The 2025 HiMCM Results

Kristin Fetherston-Arney, HiMCM Director - Wesley Hamilton, HiMCM Associate Director

Congratulations to all teams who participated in the 28th High School Mathematical Contest in Modeling (HiMCM). The HiMCM continues to provide a rich and rewarding experience for students, advisors, schools, and judges around the world. This year, a total of 1,091 teams, each consisting of up to four students and representing 498 schools across 24 countries and regions, came together to engage in this remarkable challenge.

The judges were once again deeply impressed by the dedication, creativity, and mathematical insight demonstrated by participants. We recognize the significant investment of time and resources required for those involved and we extend our sincere gratitude to all who contributed to this year's competition.

We applaud every participant for their work in the HiMCM. Among these tremendous efforts, eight teams rose to the highest distinction and were named Outstanding Winners. In addition, with the support of the National Council of Teachers of Mathematics (NCTM), two teams were honored with the NCTM Award.

Outstanding Teams

- 16390 - Borden International Education School, Hubei, China
- 16442 - Shanghai Nord Anglia Chinese International School (Secondary), Shanghai, China (NCTM Winner – Problem B)
- 16486 - BASIS International School Park Lane Harbour, Guangdong, China
- 16603 - Taipei American School, Shilin, Taiwan China (NCTM Winner – Problem A)
- 16738 - Shanghai Institute School, Shanghai, China
- 16739 - Guangdong Country Garden School, China
- 16843 - Nanjing Foreign Language School, Jiangsu, China
- 17401 - Horace Greeley High School, NY, USA

The 2025 Contest

COMAP remains committed to presenting engaging, real-world problems that challenge and inspire students. As in previous years, teams selected from two distinct problems: Problem A, which focused on modeling Emergency Evacuation Sweeps, and Problem B, which explored the Environmental Impact of Sport Mega-Events.

The 2025 contest received 1091 submissions. Of these, 585 addressed Problem A and 506 addressed Problem B. In total, 4023 students participated, representing a wide range of schools and regions including teams from Australia, Brazil, Canada, China, France, Germany, Hong Kong (SAR) China, Japan, Macau (SAR) China, Malaysia, Norway, Peru, Philippines, Portugal, Singapore, South Africa, South Korea, Taiwan China, Thailand, Turkey, United Arab Emirates, United Kingdom, USA and Vietnam. Within the United States, 230 teams represented 25 states.

Overview

For more than four decades of international modeling contests, COMAP has witnessed growing enthusiasm for mathematical modeling. In 2025, HiMCM marked its 28th year, as well as the 25th anniversary of offering two problem choices. From an initial cohort of just 115 students, participation has expanded and now over 55,000 students worldwide have engaged with complex, meaningful problems through HiMCM. The contest welcomes all high school and below students. Success is grounded not in advanced coursework, but in a solid foundation of high school mathematics and curiosity about solving real-world problems.

We also wish to recognize the invaluable role of advisors and teachers. These dedicated educators inspire students to extend beyond the traditional curriculum and embrace new ways of thinking. Their mentorship is essential to the success and continued growth of the HiMCM, and we are deeply grateful for their commitment.

Rules and Expectations

The core principle of the HiMCM is that all work must be completed solely by team members using inanimate sources such as textbooks, research articles, and other published materials. Collaboration outside the team, including consultation with teachers, experts, or external individuals, is not permitted. Similarly, conducting surveys, interviews,or accessing solutions shared electronically by others violates contest rules. Again, only the team members may contribute to the solution through their knowledge and work. Additionally, COMAP is committed to ensuring that all necessary resources for success are accessible without requiring additional purchases. The materials and information provided



HiMCM Contest

Table 1: 2025 HiMCM Results

Problem	Outstanding	%	Finalist	%	Meritorious	%	Honorable Mention	%	Successful Participant	%	Unsuccessful	Disqualified	Not Judged	Total
A	5	1%	21	4%	80	14%	158	27%	315	54%	1	2	3	**585**
B	3	1%	27	6%	63	12%	152	30%	255	50%	0	5	1	**506**
Total	8	1%	48	4%	143	13%	310	29%	570	52%	1	7	4	**1091**

by COMAP, along with a team's knowledge, skills, and research, are all that is needed for success.

The use of artificial intelligence (AI) is permitted in limited, transparent ways. For example, AI may be used for research assistance, editing, or translation. All AI use must be fully documented, including prompts and outputs, in a report that does not count toward the page limit.

Advisors and students often ask what level of mathematics is required and what special coding skills are needed for the contest. Our HiMCM problems are accessible using high school level mathematics alone and no programming or coding skills are required. As in all our contests, each of our problems is accessible on multiple levels, with many teams making use of different levels of mathematics and technology tools. Students should apply the mathematics they understand and are able to explain in their solution analysis, as the ability to clearly communicate their methods and rationale is valued in final judging.

COMAP understands that there is not one standard or body of knowledge of high school level mathematics for an international math modeling competition. The judges see a wide range of abilities and skills. Teams should note that using lower levels of mathematics properly can do just as well as higher levels of mathematics in the competition. Teams should not be focused on impressing the judges with overly complex techniques as models but rather concentrate on using good modeling techniques in putting together their solution.

COMAP uses X and Weibo to provide contest information to participants. Follow us @COMAPMath on X or COMAPCHINAOFFICIAL on Weibo for contest guidance and up to date contest information.

Judging Process

Judging is conducted blindly without identifying information about teams or schools. As open-ended problems, submissions are evaluated on their assumptions, reasoning, and modeling approach rather than adherence to a single correct answer with a rubric. All submissions are reviewed electronically, allowing for a diverse and highly qualified judging panel to easily review. Each paper is initially evaluated by at least two preliminary judges. We extend our sincere thanks to these judges for their thoughtful and thorough assessments.

After the preliminary round, papers are categorized as Finalist, Meritorious, Honorable Mention, or Successful Participant. Finalist papers advance to a final round of judging, where the most exceptional submissions are selected as Outstanding. This year, 105 papers advanced to final judging, with eight ultimately earning the Outstanding designation.

Looking Ahead

The HiMCM continues to evolve as a platform for students to apply mathematics in meaningful and collaborative ways. Additionally, COMAP has expanded opportunities through the MidMCM, now in its fifth year, offering middle school/level age students the chance to engage in mathematical modeling. Both contests emphasize the power of mathematics as a tool for understanding and representing the world. Students develop critical thinking, communication, and teamwork skills while gaining confidence in tackling complex, open-ended challenges. Importantly, no advanced mathematics or programming experience is required. Success lies in clear reasoning, effective communication, and thoughtful application of accessible mathematical tools. Judges value clarity and sound methodology over unnecessary complexity.

