Beyond the Game: Sustainability of Mega-Events

Summary

In recent years, the environmental footprint of sports mega-events has drawn increasing global concern. To enhance the sustainability of major sporting events, we established an indicator system, constructed a model for Super Bowl host city selection, identified and recommended the optimal hosts for 2029, and extended the model to the Olympics. This paper provides a comprehensive decision tool for sustainable event planning.

Firstly, to identify the Super Bowl's environmental impacts, we applied **Life Cycle Assessment (LCA)** to divide the event into **5** chronological stages, generating **13** environmental indicators. We then built a **6-dimensional** city characteristic indicator system and integrated both indicator sets for a unified assessment framework.

Secondly, to select the most environmentally suitable 2029 Super Bowl hosts, we built the **AHM-PCA** module to identify key sustainability determinants. Then, the **FCE** model evaluated **Environmental Impact Factor (EIF)** scores for 22 previous host cities; the **LightGBM** model predicted EIF values for 3 potential and 3 never-hosted cities. Finally, the **MIC** model quantified the relationships between EIF values and city characteristics.

Thirdly, using our AHM-PCA model, we identified February deplaned passengers (**0.171**), hotel energy use per guest-night (**0.115**), and event electrical load peak expected (**0.113**) as the 3 indicators contributing the most to environmental pressure. Later on, applying the **AHM-PCA-FCE** model to 22 legacy hosts, we identified Indianapolis (**0.353**) as having the lowest EIF score, followed by Pontiac (**0.362**). Furthermore, LightGBM prediction for non-legacy cities showed that Santa Clara (**0.393**) and Seattle (**0.412**) would also impose relatively low environmental impacts if hosting the Super Bowl. Finally, MIC analysis revealed that the EIFs are most related to the share of renewable energy (**0.48**).

Fourthly, to extend and reflect on the module, we adapted the system to the Olympic by adding 3 indicators and removing 3. Using the **AHM-PCA-FCE** model, we assessed the EIF values of all 22 candidate cities and conducted a 3-dimensional sensitivity analysis. Results show that Detroit has the lowest EIF (**0.353**), followed by Arlington (**0.362**), indicating that Detroit and Arlington are the most sustainable host cities of the Olympics. Meanwhile, the sensitivity analysis demonstrates that the local grid CO_2 intensity baseline exhibits the highest sensitivity.

Finally, we wrote a recommendation letter to the NFL, advising **Indianapolis**, **Seattle**, **and Santa Clara** to be the 2029 Super Bowl hosts to minimize environmental impacts.

In summary, we developed an LCA-based indicator framework, built a multi-model decision system, selected the optimal Super Bowl hosts, and extended the model to the Olympics. In the future, sports mega-events may become much greener and more sustainable.

Keywords: AHM-PCA-FCE, Sport Mega-Events, Sustainability, Environmental Impact Factor, Super Bowl Host City Selection

Contents

1 Introduction

1.1 Problem Background

Sport mega-events such as the Super Bowl generate vast economic and cultural influence, yet their environmental footprint has become increasingly alarming, as shown in **Fig.1**. Massive greenhouse gas emissions and resource consumption strain local ecosystems, while sustainability considerations often remain secondary to financial and logistical priorities. As climate change intensifies, the ecological cost of these events presents an urgent challenge to global sustainability **[1-2]**.

Addressing this issue through data-driven modeling offers crucial benefits. By integrating environmental indicators into event planning, organizers can reduce carbon emissions and promote urban resilience. A quantitative evaluation model enables sustainable host city selection, encouraging investment in green infrastructure and transforming mega-events into catalysts for ecological innovation rather than environmental degradation **[3]**.

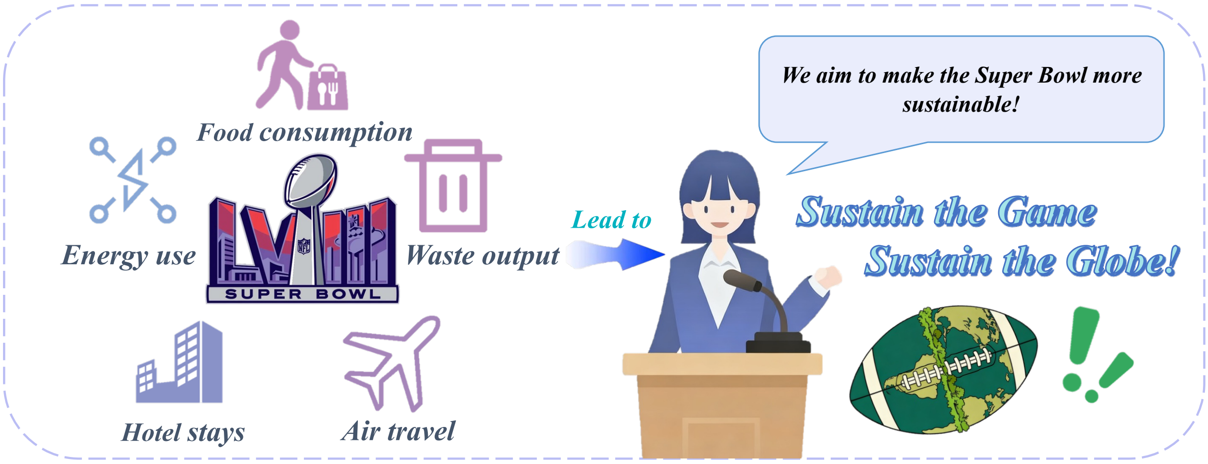


Figure 1 Environmental Footprint and Super Bowl

1.2 Restatement of the Problem

Considering the background information and restricted conditions identified in the problem statement, we need to solve the following problems:

- ✧ Quantify the key environmental impacts of hosting Super Bowl LIV in New Orleans, USA as a baseline for analysis.
- ✧ Develop a model to evaluate host cities by environmental sustainability, considering renewable energy, infrastructure, and climate factors.
- ✧ Apply the model to recommend the most environmentally sustainable city to host the 2029 Super Bowl, comparing both previous and new candidate locations across the United States.
- ✧ Extend the model to other major sporting events and propose strategies to reduce environmental impact.
- ✧ Compose a recommendation letter to the NFL, summarizing findings and justifying the proposed host city based on sustainability outcomes.

1.3 Our Work

We evaluate sport mega-events' environmental impacts, select the 2029 Super Bowl host, and generalize the framework to other events.

- Decompose the Super Bowl into 5 LCA phases, 13 indicators, and 6 urban drivers.
- Use AHM–PCA weights, FCE scoring, and MIC to link impacts with city features.
- Compute EIFs for 22 past hosts and predict EIFs for 6 candidates with LightGBM to identify the 2029 choice.
- Extend the framework to the Olympics to test transferability and robustness.
- Draft a recommendation letter to the NFL on the most sustainable 2029 host.

The full pipeline is summarized in **Fig.2**.

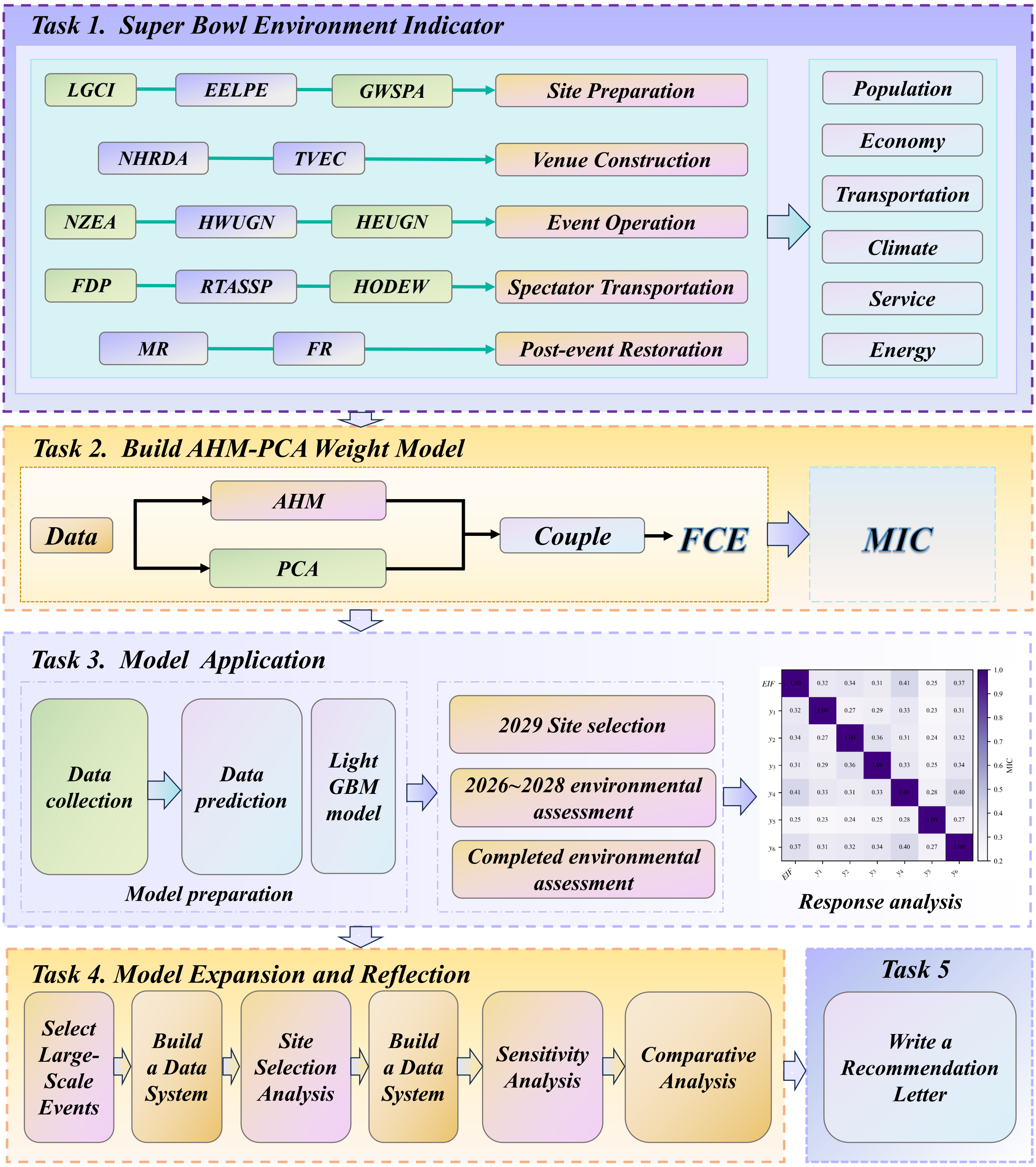


Figure 2 Model Overview

2 Assumptions

In order to solve the problems and more easily consider real-life situations, we make the following reasonable basic assumptions, each of which is properly justified.

