Prediction Of Olympic Host Countries Based on Cluster Analysis, AHP, and EWM

Xiaojia Wang *, Simeng Li #, Haoya Peng #

School of Economics and Management, Beijing Jiaotong University, Beijing, China, 100044

* Corresponding Author Email: 21711031@bjtu.edu.cn

#These authors contributed equally

Abstract. The continuation of human civilization must ensure that the Olympic Games can be held stably to maximize the world's benefits. The paper combines normalization and cluster analysis to establish an evaluation model of countries suitable for hosting the Olympic Games. Then we keep data convergence to develop an accurate evaluation scheme based on the entropy method and hierarchical analysis. The countries currently capable of stably hosting the Olympic Games were the United States, China, and Japan. The change of each index over time was predicted based on the cubic smoothing index. We anticipate the countries capable of hosting the Olympic Games after the next ten years in additional countries, such as the United Kingdom, South Korea, Mexico, etc. The political factors are simplified into functions according to the geographical location of the countries respectively. We predict the new countries that will be able to host the Olympic Games in the next twenty years, such as Switzerland, Austria, etc.

Keywords: Normalization, Cluster Analysis, Entropy Weighting, AHP.

1. Introduction

1.1. Motivation

Enthusiasm for the bid for the Olympics once swept around the world as the successful presentation of the Olympic Games hosted in Beijing [1]. Numerous people regard it as glory, and countries competed drastically for the right to host the Olympic Games. However, an increasing number of countries like Italy, Spanish, and Hungary have officially abandoned the Olympic bid, and many residents were firmly against hosting the Olympics in their countries. This phenomenon suggested that more countries have realized the negative influences behind the brilliant appearance of hosting the Olympics. These impacts include short-term and long-term aspects, including the economy, politics, and climate.

In addition, many people need to pay attention to the fact that many developed countries are unwilling to host the Olympic Games. On the contrary, many large developing countries want to gain more and more agency in global affairs by hosting major international sporting events [2].

The International Olympic Committee has long argued that hosting the Olympic Games will give each country a rare chance to celebrate the human spirit of sport with its innovative ideas. People can forever change a community, region, and government by creating their unique environmental, social, and economic heritage [3]. There will be more difficulties for the hosting countries without implementing effective measures to address the current problems. The original meaning of the Olympics that “to cultivate human beings, through sport, and contribute to world peace,” might be deteriorated [4].

1.2. The main work

We aim to find a reform measure for the Olympic Games for the IOC and encourage countries worldwide to participate in the Olympic campaign actively. This effectively promotes economic globalization and the construction of communication bridges worldwide. We find that some countries need help with financial and population shortages by referring to the data of the World Bank, which leads to their need to be qualified to host the Olympic Games. So, we consider whether a country can

successfully host the Olympic Games as a research criterion. We then investigated data from the official websites of the United Nations Statistics Division, the World Health Organization, and the Organization for Economic Cooperation and Development. We collected some valuable indicators to guide our investigation. In addition, considering the influence of time on national indicators, we divided the period into the next five years and the next ten years to make more accurate and detailed recommendations to the IOC for the Olympic reform. Another objective is rekindling the world's enthusiasm for hosting the Games. At the same time, we consider each country's unique geographical circumstances and focus on minimizing losses and maximizing benefits for the host country.

1.3. Arrangement

The introduction introduces our research motivation, the main work, and the writing arrangement. While preliminary explains what indicators we set up and how to model them. Experiments are divided into three models, and their structure and calculation method are introduced in each model. The result shows the data obtained through the experiments and analyzes the results. The conclusion is our recommendations for the IOC based on the decision.

2. Preliminary

When describing some indicators, we use specific abbreviations to represent them for ease of expression:

CI means consistency index, *CR* means consistency ratio, *BC* means cost of building.

We also simplified some nouns:

AHP means Analytic Hierarchy Process, EWM means Entropy Weight Method, SES means Holt-Winters Exponential Smoothing, DES means Double Exponential Smoothing, and ES means Exponential Smoothing.