We encourage advisors to continue fostering creativity, curiosity, and perseverance in their students. COMAP encourages all mathematics faculty to get involved, motivate your students to be problem solvers, make mathematics relevant, and open the doors to future success. We want to partner with teachers as we continually strive to improve the contest and make it accessible and impactful to all students. Together, we can ensure that mathematical modeling remains an engaging, accessible, and impactful experience for all.

2026 Contest Dates

Mark your calendars for the 6th annual MidMCM, and the 29th HiMCM, to be held November 4 - 17, 2026. Registration for the 2026 MidMCM and HiMCM will open in September. Student teams may work at any time during the contest window. At the team members' convenience, teams download and choose their problem, complete their modeling solution, and electronically submit their solution document by the deadline on November 17th. Again in 2026, one team for each problem will receive the NCTM award. Teams can learn more about COMAP's contests and registration at www.comap.org/.

MathModels.org Has Moved

Powered by COMAP content, MathModels.org is a wonderful resource for students and teachers to make mathematical modeling a year-round activity. This site includes past competition problems, exemplar solution papers, and judges' commentaries that can provide guidance for teams preparing to compete. Teachers and students may use the materials found on this site to enrich their classes and help prepare students for mathematical and interdisciplinary modeling contests. We encourage you to visit this wonderful site in its new home on the COMAP webpage: www.comap.org/mathmodels.

MATHmOdels.org

The following guidance reflects the qualities our judges consistently value in HiMCM papers. Strong submissions present a clear and comprehensive summary and address all aspects of the problem through the thoughtful development and application of one or more mathematical models. The most competitive papers distinguish themselves through clarity, organization, and well-supported reasoning, demonstrating both careful analysis and effective communication.

The strongest papers go beyond meeting the basic requirements. They present complete and well-integrated mathematical and logical analysis in a cohesive and engaging manner. These submissions are not only thorough but also accessible. They are easy to follow, logically structured, and often creative in their approach. They include clearly stated assumptions with appropriate justifications, a transparent description of the modeling process, well-explained results and analysis, and thoughtful discussion of strengths and limitations, including sensitivity analysis. Strong conclusions and appropriate references further support the overall quality of the work.

Problem Discussions

Our judges continue to emphasize that HiMCM problems are designed to be accessible using high school-level mathematics. While some teams attempt to incorporate advanced concepts or tools, these approaches are only effective when they are clearly explained and appropriately applied. Papers that rely on techniques beyond a team's ability to fully justify or interpret are often less successful. We are not seeking the use of the most advanced mathematics. Rather, the most successful papers are those in which teams develop models grounded in mathematical ideas they understand, can clearly explain, and can apply effectively.

The specific problem discussions below provide comments on the problem and how teams addressed the requirements. In the next section, we provide judges' comments about the solutions by breaking down the various parts of a submission and providing exemplars. To view the complete problem statements, visit www.comap.org/mathmodels or www.himcmcontest.org.

Problem A: Emergency Evacuation Sweeps



For the first time 25 years ago, COMAP posed two problems for High School students to choose from during the HIMCM. To honor this anniversary, we opted to revisit Problem A but with a different scenario. In 2000, student teams were asked to develop an efficient algorithm for police cars to sweep the area to force bank robbers (who were fleeing by car) into one of the established roadblocks. Today, we challenge teams to conduct a different type of sweep.

During emergencies such as fires or gas leaks, rapid and orderly evacuation of a building is critical for saving lives. Trained personnel responding to these emergencies (responders) often sweep a building – going from area to area and room to room – to ensure all occupants have evacuated safely or are actively moving towards the exits. This would include visually assessing, if possible, all spaces to confirm that they are empty with the goal of guaranteeing no one is left behind. The efficiency and strategy behind these sweeps could mean the difference between safety and true peril.

Teams were asked by the Clearing Of Multi-use Assembly Public/Private-spaces (COMAP) with designing a mathematical model to optimize sweeping strategies in multi-floor buildings during emergency evacuations. The goal is to ensure all rooms are cleared in the shortest possible time, while prioritizing occupant safety and responder efficiency. Buildings vary in layout, number of floors, stairwell placement, location of exits, and occupancy patterns while responders face pressure to act quickly with sometimes limited information. These considerations must be included in your analysis.

This problem makes teams look at emergency evacuation from different perspectives. Students have been taking part in fire drills in school throughout their lives; now they can see it from the other perspective while considering both the

buildings themselves and the responders in all of their modeling decisions, applied to three different layouts/scenarios.

Teams were asked to build a sweep model after conducting research on the aspects that they would need to incorporate. The sweep models were then applied to a simple building layout before extending the model with additional complications and applying it to different types of buildings. Teams were then asked to draft a formal letter to a local Emergency Planning Committee addressing both how to sweep buildings and assign responders including recommendations for strategies that the committee could use to classify different types of building and how to respond.

Problem B: Environmental Impact of Sport Mega-Events

Every year, the Super Bowl brings together millions of people – from fans traveling to the host city, to viewers tuning in from around the globe. This event generates enormous social and economic value, but it also comes with a significant environmental cost. With air travel, hotel stays, energy use, food consumption, waste generation, and massive broadcast operations, the Super Bowl's environmental footprint is substantial.

As we grow more concerned about sustainability, large-scale sporting events or sport mega-events like the Super Bowl are being made more environmentally responsible. External organizations have developed models and scoring systems to rate the sustainability of these large-scale events based on many factors from energy use and waste management to transportation and carbon offsets. While these environmental scores exist and a wealth of data is available on what event planners are doing to be eco-friendly, sustainability rarely drives key decisions such as where to host an event or how to design its required supporting infrastructure. This disconnect presents an opportunity to build a model where ecological impact is not just an afterthought, but a core driver of planning and location decisions. Imagine a scenario in which host cities are evaluated not only for their economic viability and fan capacity, but also for their renewable energy potential, existing infrastructure efficiency, and transportation accessibility. Such a model would enable organizing committees to balance performance, cost, and sustainability in a measurable way.

The Consulting Organization Making Athletics Planet positive (COMAP) challenged teams as sustainability consultants to assess and reduce the environmental footprint of sport mega-events. The goal is to understand the ecological impact of these events and explore how changes in event planning could make a difference. This problem is not simply about creating a model to assess the environmental impact of large-scale sporting events or sport mega-events. Instead, the focus was on bringing environmental impact to the forefront of the decision-making on where to host.