- ✧ **Short-term environmental influence.** The Super Bowl's environmental effects are assumed to occur mainly within a one-month period around the event, without lasting or permanent ecological impact on the host city.
- ✧ **Localized scope.** Environmental influences are confined to the host city and do not extend to the surrounding state or the entire nation.
- ✧ **Data accuracy and reliability.** All data used in this study are assumed to be authentic, accurate, and representative of real environmental conditions.
- ✧ **Stable conditions.** No extreme events such as natural disasters or major disruptions are presumed to occur during the event period.

3 Notations

The key mathematical notations used in this paper are listed in **Table 1**.

Table 1: Notations used in this paper

Symbol	Description
a_{ij}	how much more important criterion i is than j on a 1 to 9 scale
$\boldsymbol{w}$	principle eigenvector corresponding to the maximum eigenvalue λ_{max}
x_{ik}	raw indicator values
z_{ik}	transformed variables
r_{ij}	the similarity in variation between indicators i and j
λ_j	eigenvectors
p_j	variance contribution
P	cumulative contribution
r_{ij}	the membership degree of factor u_i to evaluation grade v_j
B	comprehensive evaluation vector
b_j	the overall degree to which the system belongs to grade v_j

Other notations instructions will be given in the text.

4 Super Bowl's Environmental Impact

To evaluate the environmental impacts of hosting the Super Bowl, we first adopt a **Life Cycle Assessment (LCA)** approach to decompose the event into **5 phases**. For each phase, we construct a set of **13 environmental indicators** that capture key pressures such as energy demand and waste recovery. Next, we identify **6 dimensions** of **urban characteristic indicators** to represent the socioeconomic and ecological conditions of host cities. By linking phase-specific environmental indicators with city-level characteristics, we establish a framework that reveals how urban attributes shape the Super Bowl's environmental footprint.

4.1 Construction of Super Bowl's Environmental Indicators

The Super Bowl generates extensive and complex environmental impacts. To systematically assess its ecological footprint, we apply the **Life Cycle Assessment (LCA)** to divide the event into **5 phases**: site preparation, venue construction, event operation, spectator transportation, and post-event restoration. As shown in **Fig.3**, each phase generates distinct environmental pressures. On this basis, we further identified **13 key impact indicators**.

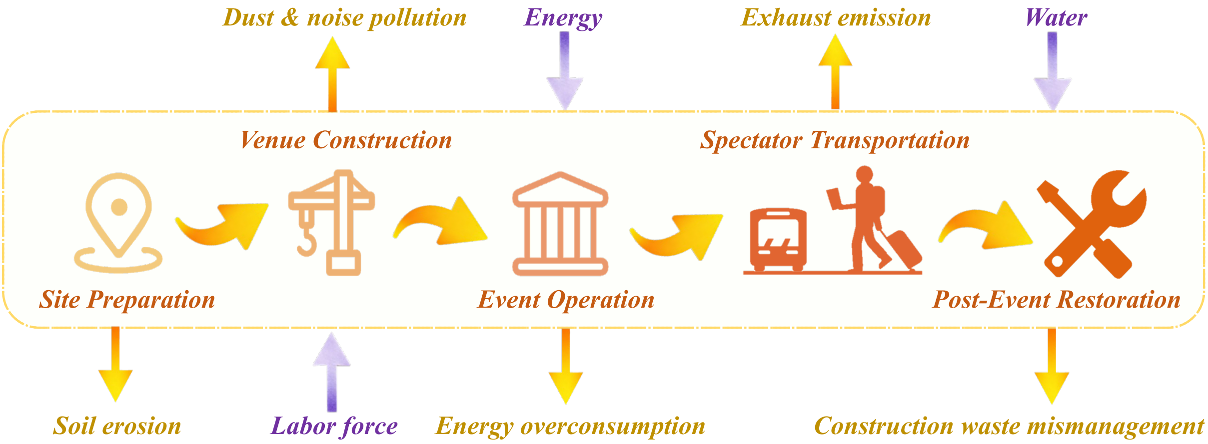


Figure 3 Five-phase LCA Structure of Super Bowl

4.1.1 Phase 1: Site Preparation

✧ Local grid CO_2 intensity baseline (LGCI, *kg CO_2/kWh*) x_1

Carbon factor of local electricity supply. Lower x_1 cuts upstream emissions for temporary works, site services, and early logistics.

✧ Event electrical load peak expected (EELPE, *MW*) x_2

Projected maximum power demand during the event. Lower peaks reduce peaking plants, curtail curtailment risk, and stabilize the grid.

✧ Game-week supply plan attribute (GWSPA, -) x_3

Strength and sustainability of the event supply plan. Higher quality plans minimize waste, secure critical materials, and buffer disruptions.

4.1.2 Phase 2: Venue Construction

✧ Number of hotel rooms in downtown area (NHRDA, *rooms*) x_4

Proxy for accommodation capacity supporting construction labor and specialist crews, shortening commutes and smoothing schedules.

✧ Total venue electricity consumption (TVEC, *MWh*) x_5

Aggregate electricity for construction and early commissioning. Efficient control of x_5 lowers embodied energy and pre-event impacts.

4.1.3 Phase 3: Event Operation

✧ Net zero energy achieved (NZEA, *%*) x_6

Share of demand met by on-site renewables and storage. Higher x_6 reduces external grid dependence and operational emissions.

✧ Hotel water use per guest-night (HWUGN, $m^3/guest\text{-}night$) x_7

Water intensity of lodging services. Lower x_7 reflects conservation, leak control, and efficient fixtures during peak occupancy.

✧ Hotel energy use per guest-night (HEUGN, $kWh/guest\text{-}night$) x_8

Energy per guest for HVAC and lighting. Lower x_8 signals efficient envelopes, controls, and load management at hotels.

4.1.4 Phase 4: Spectator Transportation

✧ February deplaned passengers (FDP, *passengers*) x_9

Scale of air arrivals in the event window. Higher x_9 raises aviation-related emissions and ground access pressures.

✧ RTA special service plan (RTASSP, -) x_{10}

Coverage and quality of dedicated public transport. Higher x_{10} shifts mode share, easing congestion and reducing per-capita emissions.

✧ Hotel occupancy during event week (HODEW, *%*) x_{11}

Concentration of visitors in local hotels. Higher x_{11} amplifies demand for utilities and services, stressing local systems.

4.1.5 Phase 5: Post Event Restoration

✧ Materials recovered (MR, *t*) x_{12}

Mass of construction and operational materials diverted from landfill. Higher x_{12} improves circularity and reduces residual impacts.

✧ Food rescued (FR, *kg*) x_{13}

Surplus food redirected to communities. Higher x_{13} lowers waste burdens while delivering measurable social co-benefits.

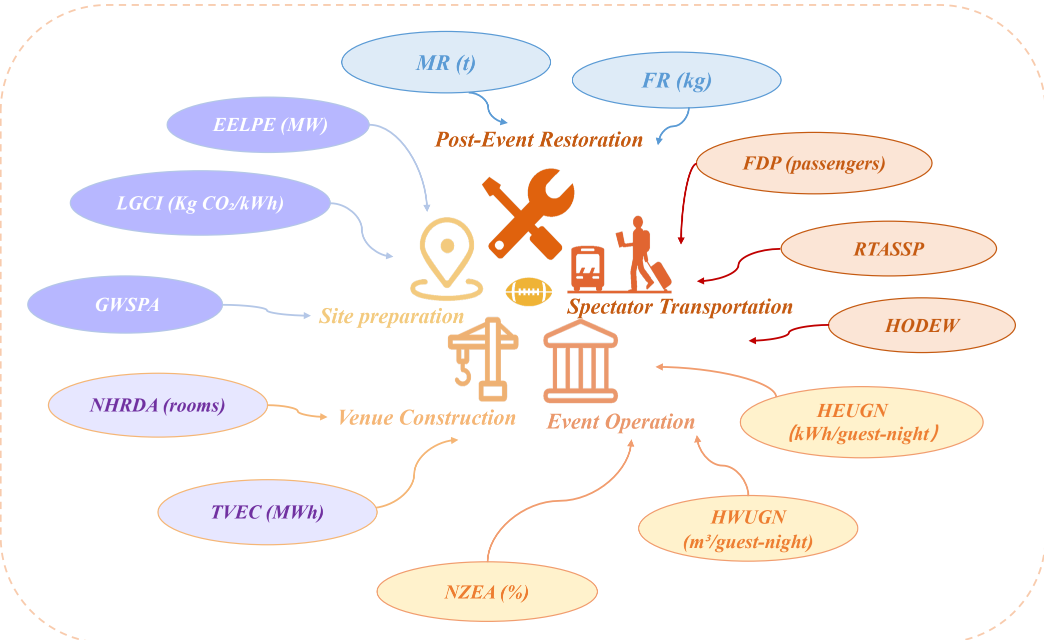


Figure 4 LCA Framework with 13 Indicators

In summary, **Fig.4** illustrates the LCA framework with 13 environmental impact indicators across 5 phases, from site preparation to post event restoration. This system effectively evaluates the Super Bowl's overall ecological footprint.

4.2 Construction of Urban Characteristic Indicators

To comprehensively evaluate the environmental sustainability of different host cities, we identify **6 dimensions** of **urban characteristic indicators**: economy, population, transportation, energy, climate, and service, as shown in **Fig.5**. These dimensions are related to a city's capacity for sustainable development and its potential to host large-scale events with minimal ecological impact.

Figure 5 Six-dimensions of Urban Characteristics

✧ **Gross domestic product per capita** (GDPPC, *USD/person*) y_1

Captures the city's economic development level. Higher values imply stronger economic vitality, greater productivity, and greater capacity to invest in infrastructure.

✧ **Total population** (TP, *persons*) y_2

Measures the demographic scale. Larger populations increase demand for resources, infrastructure, and public services, raising environmental pressure and management complexity.

