2025

Mathematical Contest In Modeling®

Certificate of Achievement

Be It Known That The Team Of

Yi Lin
WangJie Xia
Zekai Wu

With Faculty Advisor

Guochun Ma

Of

Hangzhou Normal University

Was Designated As

Honorable Mention

Benjamin J. Galluzzo, Executive Director

Administered by COMAP

Steven B. Horton, Contest Director

With support from

AMS ASA informs MAA MATHEMATICAL ASSOCIATION OF AMERICA siam

Problem Chosen	2025 MCM/ICM Summary Sheet	Team Control Number
C		**2520025**

The Numerical Code Behind Olympic Medals: Constructing a Medal Prediction Model

Summary

The Olympics' growing importance highlights a nation's prowess and culture. Medal tallies reflect advancements in sports, economic might, and societal investment. While previous research has examined various factors influencing Olympic medals, few studies have quantitatively assessed the impact of "great coaches" or event diversity on medal distribution. Our framework primarily consists of three components: **(1)** Olympic Medal Prediction System(OMPS), **(2)** Coach Influence Analysis Model (CIAM), and **(3)** Olympic Strategic Insight Model (OSIM). Extensive experiments demonstrate the effectiveness and robustness of our approach.

For the **Olympic Medal Prediction System**, we developed a novel *MedalFlow Predictor (MFP)*, combining negative binomial regression and ARIMA, to forecast gold and total medal counts for each nation at the 2028 Los Angeles Games, accounting for **overdispersion in medal counts** and **temporal dependencies**. To account for nations yet to win medals, we developed a **real-time**, **small-sample** detector, *LogiSVM*, which estimates medal probabilities and captures nonlinear relationships. By analyzing historical performance, athlete development, and policy support, the model offers a precise evaluation of emerging nations' potential. We incorporate **interaction effects between countries and events**, along with **international competition levels**, based on *MFP*, to identify event categories that significantly impact medal counts,exploring the relationship between Olympic events and medal counts. The model demonstrates good predictive ability and interpretability based on various evaluation metrics and *5K-Cross* validation.

For the **Coach Influence Analysis Model**, we propose an **interpretable** *Coach Coefficient* to quantify the impact of "great coach" on medal counts, using the exponential result to represent the multiplicative effect of coaches on medal outcomes. By analyzing the regression coefficients for various sports, we identify sports most significantly influenced by coaching. We provides targeted coach investment recommendations for key sports in countries — *China* (volleyball), *the United States* (gymnastics), and *Germany* (taekwondo).

In the **Olympic Strategy Insight Model**, we find that the combined effects of economic and population indicators significantly impact medal counts, with their interaction playing a key role. Event diversity influences medal distribution, and countries should allocate resources and select events based on their strengths. The specialization of sports disciplines correlates with medal efficiency, prompting countries to optimize resource allocation. Socio-cultural factors are also crucial, and nations should focus on fostering a strong sports culture. Long-term medal trends reflect the sustainability of sports development, and *National Olympic Committees* should adjust strategies to maintain competitiveness.

Keywords: Olympic Games, Medal Prediction, Great Coach Effect

Contents

1 Introduction

1.1 Research Background

The Olympic Games, as the most prestigious multi-sport event in the world, attracts global attention.[5] The number of medals a country wins in the Olympics is not only a symbol of its athletic achievements but also reflects its national strength, social development, and cultural influence in various aspects. With the continuous development of the Olympic Games, accurately predicting the medal tally of each country has become an important research topic. This helps nations develop more scientific sports development strategies, allocate resources effectively, and enhance their competitiveness in international sports events. Moreover, understanding the factors influencing the number of medals.

1.2 Literature Review

Previous research on Olympic medal prediction has used various methods. *Wang Guofan* [2] applied ARIMA to forecast future medal counts based on historical data, analyzing past performance trends. However, this approach often overlooks other influential factors. *Yang Yucang* [1] integrated economic indicators, such as Gross Domestic Product (GDP), recognizing that economic strength drives sports development through investments in facilities, athlete support, and sports science. *Wang Yupeng* [3] also included population factors, as a larger population can provide a broader talent pool. Literature, including *Li Fawei* [4], highlights the critical role of coaches in improving athlete performance, responsible for technical guidance, tactical planning, and motivation.

However, few studies have quantitatively analyzed the impact of "great coaches" on medal counts across multiple Olympic Games. Regarding event-related factors, research shows that the number and types of Olympic events significantly affect medal distribution, with traditional sports contributing more medals due to a higher number of competitions, while emerging or niche sports have a smaller impact.[6]

1.3 Restatement of the Problem

Despite existing research, there are still limitations. Prediction models based solely on historical data lack comprehensiveness, while those incorporating multiple factors need further refinement in accuracy and handling complex data relationships. The "great coach" effect requires more systematic and quantitative analysis to accurately assess its contribution to national medal tallies. This paper makes the following contributions:

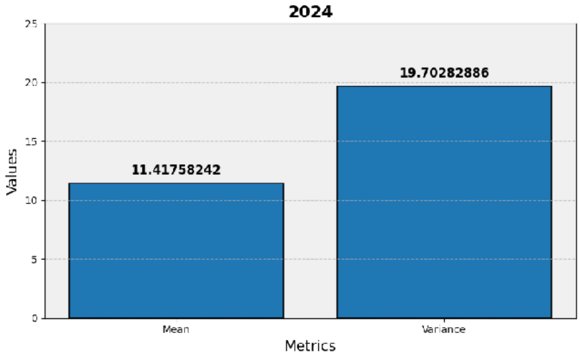
- **We construct a comprehensive mathematical model to predict the number of gold medals and total medals for each country in the 2028 Los Angeles Summer Olympics.**
- **We establish a model to quantitatively analyze the impact of "great coach" on a country's medal tally, selecting specific countries and key sports for in-depth study.**
- **Based on the above model, we explore other unique insights regarding Olympic medal counts and provide practical suggestions for national Olympic committees.**

2 Assumptions and Justification

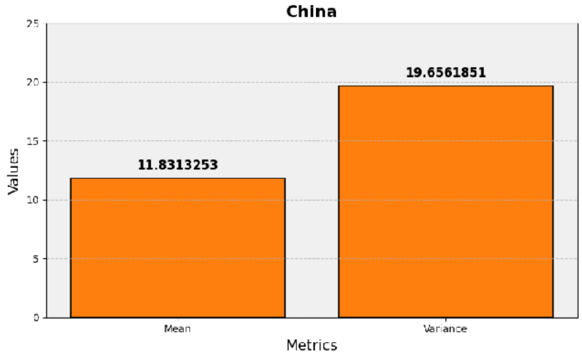
Distribution Assumption: We analyze the mean and variance of medals won by different countries, as shown in Figure 1(a), and the mean and variance of a country's performance across different Olympic Games in Figure 1(b). It is observed that the variance exceeds the mean, indicating **overdispersion**. Therefore, we assume that the number of medals follows a negative binomial distribution, i.e.,

$$G_{c,t} \sim NegBin(\mu_{c,t}, \phi),$$

where $\mu_{c,t}$ represents the expected number of gold medals for country c in the t-th Olympic.



(a)



(b)

Figure 1: **Discrete phenomenon of Olympic medals.** (a) 2024 avg. & var. of medals won by each country. (b) China's avg. & var. of medals in all Olympics.

Random Effect Independence Assumption: To model country-specific unobserved differences and systemic effects of the Olympic Games, we introduce random effects u_c and v_t. We assume that u_c and v_t are independent and normally distributed, i.e.,

$$u_c \sim N(0, \sigma_u^2) \quad \& \quad v_t \sim N(0, \sigma_v^2).$$

In practice, although there might be some underlying correlation, this simplification is reasonable as it clearly separates the country-level factors and the Olympic-level factors influencing medal counts.

Linearity Assumption: To make the model simple and interpretable while closely approximating the real-world relationships, we assume that the log of the expected number of gold medals, $\log(\mu_{c,t})$, is a linear combination of various predictors, including historical performance $HistoricalGold_{c,t}$, host country status $Host_{c,t}$, event-related factors $S_{t,k}$, and the random effects u_c and v_t.