Teams were asked to build a decision model after conducting research on factors included in current Super Bowl sustainability Metrics. The decision model was to be applied to where to host a future Super Bowl before extending the model with additional complications and applying it to a different type of sport mega-event. The addition of infrastructure in this extension was a key factor. Teams were then asked to draft a persuasion paper letter to the NFL recommending the city that should host the Super Bowl because of sustainability.



Judges' Discussion

Here we will examine in-depth what judges were looking for in each aspect of a well-written submission. For each piece, we will outline what we expected to see, possibly including some specifics related to each problem, and then highlight actual examples from this year's Finalist and Outstanding papers. Students and coaches are encouraged to review these discussions and incorporate these best practices in future competitions.



Note!
An extensive online Appendix "2025_HiMCM_Examples" containing all the examples mentioned below can be found here:
www.comap.org/2025HimcmExamples

Overall

Student participants are expected to ensure that their papers adhere to all contest rules as outlined on the official contest website. Submissions must be presented within the 25-page limit and use a minimum 12-point font. Each year, judges encounter papers that attempt to exceed these limits by reducing font size, adjusting margins, or adding additional pages. While such submissions are not automatically disqualified, they are unlikely to earn top designations.

Papers that advance to the Finalist and Outstanding levels create a positive and lasting impression and consistently demonstrate a high degree of clarity, organization and coherence. They begin with a clear and informative summary and provide a structured preview of the paper. These submissions present well-justified assumptions, clearly explain the development and application of their model, and support their results with appropriate mathematical analysis. The strongest papers also include sensitivity analysis, discuss model strengths and limitations, and conclude with clear, well-supported findings. Ultimately, it is not the length of a paper but the clarity, completeness, and effectiveness of its communication that distinguishes strong submissions.

Before submission, teams are encouraged to carefully review their work, including the use of spelling and grammar tools. For international teams, all symbols and labels in tables and figures should be presented in English. To ensure fairness in the judging process, student and school names must not appear anywhere in the paper.

Executive Summary

The executive summary is an essential component of a standout modeling paper as it requires synthesizing the most critical information from the report into a clear, concise overview of the work to follow. A memorable executive summary sets the tone for what can be expected throughout the report, while also being able to stand alone as a complete depiction of the problem from its introduction to its conclusion including implementable recommendations.

Teams are encouraged to write the Executive Summary after completing their paper so they can accurately capture the entire analysis. Once written, it should be reviewed carefully to ensure it stands alone as a representation of the full submission.

For Problem A, Team 16486's executive summary (Example 1) found an effective mix between providing high-level, lay-person overview of their model while including technical keywords to prime the reader as to what the implementation would entail. They include numbers quantifying the performance of their model that are understood in the context of the problem statement and finish their summary with a concise description of their key contributions.

For problem B, this meant addressing both the Super Bowl and another global mega-event and addressing the key environmental impact factors chosen to develop the model around. An excellent example of this is from Team #17173 (Example 2). Despite being one of the shortest executive summaries submitted, it efficiently and clearly summarized the goals, methods and results of their work in a way that was distinct from the main body of the report. This summary whets the appetite of the reader, enticing them to read further to learn more about how the locations for the events were chosen.

Demonstrated Contextual Understanding

Strong papers demonstrate a clear understanding of the complexity of the problem beyond what is explicitly stated in the prompt. This understanding is often reflected in the introduction, where the better teams move beyond simply restating the problem to provide their own unique meaningful insight into the bigger picture. Some teams incorporate background research or elements of a literature-style review, though this is not required. Teams should know that what matters most is the depth of understanding conveyed. This deeper perspective strengthens the entire paper. It enables teams to more clearly define the problem and make thoughtful modeling decisions.

Teams are encouraged to make the introduction their own by establishing a strong foundation before beginning their model development. Judges

appreciate well-developed insights, even those that are not directly incorporated into the final model, as they demonstrate deeper engagement with the problem. Clear documentation of research and reasoning further enhances the credibility and quality of the work.

This year and in the past, judges looked for teams that provided well thought out insights into the problem. Sometimes these insights had to be ignored to build the base model, but the fact that a team would recognize and communicate these insights was appreciated.

Additionally, some teams produced a literature review. Using outside references to provide understanding of the problem was resourceful. Some teams did great without the literature review. Judges looked for teams that produced insights and understanding of the problem that went beyond what was in the problem statement.

Judges saw many excellent examples of what demonstrated contextual understanding can look like in a submission. Team 16969, when tackling Problem A, produced a well referenced literature review (Example 3) that discussed the multitude of problems and considerations that should be examined when tackling an evacuation problem. Additionally, judges read many papers that provided a good understanding of the problem when they went above and beyond regurgitating the problem statements in their introduction. As an example, Team 16603 (Example 4) discusses the condition for an ``all clear'' to be given in an emergency evacuation. Papers such as this tell the narrative of what the objectives are for addressing the situation. In Problem B, Team 16521 (Example 5) conducted a literature review and then nicely consolidated this research in their section entitled "Problem Analysis". Here the team explained how the reviewed foundational works on the environmental assessment of sport mega-events and how that research drove the key environmental impact areas that they chose to focus on in their modeling. This strong foundation directly set the stage for their model development.

Judges would also like to highlight a team that demonstrated that they truly understood why they were tackling this problem in a different way than most. Team 16739 (Example 6) had a section towards the end of their paper, after they had implemented the model for the Super Bowl, adapted it for multi-sport or multi-venue events, and then implemented that model for Summer Olympics. The team discussed key differentiating mechanisms between types of sport mega-events in a section entitled "Comparative Analysis of Model Results" which truly demonstrated to judges that they understood why they were conducting the analysis they had. This section was not a problem requirement but instead demonstrated this team's deeper understanding of the bigger picture problem which was to determine what factors matter the most for different types of sport mega-events when it comes to determining the most sustainable host cities.

Assumptions

Teams are expected to clearly state all assumptions, either as a dedicated list or within the context of the paper where they are applied. Assumptions play a critical role in simplifying complex real-world problems and addressing gaps in data or knowledge. Effective assumptions are purposeful and directly connected to the model. Each assumption should be clearly justified, with reasoning that explains both why it is necessary and how it supports the modeling approach. Judges look for assumptions that are thoughtfully chosen and meaningfully integrated into the analysis, rather than included as general statements.

Teams are encouraged to consider where and how each assumption is used within their model. The strongest papers not only state and justify their assumptions but also demonstrate explicitly how those assumptions influence the development, application, and interpretation of the model.