✧ **Total road length** (TRL, *km*) y_3

y_3 reflects the level of urban transportation infrastructure development. A higher y_3 indicates a more complete and accessible road network, which facilitates mobility, logistics efficiency, and economic activity.

✧ **Share of renewable energy** (SRE, *%*) y_4

Represents the proportion of final energy from renewable sources. Higher shares signal deeper clean-energy adoption, reduced fossil dependence, and lower environmental impact.

✧ **Annual average temperature** (AAT, *°C*) y_5

y_5 reflects the climatic conditions of a city. A higher or lower y_5 can influence energy demand, agricultural productivity, and human comfort.

✧ **Hotel reception capacity** (HRC, *rooms*) y_6

Reflects service and tourism carrying capacity. Higher capacity supports large events and visitor flows while expanding tourism-related economic benefits.

4.3 Analysis of Super Bowl's Carbon Footprint

We conducted a carbon footprint analysis of the Super Bowl, using an **LCA framework**.

PHASE 1: Site Preparation

This phase depends on **population, energy, and climate**. Larger populations raise baseline demand, while cleaner grids and moderate temperatures lower initial carbon intensity.

$$\text{Phase 1: } (x_1, x_2, x_3) = f(y_2, y_4, y_5) \tag{1}$$

PHASE 2: Venue Construction

This stage depends on the **economy, energy, and service**. Cities with stronger economies and greener power systems can complete construction more efficiently and sustainably.

$$\text{Phase 2: } (x_4, x_5) = f(y_1, y_4, y_6) \tag{2}$$

PHASE 3: Event Operation

This phase is shaped by **energy, climate, and service**. High renewable penetration and mild weather enhance energy efficiency, while sufficient service capacity ensures stable and low-carbon operations.

$$\text{Phase 3: } (x_6, x_7, x_8) = f(y_4, y_5, y_6) \tag{3}$$

PHASE 4: Spectator Transportation

This phase is influenced by **transportation and service**. Well-developed road systems and hotel facilities help distribute visitor flows, reducing private travel emissions.

$$\text{Phase 4: } (x_9, x_{10}, x_{11}) = f(y_3, y_6) \tag{4}$$

PHASE 5: Post-Event Restoration

This stage is affected by the **economy, population, and energy**. Larger populations generate more post-event materials, while stronger economies and cleaner energy systems enhance recycling and recovery.

$$\text{Phase 5: } (x_{12}, x_{13}) = f(y_1, y_2, y_4) \tag{5}$$

In short, the relationship between urban and environmental indicators is shown in **Fig.6**.

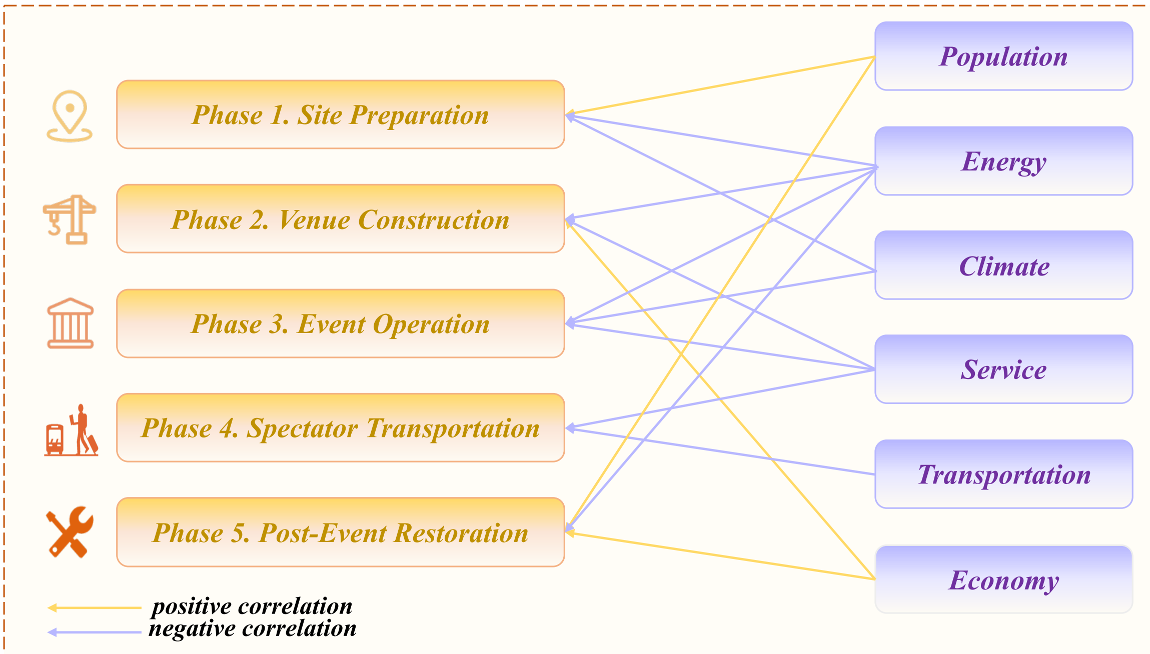


Figure 6 Urban–Environmental Indicator Relationship Structure

5 Super Bowl Host City Selection Model

We apply a **coupled weighting approach** that integrates **subjective weights (AHM)** with **objective weights (PCA)** to identify the core drivers of the Super Bowl's environmental impact. In addition, we use the **FCE** model to evaluate these environmental factors and determine the most suitable host city. Finally, the **MIC** model is employed to quantify the relationship between urban characteristics and environmental indicators.

5.1 Environmental Impact Driver Identification Model

5.1.1 Data Pre-processing

We compiled environmental and urban variables from authoritative sources. Units were standardized to common references such as kilograms of CO_2 per kilowatt-hour and cubic meters per guest-night, and values were stored with four-decimal precision.

City names were harmonized to metropolitan identifiers; for example, Los Angeles County was mapped to Inglewood. Outliers beyond three standard deviations were winsorized to protect principal-component estimates. All indicators were min–max scaled to the interval [0, 1], and records were alphabetized by city to ensure stable indexing for AHM–PCA weighting and FCE evaluation.

5.1.2 Construction of the AHM Model

The **Analytic Hierarchy Model (AHM)**, proposed by Saaty, structures problems into goal, criteria, and indicator levels. Here, AHM subjectively evaluates and weights key factors driving the Super Bowl's environmental impacts **[4]**.

STEP 1: Construct the Pairwise Comparison Matrix

$$A = [\, a_{ij} \,]_{n \times n},\ a_{ij} > 0,\ a_{ij} = \frac{1}{a_{ji}},\ a_{ii} = 1 \tag{6}$$

Decision-makers perform pairwise comparisons for all n criteria to express subjective judgments. Each a_{ij} quantifies how much more important criterion i is than j on a 1 to 9 scale. The reciprocal property ensures internal consistency and forms the foundation for calculating weights.

STEP 2: Derive the Priority Vector

$$A_{w} = \lambda_{\max} w,\ W = (w_1, \ldots, w_n)^{\top},\ \sum_{i=1}^{n} w_i = 1 \tag{7}$$

Compute the principal eigenvector w corresponding to the maximum eigenvalue $\lambda_{\max}$. Normalize the vector so that the sum of all components equals 1 to obtain the subjective weights. These weights reflect the relative importance of each criterion according to expert judgment.

STEP 3: Consistency Check

$$\mathrm{CI} = \frac{\lambda_{\max} - n}{n-1},\ \mathrm{CR} = \frac{\mathrm{CI}}{\mathrm{RI}} \tag{8}$$

The Consistency Index and Consistency Ratio assess the logical consistency of the pairwise comparisons. RI is the Random Index, depending on the matrix size n. A CR less than 0.10 indicates acceptable consistency; otherwise, the comparisons should be revised.

STEP 4: Aggregate Expert Judgments

$$\bar{A}_{ij} = \exp\left(\frac{1}{p}\sum_{k=1}^{p} \ln a_{ij}^{(k)}\right) \tag{9}$$

When multiple experts provide matrices $A^{(k)}$, their judgments are combined using the geometric mean. This ensures reciprocity is maintained and reduces the influence of outlier opinions. The aggregated matrix $\bar{A}$ is then used to recompute the priority vector $\bar{w}$.

STEP 5: Model Output

$$\boldsymbol{w} = (w_1, w_2, \ldots, w_n) \tag{10}$$

The weight vector $\boldsymbol{w}$ represents the normalized subjective importance of each criterion. A larger w_i indicates that criterion i contributes more to the overall evaluation.

5.1.3 Construction of the PCA Model

Principal Component Analysis (PCA) reduces data dimensionality by transforming correlated variables into uncorrelated principal components that retain maximum variance. Objective and data-driven, PCA removes multicollinearity and simplifies complex systems **[5]**.

STEP 1: Standardize the Raw Data

$$z_{ik} = \frac{x_{ik} - \bar{x}_i}{s_i},\ i = 1, 2, \ldots, n;\ k = 1, 2, \ldots, m \tag{11}$$

All raw indicator values x_{ik} are standardized to remove the effects of scale and units. The transformed variables z_{ik} have zero mean and unit variance.

STEP 2: Construct the Correlation Matrix

$$R = [\, r_{ij} \,]_{n\times n},\ r_{ij} = \frac{1}{m}\sum_{k=1}^{m} z_{ik}\, z_{jk} \tag{12}$$

The correlation matrix R summarizes the linear relationships among all standardized indicators. Each element r_{ij} measures the similarity in variation between indicators i and j.

STEP 3: Perform Eigenvalue Decomposition

$$Rv_j = \lambda_j v_j,\ j = 1, 2, \ldots, n \tag{13}$$

Eigenvalue decomposition yields eigenvalues λ_j and corresponding eigenvectors v_j.

STEP 4: Compute Variance Contribution and Select Principal Components

$$p_j = \frac{\lambda_j}{\sum_{i=1}^{n} \lambda_i},\ P = \sum_{j=1}^{k} p_j \tag{14}$$

The variance contribution p_j indicates the proportion of total variation explained by component j. The cumulative contribution P determines how many components k are retained. Typically, components are selected until the cumulative variance exceeds 85%, ensuring sufficient information coverage.

STEP 5: Derive Objective Weights from Component Loadings

$$w_i = \frac{\sum_{j=1}^{k} p_j |v_{ij}|}{\sum_{i=1}^{n}\sum_{j=1}^{k} p_j |v_{ij}|} \tag{15}$$

Objective weights are computed by combining loadings v_{ij} with variance contributions p_j.