Table 1. Notations

Symbol	Description
θ_i	Type of country
a_{yz}	How important is *y*/how important is *z*
ω	Index weight
S	Score of capital stock
ΔS	Changes in score after hosted game
S_{final}	Score on the capital stock after hosted game
x_{ij}	*j* index data of the *i*-th sample country
$\widehat{x_{ij}}'$	Indicator in the sample country
W_j	Weight of the *j*-th indicator
ϕ_t	Smoothed values
$\alpha, \alpha(1-\alpha)$	Weighting factors

3. Experiments

3.1. Feasibility Impact Model

After selecting various indicators, we categorized them into three significant aspects to measure better the feasibility of a country hosting the Olympics. We counted seven secondary and 13 tertiaries economic, society, and environment indicators based on the strong hands. The secondary indicators of the economy are the inside and outside economy. For the inside economy, we considered the cost of building, GDP, and unemployment rate. The external economy was divided into tourism and country deals. Society has three secondary indicators: safety, life, and technology. Each has a tertiary indicator: political changes, national prestige, and technology development. Among them, Fragile

State Index presents political instability [5]. As to the environmental indicator, we grouped it into natural factors and human factors. The use of land and air quality are the sub-indicators of human factors, while biological factors include natural disasters, water, and vegetation.

We applied cluster analysis for each country: utilizing the data of the countries that had hosted the Olympics to estimate the data of those never hosted ones under the same category. Wherein we adopted the same weight to measure the amount of capital currently owned domestically for one category. The countries are divided into three sorts based on these indicators: Economic-Priority Countries (θ_1), Society-Priority Countries (θ_2) and Environment-Priority Countries (θ_3). In this case, we can comprehensively analyze the countries around the world.

3.1.1 The Structure of Feasibility Impact Model

The importance levels for distinctive indicators vary for different types of countries. Since Analytic Hierarchy Process (AHP) is famous for practical multiple-criteria decision-making, we used it to weigh the seven secondary indicators in the 3 kinds of countries [6].

Step 1: Construct a judgment matrix.

$$a_{yz} = \text{How important is } y/\text{how important is } z$$

Step 2: Consistency inspection

Each matrix is checked for consistency, and the consistency index *CI* is first calculated:

$$CI = \frac{\lambda_{max}-7}{n-1} \tag{1}$$

Calculate the consistency ratio *CR*.

$$CR = \frac{CI}{RI} \tag{2}$$

Step 3: Calculate the weight.

According to the arithmetic mean method:

$$\omega_y = \frac{1}{7}\sum_{z=1}^{7}\frac{a_{yz}}{\sum_{k=1}^{7}a_{kz}}, (y = 1, 2, \dots, 7) \tag{3}$$

The geometric mean method:

$$\omega_y = \frac{(\prod_{z=1}^{7}a_{yz})^{\frac{1}{7}}}{\sum_{k=1}^{7}(\prod_{z=1}^{7}a_{yz})^{\frac{1}{7}}}, (y = 1, 2, \dots, 7) \tag{4}$$

The eigenvalue method obtains the weight, the consequences are averaged, and the final weight is obtained comprehensively.

Then, we calculated the current capital stock score for each of the three groups of countries.

$$S_{\theta c} = \sum_{y=1}^{7} S_{source} \times (\theta_c, \omega_y) \tag{5}$$

Weighing the indicators, we selected to measure countries' feasibility is necessary, so we chose to use Entropy Weight Method (EWM). Through EWM, we can objectively calculate the information entropy, and the weight will be more significant when it is smaller. This impartial way enables us to make the most of original data to get a result with high accuracy and objectivity [7].

3.1.2 Data normalization:

The values of different indicators are positive or negative, representing other properties. They can be divided into benefit attribute index and cost attribute index. GDP, national prestige, country deal, tourism, water, satisfaction, and technology development are grouped in the benefit attribute index. At the same time, the cost attribute index includes the cost of building, unemployment rate, political changes, use of land, and air quality.

For the benefit attribute index, the larger the value, the better, we use:

$$\widehat{x_{ij}} = \frac{x_{ij}-min(x_i)}{max(x_i)-min(x_i)} \tag{6}$$

For the cost attribute index, the smaller the value, the better, we use:

$$\widehat{x_{ij}} = \frac{max(x_i) - x_{ij}}{max(x_i) - min(x_i)} \tag{7}$$

3.1.3 EWM

Suppose there are *n* countries that have already hosted the Olympics in total.