Judges for Problem A looked for teams that used assumptions as a tool to assist in building the base model for searches to eliminate complexities so the base model could be built. Those teams would then build further models without the assumption constraints of the base model. For example, in the paper written by Team 16486 (Example 7), the team clearly states an assumption that will be in place for the base model. The team later builds upon the base model when assumption 5 is no longer in place.

Additionally, judges looked for papers that referenced where they obtained their data for certain assumptions. The judges appreciated when teams clearly referenced where they obtained the speed they used for the movement of firefighters like in Team 16691's paper (Example 8).

In Problem B, the paper written by Team 17205 (Example 9) was very readable, but they did an exceptional job of making their assumptions clear by organizing them into categories.

Model Selection

There are many valid approaches to modeling HiMCM problems, and judges appreciate appropriate and creative solutions. The most successful teams identify the key factors driving their model and explain the reasoning behind their chosen approach. This includes discussion of model inputs, structure, and data, as well as any adaptations made when using existing models. Strong papers often acknowledge the distinction between an ideal model and what is feasible given available data and time constraints.

While teams may explore multiple approaches, successful papers present a clear and focused modeling strategy. Rather than introducing several disconnected models, strong submissions develop a primary model that is refined or extended as needed. Judges consistently favor well-reasoned, clearly applied models over overly complex approaches that are difficult to justify or implement.

While the problem statement for Problem A did not make direct mention of graphs or networks, almost every team used a network model as the underlying framework. How these

graphs were adapted to finding optimal paths varied among the submissions though, with some papers using a linear programming formulation, some papers using optimal graph path algorithms like A*, traveling salesman, or Dijkstra, and/or some papers incorporating stochastic simulations to explore these relatively large state spaces.

One common approach was to formulate room allocation and path planning in one swoop using a linear program or similar optimization framework. Team 16486 (Example 10) used this approach, effectively describing their set-up in a direct manner. They formulated the sweep-plan as a room and edge assignment problem and described several constraints their solution should meet, which encoded that a sweep-plan must travel along consecutive edges and every room must be swept, among others. With these determined, they then naturally define their objective function, which allocates all room assignments and responder paths in one swoop. In their model evaluation, they also considered three other strategies for room assignment, showing that some thought went into deciding to use a linear program.

In slight contrast, many teams opted to build and describe their models in a way that highlighted properties of the networks a little more explicitly. This is not to say one approach is better than another, but rather the tools that become applicable are more easily seen in one versus another. This is the approach Team 17401 (Example 11) used, splitting the room allocation and path-finding problems into two pieces, and using graph algorithms for the second piece. In their model, rooms of a building were first clustered based on expected workload to clear them, after which an optimal path was computed using a traveling salesman problem solver they built and adapted to this problem. Team 16603 (Example 12) used a similar approach as above, splitting the problem into a room allocation and then path-finding problem, though in this case they used existing traveling salesman solvers for the initial path construction. The networks they constructed were dynamic, so it could happen that a responder's original path is no longer available; in these instances, their model used an A* algorithm to find a way around these obstacles. Moreover, they effectively used Monte Carlo simulations for the dynamic graphs to gather data and provide strong evidence for the optimality of sweep-plans they construct in the presence of noise.

These last two models naturally followed from their approach to the problem: first identify who should search which rooms, then figure out how to get them there. This contrasts with the first model mentioned, in which the two steps are performed at the same time. As stated earlier, neither approach is better than another - teams should explore the problem and potential solutions and let their work tell them which approach makes the most sense.

For Problem B, there were many modeling options and techniques available to teams participating in this competition, although many teams utilized similar approaches. What set teams apart was their ability to not only describe what they did, but why they used the approaches they chose. For example, many groups this year used Analytical Hierarchy Process (AHP) and Entropy Weight Method (EWM) together for a multi-criteria decision-making model and justified balancing the subjectivity of one with the inherent objectivity of the other. The main factor distinguishing groups using similar approaches was the ability to justify how the rankings were determined, or how the balance between methods was determined for their multi-criteria decision-making model. There were some excellent papers submitted that did not use the most powerful or unique modeling techniques but distinguished themselves by clearly and effectively explaining the reasoning behind their choices. For examples, judges chose to present the three Outstanding papers as they all had more standard modeling approaches with excellent extensions and explanations.

For their multi-criteria decision-making model, the Outstanding Team 16442 used AHP and EWM to determine the weights of their indices. This team used five core indices of environmental impact based on over 14 contributing factors. This allowed them to create their City Impact Evaluation Model (CIEM) to evaluate potential host cities for the Super Bowl followed using Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) for the determination of a sustainability score which can then be ranked. What judges appreciated the most about this standard approach was the team's explanations of the process and their adaptation of this method for their model extension (Example 13). When Team 16442 went to determine the best potential host countries for the World Cup, they adapted their initial CIEM by refining the five core indices and then added one more index to cover the environmental impact of infrastructure.

The Outstanding Team 16738 chose a similar methodology for their multi-criteria decision-making model. They utilized 12 secondary indicators and then Analytic Network Process (ANP) with EWM followed by TOPSIS. The team explained their choice of ANP over AHP (Example 14) because they wanted to capture the interdependencies between criteria. The judges appreciated the analysis that followed which allowed the team to extrapolate into the future based on what they determined would be the two indices that would most likely change significantly over a five-year period due to newly established public policy (Example 15). It was insightful for the team to note and understand that a sustainability score for 2029 might be significantly different than today's score if certain city's trends in implementing more green practices were considered.

Finally, the Outstanding Team 16739 approached their model creation a bit differently, still using techniques that are regularly seen by judges throughout this and other contests. The team chose Principal Component Analysis (PCA) but then took it further based on Pearson correlations and then a PCA-based multi-variate linear regression model (Example 16) to identify grid carbon intensity, renewable energy share, public transit coverage and the share of green or low-carbon travel as the main drivers. The next step for Team 16739 is what the judges really appreciated as the team used Autoregressive Integrated Moving Average (ARIMA) models to forecast future city characteristics for both cities that have posted and three that have not yet hosted. The identified features were then inputted back into the original analysis before conducting EWM.

Model Implementation

Teams must apply their model to generate meaningful results and demonstrate how those results address the goals of the problem. This requires more than presenting outputs as teams should interpret their findings in the context of the model's purpose and explain how the results inform the original questions. Strong papers reflect thoughtfully on their results, considering whether the model effectively captures the key aspects of the problem and whether the outcomes are reasonable and actionable. Once the theoretical development of their model is developed, teams are expected to apply their work to the problem to get tangible results that address the problem statement. In Problem A's case, teams were expected to implement their models to analyze response times for a simple six room office building and then consider at least two more building layouts.