STEP 6: Model Output

$$w = (w_1, w_2, \ldots, w_n) \tag{16}$$

The vector w represents the objective importance of each indicator determined purely by data characteristics. Higher weights correspond to indicators explaining more variation across the dataset.

5.1.4 Construction of the Coupled Weight Model

To calculate the coupling weight of AHM and PCA combined, the multiplier synthesis normalization method is adopted. This approach integrates subjective preferences with objective data variability, while preserving each method's contribution structure to the final comprehensive weight **[6]**.

$$W_{\text{AHP-PCA},k} = \frac{w_{AHM,k} \times w_{PCA,k}}{\sum_{m=1}^{n} w_{AHM,m} \times w_{PCA,m}} \tag{17}$$

5.2 Construction of the FCE Model

The Fuzzy Comprehensive Evaluation (FCE) model assesses systems with both qualitative and quantitative factors, handling ambiguity where performance levels lack clear boundaries. By mapping factor membership and synthesizing scores, FCE integrates diverse indicators **[7-9]**. Here, it evaluates and compares potential Super Bowl host cities' environmental readiness comprehensively and flexibly.

STEP 1: Establish the Factor Set and Evaluation Set

Let the factor set be

$$U = \{u_1, u_2, \ldots, u_n\} \tag{18}$$

and the evaluation set be

$$V = \{v_1, v_2, \ldots, v_m\} \tag{19}$$

STEP 2: Determine the Weight Vector

$$W = (w_1, w_2, \ldots, w_n),\ \sum_{i=1}^{n} W_i = 1 \tag{20}$$

The weight vector W expresses the relative importance of each indicator. Weights may come from AHM, PCA, or the coupled AHM–PCA model.

STEP 3: Construct the Fuzzy Membership Matrix

$$\mathrm{R} = \begin{bmatrix} r_{11} & \cdots & r_{1m} \\ \vdots & \ddots & \vdots \\ r_{n1} & \cdots & r_{nm} \end{bmatrix} \tag{21}$$

Each element $r_{ij} \in [0,1]$ denotes the membership degree of factor u_i to evaluation grade v_j. The membership matrix represents uncertainty and fuzziness in human evaluations..

STEP 4: Perform Fuzzy Synthesis

$$B = W \circ R = (b_1, b_2, \ldots, b_m) \tag{22}$$

The fuzzy synthesis operator aggregates weights and membership values to obtain a comprehensive evaluation vector B. Each b_j indicates the overall degree to which the system belongs to grade v_j.

STEP 5: Determine the Final Evaluation Outcome

$$\text{Result} = \arg\max(b_j) \tag{23}$$

The evaluation grade corresponding to the largest membership value becomes the final classification. This step translates fuzzy assessments into a clear qualitative decision.

5.3 Construction of the MIC Model

The Maximal Information Coefficient (MIC) detects both linear and nonlinear associations by maximizing mutual information across all data. It captures complex dependencies without assuming functional forms **[10-12]**. In this study, MIC measures how urban characteristics influence Super Bowl Environmental Impact Factors. The steps constructing MIC are as follows:

STEP 1: Data Preparation

Collect paired data (X_i, Y_i) for EIFs and urban characteristic indicators for all candidate cities. Normalize the variables if needed to ensure comparability.

STEP 2: Grid Partitioning and Mutual Information Calculation

For each variable pair, construct a set of grids with different numbers of bins. Compute the **mutual information** $I(X, Y)$ for each grid:

$$I(X, Y) = \sum_{x\in X}\sum_{y\in Y} p(x,y)\log\frac{p(x,y)}{p(x)p(y)} \tag{24}$$

where $p(x)$ and $p(y)$ are marginal probabilities, and $p(x, y)$ is the joint probability.

STEP 3: Maximal Information Coefficient Determination

Identify the grid that maximizes normalized mutual information, yielding the MIC score:

$$\mathrm{MIC}(X, Y) = \max\frac{I_G(X,Y)}{\log\min\{|X|,|Y|\}} \quad (G\in \text{ all grids}) \tag{25}$$

In summary, the roadmap for the Super Bowl site selection model is shown in **Fig.7**.

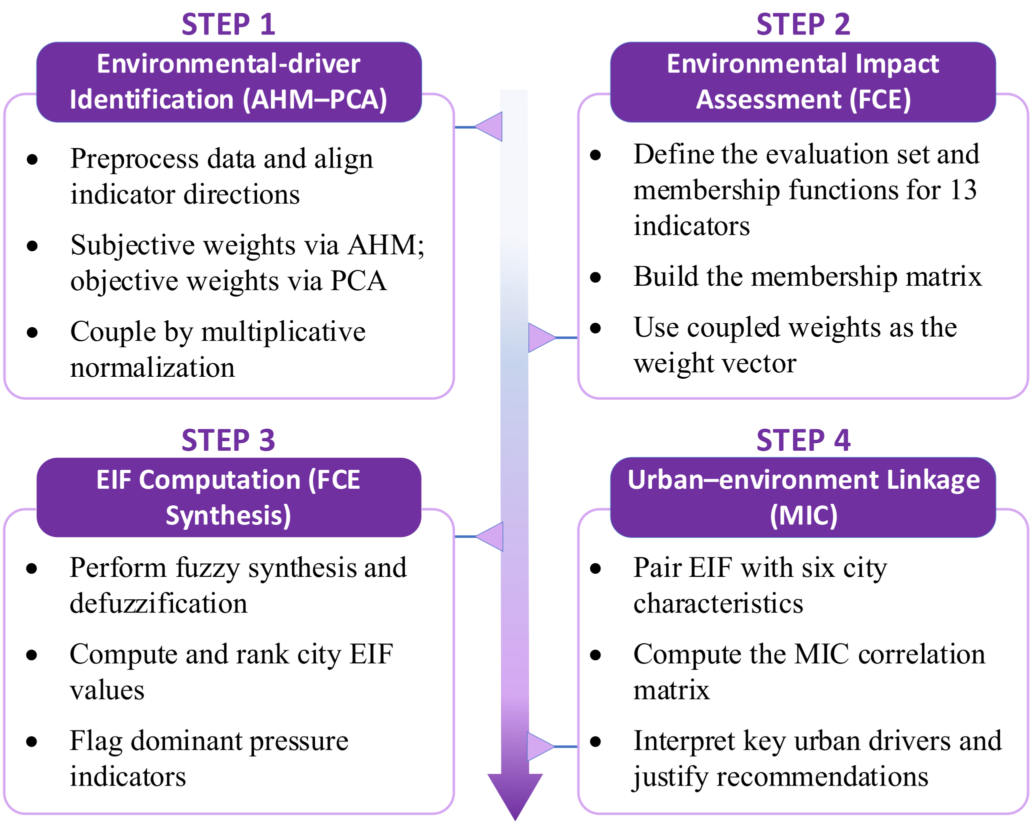


Figure 7 Super Bowl Site Selection Model

6 Super Bowl Host Selection Calculation

6.1 Data Collection

All data are drawn from state-level or international statistical sources. When city-level data are unavailable, we use state-level series adjusted by city population shares. Energy and emission factors come from the U.S. Energy Information Administration, EIA. Macroeconomic and demographic baselines are from the World Bank. Aviation traffic is from the U.S. Bureau of Transportation Statistics. Metropolitan indicators are from FRED. These steps ensure consistency in constructing the Energy Information Factor, EIF, and in subsequent analyses, as shown in **Table 2**.

Table 2 Data Link

Link	Collected data
https://www.eia.gov/environment/emissions/state/	State CO_2 emissions factors, electricity generation mix, grid carbon intensity (*kg CO_2/kWh*)
https://data.worldbank.org/	GDP per capita (current US$), total population, urbanization rate, CO_2 emissions per capita
https://www.bts.gov/	Airport passenger throughput, enplaned and deplaned passengers, aviation activity by airport
https://fred.stlouisfed.org/	Metropolitan GDP, employment in leisure and hospitality, regional price indices
https://www.noaa.gov/climate	Annual average temperature, precipitation, climate normals for U.S. cities and regions

6.2 Super Bowl Host City Selection Analysis

6.2.1 Environmental Impact Driver Analysis (AHM-PCA)

Based on the AHM-PCA model, the coupling is calculated as shown in **Fig.8**, and the weight comparison of AHM-PCA for the environmental indicators is shown in **Fig.9**.

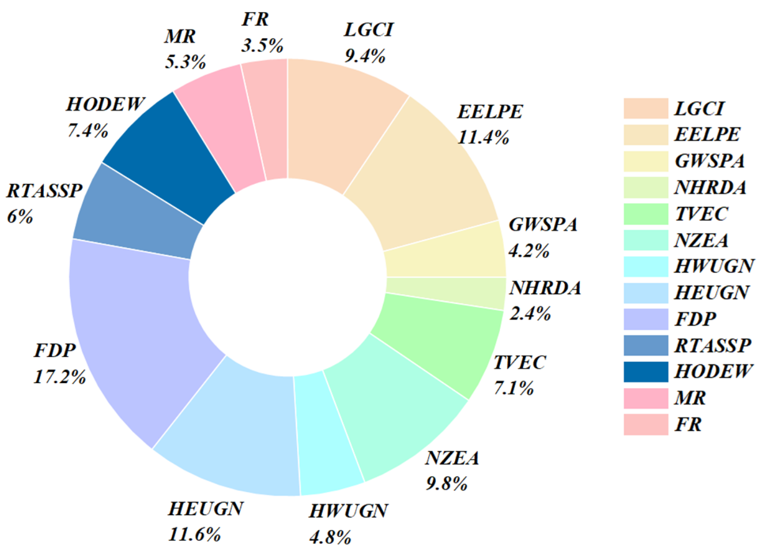
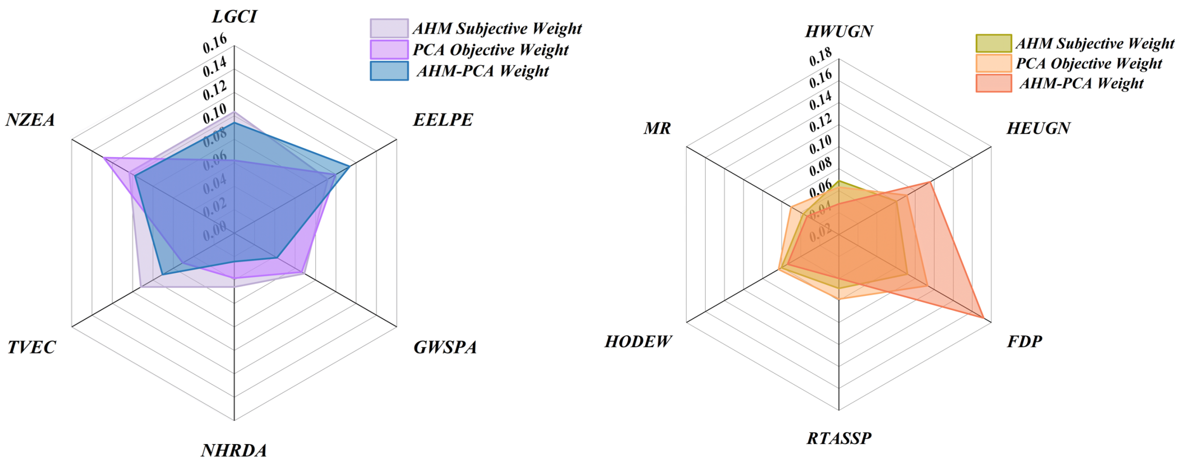


