Evaluation and Analysis Model of Economic Development Level for Latin American Countries

Hao Cui

School of Institute for Research on Portuguese-Speaking Countries, City University of Macau, Taipa 999078, Macau, China

Correspondence should be addressed to Hao Cui; a19092105151@cityu.mo

Received 25 June 2022; Revised 30 July 2022; Accepted 3 August 2022; Published 24 August 2022

Academic Editor: Lianhui Li

Copyright © 2022 Hao Cui. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study is devoted to the analysis of high-quality economic development in Latin American countries. Under the guidance of the new development concept strategy, the evaluation index system for the high-quality economic development level of Latin American countries is constructed from the six dimensions of economic strength, innovation, coordination, green, openness, and sharing, and the entropy weight method is used to calculate the weight of each evaluation index. Based on UTM projection coordinate quantification and K-Means clustering algorithm, the spatial effect clustering model of Latin American countries is established. According to the two dimensions of geographical location and comprehensive index, Latin American countries are divided into different economic zones, and the agglomeration effect in the process of economic development of Latin American countries is analyzed from a spatial perspective.

1. Introduction

Latin America is the region with the largest concentration of middle-income economies. It is also the region that first became independent, started the modernization process, crossed the low-income stage, and entered the middle-income stage [1–6]. As early as the early 1960s and 1970s, most Latin American countries have got rid of the poverty trap and entered the middle-income stage.

Latin American countries have always been the pioneers of the developing world in actively exploring and trying to use new economic development theories to guide practice [3]. Therefore, Latin American countries are called the experimental field of development economics. Since their independence in the early nineteenth century, Latin American countries have successively explored and applied the theory of comparative advantage, the theory of structuralism, the theory of neoliberalism, and the theory of socialism in the twenty-first century to guide their economic development for a period of time. Among them, the theory of structuralism and the theory of socialism in the twenty-first century were put forward and founded by the people of Latin American countries according to the development difficulties of Latin American countries, which not only affected Latin America itself, but also has had a broad impact on other developing regions [5]. In the developing world, Latin American countries are also the first to implement the neoliberal theory.

However, after guiding the practice of Latin American countries for a period of time, these development theories ended in failure. At present, the mainstream economic academia mainly analyzes the growth and development dilemma of Latin American countries from the perspective of total factor productivity and innovation emphasized by the new growth theory and from the perspective of institution, corruption, and democracy emphasized by the new institutionalism theory. Either these analyses point out the problem but are difficult to go deep into the root cause of the problem, and then it is difficult to give constructive suggestions (referring to the analysis of innovation), or they reverse the results and causes of development and make it difficult to convince people (referring to systems and democracy).

Evolutionary economics is another set of economic theory system juxtaposed with mainstream economics. Since the international financial crisis in 2008, evolutionary

2.1.1. Economic Strength. Improving economic strength is still the key task of high-quality development of Latin American countries. Under this dimension, this article sets up three secondary indicators of the economic scale, development level, and investment intensity. Under the economic scale, this article sets up two indicators of regional GDP and industrial output above the designated size. Under the development level, this article sets up two indicators of general public budget revenue and regional average GDP. Under the investment intensity, this article sets up two indicators of fixed asset investment and total industrial investment. The economic scale of country-level high-quality development is measured from three aspects: economic scale, development level, and investment intensity.

2.1.2. Innovative Development. Innovative development belongs to the key content of high-quality development in the Latin American countries. Under the dimension of innovation and development, this article sets up two secondary indicators of innovation vitality and innovation potential. Under the vitality of innovation, this article sets up two indicators for the number of invention patents authorized and the R&D investment in industrial enterprises above the scale. Under the potential of innovation, this article sets up three indicators: the single-digit number of industrial enterprises above the designated size, the balance of various deposits in financial institutions, and the number of full-time teachers per 100 primary and secondary school students. These five indicators are used to measure the degree of innovation in the country’s economic development, which also constitutes a representative indicator of the innovation-driven development of the country.