Many teams defined a dynamic weighted graph representing the geometry and logistics for sweeping a building, allowing for dynamic hazards and the subsequent effects on the network and first responders' possible sweep-plans. Team 16843's implementation (Example 17) stood out to the judges for the in-depth and explicit implementation framework for their optimal paths, which took shape as five algorithms written in pseudocode accompanying their discussion. This helped bridge the gap between theory and implementation, which made analysis of their model's results easier to digest and understand.

Additionally, Problem A judges enjoyed the creativity displayed by teams in the variety of examples considered with their two additional buildings, especially when teams used building layouts based on buildings relevant to their lives. This was one of the reasons Team 17402 stood out during final judging: they considered actual layouts for a nearby kindergarten school (Example 18), as well as a floor of a high school building (Example 19), and performed an extensive analysis of their model applied to each (resulting in an optimal number of 3 and 4 firefighters to sweep the kindergarten and high school floor respectively).

In Problem B, the first application was to finding the most sustainable host city for the Super Bowl. Many teams simply took their model and ran it to determine the best host city for 2029. Judges wanted to highlight two teams that went further on their implementation by providing a differentiation and expansion that made their analysis more trustworthy and demonstrated to judges that they understood the details of this bigger problem. Team 16947 (Example 20) understood the data differences between teams that have previously hosted the Super Bowl and teams that have never hosted. This creates difficulties in the model when attempting to determine who should host in the future. This team implemented them separately and then nicely brought the lists together for their results discussion. Judges also appreciate the work that Team 17173 (Example 21) conducted to add time adjustments. They recognized that cities hosting in earlier years might obtain lower scores using their model because of technological limitations. As they did not want the temporal effect to penalize host cities from the earlier Super Bowl years, they implemented this time adjustment methods after the initial model results and then compared the differences. They then appropriately made their determination based on the updated model.

Model Cohesion & Integration

Teams should guide the reader logically from assumptions to model development, application, and results, ensuring that each step is connected. Papers that demonstrate this coherence are the most effective. While teams may analyze multiple aspects of a problem, the greatest value comes from integrating these components into a cohesive narrative that addresses the overall question. Strong papers demonstrate how the requirements connect, whether by building upon an initial model or adapting it to address additional dimensions of the problem. When multiple models or approaches are used, teams should clearly justify these choices and explain how each contributes to a deeper understanding of the problem. The most successful papers move beyond isolated analyses to present a connected line of reasoning that adds meaning and insight to the bigger picture problem teams are trying to solve.

In Problem A, teams overwhelmingly adapted graph-based approaches to build their model and apply it to the scenarios given in the problem statement. Judges particularly appreciated the discussion in Team 16390's development (Example 22), wherein they built their model while including pseudocode and enough details that it was clear how the model was to be applied, and how their unique approach to task allocation was to be integrated in their graph search and sweep model. This was followed by a brief model validation, then presentation of results, and finally a discussion of sensitivity. They presented a cohesive model discussion (from development to results) in this section, which gave judges a clear view of their approach for a single level building before extending their model to multiple floors, and other more complex scenarios.

Problem B judges viewed this aspect critical to identifying the best papers. If a team was able to integrate all their modeling into a cohesive narrative where subsequent work is linked and dependent on the initial work, then they were awarded a higher designation. The three Outstanding teams, Team 16442, Team 16738, and Team 16739 were able to do exactly that. Judges would also like to highlight the work of Team 16831 who had unique constructs for analysis of Super Bowl hosts where they examined those that have hosted, will host and have not yet hosted as well as the sustainability factors pre-event, during the event and then post-event. Additionally, the judges truly appreciated the approach Team 16831 (Example 23) had when developing their multi-venue, multi-day model. The team used the same model and added new sustainability factors that mattered with the change in scope like transportation and infrastructure. This idea of taking an initial model and then continually refining it to add the complexities that a team wishes to analyze is highly sought after by judges.

Sensitivity Analysis

Sensitivity analysis is an important component of the modeling process, as it examines how changes in assumptions, parameters, or input data affect the results of a model. It helps determine the robustness, reliability, and overall usefulness of the model. Effective sensitivity analysis often includes revisiting simplifying assumptions and evaluating how variations in key inputs influence outcomes. While there is no single required method, judges look for clear rationale behind the approach taken and thoughtful interpretation of the results.

Strong papers use sensitivity analysis to justify the level of confidence in their model and its conclusions. While some teams present this analysis as a dedicated section, those that integrate it throughout their work, connecting it directly to model development and implementation stand out. Judges feel that a well-executed sensitivity analysis demonstrates a deeper understanding of the model's behavior and its practical implications.

Solutions to this year's Problem A, Emergency Evacuation Sweeps, included a variety of factors and objectives. In reflecting on the stability of their models, most teams needed to consider the relative importance of a number of parameters describing the effects of such things as the walking speed of responders, changes due to the presence of smoke or clutter, or the number and/or condition of possible victims in the building. In general, such explorations involve systematic variation of parameter values so that the influence of that parameter may be isolated. Ideally, such influences would then be compared and interpreted in the context of the overall objective of the emergency sweep model.

Team 17460's sensitivity analysis discussion (Example 24) of its process is particularly strong. They describe exactly how the analysis was conducted, identifying the parameters to be tested, the range of values to be used, the method of testing, and the key output of interest, all clearly enough that the reader could duplicate their work independently. They report their results in easy-to-read, easy-to-compare graphical form, and the final discussion provides practical implications of these comparisons for planners for emergency response teams.

Judges from Problem B were looking for an extensive sensitivity analysis as there were many assumptions teams should make, data that had to be researched and incorporated, as well as a variety of parameters for most models. The presentation of the sensitivity analysis by Team 16442 (Example 25) was notable for its depth followed by the explanations of the meaning. The team conducted two examinations, and each was followed by graphics with results and then an analysis of results. Additionally, judges were impressed that the team brought it all together with a conclusion that reinforced their confidence in their selection for the most environmentally sustainable choice for the Super Bowl host city in 2029.

Strengths, Limitations and Conclusions

After completing the modeling and analysis, teams should include a clear discussion of the strengths and limitations of their model, along with concluding insights. This process pushes students to step back and look at the bigger picture of what their model can achieve and where there may be room for improvements, which is entirely reasonable and expected given the time constraints of the contest. It is easy to fall into the trap of saying "Our model is amazing and can solve any problem!" but judges are looking for a realistic assessment. Effective discussions highlight what the model does well by identifying the types of questions it is best suited to address while also acknowledging its limitations. These limitations may stem from simplifying assumptions, data constraints, or modeling choices. Strong papers use this section to reflect on the entire modeling process, offering insight into what worked well, where improvements could be made, and how the model might be extended with additional time or data. In many cases, judges are already familiar with the strengths and limitations of a particular modeling approach, and it is refreshing to see teams arrive at these same conclusions independently.