Figure 8 AHM-PCA Coupling Weight

As can be seen from **Fig.8**, x_9 (February deplaned passengers) has the highest coupling weight (0.1719), with x_8 (hotel energy use per guest-night) and x_2 (event electrical load peak expected) in the next spots with 0.1158 and 0.1136, respectively. These 3 indicators dominate the environmental impact assessment, reflecting the importance of traffic intensity and energy use. From what has been discussed, it is obvious that direct, quantifiable environmental factors, when compared to social or cultural variables, are more important in determining city-level sustainability.



(a) First Six Indicator Weights **(b) Following Six Indicator Weights**

Figure 9 Comparison of Weight Results

As seen in **Fig.9**, the AHM–PCA combined procedure yields smooth and consistent weight distributions compared to the use of a single method. The radar plots show high peaks at x_6, x_8, and x_9, reflecting their dominant influence on sustainability. For the moderate indicators, x_1, x_2, and x_{10}, the effects show balance but are secondary, while the low-weight indicators, namely x_4, x_7, and x_{13}, reflect a weaker contribution to local environmental outcomes.

6.2.2 Super Bowl Host Selection Analysis

Based on the AHM-PCA-FCE model, we calculated the EIFs for 22 host cities, which can be seen in **Fig.10**.

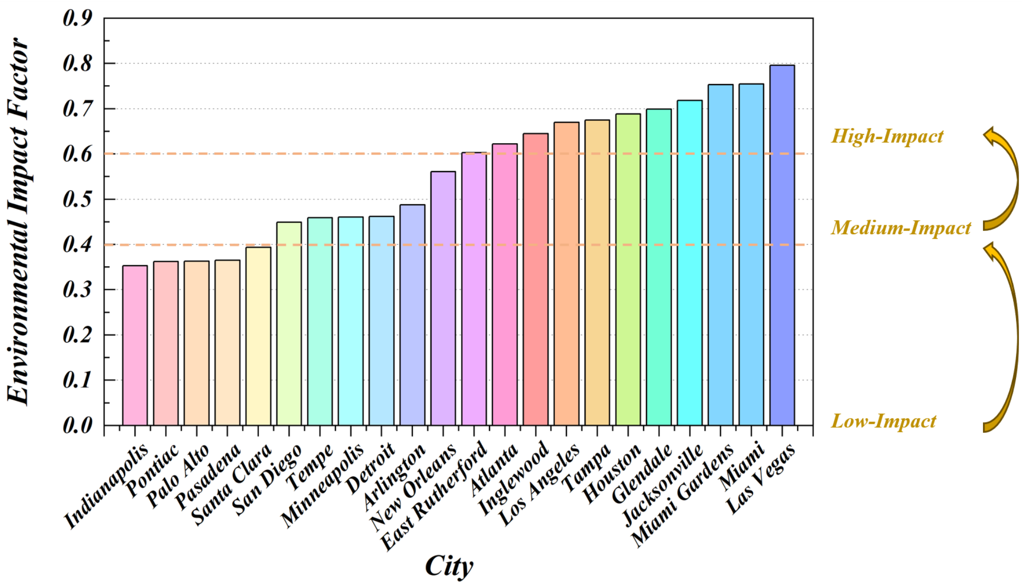


Figure 10 EIFs for 22 Host Cities

Calculation of the environmental impact factor for 22 host cities shows clear variations in sustainability performance. From **Fig.10**, the lowest EIFs were recorded for Indianapolis **(0.353)**, Pontiac **(0.363)**, and Palo Alto **(0.363)**, reflecting low ecological impact and high sustainability efficiency. Las Vegas has the highest impact value of 0.796, followed by Miami at 0.755 and Miami Gardens at 0.753.

Atlanta represents middle-ranking cities with a score of 0.622, while Inglewood scored 0.645, indicating a moderate environmental efficiency. Obviously, a north-south gap can be seen in the results, with northern and inland cities having a low environmental burden and southern coastal cities facing greater energy and emission pressures.

6.2.3 Host Cites Clustering

To further analyze the environmental impacts caused by potential host cities, we apply the **K-Means clustering** model to group cities based on their environmental impacts. Since K-Means is a widely used clustering method, we do not elaborate on its algorithmic details here for brevity. The 22 cities are classified into three categories: **environmentally sustainable**, **moderately sustainable**, and **environmentally unsustainable**, as illustrated in **Fig.11**.

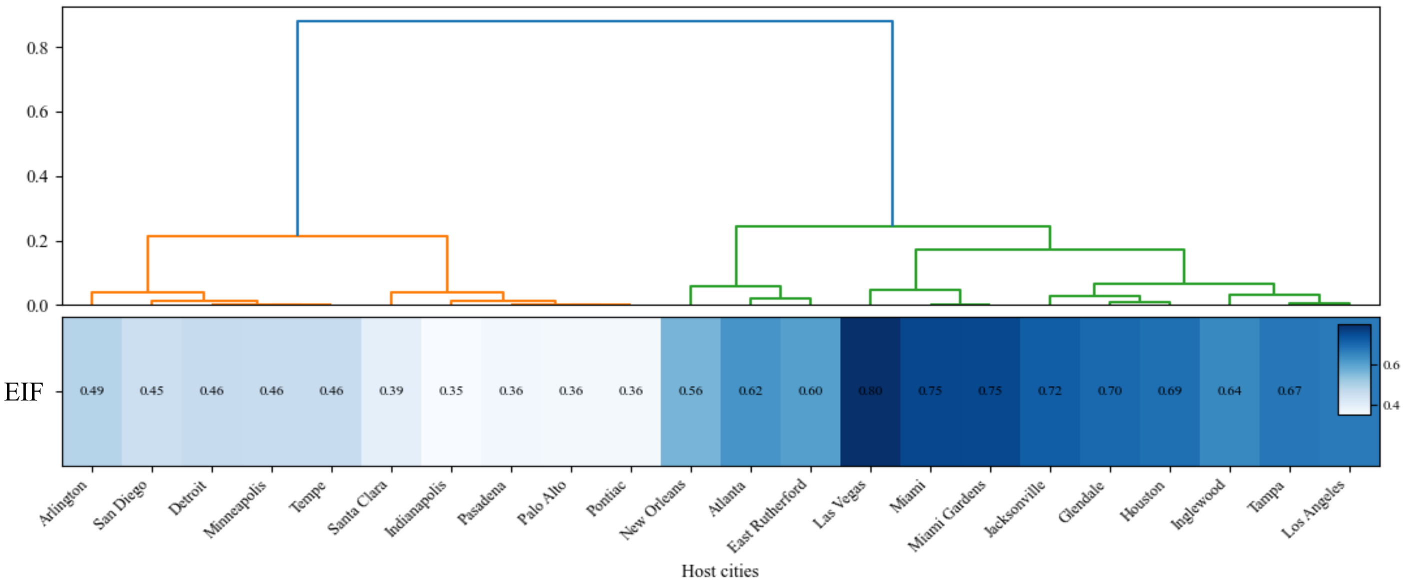


Figure 11 Clustering of Super Bowl Host Cities

- ✧ The first cluster shows cities characterized by high sustainability, such as **Indianapolis (0.35)** and **Pontiac (0.36)**. These cities show the smallest EIF scores principally because of their compact urban structure and efficient energy systems.
- ✧ The second cluster includes those with moderate sustainability, such as **Arlington (0.49)** and **San Diego (0.45)**. These cities manage energy and waste relatively well within an acceptable balance but also produce medium-to-high emissions concerning transportation.
- ✧ The third cluster is made up of low-sustainability cities, such as **Las Vegas (0.80)** and **Miami (0.75)**. These cities exhibit the highest EIF scores due to heavy energy demands, high dependence on air conditioning, and extensive travel-related carbon emissions.

6.3 Predicting the EIFs for Super Bowl Host Cities

6.3.1 Construction of the LightGBM Module

Light Gradient Boosting Machine (LightGBM) is a fast, tree-based ensemble learning algorithm optimized for large-scale and high-dimensional data. Its efficient leaf-wise growth and built-in regularization enhance prediction accuracy **[13-15]**. In our analysis, we apply LightGBM to forecast the environmental impact factors of 6 Super Bowl candidate cities.

STEP 1: Data Normalization

$$x'_i(k) = \frac{x_i(k) - \min x_i}{\max x_i - \min x_i},\ i = 1, 2, \ldots, n;\ k = 1, 2, \ldots, m \tag{26}$$

All raw indicator sequences $x_i(k)$ are normalized to eliminate scale and unit inconsistencies. This process ensures that all variables fall within [0,1].

STEP 2: Construct the Weighted Feature Matrix

$$Z = W \cdot X' \tag{27}$$

Here, X' is the normalized data matrix and W represents the vector of indicator weights derived from the objective–subjective coupling process.

STEP 3: Model Training Using LightGBM

The weighted matrix Z is divided into training and validation sets. LightGBM constructs gradient-boosted decision trees through leaf-wise growth with depth constraints, minimizing the objective function:

$$\mathcal{L} = \sum_{k=1}^{K} \ell(y_k, \hat{y}_k) + \Omega(T) \tag{28}$$

where $\ell(\cdot)$ is the loss function and $\Omega(T)$ is a regularization term controlling model complexity.