Poisson Distribution Assumption: Since medals won by a country in a specific sport event are rare events, we assume that the number of medals $M_{c,s,t}$ won by country c in sport s at the t-th Olympic Games follows a Poisson distribution, i.e.,

$$M_{c,s,t} \sim Poisson(\lambda_{c,s,t}).$$

Linear Relationship Assumption: To intuitively reflect the relationship between the coach variable, fixed effects, random effects, and the expected medal count, we assume that the log of the expected medal count, $\log(\lambda_{c,s,t})$, has a linear relationship with the coach variable

$Coach_{c,s,t}$, country fixed effects γ_c, event fixed effects δ_s, Olympic Games fixed effects ϵ_t, and the country-event random effects $\eta_{c,s}$, i.e.,

$$\log(\lambda_{c,s,t}) = \alpha + \beta \cdot Coach_{c,s,t} + \gamma_c + \delta_s + \epsilon_t + \eta_{c,s}.$$

In practice, the introduction of a great coach is expected to have a relatively stable impact on the expected number of medals, with other fixed and random effects serving as adjustments to this impact.

3 Notation

Table 1: **Notations Table**

Notation	Definition
$G_{c,t}$	Gold medals for country c in the t-th Olympics
$\mu_{c,t}$	Expected gold medals for country c
ϕ	Overdispersion parameter
u_c	Country-specific random effect
v_t	Year-specific random effect
$HistoricalGold_{c,t}$	Past gold medal count
$Host_{c,t}$	Host country indicator (1 or 0)
$S_{t,k}$	Number of events in category k
$ARIMA(G_{c,t})$	ARIMA forecast for $G_{c,t}$
$M_{c,s,t}$	Medals in event s for country c
$\lambda_{c,s,t}$	Expected medals in event s
$Coach_{c,s,t}$	Great coach indicator (1 or 0)
γ_c	Country fixed effect
δ_s	Event fixed effect
ϵ_t	Year fixed effect
$\eta_{c,s}$	Random effect for country c and event s
β	Regression coefficient for coach effect
e^{β}	Coach effect on medals
$\beta_{3,k}$	Effect of event category on medals
$Tech_k$	Technical difficulty of event k
$Comp_k$	Competitiveness of event k
$\gamma_{3,k,c}$	Interaction effect for event k and country c
Y_c	First-time medal in 2028 (1 or 0)
$P(Y_c = 1)$	Probability of first-time medal in 2028
$X_{c,2028}$	Features for country c in 2028
w	SVM weights
b	SVM bias
θ	Model parameters
$p(\theta \mid D)$	Posterior of parameters
$G_{c,2028}(n)$	Predicted medals for country c in 2028
L, U	Prediction interval bounds
τ	Threshold for first-time medal

where we define the main parameters while specific value of those parameters will be given later.

4 Model Overview

This work builds a comprehensive modeling system around Olympic medal counts, aiming to predict the number of medals, analyze the impact of "great coach" and reveal other related influencing factors. The goal is to provide in-depth insights and practical guidance for sports strategies.

4.1 Olympic Medal Prediction System

To forecast gold and total medals for countries at the 2028 Los Angeles Summer Olympics, we utilized the multi-level *MedalFlow Predictor (MFP)*. The logarithm of expected gold medals ($\log(\mu_{c,t})$) is modeled as a linear combination of various factors. The model also incorporates random effects (u_c and v_t) to account for country-specific differences and Olympic-specific systematic changes. By fitting historical data from 1896 to 2024, we obtain parameter estimates and confidence intervals for future predictions. The MFP model captures the complex dynamics of Olympic medal distribution, accounting for both country-specific and event-specific factors.

4.2 Coach Influence Analysis Model

To assess the impact of "great coach" on medal counts, we constructed a multi-level Poisson regression model. The medal count $M_{c,s,t}$ for a specific sport event and Olympic Games follows a Poisson distribution, i.e., $M_{c,s,t} \sim Poisson(\lambda_{c,s,t})$. The log of expected medal count $\log(\lambda_{c,s,t})$ is linearly related to the presence of a "great coach" $Coach_{c,s,t}$, with additional country fixed effects γ_c, sport-specific fixed effects δ_s, Olympic year fixed effects ϵ_t, and country-sport random effects $\eta_{c,s}$. Model parameters are estimated through Maximum Likelihood Estimation (MLE). The effect of "great coach" on medal counts is quantified through the exponential of the regression coefficient β, i.e., e^{β}, representing the **multiplicative effect** of having a "great coach" on medal counts compared to not having one. The effect can be specifically analyzed using examples like Chinese volleyball, American gymnastics, and German taekwondo.

4.3 Insights from Other Influencing Factors

Through an in-depth analysis of the models, several key factors influencing Olympic medal counts have been identified. Economic and demographic indicators have a compound effect, with the interaction between economic strength and population size significantly impacting medal outcomes. National Olympic Committees should recognize the synergistic effects of these factors and allocate resources effectively. Event diversity greatly affects medal distribution, with different categories contributing variably to the total count. Countries should strategically allocate resources based on their strengths in specific sports. The host country effect not only boosts medals in the hosting year but can also have a lasting impact on subsequent Olympics. Countries can leverage the experiences of host nations to enhance their own sports development. Specialization in certain sports is closely tied to medal efficiency, and countries should identify efficient sports to optimize training and resource distribution. Social and cultural factors also play a critical role in enhancing competitiveness in specific events. Countries should foster a strong sports culture and capitalize on cultural advantages. Long-term medal trends reflect the sustainability of sports development, and National Olympic Committees should monitor these trends to adjust strategies and maintain international competitiveness.

5 Olympic Medal Prediction Model System

5.1 MedalFlow Predictor (MFP)

To predict the number of gold and total medals for each country in the Olympics, we propose a multi-level negative binomial regression model, *MedalFlow Predictor (MFP)*. The gold medal count $G_{c,t}$ for country c in the t-th Olympics is assumed to follow a negative binomial distribution: $G_{c,t} \sim \text{NegBin}(\mu_{c,t}, \phi)$, where $\mu_{c,t}$ represents the expected gold medal count and ϕ denotes the overdispersion parameter. Using a log-link function, the logarithm of $\mu_{c,t}$ is expressed as a linear combination of relevant predictor variables:

$$\log(\mu_{c,t}) = \alpha + \beta^{\top} X_{c,t} + u_c + v_t,$$

where the predictor variables $X_{c,t}$ include historical performance $HistoricalGold_{c,t}$, whether the country is the host $Host_{c,t}$, and other factors. Considering the time-series nature and trends in medal counts, we apply the **ARIMA (AutoRegressive Integrated Moving Average)** to capture the trend and cyclic changes in historical data. ARIMA is used to predict the gold medal count $\hat{G}_{c,t}$ for country c in the t-th Olympics:

$$\hat{G}_{c,t} = ARIMA(G_{c,t}) + \alpha + \beta^{\top} X_{c,t} + u_c + v_t$$

Here, $ARIMA(G_{c,t})$ represents the ARIMA model's prediction of the gold medal time series, while $\alpha + \beta^{\top} X_{c,t} + u_c + v_t$ combines the predictor variables from the basic regression model. Figure 2 shows the first and second order differences of historical medal data for *Japan*, *France*, and *the United States*.

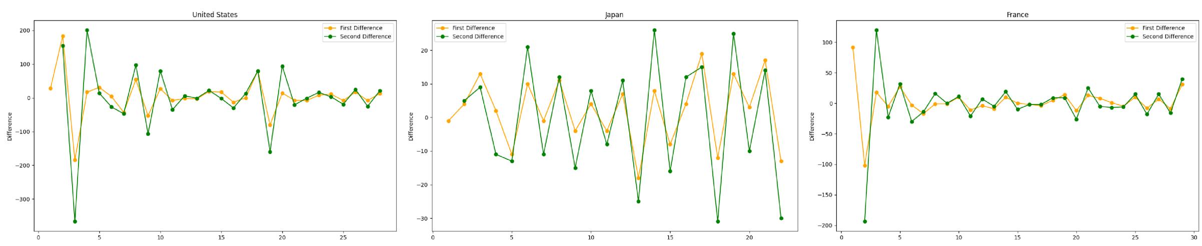


Figure 2: **The first and second order differences for Japan, France, and the United States in the ARIMA.**

To evaluate the model's performance, we split the dataset into K non-overlapping subsets $D_1, D_2, \ldots, D_K$. In each iteration, one subset D_i serves as the test set, and the remaining $K-1$ subsets are used for training. The model is trained and evaluated on the test set using the performance metric S_i.