Since the time and space distribution of each country is relatively discrete, there are still indicators with a value of 0 after standardization, and the translation is performed:

$$\widehat{x_{ij}}' = \widehat{x_{ij}} + 0.0001 \tag{8}$$

$$P_{ij} = \frac{\widehat{x_{ij}}'}{\sum_{i=1}^{n} \widehat{x_{ij}}'} \tag{9}$$

$$e_j = -\ln n \sum_{j=1}^{n} \widehat{x_{ij}}' \ln \widehat{x_{ij}}' \tag{10}$$

$$W_j = \frac{1 - e_j}{\sum(1 - e_j)} \tag{11}$$

After analyzing the existing data, we got the weights for each indicator, and we concluded the weighted summation formula of the composite value:

$$\Delta S = \sum W_j X_{ij} \tag{12}$$

$$S_{final} = S_{\theta c} + \Delta S \tag{13}$$

Based on the calculated scores, we defined the top 8 countries as the “Most Feasible”, the 8 to 20 countries as “Relatively Feasible”, the 20 to 50 countries as “Mildly Feasible”, and after 50 countries are defined as “Infeasible”. Then, the ratings of countries around the threshold of each rank were averaged to obtain the criteria score for evaluating the “impact on indicators.”

In addition, we used Happiness Index to evaluate residents’ satisfaction and attitudes about hosting the Olympics [8]. This involved the evaluation of residents’ freedom to make life choices, social support, healthy life expectancy, etc.

Apart from residents’ satisfaction, the degree of satisfaction of athletes plays a significant role in the Olympics’ host. According to Maslow’s Hierarchy of Needs, we divided the five levels of needs into two parts: material needs, which include the demands for belongingness, safety, and physiological needs [9]. The other is about the spiritual demands of pursuing self-actualization and gaining esteem. We have already weighted the cost of building (BC) indicator for the physical needs, which includes constructions, diets, services, and necessary facilities for hosting the Games. Regarding athletes’ spiritual needs, the number of medals they received was considered. Therefore, aimed at the Winter and Summer Olympics, we selected all countries that have hosted the Olympics. Then we calculated the difference between the number of medals they won for the rest of the Games and the number of medals they won as hosts. Furthermore, weights of 0.5, 0.2, and 0.3 were assigned respectively to the gold, silver, and bronze medals according to psychological theory.

3.2. Time Impact Model

The triple exponential smoothing method can predict the time series with trend and seasonality, so we utilized the Holt-Winters Exponential Smoothing (SES) and the secondary curve trend of Double Exponential Smoothing (DES) to counteract the lag of the Simple Exponential Smoothing. We employ a triple exponential smoothing forecasting method since the changes of various indicators in most countries show a non-linear increasing trend [10].

We assess the scores by bringing these values to the model above after averaging each indicator of current countries on the brink of the first 10 and 25 countries. The lowest score to be a “Most Feasible” country is 0.75, while the lowest score for “Relatively feasible” is 0.29. Therefore, we use

Exponential Smoothing (ES) for the previous *n* years of data to make predictions. However, we only predicate the next ten years' data since this method can only forecast short-term data.

Data from *n* periods ago do not affect future discounts if the historical data of the recent *n* year has the same effect on future values.

$$\begin{cases} \Phi_t^{(1)} = \alpha y_t + (1-\alpha)\varphi_{t-1}^{(1)} \\ \Phi_t^{(2)} = \alpha \varphi_t^{(1)} + (1-\alpha)\varphi_{t-1}^{(2)} \\ \Phi_t^{(3)} = \alpha \varphi^{(2)} + (1-\alpha)\varphi^{(3)} \end{cases} \tag{14}$$

we set the average value of the first three data for the initial value of the smoothing annotation once, twice, and three times, to determine the initial value and the smoothing coefficient.

The weighting coefficient α is 0.5 for the increasing index GDP, technological development, etc. The weighting coefficient α is 0.25 when the index with large fluctuations, such as FSI, air quality, tourism, etc.

3.3. Policy Impact Model

We searched and encapsulated the policies enacted in "the Most Feasible" countries over the recent years based on the impact of the timeline on each indicator. The lands were divided into three categories based on their geographical location based on setting these policies: island countries, coastal countries, and landlocked countries.

Policies are essential to countries, and they can permeate and change society. We added another 10-year forecast to the next 10-year forecast under the effect of the timeline since the impacts of policies are long-term and have lag [11]. Hence, a holistic assessment shows how these policies influence the indicators and the final score.

In coastal countries, most countries are taking measures to strengthen the protection of oceans and forests to achieve sustainable development. For example, America has adopted the Global Program of Action to protect the Marine Environment from Land-Based Activities [12]. Although many countries need to develop tourism, preserving the environment is more important. We approximate the impact of the policy on coastal countries as a model. The growth rate will reach a peak and gradually remain stable, like the logistic regression curve.