STEP 4: Hyperparameter Optimization

$$\theta = \arg\min(\mathcal{L}_{val}(\theta) \text{over all } \theta) \tag{29}$$

Key hyperparameters are optimized via cross-validation to enhance predictive stability and prevent overfitting.

STEP 5: Environmental Impact Prediction

$$\hat{y}_t = f\ \theta\ (Z_t),\ t = 2026, 2027, 2028 \tag{30}$$

The predicted values reflect interactions across all weighted indicators.

STEP 6: Model Evaluation

$$\text{RMSE} = \sqrt{\frac{1}{m}\sum_{k=1}^{m}(y_k - \hat{y}_k)^2} \tag{31}$$

6.3.2 Prediction Results and Analysis

Since the future host cities have not yet held a Super Bowl, we use the LightGBM model to predict the environmental impacts for the three scheduled host cities in 2026–2028. In addition, we select ***Seattle, Kansas City, and Nashville*** as candidate cities for 2029. Using the previously computed environmental indicators and urban characteristics, we train the model and then generate predictions of environmental impacts for these six cities. The training accuracy is shown in **Fig.12**, the 2026–2028 scheduled host cities are presented in **Fig.13**, and the three 2029 candidate cities are shown in **Fig.14**.

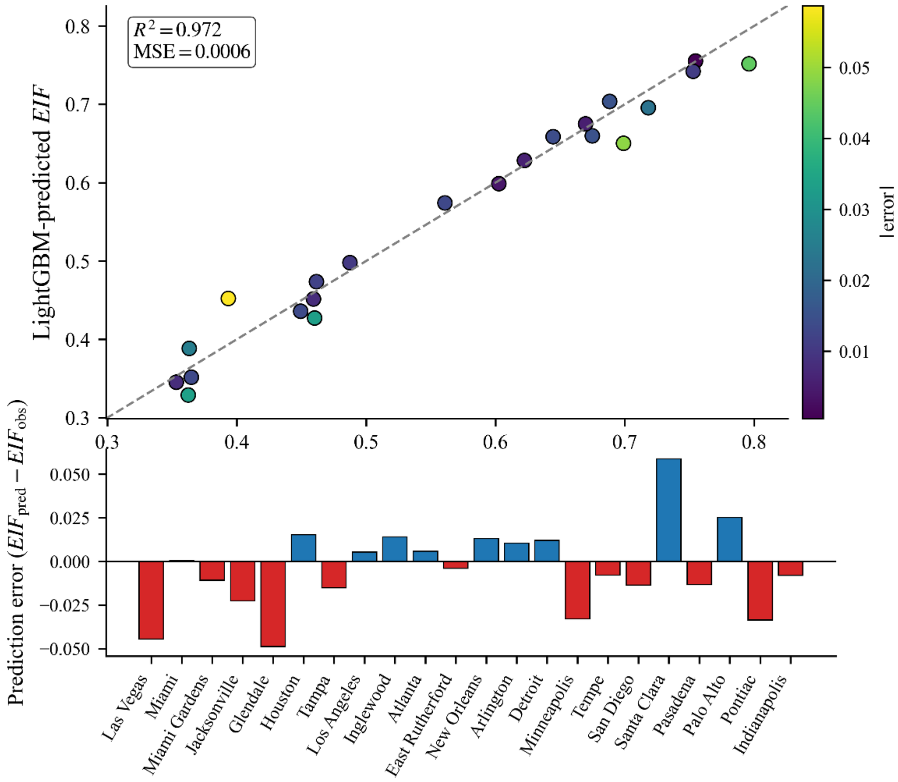


Figure 12 Training Accuracy

The upper panel compares observed EIF values derived from the AHM–PCA–FCE framework with LightGBM predictions for the 22 historical host cities. Points cluster around the 1:1 line and the reported R^2 and MSE indicate that the model reproduces the overall variation in environmental impact reasonably well. The lower panel shows city-level prediction errors, highlighting where LightGBM tends to overestimate or underestimate EIF.

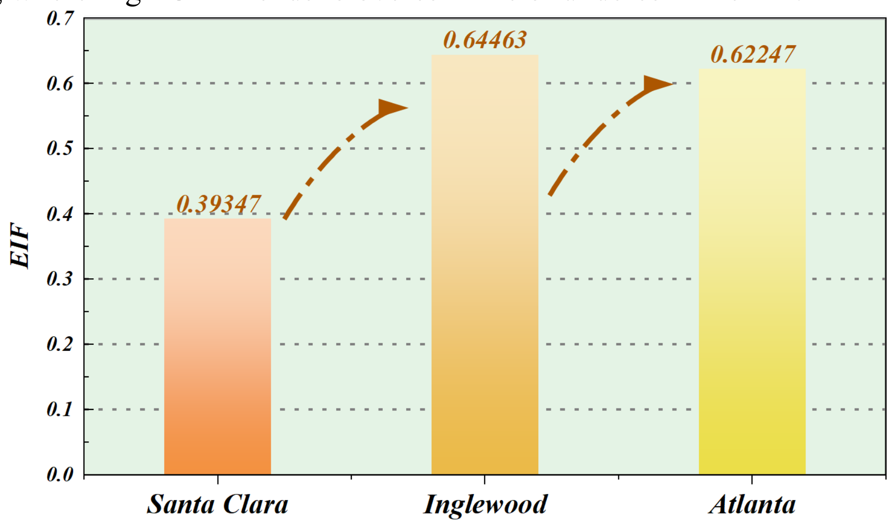


Figure 13 Predicted EIFs of scheduled host cities(2026-2028)

According to the predicted EIF index, Inglewood**(0.6446)** shows the highest environmental impact, followed by Atlanta**(0.6225)**, indicating higher carbon intensity and lower sustainability. In contrast, Santa Clara**(0.3935)** records the smallest value, suggesting superior environmental efficiency and infrastructure sustainability. These results are consistent with the model's feature importance ranking, where energy use and transportation remain the primary contributors to city-level impact.

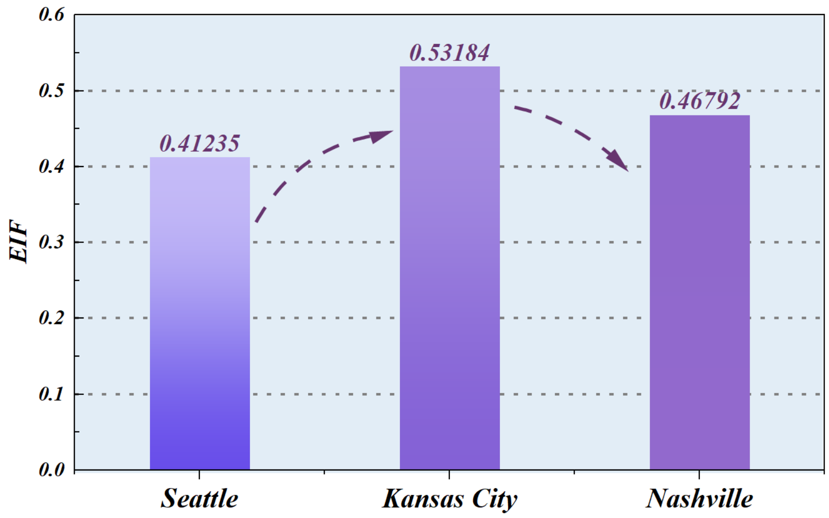


Figure 14 Predicted EIFs of 2029 candidate cities

Based on the predicted EIF index values, Kansas City**(0.5318)** presents the highest environmental impact out of the three, meaning relatively higher energy and emission intensity. Nashville**(0.4679)** is middle-ranked for moderate sustainability performance. Seattle shows the lowest EIF value of **0.4124** and, therefore, the strongest environmental efficiency and most sustainable conditions to host the 2029 Super Bowl.

6.4 Relation Analysis Between EIFs and Urban Characteristics

Based on the MIC model, we classify cities into four groups for comparative analysis: **legacy hosts without explicit sustainability focus**(Indianapolis, Pontiac, Palo Alto, Pasadena, San Diego, Tempe, Minneapolis, Detroit, Arlington, Miami, Miami Gardens, Jacksonville, Houston, East Rutherford, Los Angeles, Tampa, and Atlanta), **sustainability-focused hosts**(Inglewood, Glendale, Las Vegas, and New Orleans), **scheduled hosts**(Santa Clara, Inglewood, and Atlanta), and **candidate NFL cities**(Seattle, Kansas City, and Nashville). We then examine how each type of city's environmental impact is influenced by its urban characteristics under different indicator dimensions. The results are shown through **Fig.15** to **Fig.18.**