$$S_i = w_1 \times \left(1 - \frac{\sum_{G_{c,t}\in D_i}(\hat{G}_{c,t} - G_{c,t})^2}{\sum_{G_{c,t}\in D_i}(G_{c,t} - \bar{G}_c)^2}\right) + \frac{1}{|D_i|}\sum_{G_{c,t}\in D_i}(\hat{G}_{c,t} - G_{c,t})^2$$

where w_1, w_2 are weight factors satisfying $w_1 + w_2 + w_3 = 1$ and $w_1, w_2, w_3 \geq 0$. In the experiments, we set $K = 5$, $w_1 = w_2 = \frac{1}{2}$. For uncertainty analysis, we use a **Bayesian approach** with **MCMC (Markov Chain Monte Carlo)** sampling to obtain the posterior distribution of the model parameters. Let the parameter vector be θ and the observed data be $D = \{G_{c,t}\}$. By **Bayes' theorem**, the posterior distribution of θ is:

$$p(\theta|D) = \frac{p(D|\theta)p(\theta)}{p(D)},$$

Here, $p(D|\theta)$ is the likelihood function, $p(\theta)$ is the prior distribution, and $p(D)$ is the normalization constant. Using MCMC sampling, we draw N samples $\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(N)}$ from the posterior distribution $p(\theta|D)$:

$$\log(\mu_{c,t}) = \alpha + \beta^\top X_{c,t} + u_c + v_t,$$

Let $X_{c,2028}$ be the corresponding predictor variables for 2028. For each $\theta^{(n)}$ ($n = 1, 2, \ldots, N$), we calculate:

$$\mu_{c,2028}^{(n)} = \exp\left(\alpha^{(n)} + (\beta^{(n)})^\top X_{c,2028} + u_c^{(n)} + v_{2028}^{(n)}\right).$$

The values $\mu_{c,2028}^{(1)}, \mu_{c,2028}^{(2)}, \ldots, \mu_{c,2028}^{(N)}$ form the empirical distribution of $\mu_{c,2028}$. Since $G_{c,t} \sim NegBin(\mu_{c,t}, \phi)$, for 2028, given $\mu_{c,2028}^{(n)}$ and the fixed overdispersion parameter ϕ, a sample $G_{c,2028}^{(n)}$ is drawn from $NegBin(\mu_{c,2028}^{(n)}, \phi)$, repeated N times to obtain the predicted distribution of $G_{c,2028}$:

$$G_{c,2028}^{(1)}, G_{c,2028}^{(2)}, \ldots, G_{c,2028}^{(N)}, \; G_{c,2028}^{((1))} \leq G_{c,2028}^{((2))} \leq \cdots \leq G_{c,2028}^{((N))}.$$

The 95% prediction interval is given by the lower bound $L = G_{c,2028}^{(\lfloor 0.025N \rfloor)}$ and the upper bound $U = G_{c,2028}^{(\lceil 0.975N \rceil)}$. This interval $[L, U]$ captures the uncertainty in predicting the gold medal count for 2028.Our prediction results are shown in Figure 3.

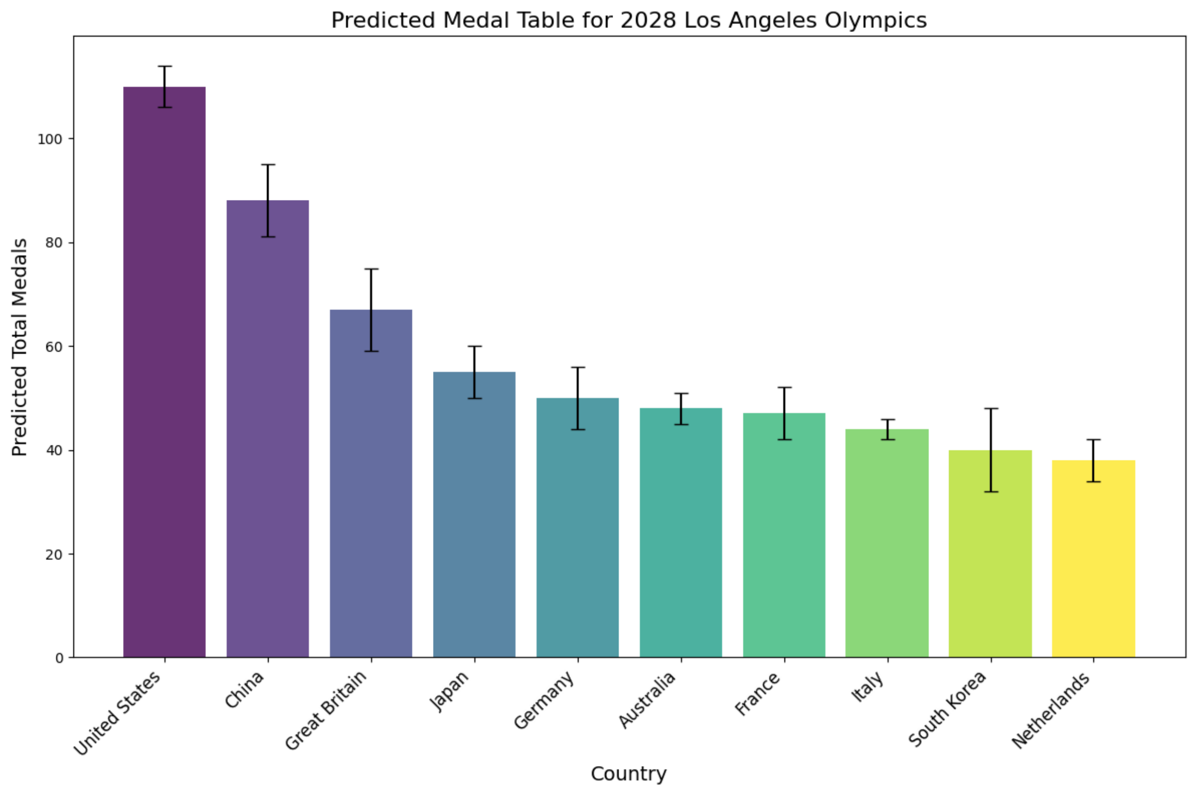


Figure 3: **2028 Olympics top 10 countries' predicted medal counts**

5.2 Analysis of National Gold Medal Trends

To predict gold medal performance for the 2028 Olympics and identify countries likely to improve or decline, we propose a systematic approach. This includes defining the rate of change in gold medal counts, integrating historical trends with future predictions, analyzing host country effects, and assessing parameter uncertainty. The rate of change in the gold medal count $\Delta G_{c,t}$ for country c at the t-th Olympics is defined to measure the trend in its gold medal performance:

$$\Delta G_{c,t} = \frac{\hat{G}_{c,t} - G_{c,t-1}}{G_{c,t-1} + \epsilon},$$

where $\hat{G}_{c,t}$ represents the predicted gold medal count for country c at the t-th Olympics, $G_{c,t-1}$ represents the actual gold medal count from the previous Olympics, and ϵ is a small positive constant (e.g., 10^{-6}) to avoid division by zero. If $\Delta G_{c,2028} > 0$, it indicates that country c is expected to improve its gold medal performance, while $\Delta G_{c,2028} < 0$ suggests a decline.

To more comprehensively assess national performance trends, we propose a composite evaluation metric $T_{c,2028}$ that combines results from the time series model and the Basic Medal Prediction Model (BMPM):

$$T_{c,2028} = w_1 \cdot \Delta G_{c,t}^{ARIMA} + w_2 \cdot \Delta G_{c,t}^{BMPM},$$

Here, $\Delta G_{c,t}^{ARIMA}$ represents the rate of change in gold medal counts from the ARIMA model, $\Delta G_{c,t}^{BMPM}$ is derived from the BMPM model, and w_1 and w_2 are weight parameters such that $w_1 + w_2 = 1$. Typically, $w_1 = w_2 = 0.5$ is used to balance the influence of both methods. When $T_{c,2028} > 0$, country c's gold medal performance is expected to improve, with larger values indicating greater improvement. Conversely, when $T_{c,2028} < 0$, it signals a decline, with smaller values indicating more significant deterioration.

To better analyze the performance of MFP, we conduct a comprehensive evaluation.

- **Analysis of Host Country Effects.** The host country effect significantly influences changes in gold medal counts. Let $Host_{c,t}$ be the host country indicator, and β_{Host} be its corresponding coefficient, representing the strength of the host country effect. The incremental gold medal count due to hosting, $\Delta G_{c,t}^{Host}$, is expressed as,$\Delta G_{c,t}^{Host} = \exp(\alpha + \beta_{Host} + \sum_j \beta_j X_{j,c,t}) - \exp(\alpha + \sum_j \beta_j X_{j,c,t})$, where $X_{j,c,t}$ represents other feature variables, α is the intercept, and β_j are the regression coefficients for the features. When $\beta_{Host} > 0$, the host country's gold medal count is expected to increase significantly.