2.1.3. Coordinated Development. Coordinated development is the central link in the high-quality development of Latin American countries. Under the dimension of coordinated development, this article sets up two secondary indicators of social development coordination and people’s life coordination. Under the coordination of social development, this article sets up three three-level indicators: the secondary industry’s added value as a proportion of GDP, the added value of the tertiary industry as a proportion of GDP, and general public budget revenue as a share of GDP. Under the coordination of people’s livelihood, this article sets up three three-level indicators, retail sales of consumer goods per capita, urbanization rate, and highway density. These indicators can not only fully measure the coordination between people’s life and social development, but also reflect the reality of the current country development.

2.1.4. Green Development. Achieving green development plays a vital role in promoting the high-quality development of Latin American countries. Under the green development dimension, this article sets up two secondary indicators which are livability and energy conservation and environmental protection. Under livability, this article sets up two indicators: the ratio of good days and the green area...
coverage. Under energy conservation and environmental protection conditions, this article sets up two indicators, the comprehensive utilization rate of solid waste and the comprehensive sewage treatment. It is mainly used to measure the green development level in the current high-quality development of counties.

2.1.5. Open Development. Open development is an indispensable part of the high-quality development of Latin American countries and plays a vital role in the development of Latin American countries. This article sets up two indicators, the total import and export volume and the level of actual utilization of foreign capital, which can have high representativeness for the opening level of the Latin American countries.

2.1.6. Shared Development. Shared development is the ultimate destination of high-quality development in the Latin American countries. Under the economic sharing, this article sets up three indicators of per capita annual GDP, per capita deposit balance, and per capita disposable income of residents. Under the living security, this article sets up three indicators: the rate of the urban registered unemployment, the number of beds per 10,000 people, and the number of health technicians per 10,000 people. The above indicators can not only mainly cover the shared development level of the country but also confirm the high-quality development level of the country to a certain extent.

To sum up, from the six dimensions of economic strength, innovative development, coordinated development, green development, open development, and shared development, a country-level high-quality development level evaluation system was constructed, including 12 secondary and 29 three-level indicators, as shown in Table 1.

2.2. Data Processing and Computation

2.2.1. Entropy Weight Method. At present, there are many references for multi-index evaluation methods in academic research, which are mainly divided into two categories: subjective evaluation and objective evaluation. The subjective evaluation method mainly uses the evaluator to determine the weight of each index subjectively through his psychological cognition, education level, life experience, etc. The objective evaluation method is to analyze the fluctuation of data through scientific methods to find the relationship between the data, determine the weight of each index, and evaluate it. To avoid the subjectivity and some objective limitations of artificially determining the weight of indicators and further improve the validity and reliability of this paper, this paper uses the entropy weight method in the objective evaluation method to calculate the weight of each indicator in the evaluation system of high-quality development in the country area. The specific calculation steps [13–16] are as follows:

1. Data standardization processing

   Positive indicator: \( x'_{ij} = \frac{x_{ij} - \min\{x_{ij}\}}{\max\{x_{ij}\} - \min\{x_{ij}\}} \) (1)

   Negative indicator: \( x'_{ij} = \frac{\max\{x_{ij}\} - x_{ij}}{\max\{x_{ij}\} - \min\{x_{ij}\}} \)

   In the formula, \( x_{ij} \) is the original value of the \( i \)-th sample and the \( j \)-th index, and \( x'_{ij} \) is the standardized index value.

   Since the normalized value will appear logarithmic, the standardized data needs to be further calculated later, so this paper adopts the translation value to avoid the influence of the logarithm.

   \( L_{ij} = x'_{ij} + a \) (2)

   In the formula, \( L_{ij} \) is the value after translation, and the translation amplitude \( a \) is set to 0.0001.

2. Calculate the proportion of the \( i \)-th country in this indicator under the \( j \)-th indicator \( (p_{ij}) \):

   \( p_{ij} = \frac{L_{ij}}{\sum_{i=1}^{n} L_{ij}} \) (i = 1, 2, . . . , n; j = 1, 2, . . . , m). (3)

   In the formula, \( n \) is the number of samples (counties), and \( m \) is the number of indicators.

3. Calculation of index information entropy:

   \( e_i = -k \sum_{i=1}^{n} p_{ij} \ln(p_{ij}) \) (4)

   In the formula, \( k = 1/\ln(n) \), \( 0 \leq e_i \leq 1 \).