Clear, well-supported concluding thoughts that tie back to the problem and results are essential, ensuring the paper ends with a sense of purpose. With so many demands asked for in a modeling problem like this, it is easy to overlook the importance of including a conclusion in the report. The conclusion should contain direct recommendations as we solve these real-world problems to provide insight to interested parties. Team 16691 in Problem A (Example 26) did

an excellent job highlighting the capabilities of their model as it represented real safety and rescue priorities. They also chose robust optimization tools. This team also gave an honest reflection by relating their limitations directly back to specific assumptions. The judges also appreciate that their strengths and weaknesses were not presented as a bullet list but written out clearly as a cohesive discussion. It gave the impression that the students really took time to justify each of their assumptions and saw the connection to their model's value.

In Problem B, Team 16624 (Example 27) recognized that the environmental performance of a city does not remain stagnant from 2023 to 2029, as assumed in their model. The team acknowledged that future environmental performance could be forecast by incorporating data from 2023 and earlier. They identified the inclusion of planned sustainability upgrades for a city as a direction for future work.

Non-Technical Document

The purpose of a letter or other non-technical document is to communicate model-based recommendations in a way that is accessible to a non-technical audience. This section demonstrates a team's ability to translate complex mathematical analysis into practical, understandable insights. Since the intended audience often lacks technical expertise, teams must present their findings in clear, concise language, using an appropriate format such as a letter, memo, infographic, or article, as specified in the prompt. Rather than simply stating general practices, strong submissions explain what the model reveals about the implications of a particular course of action.

While some teams may think of it as a secondary component, the non-technical document can be a distinguishing feature of a strong paper. The most effective submissions reflect a clear understanding of the modeling results and communicate them in a way that is both informative and actionable for decision-makers. Judges thought that Team 16846 effectively conveyed their approach to non-technical audiences for Problem A (Example 28), which was amplified by the team publishing the letter at the start of their paper right after their executive summary. They provided a brief overview of their approach before spending the most space discussing their results in context, including some diagrams of actual plans produced by their model.

For Problem B, judges really focused on the omission of specialized mathematical terminology and the avoidance of acronyms without clear definitions. Because the letter is addressed to executives in the NFL rather than to students or instructors in mathematical modeling, accessibility and simplicity were essential. Many papers met or exceeded this standard of proficiency.

However, Team #16739 (Example 29) provided one of the strongest templates for a non-technical letter. The letter began with a concise summary of the model considerations, followed by the recognition of the winning city. To persuade the NFL of their recommendation, the team explained the reasons for their choice in clear, common language rather than technical detail. The letter concluded by identifying other potential host cities and offering an inspiring rationale for making environmental impact an important factor in the selection of a host city.

Appropriate Explanations

Discussions throughout the paper should be clear and detailed to allow the reader to follow the team's modeling process from beginning to end. This includes not only describing what was done, but also explaining why choices were made, with connections between modeling decisions and the requirements of the problem. Every year, judges seek to highlight submissions that offer a balance of sound mathematics with well written justifications of their approach. Judges looked for implementable measures that were backed by research and then well explained.

The Emergency Evacuation Sweeps problem of Problem A is closely related to the well-known Traveling Salesman Problem of graph theory. Essentially all student solutions to Problem A began with converting a floor plan and frequently other important information into a vertex-edge graphical representation. Starting with such a representation the task then becomes two-fold: (i) generate a set of paths that covers all rooms, and (ii) minimize an appropriate "cost" function associated with such paths. While it is possible to formulate a "brute force" solution to these tasks, such an approach becomes computationally impractical very quickly. The examples below illustrate (portions of) two different approaches to addressing this difficulty. In each case the teams made it easy for the reader to follow their work, using clear language to explain and provide essential information such as values of key parameters and references for further information.

Team 17401 (Example 30) addressed the computational difficulty by breaking their task into two steps, first partitioning rooms into localized groups, then building appropriate paths. For each task, the students provided clear instructions on exactly how the work is carried out. In addition to these instructions, they provided citations for each step of the process, making it possible for the reader to carry out their work independently but also to find additional theoretical discussion of the methods used. Team 16691's approach (Example 31) to solving the problem of computational complexity uses Ant Colony Optimization, a probabilistic simulation. Even though this approach may not be as familiar as more direct methods, the students provided first a general overview of the ideas behind the method, then gave the details of their actual set-up, including their selected parameter values and stopping rule. The reader would easily "follow along" with the solution process. Again, as is the case in Team 17401's solution, reference to the method is cited.

For Problem B, judges believe that what set teams apart was their ability to not

only describe what they did in their modeling process, but why they used the approaches they chose. There were some excellent papers submitted that did not use the most powerful or unique modeling techniques but distinguished themselves by clearly and effectively explaining the reasoning behind their choices. In addition to narrowing down the key factors driving their model, Team #16738 (Example 32) did an excellent job of highlighting the disparity between potential host cities' performance on these different benchmarks to emphasize the reason for their inclusion in their model.

Organization & Writing

Strong papers are well-structured, clearly written, and easy to follow, guiding the reader smoothly through the team's approach and analysis. A clear organizational framework allows the reader to understand how each component of the paper contributes to the overall solution. Effective submissions consistently connect assumptions, methods, and results back to the context of the problem. The strongest papers enhance understanding through thoughtful presentation. Equations are paired with clear explanations, and visual elements such as graphs or tables are used when appropriate to communicate complex ideas. Overall, strong communication demonstrates a deliberate effort to make the modeling process accessible to the reader.

Team # 16603 (Example 33) submitted a well-organized and well-written paper for their solution to Problem A. Note that even though it was outlined similar to the template, it had a natural flow between sections and information was presented in a reasonable order. One judge in particular appreciated the way that the definitions of the variables were well-organized with descriptions of why and how they were used with units included. When they presented a relationship between variables or a piece of their model, they explained why it made sense. They also summarized all of their findings together and included sections on validating their model. In Problem B, papers that impressed the judges had clear explanations of the models. Judges valued knowing why choices were made. Explaining abbreviations is helpful to the reader. When diagrams or graphs were used, the explanations of them should immediately precede or follow the graph or diagram. Judges appreciated papers that had been proofread and were free of spelling errors or incomplete sentences. We recognize that teams used AI platforms to help with translation and clarity, but teams should make sure their final papers use authentic language. Team #16442 (Example 34) wrote a very readable and organized paper throughout.