Island countries desire to avoid economic stagnation and recession. So, most countries try to develop nearby islands as tourist attractions and even build islands for reclamation [13]. In addition, the impact of the pandemic, such as Covid, is weakening, the tourism industry is gradually rising, and the frequency of international trade is also increasing progressively [14]. Some countries, like Japan, have also adopted the pollution-control agreement to reduce the impact of environmental pollution on national development under constraints [15]. Therefore, we approximate the changing trend of each index of the island countries as an exponential growth curve.

Most landlocked countries with the competence to host the Olympics have relatively small land areas with numerous mountains. For instance, Switzerland's Law on Spatial and Regional Planning has been revised since its implementation constantly. This policy assures these landlocked countries to raise the land utilization rate. Nevertheless, the upper limit of benefits that land can bring was reduced [16]. Furthermore, the United Nations has issued an appeal to the international community. The request will support the implementation of the World Trade Organization's Trade Facilitation Agreement with neighboring transit countries because the landlocked countries face challenging transportation conditions. This will reduce landlocked countries' costs in their international trade activities [17]. Landlocked countries enable a short-term increment in the benefits of international trade. However, the economy in these regions may stagnate or even backfire under the timeliness of policy [18]. Therefore, the changing trend of each indicator in landlocked countries is approximated to an exponential decay function.

4. Result

4.1. The establishment of the simulation model

Through AHP, we got the result that in Economic-Priority Countries, natural factors of the environment have the most weight, which is 0.3217. They emphasize the inside economy more than the outside one: the consequences are 0.1720 and 0.1501, respectively. Safety and technology are less critical for those countries. For Society-Priority Countries, the protection weight is 0.0622, more than the other two types of countries. Besides, the life indicator weighs 0.1399 in this type of country. For Environment-Priority Countries, the weight of natural environmental factors is 0.3338, and the human aspects of the environment consider 0.1167. The inside economy indicator weighs 1.677, while the outside one weighs 0.1297. In short, all these countries have greater weight in economy and environment, but they have their priorities.

We got the weights for each indicator in Table 2 based on the EWM.

Table 2. Indicators' weights

Level 1	The weight of Level 2/Level 1	Level 2	The weight of Level 3/Level 2	Level 3	The weight of Level 3/Level 1
Economy	0.1699	Inside	0.3348	BC	0.0569
			0.3483	GDP	0.0592
			0.3169	UR	0.0539
	0.1427	Outside	0.5245	TR	0.0748
			0.4755	CD	0.0678
Society	0.0579	Safety	1.0000	PC	0.0579
	0.1300	Life	1.0000	NP	0.1300
	0.0627	Technology	1.0000	TD	0.0627
Environment	0.1152	Human factors	0.4715	LU	0.0543
			0.5285	AQ	0.0609
	0.3214	Natural factors	0.6295	ND	0.1058
			0.3705	WA	0.0623
			0.4769	VG	0.1533

We figured out that the current average Happiness Index of the world is 0.5305 after analyzing the data.

Regarding countries that have never hosted the Olympics, we used their medal count from previous games as a base to predict how much satisfaction a country's athletes would get from hosting the Games. Based on the trend, MEDALS won by many host countries and ended up with an average score of 0.5347.

Therefore, we will only consider countries that satisfy the satisfaction of residents and athletes as the criteria for hosting the Olympic Games. This ensures that the Olympic Games can help and inspire all humanity spiritually.

Almost all countries get higher scores in our assessment system with time flowing. After ten years, among the countries ranked 11th to 25th, the ratings of the UK, Mexico, and South Korea will be 1.103, 0.851, and 0.864, respectively, based on the ratings and rankings obtained from Model 1, satisfying to become "Most Feasible" country conditions. These three countries will also host the Olympic Games.

After 20 years of different policies promoting countries' development, only a few countries are infeasible to host the Olympics. They are Austria, Kazakhstan, Czech, Austria, and Peru (scores 0.699, 0.469, 0.644, 0.347, and 0.526, respectively).

4.2. Analysis of experimental results

We estimated when the number of emergences occurring is *d* according to the Poisson distribution, taking annual contingencies (e.g., natural disasters, terrorist attacks, etc.) into account:

$$P(x = d) = \frac{e^{-\lambda}\lambda^{d}}{d!} \tag{15}$$

An average increase in economic indicators of G=1.22 per unexpected situation was forecasted after searching the UN's information.