4. Calculation of difference coefficient:

   \( g_i = 1 - e_i \) (5)

5. Weights of indicators:

   \( w_j = \frac{g_i}{\sum_{j=1}^{m} g_i} \) (j = 1, 2, . . . , m). (6)

2.2.2. Determination of Indicator Weights. According to the above calculation steps of the entropy weight method [13–16], this paper uses the relevant indicator data from 2015 to 2018 of the Latin American countries. The weights for each year are obtained separately, and the average of the four-year weights is used as the comprehensive weight of each indicator. The calculation results are shown in Table 2.

2.2.3. Calculation of Composite Index. To further compare the high-quality development level of Latin American countries and the gap with the high-quality development level of the top counties, this paper further calculates the comprehensive development index of each Latin American
country by using the weight of each index calculated in the previous section.

Since the units of each indicator in the indicator system are different (such as urbanization rate: 35%, GDP: 100 billion), it cannot be directly calculated, so we need to perform dimensionless processing on it. The specific processing formula is as follows:

\[ f(x) = \frac{X}{X^*} \quad (X^* \neq 0). \]  

Among them, \( X \) is the specific indicator that needs to be dimensionless in the indicator system, \( X^* \) is the indicator quantification benchmark, and the value of \( X^* \) is the median of each indicator in the 2018 data of the top 100 counties. The median is used here as the representative level of high-quality development in the country. Compared with the average, the median is determined by its middle position in the dataset to avoid being affected by extreme values. To a certain extent, the representativeness of the median to its dataset is improved, so the median can better represent the general level of this dataset than the mean.

Therefore, the final calculation formula of the comprehensive development index \( S \) of each country is as follows:

\[ S = \sum_{i=1}^{n} \omega_i x_i \quad (n = 1, 2, \ldots, 29). \]

### 3. K-Means Clustering Model of Spatial Effects of Latin American Countries

#### 3.1. Principle of K-Means Clustering Algorithm

K-Means defines the centroid of a cluster as the mean value of the points in the cluster. It is a complex clustering algorithm.
based on dividing. The model is relatively simple, easy to operate, and has good applicability to larger datasets. All K-Means receives are unlabeled datasets, which is an unsupervised learning algorithm. During the clustering process, K is used as a parameter to select \( n \) objects randomly, and each object represents a cluster’s initial average value. These objects are then divided into K clusters. And calculate the center distance of the remaining objects from each cluster and assign it to the nearest cluster to achieve similarity clustering. At this time, the mean value of each cluster will change, and it is necessary to recalculate the mean value of the corresponding cluster and iterate until the test function begins to converge.

Define a sample set containing \( n \) data as \( \Omega \), shown as

\[
\Omega = \{x_i | x_i = (x_{i1}, x_{i2}, \ldots, x_{is}), \quad i = 1, 2, \ldots, n\},
\]

where \( x_i = (x_{i1}, x_{i2}, \ldots, x_{is}) \) is a \( s \)-dimensional vector representing the \( s \) different features of the \( i \)-th data.

The cluster center point \( C \) is shown as

\[
C = \{c_j | c_j = (c_{j1}, c_{j2}, \ldots, c_{js}), \quad j = 1, 2, \ldots, k\},
\]

where \( c_j = (c_{j1}, c_{j2}, \ldots, c_{js}) \) is the center point of the \( j \)-th cluster, and each center point \( c_j \) contains \( s \) different features, where \( k \) is the total number of clusters.