Tables & Figures

Effective visualizations, such as charts, graphs, and diagrams, play an important role in enhancing understanding and supporting key points within a report. When used thoughtfully, they can clarify complex relationships and make results more accessible to the reader.

However, visualizations should not stand alone. Each must be clearly explained within the body of the paper, with explicit discussion of what it shows, why it is included, and how it supports the team's analysis. Tables and figures should complement the written analysis rather than replace it, helping to communicate complex ideas more clearly and reinforcing the team's reasoning and findings.

For Problem A, the paper written by Team 16486 had so many exemplars for figures. The floor plan conversions in Figures 3, 5, and 7 (Example 35) were clear, showing side-by-side comparisons of architectural layouts and their corresponding node-graph representations with explicit node labels that make the abstraction process transparent and reproducible. Figure 6 and Figure 8 (Example 36) which were the optimal sweep routes for the school and house layouts, made effective use of color-coding, assigning distinct colors to each responder team's route, clarifying how the optimizer naturally partitions buildings into intuitive zones. Table 3 and Table 4 (Example 37) communicated the four-factor priority system effectively by showing both the raw hazard ratings and the resulting normalized weights, allowing the reader to understand why certain rooms received priority over others. Figure 9's (Example 38) dual presentation bar charts compared four strategies and line graphs showing sweep time versus responder count, provided compelling visual evidence of the MILP model superiority over random, greedy, and manual-split approaches. Furthermore, all numerical charts had clear and readable axis labels. For Problem B, there were so many teams that presented excellent tables and figures with descriptive labels and explanations embedded in the accompanying section's text. Judges wanted to highlight the way Team #16517 used (Example 39) tables to summarize information. For example, Table 20 conveys the comparison of the important factors between hosting a Superbowl and the Olympics.

Creativity

Creativity is an important element of strong modeling and may appear at any stage of the process. This creativity may be reflected in how a problem is framed, how key factors are identified, or how models are constructed and adapted. It can also appear in the way teams combine methods, incorporate real-world considerations, or communicate their results. Judges value originality when it is purposeful. It is not about unnecessary complexity, but about choices that improve the model and its applications. The most successful teams demonstrate creativity by approaching the problem with curiosity and a willingness to explore ideas beyond the most obvious solution paths.

Emergency evacuation sweep problems like Problem A are challenging and involve reconciling competing objectives under uncertainty. Unlike optimization problems with clear single objectives, these problems require consideration of factors such as: minimizing total clear-

ance time, prioritizing vulnerable occupants, balancing responder workloads, and accounting for unpredictable hazards like spreading fire, structural collapse, and occupant behavior under stress. Creative modeling approaches move beyond deterministic models to capture the tail-risk (unlikely) scenarios that are important in safety-critical applications. Creativity comes from recognizing which real-world complexities need to be explicitly modeled (such as dynamic path real-time re-routing or fatigue affecting responder speed) versus which can be abstracted away; and from designing computational frameworks that remain tractable while preserving the stochastic, time-dependent nature of actual emergencies.

Paper 16603 (Example 40) demonstrates creativity through an innovative "Optimize-then-Simulate" hybrid framework that combines deterministic optimization with stochastic simulation. Rather than treating the evacuation sweep problem as purely deterministic or probabilistic, the team separates the problem into three computational stages: K-Means clustering for zone assignment, TSP (Traveling Salesperson Problem) heuristics for path planning, and Monte Carlo simulation with pathfinding for uncertainty modeling (referred to as A* modeling in the paper). They recognize that while optimal routes can be pre-computed under ideal conditions, actual emergency sweeps require dynamic adaptation to blockages, smoke, and rescue needs. The model's capability to distinguish between fire scenarios (higher baseline speed, smoke penalties, zero blockages) and earthquake scenarios (lower speed, structural blockages, clear visibility) through straightforward parameter toggling demonstrates an elegant design for reusability across emergency types.

For Problem B, the judges expected certain types of models (such as AHP, EWM, PCA and TOPSIS) to be used and were impressed when teams employed unique strategies or model types. Team 16785 (Example 41) used a two-stage machine learning framework to create their model. Team 16911 (Example 42) created three machine learning models to determine the best city to host in 2029. While Team 16521 (Example 43) used PCA and they added a geographic factor to their ecological footprint index to incorporate regional differences. Additionally, they inventively used a Borda count to establish a final ranking.

The judges also enjoyed seeing teams use novel techniques in other steps of the modeling process. For example, Team 16947 (Example 44) explained steps a city could take to improve their ability to be chosen as a sustainable host but then went further and picked a city that had ranked in the middle and showed how the city's ranking would improve if they applied some of the suggested strategies. Team 16739 (Example 45) employed ARIMA to predict the characteristics of the cities in 2029 to plug into their model.

Citations, References, AI Usage Report

Teams must give credit where credit is due. It is okay to use existing models found elsewhere, but teams that do this but do not cite, lose the trust of the judges. Teams that use existing models must cite their source(s) within the paper at the point they present the model and include a reference citation in the back of the paper. This is also true for all graphs and tables taken from literature. All data, figures, graphs, and tables that come from outside sources require documentation at the point in the paper that they appear. Lack of documentation will result in a lower designation. Teams should use "in line" documentation with footnotes or endnotes to give proper credit to outside sources.

With the increasing availability of AI tools, COMAP recognizes that teams may use these resources to support their modeling process and the preparation of their submissions. Any use of AI must be treated as an external source and clearly documented in an AI Use Report. This documentation should provide transparency into how the tool was used by including the excerpts of prompts and outputs. Teams have used AI tools in a variety of ways, such as gathering or summarizing data, exploring coding approaches, or generating potential modeling strategies. When used thoughtfully, these tools can support and enhance the modeling process. However, reliance on AI without careful evaluation can weaken a submission. Like any resource, AI outputs should be approached critically. Teams are responsible for any AI-generated content they choose to incorporate. Thoughtful and responsible use of AI strengthens both the credibility and quality of a submission.

Final Thoughts

On behalf of COMAP and the contest judges, we would like to again thank all advisors and students for their participation in the HiMCM. Each year the quality and level of mathematics demonstrated by the high school student teams amazes and impresses our judges. We truly enjoy reading all solution papers. Successful teams use a wide variety and level of mathematics. While teams using post-high school/undergraduate level mathematics are in a league of their own if they understand and can explain their work, teams using basic high school mathematics and much simpler approaches are often among our Outstanding designees. We encourage students of all levels to compete in future HiMCM competitions as well as our ICM/MCM contests targeted to undergraduates. To be successful, read the comments and guidance provided in this article, see the TIPS article on the COMAP website, and visit: www.comap.org/mathmodels to review previous problems.