- **Combining Historical Performance and Time Series Trends.** The regression coefficient for historical gold medal counts ($HistoricalGold_{c,t}$), β_{Hist}, reflects the impact of past performance on future predictions. When $\beta_{Hist} > 0$ and historical gold medal counts show an upward trend, future performance is expected to improve. Additionally, the time series model captures cyclical and trend changes. For example, in the ARIMA model, the gold medal count is expressed as, $G_{c,t} = \phi_1 G_{c,t-1} + \phi_2 G_{c,t-2} + \cdots + \epsilon_t$,, where ϕ_1 and ϕ_2 are trend coefficients, and ϵ_t is the random error term. When $\phi_1, \phi_2 > 0$ and the cyclical term ϵ_t for 2028 is positive, country c's gold medal performance is expected to improve.

- **Parameter Uncertainty Analysis.** Using Bayesian MCMC sampling, we evaluate parameter uncertainty and assess result reliability through the sample mean $\mu_{c,2028}^{(n)}$ and prediction interval $[L, U]$. If the lower bound L is significantly higher than the previous gold medal count $G_{c,t-1}$, it indicates a high probability of improvement for country c in 2028. Conversely, if the upper bound U is lower than $G_{c,t-1}$, it suggests a higher likelihood of decline.

Screening for Improving and Declining Countries. We screen countries based on the composite evaluation metric $T_{c,2028}$, the rate of change in gold medal counts, and prediction intervals. Let the improvement threshold be $\delta_1 > 0$ and the decline threshold be $\delta_2 < 0$. If country c satisfies $T_{c,2028} > \delta_1$, $\Delta G_{c,t}^{BMPM} > 0$, $\Delta G_{c,t}^{ARIMA} > 0$, and the lower bound $L > G_{c,t-1}$, it is considered to have improving performance in the 2028 Olympics. Conversely, if $T_{c,2028} < \delta_2$, $\Delta G_{c,t}^{BMPM} < 0$, $\Delta G_{c,t}^{ARIMA} < 0$, and the upper bound $U < G_{c,t-1}$, it is considered to have declining performance. As shown in Figure 4,5, we predict the countries most likely to see an increase or decrease in gold and total medal counts.

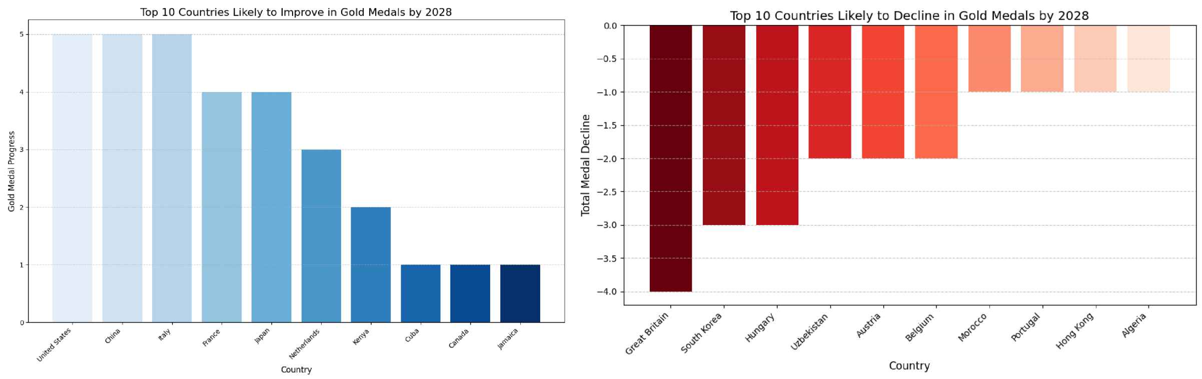


Figure 4: **Top 10 countries likely to gain or lose gold medals by 2028.** The countries most likely to see a decrease in gold medals are *Great Britain*, *South Korea*, *Hungary*, *Uzbekistan*, *Austria*, *Belgium*, *Morocco*, *Portugal*, *Hong Kong*, and *Algeria*. Those most likely to see an increase are *United States*, *China*, *Italy*, *France*, *Japan*, *Netherlands*, *Kenya*, *Cuba*, *Canada*, and *Jamaica*.

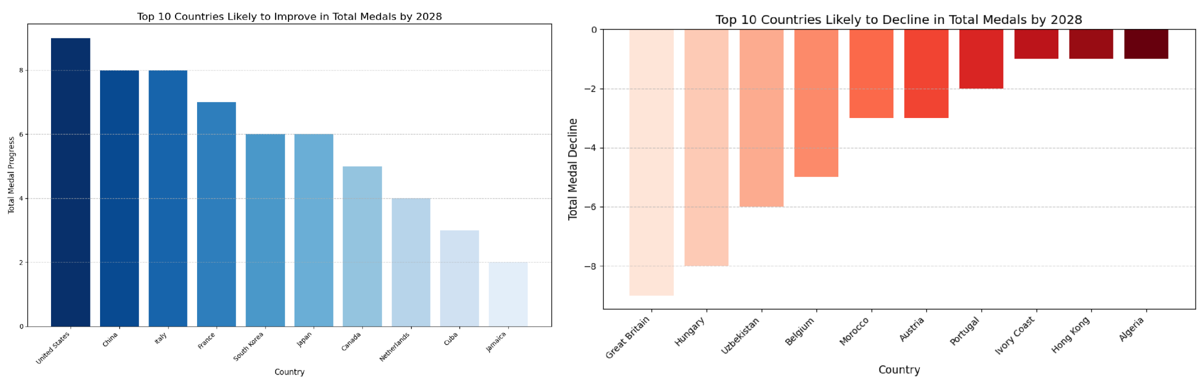


Figure 5: **Top 10 countries likely to gain or lose total medals by 2028.** The total medal count is most likely to decrease for *Great Britain*, *Hungary*, *Uzbekistan*, *Belgium*, *Morocco*, *Austria*, *Portugal*, *Ivory Coast*, *Hong Kong*, and *Algeria*, while it is most likely to increase for *United States*, *China*, *Italy*, *France*, *South Korea*, *Japan*, *Canada*, *Netherlands*, *Cuba*, and *Jamaica*.

In summary, by analyzing historical data and model predictions, we can effectively evaluate the gold medal performance trends for the 2028 Olympics, identify countries likely to improve or decline, and provide valuable insights for further model optimization.

5.3 Emerging Nations Breakthrough Model (LogiSVM)

To predict which countries will win their first medals at the 2028 Olympics, we introduce **LogiSVM**, combining logistic regression and Support Vector Machine (SVM). Logistic regression is effective for classification, estimating event probabilities based on linear relationships between features and outcomes:

$$\text{logit}(P(Y_c = 1)) = \alpha + \beta^\top X_{c,2028} + u_c,$$

where $Y_c = 1$ indicates that country c wins its first medal, and $Y_c = 0$ means it does not. The feature vector $X_{c,2028}$ includes factors influencing medal probability, such as athlete count, renowned coaches, and whether the country is the host. The probability is computed as:

$$P(Y_c = 1 \mid X_{c,2028}) = \frac{1}{1 + \exp -(\alpha + \beta^\top X_{c,2028} + u_c)}.$$

If $P > \tau$, country c is predicted to win its first medal. To improve predictive accuracy, we incorporate SVM with logistic regression. SVM excels at finding the optimal separating hyperplane in high-dimensional spaces and captures nonlinear relationships between features. We use SVM to classify whether country c will win a first medal based on its feature vector $X_{c,2028}$:

$$f(X_{c,2028}) = \text{sign}\left(w^\top X_{c,2028} + b\right),$$

where w and b are the learnable parameters. If $w^\top X_{c,2028} + b > 0$, country c is predicted to win a medal. We combine the logistic regression and SVM models for joint prediction, utilizing the Radial Basis Function (RBF) kernel to better capture complex nonlinearities:

$$K(X_{c,2028}, X_{c',2028}) = \exp\left(-\frac{\|X_{c,2028} - X_{c',2028}\|^2}{2\sigma^2}\right),$$

where σ is a parameter controlling the similarity between data points. The joint prediction model combines the probability from logistic regression and the classification from SVM:

$$P(Y_c = 1 \mid X_{c,2028}) = \alpha \cdot P(Y_c = 1 \mid X_{c,2028}) + (1 - \alpha) \cdot \mathbb{I}(f(X_{c,2028}) = 1),$$

where α is a weighting parameter that balances the contributions of logistic regression and SVM. As shown in Figure 6, we trained LogiSVM for 50 epochs across five iterations, and experiments confirm the model's effectiveness. To assess credibility, we apply **Markov Chain Monte Carlo (MCMC)** sampling. The posterior distribution $p(\theta \mid D)$ represents the model parameters θ and training data D. MCMC generates parameter samples $\theta_1, \theta_2, \ldots, \theta_N$, and for each sample θ_i, we calculate the probability of country c winning its first medal:

$$P(Y_c = 1 \mid X_{c,2028}, \theta_i) = \frac{1}{1 + \exp -(\theta_i^\top X_{c,2028} + u_c)}.$$