**Definition 1.** Let the Euclidean distance between two data pieces \( x_i \) and \( c_j \) be \( \text{dis}(x_i, c_j) \); then \( \text{dis}(x_i, c_j) \) can be expressed as

\[
\text{dis}(x_i, c_j) = \sqrt{\sum_{s=1}^{d} (x_{is} - c_{js})^2}.
\]
shown as follows, where $N(\varphi_j)$ represents the number of data pieces in the same cluster.

$$c_{jk} = \frac{1}{N(\varphi_j)} \sum_{x \in \varphi_j} x_{k}.$$  

(12)

**Definition 3.** The criterion function is a parameter to measure the effect of the entire clustering. Only when it converges can each cluster achieve the best independence, which is shown as

$$E = \sum_{j=1}^{k} \sum_{x \in \varphi_j} \text{dis}(x_i, c_j).$$  

(13)

3.2. Selection of the Optimal K Value of K-Means Clustering Algorithm. Now we will analyze the spatial distribution characteristics of high-quality development levels in Latin American countries, that is, the relationship between the relative position of each country and the level of development. In this paper, the horizontal position coordinates, vertical position coordinates, and comprehensive index of each country are used as the three characteristics of the cluster analysis datasets for K-Means cluster analysis [17–23]. At the same time, the K value needs to be selected. K is a parameter the user specifies, that is, the number of expected clusters. The selection of the K value directly determines the effect of clustering. Currently, there are two different methods to determine the K value, external and internal, and the division is based on whether there is a benchmark for reference. An irrelevant method to determine the K value when there is a benchmark for reference is called an unsupervised method. In the absence of a benchmark reference, an intrinsic method is used to determine the K value. The effect of clustering is judged by the degree of cluster dispersion, called an unsupervised method.

3.2.1. External Method. Many metrics can be used to evaluate the performance of clustering models in extrinsic methods. Entropy and F-measure are typical representatives [24–26]:

Entropy: it indicates the degree of chaos in the system; the greater the entropy, the more chaotic the system. The specific calculation process of the entropy of each cluster $i$ is shown as follows, where $L$ is the number of classes, $p_{ij} = m_{ij}/m_i$ represents the probability that the data in cluster $i$ belongs to class $j$, $m_i$ represents the number of datasets in cluster $i$, and $m_{ij}$ is the sum of the number of values in class $j$ in cluster $i$, $e$ represents the total entropy of all clusters, $K$ represents the sum of the number of clusters, and $m$ represents the total number of datasets.

$$e_i = \sum_{j=1}^{K} p_{ij} \log_2 p_{ij},$$

$$e = \sum_{i=1}^{K} \frac{m_i}{m} e_i.$$  

(14)

F-measure: it describes the probability that a cluster contains an object of a single specified class and contains all objects of that class. The calculation process of cluster $i$ with respect to class $j$ is shown as

$$F(i, j) = \frac{2 \times \text{precision}(i, j) \times \text{recall}(i, j)}{\text{precision}(i, j) + \text{recall}(i, j)}.$$  

(15)

3.2.2. Internal Method. When the benchmark of the dataset is lacking, it is necessary to use an internal method to judge the quality of the clustering effect. The silhouette coefficient and the elbow method are the representative calculation methods.

Silhouette coefficient: Assuming that there are $n$ objects in the datasets $D$, clusters $C_1, C_2, \ldots, C_k$ indicate that the datasets are divided into $k$ clusters, and $o \in C_i (1 \leq i \leq k)$, $s(o)$ represents the silhouette coefficient of the object $o$, $a(o)$ represents the average distance between objects in the cluster where $o$ is located, and $b(o)$ represents the minimum average distance between objects except for the cluster where $o$ is located; then:

$$s(o) = \frac{b(o) - a(o)}{\text{max}(a(o), b(o))}$$

$$a(o) = \sum_{0 \in C_o, o \neq o'} \text{dist}(o, o') / |C_o| - 1$$

$$b(o) = \min_{C_j: 1 \leq j \leq k} \left\{ \frac{\sum_{0 \in C_j} \text{dist}(0, 0')}{|C_j| - 1} \right\}.$$  

(16)

Elbow method: The core idea is that as the value of $K$ increases, the similarity of the number of samples in each cluster will also increase, and the clustering error ($SSE$) of all samples will gradually decrease. And when $K$ gradually increases and does not reach the optimal value, the aggregation degree of each cluster will be greatly improved so that the value of $SSE$ will also be significantly reduced; when $K$ gets closer to the optimal value, the decrease of $SSE$ will gradually decrease until $K$ reaches the optimal value. Continuing to increase the value of $K$ at this time will gradually flatten the decline of $SSE$, and the inflection point will be the optimal point. The core indicator $SSE$ is shown as

$$SSE = \sum_{i=1}^{K} \sum_{p \in C_i} \|p - m_i\|^2,$$  

(17)

where $C_i$ is the $i$-th cluster, $p$ is the sample point in $C_i$, and $m_i$ is the centroid of $C_i$ (the mean of all samples in $C_i$).