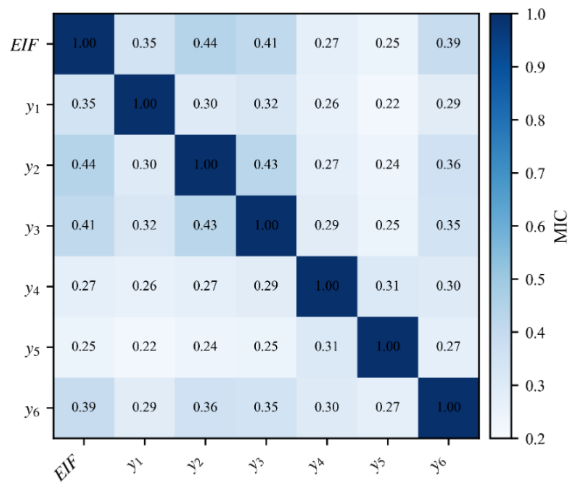


Figure 15 MIC Matrix of Legacy Hosts

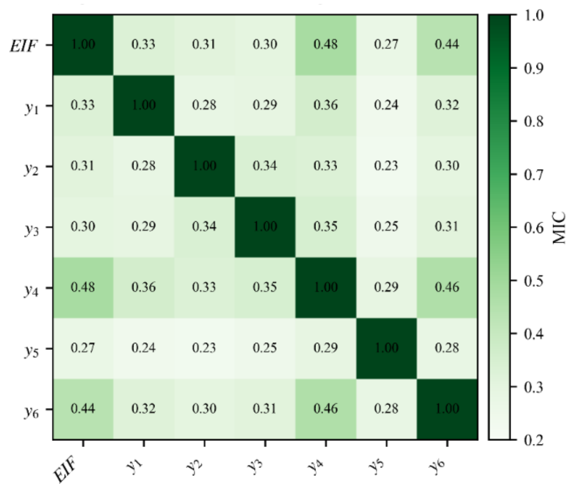


Figure 16 MIC Matrix of Sustainability-focused Hosts

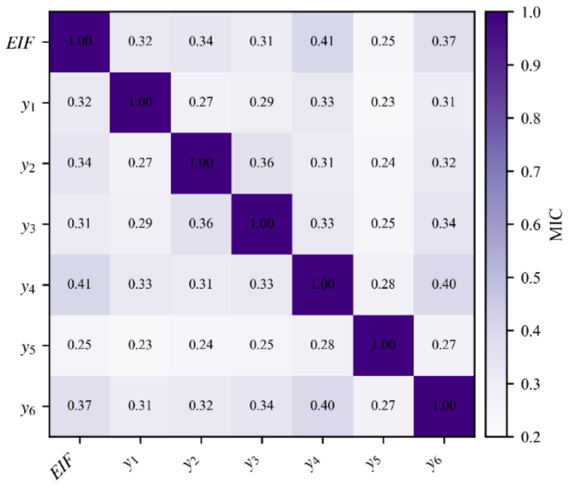


Figure 17 MIC Matrix of Scheduled Hosts

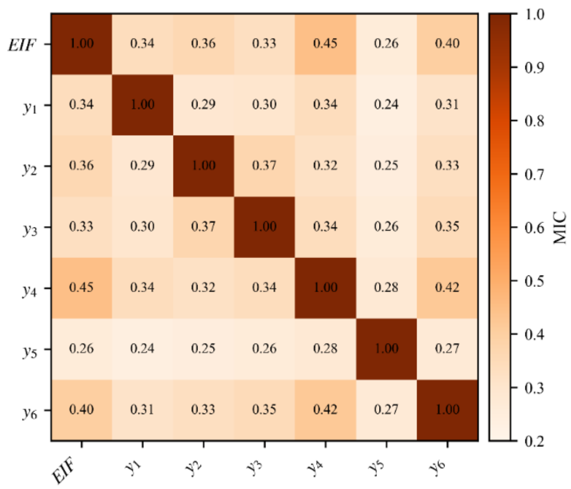


Figure 18 MIC Matrix of Candidate NFL Cities

- From **Fig.15**, it can be seen that legacy host cities are moderately associated with the demographic and infrastructure variables of total population y_2, road length y_3, and hotel capacity y_6, as indicated by a MIC of approximately **0.36-0.44**. The economic level y_1 is relevant but slightly weaker, while the share of renewable energy y_4 and climate y_5 show only a modest link to EIF. This implies that in early Super Bowls, environmental impact was driven more by the scale of the city and tourism infrastructure than by energy or climate considerations.

- As shown in **Fig.16**, EIFs of the four sustainability-oriented hosts are most strongly related to renewable energy share y_4 and hotel capacity y_6 (**0.44-0.48**), while the influence of population and road length becomes comparatively weaker. This indicates that, once explicit sustainability strategies are adopted, energy structure and accommodation management play a more central role in shaping environmental impact. The stronger coupling between EIF and y_4 reflects the effectiveness of integrating low-carbon energy into event planning.

- **Fig.17** illustrates that for the 2026-2028 upcoming hosts, EIFs are moderately correlated with renewable energy share y_4 and hotel capacity y_6 (**0.37-0.41**), while the economic and demographic indicators remain secondary. In summary, this pattern between legacy and fully sustainability-focused hosts also means that energy and service planning are already crucial but not yet dominant. This also implies that further reducing the environmental burden of the future Super Bowls could be effectively achieved.

- From **Fig.18**, it can be seen that EIFs of three candidate cities present a mixed pattern: while correlations with y_4 and y_6 are still relatively high(**0.40-0.45**), links with y_2 and y_3 remain non-negligible. In other words, for them, both energy structure and city scale are going to matter toward their future hosting potential. Compared with the sustainability-focused hosts, their MIC profile is similar but a bit weaker, which implies room for further improvement of renewable energy penetration and hospitality efficiency.

7 Olympics Host City Selection

To extend and reflect on our model, we first apply it to the Olympic Games as a representative mega-event. We adjust the indicator system by adding three indicators and removing three others to better align with the environmental characteristics of the Olympics. Using the AHM–PCA–FCE framework, we then evaluate 22 cities to identify the most suitable Olympic host. Finally, we conduct sensitivity analyses across three dimensions to examine the robustness of our extended model.

7.1 Event Selection

We choose to extend our model to **the Olympic Games** because their global scale, long duration, and substantial environmental footprint make them comparable to the Super Bowl in sustainability challenges. As a multi-venue, multi-sport mega-event, the Olympics feature complex energy use, transportation flows, and infrastructure demands, providing a rigorous test of the model's generalizability. Furthermore, the availability of its standardized environmental data and established sustainability frameworks supports reliable multi-event validation and strengthens the robustness of our approach.

Besides, to ensure that the constructed environmental indicator system aligns with the characteristics of the Olympic Games, we **remove** x_3 game-week supply plan attribute, x_9 February deplaned passengers, x_{11} Hotel occupancy during event week; and **incorporate** z_1 Cluster Compactness Index(CCI), z_2 Inter-Venue Transit Share(IVTS), and z_3 Athleter Village Intensity as additional indicators. The adjusted indicator framework is shown in **Fig.19**.

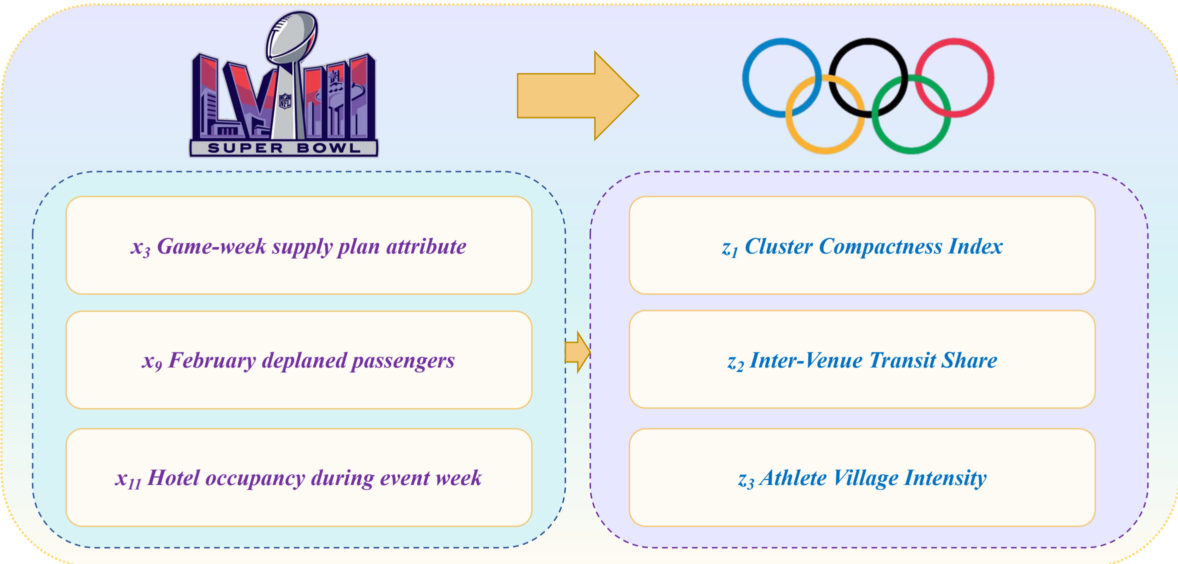


Figure 19 Olympic Adjusted Indicator Framework

7.2 Model Extension and Computational Evaluation

Based on the previously constructed AHM–PCA–FCE model, we evaluate the environmental impacts of hosting the Olympic Games across 22 cities. The results are presented in **Fig.20**.

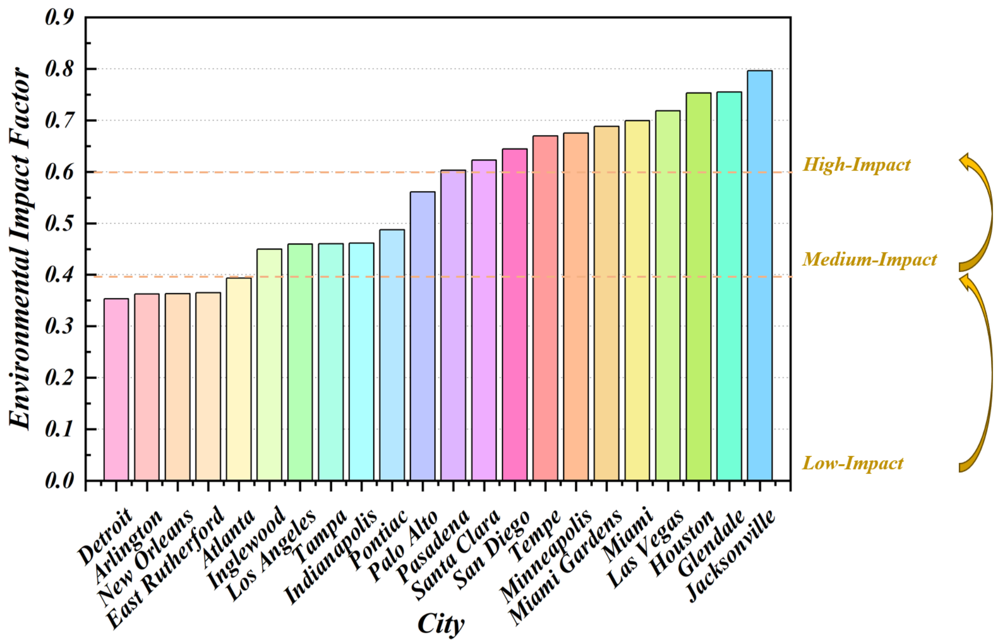


Figure 20 Selection of Olympic Host Cities Based on EIFs

The EIFs among the 22 cities exhibit a great difference. Indianapolis, Pontiac, and Palo Alto fit into the low-impact category, with EIFs less than 0.4, reflecting strong sustainability and low ecological pressures. Atlanta, Inglewood, and Los Angeles fall into the middle range of impact, with their respective values between 0.6 and 0.7, recording average energy and transportation burdens. However, the highest values of EIFs greater than 0.75 are recorded for Miami, Miami Gardens, and Las Vegas, reflecting higher carbon emissions and resource consumption.