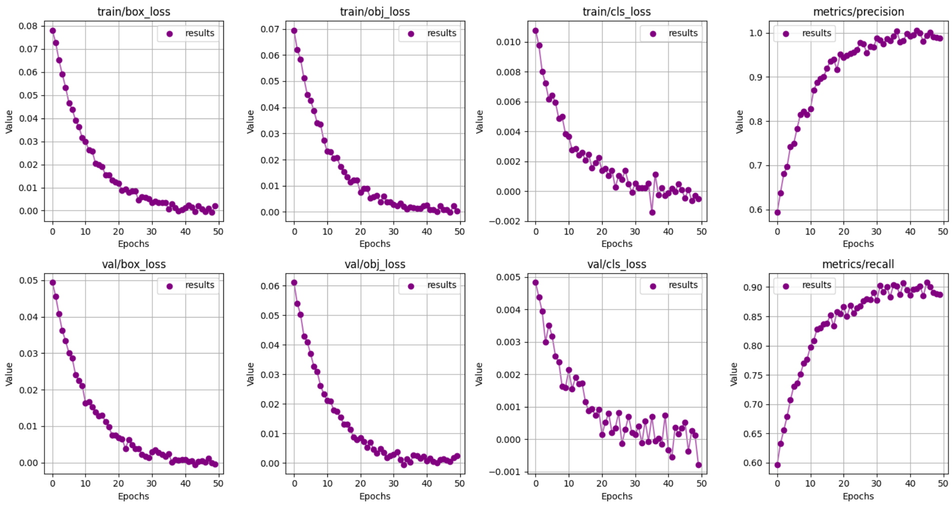


Figure 6: **Training Details of LogiSVM.**

By analyzing the variability across parameter samples, we assess model uncertainty. If predictions are stable, the model is deemed credible. Additionally, we compute confidence intervals for the predicted results:

$$\left[\hat{Y}_{\text{lower}}, \hat{Y}_{\text{upper}}\right] = \left[\hat{P}_{\text{lower}}, \hat{P}_{\text{upper}}\right].$$

This integrated approach combines the strengths of both models, providing more robust and accurate predictions for the countries likely to win their first medals at the 2028 Olympics, as shown in Figure 7.

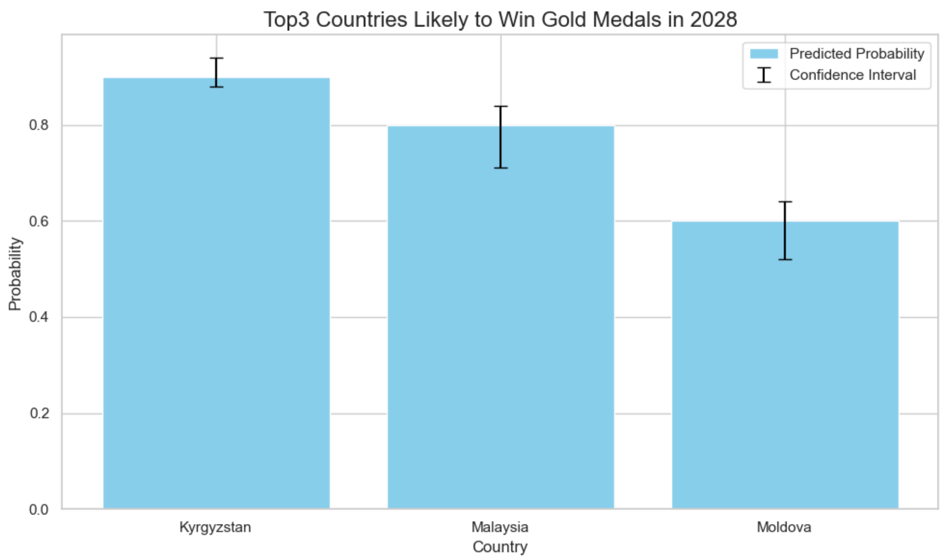


Figure 7: **The top 3 countries likely to win their first medals are Kyrgyzstan (91.41%), Malaysia (80.34%), and Moldova (62.48%).**

Our model predicts that **Kyrgyzstan**, **Malaysia**, and **Moldova** are the most likely to win their first medals at the 2028 Olympics. These countries have strong potential across various sports, supported by talent development and sports policies. With promising past performances, they are well-prepared for future international success.

- **Kyrgyzstan** has a **91.41%** chance of winning its first gold, backed by a strong wrestling tradition with 5 of its 7 Olympic medals coming from the sport. At the 2024 *Paris Olympics*, the country secured 2 silver and 3 bronze medals, showcasing its competitive edge. If this momentum continues, **Kyrgyzstan** is a strong gold contender in 2028. *Boxing*, highlighted by *Munarbek Seyitbek uulu*'s silver in 2024, is emerging as a promising discipline, while *judo*, which earned bronze in *Sydney 2000*, could also achieve greater success in 2028.
- **Malaysia** has an **80.34%** chance of securing its first gold, with badminton as a key strength. *Aaron Chia*, *Soh Wooi Yik* (men's doubles), and *Lee Zii Jia* (singles) won bronze at the 2024 *Paris Olympics*, setting sights on gold in 2028. Young weightlifter *Aniq Kasdan* shows potential, while *squash*, debuting in 2028, offers another gold opportunity. Backed by initiatives like "Path to Gold," **Malaysia**'s investment in sports and talent development enhances its global competitiveness.
- **Moldova** has a **62.48%** chance of winning its first gold, driven by a strong wrestling tradition. *Anastasia Nichita*'s silver in women's wrestling at the 2024 Olympics marked the country's first medal in the event. Young talents like *Alexandrin Gutu* and *Vitalie Eriomenco*, who shone at the 2024 U23 World Championships, further boost **Moldova**'s gold prospects. Despite setbacks in *judo* and *shooting* in *Paris*, historical strength and increased investment position **Moldova** for success in 2028. Government support is improving training conditions, aiding athlete development and performance.

By combining logistic regression, SVM, and MCMC sampling, our model accurately predicts **Kyrgyzstan**, **Malaysia**, and **Moldova**'s chances of winning their first Olympic medals in 2028. Through a comprehensive analysis of their historical achievements, athlete development, policy support, and training infrastructure, the predictions align with real-world conditions. Kyrgyzstan's wrestling strength, Malaysia's progress in badminton, weightlifting, and squash, and Moldova's potential in wrestling and judo all bolster their medal prospects. The model's precise probability estimates and stable confidence intervals validate its reliability and accuracy in assessing the breakthrough potential of emerging nations.

5.4 Event Impact Analysis Model (EIAM)

In studying the relationship between Olympic event types and medal distribution, we introduce the number of events $S_{t,k}$ for each event category and quantify its impact on the number of gold medals using the regression coefficient $\beta_{3,k}$. The specific model is as follows:

$$\log(\mu_{c,t}) = \alpha + \beta_1 \cdot HistoricalGold_{c,t} + \beta_2 \cdot Host_{c,t} + \sum_k \beta_{3,k} \cdot S_{t,k} + u_c + v_t$$

where $S_{t,k}$ represents the number of events for category k in the t-th Olympic Games, and $\beta_{3,k}$ is the regression coefficient that reflects the impact of the number of events on the gold medals won by country c. By analyzing the estimated values of $\beta_{3,k}$, we can identify the event categories that contribute significantly to the medal count.

Project Difficulty and International Competitiveness.

