LIU B. Geochemical weathering of aeolian sand and its palaeoclimatic implications in the Mu Us Desert, northern China, since the Late Holocene. Journal of Arid Land, **8** (5), 647, **2016**.

12. TONG S., MA Y. Evaluation of ecological vulnerability of Yulin City under, the background of returning farmland to forests. Materials Science and Engineering, **768**, 052, **2020**.

13. ZHANG C., WANG X., LI J., HUA T. Identifying the effect of climate change on desertification in northern China via trend analysis of potential evapotranspiration and precipitation. Ecological Indicators, **112**, 106141, **2020**.

14. FU T. Quantitative analysis of ground penetrating radar data in the Mu Us Sandland. Aeolian Research, **32**, 218, **2018**.

15. LI S., WANG T., YAN C. Assessing the role of policies on land-use/cover change from 1965 to 2015 in the Mu Us Sandy Land, northern China. Sustainability, **9** (7), 1164, **2017**.

16. GE Q. Temporal and spatial variation of vegetation analysis based on MOD13 in Mu Us Sandy Land (RAC-2016). Atlantis Press, **2016**.

17. WANG X. Key driving forces of desertification in the Mu Us Desert, China. Scientific reports, **7** (1), 1, **2017**.

18. ZHANG D., DENG H. Historical human activities accelerated climate-driven desertification in China's Mu Us Desert. Science of The Total Environment, **708**, 134771, **2020**.

19. LI N. Cultivated-land change in Mu Us Sandy Land of China before and after the first-stage grain-for-green policy. Sciences in Cold and Arid Regions, **10** (4), 347, **2018**.

20. ZHAO J., CAO C., YOMWAN P. Aboveground shrub biomass estimation based on landsat data in MU US sandy land, China. **2012**.

21. ZHENG Y. Effects of revegetation on climate in the Mu Us Sandy Land of China. Science of The Total Environment, **739**, 139958, **2020**.

22. ZHOU J., XU Y., GAO Y., XIE Z. Land use model research in agro-pastoral ecotone in northern China: A case study of Horqin Left Back Banner. Journal of Environmental Management, **237**, 139, **2019**.

23. YE Q., LIU H., LIN Y. Study of Maowusu sandy land vegetation coverage change based on Modis Ndvi. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **3**, 2127, **2018**.

24. WANG F. Combating desertification in China: past, present and future. Land Use Policy, **31**, 311, **2013**.

25. MOLINA I., MARTÍN N. Empirical best prediction under a nested error model with log transformation. The Annals of Statistics, **46**, 1961, **2018**.

26. GAO A., FISCHER-JENSSEN J., WROBLEWSKI C., MARTOS P. Interpretation and implications of lognormal linear regression used for bacterial enumeration. Journal of AOAC International, **103**, 1105, **2020**.

27. SHI P. Leaf area–length allometry and its implications in leaf shape evolution. Trees, **33**, 1073, **2019**.

28. BROOTHAERTS W., CORDEIRO F., CORBISIER P., ROBOUCH P., EMONS H. Log transformation of proficiency testing data on the content of genetically modified organisms in food and feed samples: is it justified? Analytical and Bioanalytical Chemistry, **412**, 1129, **2020**.

29. SAULNIER E. Seasonality in coastal macrobenthic biomass and its implications for estimating secondary production using empirical models. Limnology and Oceanography, **64**, 935, 2**019**.

30. MIN H., PARK J., KIM H. J. Common method bias in hospitality research: A critical review of literature and an empirical study. International Journal of Hospitality Management, **56**, 126, **2016**.

31. KLINE R. Principles and practice of structural equation modeling. Guilford publications, 2, **2015**.

32. BYRNE B. Basic concepts, applications, and programming. Structural equation modeling with AMOS, **396**, 227, **2010**.

33. HAIR J., JOSEPH F. A primer on partial least squares structural equation modeling (PLS-SEM). Sage publications, **2016**.

34. AHMADINEJAD M. Factors Affecting on research practices development regarding entrepreneurship in agricultural higher education: A two-step approach to structural equation modeling. International Journal of Agricultural Management and Development, **9**, 319, **2019**.

35. KIM H. Y. Statistical notes for clinical researchers: Analysis of covariance, ANCOVA. Restorative Dentistry & Endodontics, **43**, **2018**.

36. LI G., WANG X. Prediction accuracy measures for a nonlinear model and for right-censored time-to-event data. Journal of the American Statistical Association, **114**, 1815, **2019**.

37. LAMOTTE L.R. Proportional subclass numbers in two-factor ANOVA. Statistics, **52**, 228, **2018**.

38. ROUDER J.N., ENGELHARDT C.R., MCCABE S., MOREY R.D. Model comparison in ANOVA. Psychonomic bulletin and review, **23**, 1779, **2016**.

39. ZHU W., GAO Y., ZHANG H., et al. Optimization of the land use pattern in Horqin Sandy Land by using the CLUMondo model and Bayesian belief network. Science of The Total Environment, **739**, 139929, **2020**.

40. GUO M., DING G., GAO G., et al. Community composition of ectomycorrhizal fungi associated with Pinus sylvestris var. mongolica plantations of various ages in the Horqin Sandy Land. Ecological Indicators, **110**, 105860, **2020**.

41. FAN M., PENG Y., WANG Q., et al. Correlations between landscape pattern and plant diversity at multiple spatial scales: a case study of Hunshandak Sandland. Acta Ecologica Sinica, **7**, 22, **2018**.

42. DUAN H., WANG T., XUE X., et al. Dynamic monitoring of aeolian desertification based on multiple indicators in Horqin Sandy Land, China. Science of The Total Environment, **650**, 2374, **2019**.

43. MAO D., HE X., WANG Z., et al. Diverse policies leading to contrasting impacts on land cover and ecosystem services in Northeast China. Journal of Cleaner Production, **240**, 117961, **2019**.

44. XING L., XUE M., WANG X. Spatial correction of ecosystem service value and the evaluation of eco-efficiency: A case for China's provincial level. Ecological Indicators, **95**, 841, **2018**.