7.3 Sensitivity Analysis

We conduct sensitivity analyses on indicators x_1 Local grid CO_2 intensity baseline**,** x_{10} RTA special service plan, and x_{12} Materials recovered to examine how changes in these factors affect the environmental impact of hosting the Olympic Games. Each indicator is adjusted by **5%**, **15%**, and **30%**, and the corresponding results are shown in **Fig.21**.

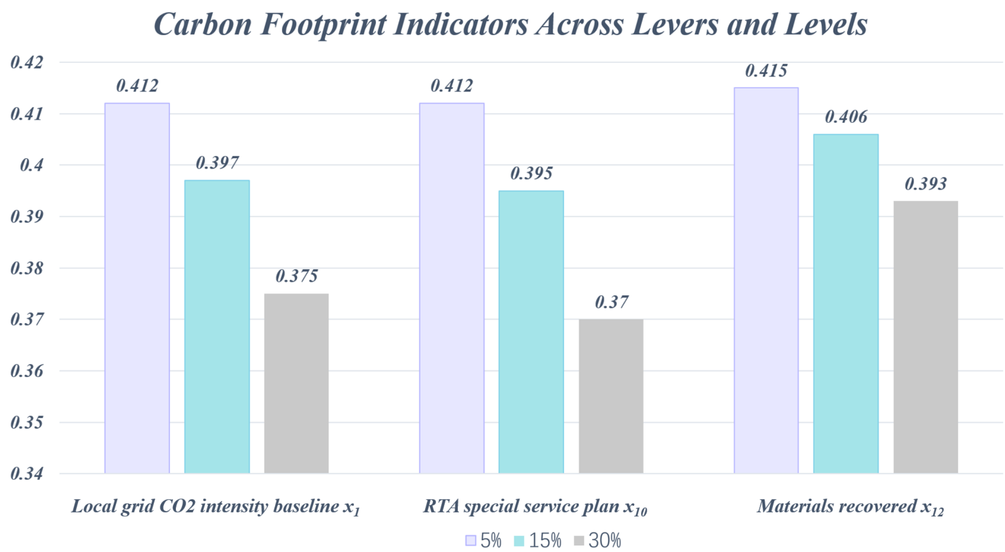


Figure 21 Olympic Environmental Impact Under Indicator Changes

According to sensitivity analysis, as these three parameters increase from 5% to 30%, the mean EIF decreases gradually for x_1**(0.412→0.375)**, x_{10}**(0.412→0.370)**, and x_{12}**(0.415→0.393)**. Among them, x_1 has the highest sensitivity, which means improvements in energy structure have the most substantial effect on sustainability. We further compare the environmental impact patterns between the Olympic Games and the Super Bowl as follows:

✧ **Event Duration and Scale**

Duration and scale: the Olympics span multiple weeks and many venues, accumulating base load and infrastructure emissions; the Super Bowl concentrates impacts within a single day with strong local intensity.

✧ **Distribution in space and structure**

Spatial structure: the Olympics are spatially dispersed, elevating inter-venue transport emissions; the Super Bowl is centralized, increasing local pressure on energy and waste systems.

✧ **Sustainability Strategies**

Strategy emphasis: the Olympics prioritize long-horizon green construction and legacy reuse; the Super Bowl emphasizes operational efficiency and near-term carbon reduction.

8 Model Evaluation

8.1 Strengths

- ✧ **Indicators are selected from multiple dimensions**, including environment, infrastructure, climate, resources, and economy. This multi-aspect system provides a comprehensive framework that reflects the characteristics of each host city and captures its environmental performance.
- ✧ **Life Cycle Assessment (LCA) based analysis.** We divide Super Bowl related environmental impacts into 5 stages chronologically. This ensures that the evaluation process aligns closely with actual event operations.
- ✧ **Multiple models are integrated,** including LCA, MIC, PCA, AHM, and LightGBM, forming a complete and coherent analytical pipeline. This combination enhances robustness and supports both qualitative interpretation and quantitative prediction.
- ✧ **City-level modeling is performed separately,** allowing the model to account for regional heterogeneity in climate, infrastructure, and resource availability. This ensures that results remain aligned with real-world urban differences.

8.2 Weakness

- ✧ **Data quality partly depends on third-party sources,** which may introduce uncertainties or inconsistencies.
- ✧ **Historical data may not fully represent future urban development,** especially for rapidly evolving cities. Changes in infrastructure, policies, or environmental technologies could weaken the predictive validity of models trained on past information.

9 Recommendation Letter

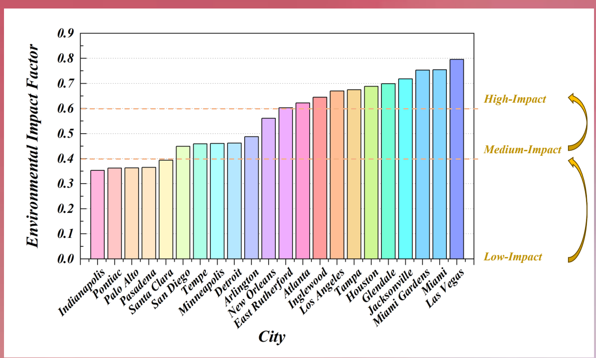
Super Bowl

To the National Football League:

We are writing to share our recommendation for the 2029 Super Bowl host city based on our environmental impact analysis. Using an evaluation framework and prediction methods, we identify a city that best minimizes ecological impact while meeting the requirements of a large-scale sport mega-event in three categories. Here are the details:

✧ Recommendation for Past Host Cities

Among the 22 cities that have previously hosted the Super Bowl, **Indianapolis** shows the lowest overall environmental impact among legacy hosts, supported by efficient energy use, compact urban travel patterns, and strong service capacity that reduces emissions during major events.

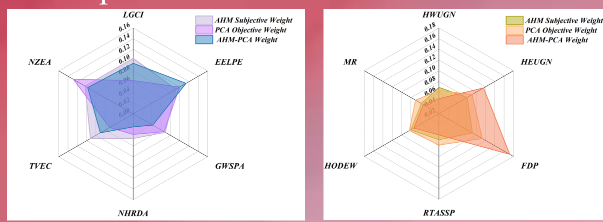


✧ Recommendation for Scheduled Host Cities (2026-2028)

With the scheduled host cities in 2026-2028, **Santa Clara** offers the most favorable environmental performance.

✧ Recommendation for 2029 Candidate Cities

Among the new candidate cities for 2029 Super Bowl, **Seattle** performs best across multiple sustainability measures. It demonstrates the lowest predicted environmental impact, driven by strong performance on key environmental drivers and high renewable-energy availability that reduces emissions and supports sustainable event operations.



✧ Method Extension and Robustness

We further applied our evaluation approach to the Olympic Games to test its adaptability beyond the Super Bowl. The model produced stable and reliable results across the larger, multi-venue Olympic framework, demonstrating the robustness and generalizability of our methodology.

✧ Conclusion

Indianapolis, Santa Clara, and Seattle each offer the strongest environmental advantages within their respective groups. Selecting these cities would meaningfully reduce the ecological footprint of future Super Bowls and support the NFL's commitment to sustainable event management.

References

[1] Babiak, K., Trendafilova, S., & Nite, C. (2023). *Sustainability in sport mega-events: Emerging frameworks and future directions*. Sport Management Review, 26(4), 567–582.

[2] Su, X. Z., Chen, L., & Xu, X. L. (2025). Carbon emission and energy risk management in mega sporting events: challenges, strategies, and pathways. *Frontiers in Environmental Science*, 12, 1513365.

[3] Lu, S., Yu, Q., & Deng, C. (2024). The risk of large-scale sports events and its prevention: based on the perspective of urban ecological environment protection. *Journal of Biotech Research*.

[4] Rusyn, B., Alokhina, O., Korus, M., Ivchenko, D., & Pits, N. (2023). Analytical hierarchy method for evaluating the environmental sustainability in the context of ecosystem services enhancement. *Environmental and Climate Technologies, 27*(1), 912–927.

[5] Greenacre, M., Groenen, P. J. F., & Tuzhilina, E. (2022). Principal component analysis. *Nature Reviews Methods Primers, 2*, Article 101.

[6] Guo, A., & Others. (2024). *AHP–EWM weight coupling model used for assessing structural damage in heritage architecture.* Heritage Science, 12, 1513365.

[7] Zhang, H. (2023). Fuzzy comprehensive evaluation and quantitative weight analysis in structure management of human resources. *PLoS ONE, 18*(7), e0288795.

[8] Bhattacharya, P. (2024). *A fuzzy synthetic evaluation approach to assess the risks associated with municipal waste management: Implications for sustainability.* Resources, Conservation & Recycling, 213, Article 106605.

[9] Aziz, A. (2025). *A fuzzy logic-based risk evaluation and precaution level model for handling hazardous chemicals.* Journal of Loss Prevention in the Process Industries, 76, Article 104445.

[10] Lazarsfeld, J., Ugander, J., & Shah, R. (2022). Differentially private maximal information coefficients. *Journal of Privacy and Confidentiality, 12*(1), Article 5.

[11] Zhao, Z., Wang, H., & Liu, X. (2025). A population spatialization method based on the integration of the maximum information coefficient. *PLoS ONE*, 20(2), e0321263.

[12] Gamage, H. N., Chetty, M., Lim, S., & Hallinan, J. (2023). MICFuzzy: A maximal information content-based fuzzy approach for reconstructing genetic networks. *PLOS ONE, 18*(7), e0288174.

[13] Constantin, R.-S., Davidescu, A. A., & Manta, E. M. (2025). Time series forecasting with LightGBM under data scarcity: An application to Romania's inland gas consumption. *Proceedings of the International Conference on Business Excellence, 19*(1), 1518–1531.

[14] Zhou, J., Liu, S., & Peng, H. (2025). A LightGBM-based power grid frequency prediction model incorporating dynamic significance-correlation weighting. *Energies, 18*(13), 3308.

[15] Oikonomou, K., & Damigos, D. (2024). Short-term forecasting of base metals prices using a LightGBM and a LightGBM-ARIMA ensemble. *Mineral Economics, 38*(1), 37–49.