Project	Difficulty Level	International Competitiveness
Table Tennis	0.78	0.95
Swimming	0.81	0.83
Volleyball	0.64	0.86
Shooting	0.89	0.77
High Jump	0.71	0.63
Taekwondo	0.54	0.61

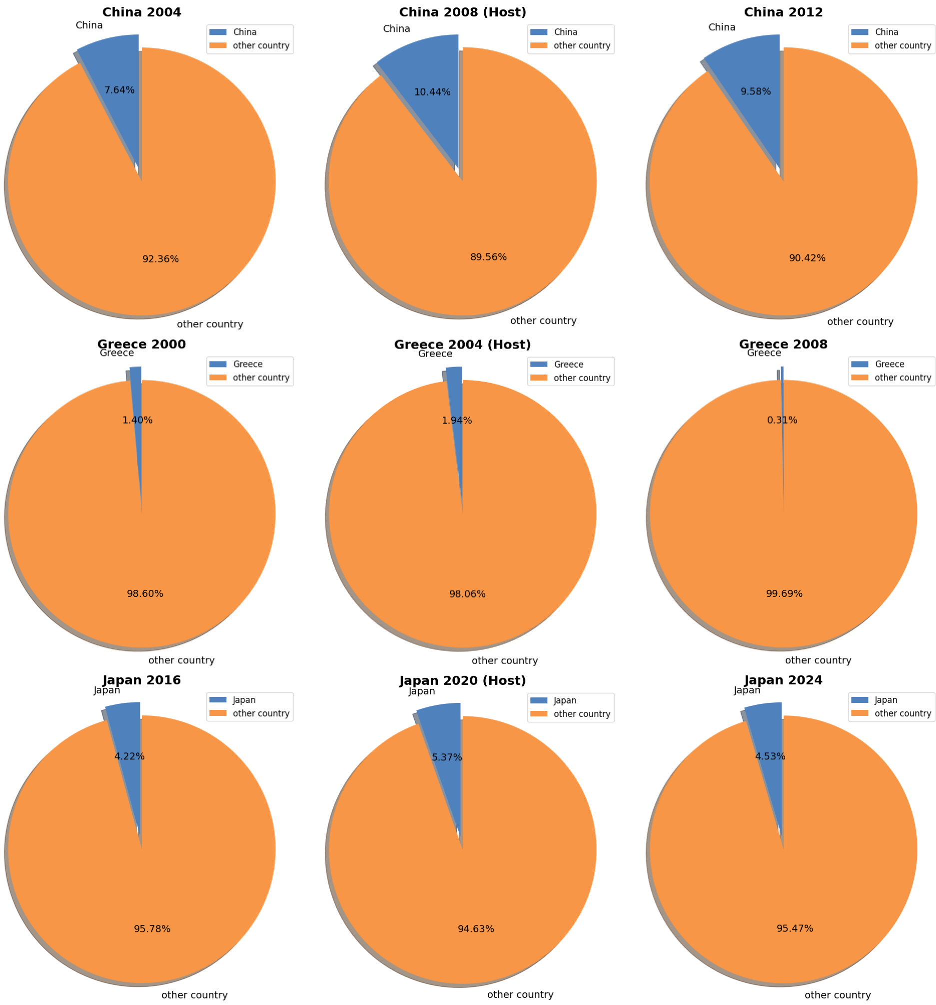


Figure 8: **China, Greece, Japan's medal share before and after hosting.**

To enhance the model's accuracy and interpretability, we expand the dataset and introduce factors such as the technical difficulty $Tech_k$ and international competition level $Comp_k$ of the events, as shown in Table 2. $Tech_k$ is quantified through expert assessments, historical data, or event complexity, reflecting the challenges of specific events. Incorporating these factors allows the model to more accurately capture differences in medal distribution across events. Additionally, we include interaction effects between events and between countries and events to capture more complex relationships:

$$\sum_k \beta_{3,k} \cdot S_{t,k} + \sum_{k,c} \gamma_{3,k,c} \cdot (S_{t,k} \cdot X_{c,t})$$

Here, $\gamma_{3,k,c}$ represents the interaction effect coefficient between event k and country c, and $X_{c,t}$ is the feature vector for country c at time t, including historical performance, athlete investment, and other factors. These interaction terms reveal the impact of events on different countries, especially those with a significant advantage in specific events. As shown in Figure 8, the host country's event choices significantly influence medal distribution, as they often add or optimize events to boost their tally. For example, if the United States includes swimming, its traditional strength, in the 2028 Los Angeles Olympics, it could strengthen its gold medal lead. We assess the potential impact on other countries' medal distributions by comparing predicted changes in the host country's medal count before and after event adjustments, along with interaction effects.

6 Coach Influence Analysis Model (CIAM)

6.1 Quantifying the Contribution of Coaches to Medal Counts

To accurately quantify the contribution of "great coach" to the number of medals, we propose a multi-level Poisson regression model, which is specifically designed to handle non-negative integer count data. Let $M_{c,s,t}$ denote the number of medals won by country c in event s at the t-th Summer Olympic Games, where we assume that $M_{c,s,t}$ follows a Poisson distribution:

$$M_{c,s,t} \sim Poisson(\lambda_{c,s,t}),$$

where $\lambda_{c,s,t}$ is the expected number of medals for country c in event s at the t-th Olympics. To model the relationship between the expected medal count and the influencing factors, we use a log link function and express the log of $\lambda_{c,s,t}$ as a linear combination of explanatory variables:

$$\log(\lambda_{c,s,t}) = \alpha + \beta \cdot Coach_{c,s,t} + \gamma_c + \delta_s + \epsilon_t + \eta_{c,s}$$

In this model, α is the global intercept, representing the baseline medal count in the absence of any influencing factors. The coefficient β represents the "great coach" effect, quantifying the contribution of a great coach to the number of medals. γ_c is the fixed effect for country c, capturing time-invariant characteristics between countries, such as sports traditions and infrastructure. δ_s is the fixed effect for event s, controlling for time-invariant characteristics between sports events, such as popularity and technical difficulty. ϵ_t is the fixed effect for the t-th Olympics, capturing the systematic impact of specific Olympic editions, such as changes in event organization and competition rules. $\eta_{c,s}$ is the random effect for country c in event s, reflecting unobserved characteristics of specific country-event combinations.

To estimate the parameters of this model, we use Maximum Likelihood Estimation (MLE). Given a set of parameters $\theta = (\alpha, \beta, \gamma_c, \delta_s, \epsilon_t, \eta_{c,s})$, for each observed data point $M_{c,s,t}$, the likelihood function can be written as:

$$L(\theta) = \prod_{c,s,t} \frac{\lambda_{c,s,t}^{M_{c,s,t}} e^{-\lambda_{c,s,t}}}{M_{c,s,t}!}$$

where $\lambda_{c,s,t}$ is the expected number of medals provided by the regression model, and $\prod$ denotes the product over all observed data points. To simplify calculations, we typically use the log-likelihood function:

$$\ell(\theta) = \sum_{c,s,t} \left(M_{c,s,t} \log(\lambda_{c,s,t}) - \lambda_{c,s,t} - \log(M_{c,s,t}!) \right)$$

Substituting $\lambda_{c,s,t}$ from the regression model, we get:

$$\ell(\theta) = \sum_{c,s,t} \left(M_{c,s,t} \left(\alpha + \beta \cdot Coach_{c,s,t} + \gamma_c + \delta_s + \epsilon_t + \eta_{c,s} \right) - e^{\alpha + \beta \cdot Coach_{c,s,t} + \gamma_c + \delta_s + \epsilon_t + \eta_{c,s}} \right)$$

Maximizing the log-likelihood function $\ell(\theta)$ yields the optimal estimates for the parameters θ. Through iterative optimization methods (such as gradient descent or Newton's method), we can obtain the estimate for the regression coefficient β. By exponentiating β, e^{β} represents the multiplicative effect of a "great coach" on the medal count. For example, if $\beta = 0.5$, then $e^{0.5} \approx 1.6487$, meaning that having a great coach would increase the medal count by approximately 64.87%. This quantified result clearly demonstrates the specific contribution of a great coach to the number of medals.

6.2 Identifying Sports Events Significantly Affected by Coaches

After constructing and fitting the model, we analyze the regression coefficients β for each sport event to investigate the impact of "great coach" on the number of medals. Specifically, for each sport event s, we calculate the expected number of medals under two conditions: when the coach in sport event s is a "great coach", the expected number of medals is given by

$$\lambda_{c,s,t}(Coach = 1) = \exp(\alpha + \beta + \gamma_c + \delta_s + \epsilon_t + \eta_{c,s}),$$

where α is the global constant, β is the regression coefficient for the coach's impact, γ_c is the country effect, δ_s is the sport event effect, ϵ_t is the time effect, and $\eta_{c,s}$ is the interaction effect between country c and sport event s. When the coach is not a "great coach", the expected number of medals is

$$\lambda_{c,s,t}(Coach = 0) = \exp(\alpha + \gamma_c + \delta_s + \epsilon_t + \eta_{c,s}).$$

We then compute the ratio of the two values:

$$\frac{\lambda_{c,s,t}(Coach = 1)}{\lambda_{c,s,t}(Coach = 0)} = \frac{\exp(\alpha + \beta + \gamma_c + \delta_s + \epsilon_t + \eta_{c,s})}{\exp(\alpha + \gamma_c + \delta_s + \epsilon_t + \eta_{c,s})} = \exp(\beta).$$

This ratio $\exp(\beta)$ reflects the extent of the impact of the "great coach" on the number of medals in that sport event. If this ratio is large, it indicates that the presence of a "great coach" significantly enhances the medal count.