4. K-Means Cluster Analysis of High-Quality Development in Latin American Countries

4.1. Selection of the Optimal K Value for the K-Means Clustering Algorithm in Latin American Countries. The method for selecting the value of $K$ has been introduced in the previous article, which is mainly divided into two categories:
the internal method and the external method. The external method is suitable for the situation with the benchmark, and the internal method is suitable for the case without a benchmark.

The clustering of Latin American countries belongs to no benchmark clustering [27–30]. Therefore, this paper chooses the internal method in the clustering method to determine the K value. The internal methods are mainly divided into the contour coefficient method and the elbow method. Compared with the elbow method, the silhouette coefficient not only introduces the degree of separation but also limits the degree of clustering. The obtained K value is not necessarily the optimal result, and it is even cumbersome to select the optimal K value again with the help of SSE. This paper uses the elbow method to choose the optimal clustering number K value after preprocessing the data. Figure 1 shows using the elbow method to select the optimal K value.

It can be seen from Figure 1 that the K value decreases gradually from 1 to 8. When K = 5, an inflection point occurs, and the decline rate tends to be stable eventually. Then an inflection point is K = 5 which should be the optimal K value point. So, for the number of clusters in Latin American countries dataset, the optimal number of clusters should be 5. Therefore, when K-Means clustering is performed on Latin American countries in this paper, K = 5 is substituted.

4.2. Spatial Effect Analysis of Country Development in Latin American Countries. This paper uses python to perform a K-Means clustering analysis on the dataset. It obtains the scatter effect diagram of the K-Means clustering algorithm, as shown in Figure 2.

It can be seen from Figure 2 that the clustering effect of the scatter plot based on the K-Means clustering algorithm is relatively apparent. Based on the geographical location and comprehensive index, the quantitative clustering of Latin American countries is divided into five regions, among which the points in the regions are relatively points outside the region that have more similar features. This figure shows that the high-quality development characteristics of countries in the same cluster have a certain similarity and regionality. To a certain extent, it reflects that the development level of each country in Latin American countries has a specific regional clustering, but the effect of clustering needs further analysis.

From the K-Means clustering results calculated here, the clustering distribution of Latin American countries based on K-Means is obtained, as shown in Figure 3. It can be seen from the figure that Latin American countries can be divided into five development clusters, namely, five economic belts. At the same time, most agglomeration areas have one or more countries with a relatively high level of development. Because of the influence of countries with higher levels of development, these communities come together to form development clusters.
5. Conclusions
This paper attempts to carry out systematic research and discussion by following the logic of “the theoretical analysis framework of the high-quality economic development of Latin American countries, the construction of the evaluation system of the high-quality economic development of Latin American countries, the establishment of the spatial effect clustering model of Latin American countries, the evaluation of the high-quality economic development level of Latin American countries, and the spatial effect analysis of the development of Latin American countries based on K-Means clustering algorithm.”

The interpretation of this study on the high-quality economic development of Latin American countries needs to be improved. In the process of evaluating the high-quality economic development of Latin American countries, the construction of the index system is not comprehensive enough and lacks more data support. In the future, with the progress of statistical technology, the index system can be improved through more channels. At the same time, this paper makes a cluster analysis on the high-quality economic development level of Latin American countries based on the spatial distance and comprehensive index, which can be further subdivided in the later research. According to the regional development characteristics, this paper makes a cluster evaluation and analysis on the high-quality economic development level of Latin American countries and further analyzes the complementarity of their development.

Data Availability
The dataset can be accessed upon request.

Conflicts of Interest
The authors declare that they have no conflicts of interest.

References


[27] L. Li, B. Lei, and C. Mao, "Digital twin in smart manufacturing," *Journal of Industrial Information Integration*, vol. 26, no. 9, Article ID 100289, 2022.