Random factors have little relevance in evaluating the total score, so our model is reliable and stable.

5. Conclusion

Some countries, like the developed nations America and Japan, and the developing nations Russia and China, have significant potential to host the Olympic Games with prominent scores and advantages. In the near term, Paris and Los Angeles are favored for the Summer Olympics, while Milan and Cortina are selected for the Winter Olympics as expected.

We recommend selecting and notifying two countries based on our analysis of viable countries. They will be the Olympic Games' regular hosts. International Olympic Committee and the Allies can help them prepare for the negative economic, social, and climate impacts that may arise from hosting such events. This should be done before the 2030s after these two fixed countries should host the Summer and Winter Olympics.

Between the 2030s and 2040s, International Olympic Committee should give the Olympic Games host countries four seasons. Two additional selected countries to host the Spring and Autumn Olympics regularly. These two countries should be reminded to prepare their venues and finances adequately.

The four countries will be expected to host all four Olympic Games regularly after the 2040s. If global interest in the Olympic Games continues to grow, or if the world economy continues to develop positively, the number of countries capable of hosting the Olympic Games is expected to increase to eight.

References

[1] Dongfang Qiao. Analysis of risk countermeasures for hosting the Olympic Games - Insights based on the abandonment of hosting the Olympic Games by cities such as Hungary [J]. Contemporary Sports Science and Technology,2018,8(35):255-257.

[2] Jonathan Grix, Donna Lee. Soft Power, Sports Mega-Events and Emerging States: The Lure of the Politics of Attraction. Global Society, 2013,27(4), 521–536.

[3] Lucy Schouten. Should the Olympics Stop Moving Around? The Christian Science Monitor,2016, August 3.

[4] Brown, J. (n.d.). What is the primary purpose of the Olympic Games? The Knowledge Burrow, 2021, May 21.

[5] Carlsen, Lars, and Rainer Bruggemann. "Fragile State Index: Trends and Developments. A Partial Order Data Analysis." [J]. Social Indicators Research, 2017, 133(1): 1–14.

[6] Lipovetsky, Stan. Understanding the Analytic Hierarchy Process [J]. Technometrics: A Journal of Statistics for the Physical, Chemical, and Engineering Sciences, 2021, 63(2): 278–279.

[7] Kumar, R., Singh, S., Bilga, P. S., Jatin, Singh, J., Singh, S., Scutaru, M.-L., & Pruncu, C. I. Revealing the benefits of entropy weights method for multi-objective optimization in machining operations: A critical review [J]. Journal of Materials Research and Technology, 2021, 10: 1471–1492.

[8] Mevawala Jency. "Happiness Index"-The Footsteps towards Sustainable Development [J]. International Research Journal of Engineering and Technology, 2019,6(12): 615-619.

[9] Taormina, Robert J., and Jennifer H. Gao. “Maslow and the Motivation Hierarchy: Measuring Satisfaction of the Needs.” [J]. The American Journal of Psychology, 2013, 126(2): 155–177.

[10] Marek Lubos, Vrabec Michal. ARIMA models and exponential smoothing [J]. 33rd International Conference Mathematical Methods in Economics, 2015, 502-507.

[11] A. W. Phillips. Stabilisation Policy and the Time-Forms of Lagged Responses [J]. The Economic Journal, 1957, 266(67): 265–277.

[12] U.S. Department of Commerce. FACT SHEET: 2022 National Travel and Tourism Strategy, 2023.

[13] Henderson, J. C. Developing and Managing Small Islands as Tourist Attractions [J]. Tourism and Hospitality Research, 2001, 3(2): 120–131.

[14] Lin L Y, Tsai C C, Lee J Y. A study on the global cruise tourism industry trends, sustainable development, and the impacts of the COVID-19 pandemic [J]. Sustainability, 2022, 14(11): 6890.

[15] Fernando Ortiz-Moya. Green growth strategies in a shrinking city: Tackling urban revitalization through environmental justice in Kitakyushu City [J]. Journal of Urban Affairs, 2020, 42(3), 312-332.

[16] Switzerland, C. F. S. The governance of land use. Oecd.org [EB]. 2023.

[17] Unctad.org. United Nations Conference on Trade and Development, Fifteenth Session [EB], 2021.

[18] Bouzas R, Keifman S. Making trade liberalization work [J]. After the Washington Consensus: Restarting Growth and Reform in Latin America, 2003: 157-79.