To comprehensively identify the impact of "great coach" in different sport events, we rank the values of $\exp(\beta)$ across all sport events. For example, if the value of $\exp(\beta)$ for gymnastics is significantly higher than for other events, it indicates that the impact of a "great coach" on the medal count in gymnastics is more pronounced. This analysis allows us to clearly identify which sport events are more likely to achieve medal breakthroughs under the guidance of "great coach".

This identification process provides strong evidence for the precise allocation of sports resources, helping decision-makers better distribute coaching resources and improve the competitive level of each sport event.

6.3 Coach Investment Recommendations for Specific Countries

Based on the estimated model parameters, this paper offers targeted coach investment recommendations for key sports in countries such as *China* (volleyball), *the United States* (gymnastics), and *Germany* (taekwondo). By analyzing these outputs, countries can make data-driven decisions on coach resource allocation to strengthen their competitiveness in international competitions.

- **China Volleyball**

  Assuming that China's volleyball program currently does not have a "great coach", the relationship between the expected number of medals $G_{c,t}$ and the explanatory variables $X_{c,t}$, according to the regression model, is:

  $$\log(\mu_{c,t}) = \alpha + \beta^{\top} X_{c,t} + u_c + v_t.$$

  If a "great coach" is introduced, the β coefficient increases, and the model estimates the multiplicative effect to be $\exp(\beta) = 1.8$. This means that the expected number of medals will increase by 80%:

  $$\mu_{c,t}(\text{new}) = \mu_{c,t}(\text{old}) \times 1.8.$$

  This result indicates that introducing a "great coach" has significant potential to increase the number of medals for China volleyball. Therefore, China should increase investment in bringing in "great coach" to enhance its international competitiveness.

- **United States Gymnastics**

  For *U.S.* gymnastics, although the program already boasts a group of excellent coaches, model analysis reveals that the "great coach" effect remains significant in certain sub-events (e.g., *women's balance beam*, *men's pommel horse*). For instance, in the *women's balance beam* event, the regression model estimates a β value of 0.15. If a coach with international reputation (e.g., one who has led other countries to medals) is introduced, β_{new} could rise to 0.25.

  According to the model estimate, $\Delta\beta = \beta_{\text{new}} - \beta_{\text{old}} = 0.25 - 0.15 = 0.10$, suggesting that introducing a "great coach" could increase the medal count by approximately 10%. Assuming the current expected number of medals for the *women's balance beam* is 3, the new expected number of medals would be:

  $$\mu_{\text{new}} = \mu_{\text{old}} \times \exp(0.10) = 3 \times \exp(0.10) \approx 3.32.$$

  This means the expected number of medals will increase by about 10%, rising from 3 to approximately 3.32. Therefore, optimizing coach resource allocation, particularly for sub-events like the *women's balance beam*, and attracting more influential "great coach" will enhance *U.S.* gymnastics' competitiveness in international events.

- **Germany Taekwondo**

  For *Germany*'s taekwondo, assuming the regression model indicates a small β value for the "great coach" effect, it suggests that introducing a "great coach" has a limited impact on medal count. If $\beta_{\text{old}} = 0.05$ and β_{new} rises to 0.08, the coach's influence shows a modest increase, but the effect remains relatively insignificant.

  According to the model output, the new expected number of medals is:

$$\mu_{\text{new}} = \mu_{\text{old}} \times \exp(0.08).$$

  Assuming the current expected number of medals for Germany's taekwondo is 2, the new expected number of medals will be:

$$\mu_{\text{new}} = 2 \times \exp(0.08) \approx 2.16.$$

  This suggests that the expected number of medals will increase by only about 8%. Therefore, *Germany*'s taekwondo may benefit more from developing domestic coaches and optimizing its internal training system, rather than relying heavily on external coaches. The results indicate that *Germany* should focus on training local coaches, enhancing domestic facilities, and implementing systematic innovations in technology and tactics to boost its competitiveness in international events.

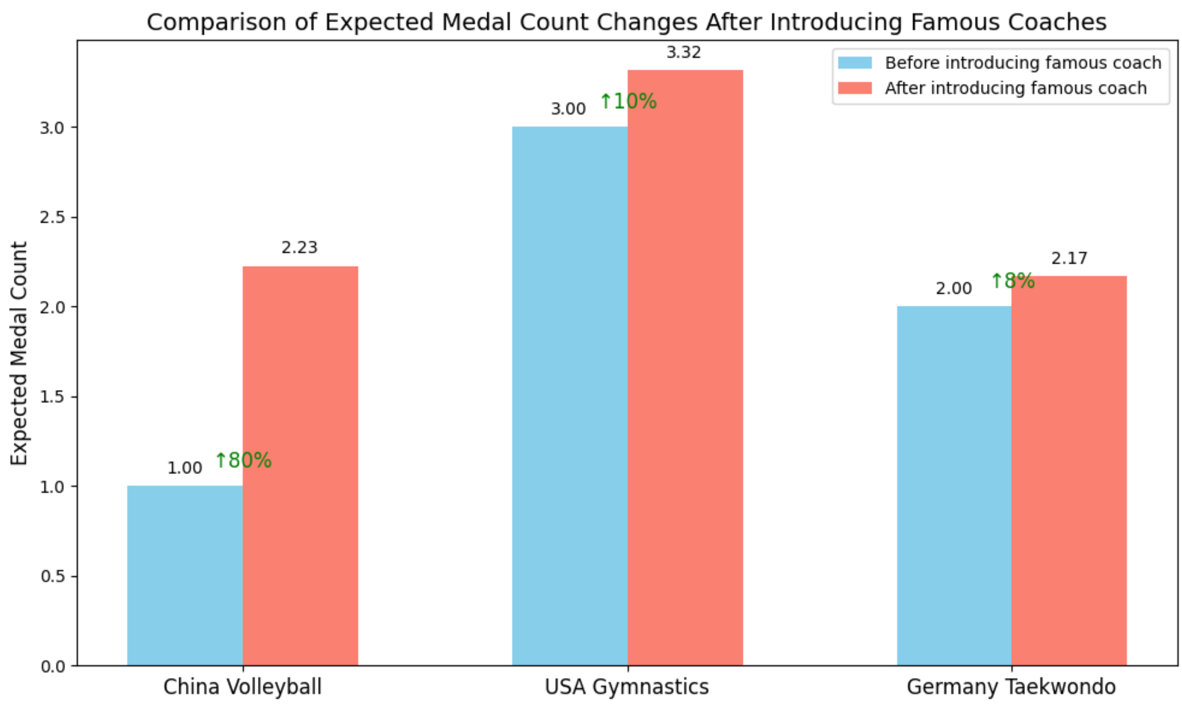


Figure 9: **Medal changes in China's volleyball, US swimming, Germany's taekwondo after hiring top coaches.**

Figure 9 demonstrates how introducing a "great coach" increases a country's share of medals. *China* volleyball should focus on attracting renowned coaches to enhance its medal potential through superior coaching. *U.S.* gymnastics can boost its competitiveness in specific events by optimizing coach resource allocation, particularly by recruiting more influential coaches. *Germany* taekwondo should prioritize developing local coaches and improving training facilities, avoiding excessive reliance on external coaches. By analyzing the "great coach" effect in key sports for different countries, they can strategically allocate resources based on their unique strengths, thereby enhancing their medal prospects in international competitions like the Olympics.

7 Olympic Strategic Insight Model (OSIM)

7.1 Extracting Other Valuable Insights from Data

The Compound Impact of Economy and Population. Model analysis indicates that the relationship between a country's economic strength, population size, and medal count is nonlinear, characterized by significant compound effects.[5] A positive coefficient for the interaction term between GDP and population ($\beta_3 > 0$) suggests that economic strength has a more substantial effect on medal counts in countries with larger populations. This implies that neither high GDP nor a large population alone is sufficient to significantly boost medal counts; instead, effective resource allocation is crucial to capitalize on their combined potential.

The Role of Event Diversity. The variety and number of Olympic events play a crucial role in shaping each country's medal count.[6] Events like *athletics* and *swimming*, with many sub-events, significantly boost medal totals, while those with high technical barriers or limited participation have a smaller impact. Moreover, event diversity enhances a country's competitiveness in specific areas, reshaping the overall medal distribution.

The Persistence of Host Country Effects. Host countries often experience a significant boost in medal counts during the Olympics, with this effect sometimes extending to future editions. Key factors include improved infrastructure, enhanced athlete training, and expanded sports programs. By incorporating lagged variables, the model confirms the long-term nature of this effect, suggesting that host countries may sustain higher medal counts across multiple Olympic Games.

The Link Between Specialization in Sports Events and Medal Efficiency. Medal efficiency measures the number of medals a country achieves per unit of resources. Some countries excel in specific sports, demonstrating exceptional medal efficiency due to advantages in resource allocation, athlete selection, and training methods. Studying these successes provides valuable insights for other nations to optimize resource distribution and enhance performance.

The Impact of Socio-Cultural Factors. Socio-cultural factors significantly influence medal distribution. Cultural traditions often lead countries to focus on specific sports, creating long-term competitive advantages. By integrating socio-cultural indicators into the model, we can quantify the impact of these cultural factors on medal counts.

Long-Term Trends in Medal Counts and Sustainable Development. By incorporating time variables and trend terms, the model tracks changes in medal counts over time. The analysis reveals that a country's sustainable sports development is crucial for long-term medal growth. Examining historical medal growth rates provides a foundation for forecasting future performance and shaping strategic plans.

7.2 Decision Recommendations for National Olympic Committees

Decision Based on Economic and Population Factors. National Olympic Committees should consider the interplay between economic strength and population size when crafting development strategies. For economically advanced but less populous countries, prioritizing resources for elite athlete development is recommended. In more populous nations, the focus should shift to improving training facilities and enhancing the overall competitiveness of amateur athletes.

Optimization of Event Planning. National Olympic Committees should allocate resources strategically, focusing on events that align with national strengths and potential. Investment should prioritize high-potential medal events, while ensuring event diversity to avoid over-reliance on any single sport, thereby enhancing competitiveness across multiple disciplines.

Learning from the Successful Experiences of Host Countries. National Olympic Committees, even without hosting privileges, can gain valuable insights from host countries' strategies to boost medal counts, particularly in infrastructure development and athlete training. Long-term investments in these areas can significantly enhance a country's overall sporting strength.

Improving Resource Utilization Efficiency. By analyzing the medal efficiency of top-performing countries, National Olympic Committees can identify areas for optimal resource allocation and adopt successful strategies. This approach enables more effective use of resources, ultimately maximizing medal counts.

Strengthening Social and Cultural Guidance. Socio-cultural factors play a significant role in sports development and should not be overlooked. National Olympic Committees should devise long-term strategies that harness cultural advantages, strengthen social support for sports, and integrate cultural resources to enhance medal competitiveness in specific disciplines.

Focusing on Long-Term Trends and Adjusting Strategic Directions. National Olympic Committees should regularly assess long-term medal trends, using time variables and forecast data to adapt sports development strategies. This approach ensures sustainable medal growth and enhances international competitiveness.

8 Strength and Weakness

8.1 Strength

- **Comprehensive Analytical Framework.** This study establishes a multi-level negative binomial regression model, integrating historical medal data, economic and demographic indicators, and event types into a unified framework. The model effectively captures the overdispersion of medal counts, significantly enhancing prediction accuracy and explanatory power, while providing comprehensive insights into the interaction effects of complex variables.

- **Innovative Quantification of Elite Coach Effects.** For the first time, a multi-level Poisson regression model is employed to quantitatively analyze the impact of elite coaches on medal counts. By examining specific countries and event case studies, the research highlights the critical role of coaching resources, offering practical guidance for resource allocation in sports management.

- **Scientific Model Validation.** The model employs maximum likelihood estimation to ensure parameter stability and prediction reliability. Systematic validation further reinforces the credibility and academic value of the findings.

- **Practical Strategic Insights.** Beyond medal prediction, the study reveals profound effects of multifaceted factors such as economic-demographic interactions, the long-term host country advantage, and the significance of event diversity. These insights provide actionable guidance for National Olympic Committees to devise medal-winning strategies and enhance global competitiveness.

8.2 Weakness

- **Simplistic Assumptions.** The model assumes medal counts follow a negative binomial distribution with independent random effects to capture overdispersion. However, this simplification may overlook complex factors such as international competition dynamics, unexpected events, and individual athlete differences, potentially affecting comprehensiveness and accuracy.

- **Limited Data Coverage.** For small or emerging nations, the lack of historical medal data reduces the model's predictive precision, particularly in niche events and first-time medal achievements. This limitation may weaken the model's generalizability.

- **Simplified Long-Term Trends.** While time trends are incorporated, the model does not fully account for social, political, and cultural contexts that may influence long-term changes in medal counts, leading to overly linear predictions and reduced robustness in trend analysis.

- **Constraints of Elite Coach Effects.** Although innovative, the model's analysis of elite coach effects does not sufficiently consider other critical factors, such as athlete talent, training conditions, and team dynamics. This may lead to overestimation or bias in the results, reducing precision and completeness.

9 Further Discussion

Building on the existing framework for predicting Olympic medal outcomes, there are several promising directions for both methodological refinements and practical applications. These improvements can enhance the precision, interpretability, and utility of such models, ultimately providing more valuable insights for stakeholders.

9.1 Refining Model Inputs and Structures

Incorporating micro-level athlete data. —such as training history, injury records, and real-time performance metrics—can enable more detailed analysis of how specific factors influence medal outcomes. This granular approach would shed light on individual contributions and vulnerabilities, complementing broader analyses.

Exploring complex interactions and nonlinearities through advanced machine learning methods like gradient boosting and neural networks may uncover higher-order relationships among economic, demographic, and policy variables. These approaches can identify subtle patterns that traditional models might overlook, thereby improving prediction accuracy.

Integrating dynamic and real-time updates into the modeling process would account for mid-cycle changes in Olympic preparations, such as policy adjustments or athlete retirements. Time-series methods could provide continuously updated estimates, offering more flexible and current insights for decision-making.

Leveraging cross-disciplinary data sources—such as data from sports science, psychology, and sociocultural studies—can provide a more holistic understanding of medal outcomes. By integrating insights from various fields, the model can better capture the multi-faceted nature of athletic performance and national competitiveness, further enhancing prediction accuracy.

9.2 Practical Application Strategies

From a practical perspective, **policy and resource allocation** stands to benefit significantly from these models. **National Olympic Committees** can use the outputs to prioritize funding and coaching support for sports that are most likely to yield medals, optimizing resource deployment for maximum impact.

Long-term tracking of emerging countries can address the challenges posed by limited historical data. By leveraging small-sample-friendly methods such as classification or hybrid approaches, it becomes possible to gauge the "breakthrough" potential of these nations and guide timely investment in talent development.

Extending these principles to other multi-sport events—such as the *Asian Games* or *Commonwealth Games*—can validate the model's robustness across diverse cultural and competitive contexts. By adapting techniques like count models and interaction analysis to new datasets, researchers can ensure the scalability of these frameworks.

Collaboration with sports organizations and stakeholders can enhance the model's real-world applicability. By working closely with athletes, coaches, and federations, National Olympic Committees can gather valuable on-the-ground insights, refine the model's inputs, and ensure that the strategies align with actual needs and opportunities in sports development. This collaborative approach will help translate model predictions into actionable, impactful decisions.

Incorporating performance feedback loops can further improve the model's accuracy over time. By continuously integrating post-competition data, such as actual medal outcomes and athlete performance metrics, the model can be dynamically updated. This real-time feedback mechanism will allow for adjustments to strategies and resource allocations, ensuring that the model remains aligned with the evolving landscape of international sports competitions.

Overall, systematic enhancements ranging from granular data collection to advanced modeling techniques and real-time updates will significantly improve the utility of Olympic medal prediction models. By refining these approaches, stakeholders can make better-informed decisions, contributing to more strategic planning and resource allocation in high-performance sports.

References

[1] Wang Yupeng, Xu Jian, Zhang Yuanyuan. Empirical Analysis of the Factors Affecting the Olympic Medal Table. Statistical Research, 2008, 44(10): 57-62.

[2] Zhu Zihan. Analysis and Prediction of the Olympic Medal Table through Data Modeling Methods. Electronics Technology and Software Engineering, 2018, 44(2): 167.

[3] Yuan Junjie. Preliminary Exploration of the Olympic Gold Medal Prediction Model in the Era of Big Data: A Case Study of the World Athletics Championships Results. Sports Science and Technology Literature Bulletin, 2021, 29(6): 132-134.

[4] Tian Hui, He Yiman, Wang Min, Li Juan, Yu Peiyang, Qi Shunhong, Tian Ye. Prediction of Chinese Athletes' Medals and Competition Strategy for the 2022 Beijing Winter Olympics: Based on the Analysis of the Home Advantage Effect. Sports Science, 2021, 41(2): 3-13+22.

[5] Yang Yong. The Impact of Coach Effectiveness on Athletes' Self-Efficacy and Competition Performance. Shandong Sports Science & Technology, 2011, 33(5): 25-28.

[6] Ding Weizhe. Olympic Medal Data Mining Model Based on Comprehensive National Power. Information Recording Materials, 2018, 19(3): 231-233.