Follow us @COMAPMath on X, LinkedIn, Facebook and/or COMAPCHINAOFFICIAL on Weibo for information about all COMAP contests.

HiMCM Final Judge Contributing Authors

Karen Bryant
Thomas Fitzkee
Alton (Landy) Godbold
Katie Kavanagh
Bethany Kubik
Veena Mendiratta
Greta Mills
Kevin Murphy
Richard West

AWARDS

Outstanding
Finalist
Meritorious
Honorable Mention
Successful Participant

After final judging, HiMCM papers receive a designation in one of the categories above. Depending upon the quality of the papers, the top 20–25% of submitted papers receive a designation of Meritorious or above, with approximately the top 1% designated as Outstanding.



List of Examples

To view unabridged versions of the examples below visit:
www.comap.org/2025HimcmExamples

Team 16390 — Borden International Education School, China (Example 22)
Team 16442 — Shanghai Nord Anglia Chinese International School (Secondary), China (Examples 13, 25 & 34)
Team 16486 — BASIS International School Park Lane Harbour, China (Examples 1, 7, 10, 28 & 35 – 38)
Team 16517 — Longcheng High School, China (Example 39)
Team 16521 — Edu Plus Academy, CA, USA (Examples 5 & 43)
Team 16603 — Taipei American School, Taiwan China (Examples 4, 12, 33 & 40)
Team 16624 — Shanghai Linstitute School, China (Example 27)
Team 16691 — Shanghai Starriver Bilingual School, China (Examples 8, 26 & 31)
Team 16738 — Shanghai Linstitute School, China (Examples 14, 15 & 32)
Team 16739 — Guangdong Country Garden School, China (Examples 6, 16, 29 & 45)
Team 16785 — Jinan Foreign Language School, China (Example 41)
Team 16831 — High School Affiliated to Nanjing Normal University, China (Example 23)
Team 16843 — Nanjing Foreign Language School, China (Example 17)
Team 16911 — Xian Tieyi International Curriculum Center, China (Example 42)
Team 16947 — Shanghai Mingsui Creative School, China (Examples 20 & 44)
Team 16969 — Beijing National Day School, China (Example 3)
Team 17173 — Shanghai Mingsui Creative School, China (Examples 2 & 21)
Team 17205 — Chongqing Yucai Secondary School, China (Example 9)
Team 17401 — Horace Greeley High School, NY, USA (Example 30)
Team 17402 — Avenues School, Shenzhen, China (Examples 18 & 19)
Team 17460 — Moonshot Academy, China (Example 24)

COMAP, Inc.
175 Middlesex Turnpike
Suite 3B
Bedford, MA 01730
781-862-7878
FAX 781-863-1202
EMAIL: info@comap.org

HiMCM®/MidMCM



Compete. Solve. Create.

What is COMAP?

COMAP (Consortium for Mathematics and Its Applications) is a nonprofit that empowers students to tackle real-world problems through math modeling contests and creative learning experiences.

What is the contest?

COMAP's HiMCM®/MidMCM are global contests where students use math and teamwork to solive real-world problems and make an impact.

Ready to compete?

Don't miss your chance to solve real-world problems and compete with students from around the globe! Form your team and register now at: www.himcmcontest.org

@COMAPMath

thank you to our contest sponsors:

MAA MATHEMATICAL ASSOCIATION OF AMERICA
informs Smarter Decisions for a Better World
NCTM

www.contest.comap.org

2025 High School Mathematical Contest in Modeling (HiMCM)
Consortium 130: Example Appendix

16390 - Borden International Education School , China (Example 22)

16442 - Shanghai Nord Anglia Chinese International School (Secondary), China (Examples 13, 25 & 34)

16486 - BASIS International School Park Lane Harbour, China (Examples 1, 7, 10, 28 & 35-38)

16517 - Longcheng High School, China (Example 39)

16521 - Edu Plus Academy, CA, USA (Examples 5 & 43)

16603 - Taipei American School, Taiwan China (Examples 4, 12, 33 & 40)

16624 - Shanghai Linstitute School, China (Example 27)

16691 - Shanghai Starriver Bilingual School, China (Examples 8, 26 & 31)

16738 - Shanghai Linstitute School, China (Examples 14, 15 & 32)

16739 - Guangdong Country Garden School, China (Examples 6, 16, 29 & 45)

16785 - Jinan Foreign Language School, China (Example 41)

16831 - High School Affiliated to Nanjing Normal University, China (Example 23)

16843 - Nanjing Foreign Language School, China (Example 17)

16911 - Xian Tieyi International Curriculum Center, China (Example 42)

16947 - Shanghai Mingsui Creative School, China (Examples 20 & 44)

16969 - Beijing National Day School, China (Example 3)

17173 - Shanghai Mingsui Creative School, China (Examples 2 & 21)

17205 - Chongqing Yucai Secondary School, China (Example 9)

17401 - Horace Greeley High School, NY, USA (Example 30)

17402 - Avenues School, Shenzhen, China (Examples 18 & 19)

17460 - Moonshot Academy, China (Example 24)

Problem Chosen	2025 HiMCM Summary Sheet	Team Control Number
A		16486

Building sweep is the process that responders verifying each area in a building is clear of occupants during emergencies [1] such as fires or gas leaks. This process is complex due to the variation of buildings' size, layout, occupants' distribution. Smoke, blocked corridors, and limited communication further complicate search operations. Our goal is to build a quantitative model that assigns responder teams to routes which clear all required rooms quickly and safely, while accounting for room priorities and possible redundancy.

To mathematically model building sweep problem, we convert each building layout to a node-edge graph: rooms, corridor junctions, stair landings and exits are nodes; walkable connections are edges with traversal times. Required rooms are evaluated on sweep times and priority weights through 0–3 rating scores in four dimensions of room function, hazard proximity, occupant vulnerability, and lingering tendency. A mixed-integer linear program (MILP) [2] is designed to solve our model, which decides each team's route and the time each room becomes "all clear", to minimize a combination of sweep finishing time, total traversal time, and a priority-weighted summation of room clearance times.

Our model framework is general that is able to take any floor plan and team number into considerations and return optimal sweep routes. We apply it to three building layouts: a basic one-storey office, a five-floor school, and a single-family house. In both the school and house, the optimizer naturally divides the layout into intuitive "zones" (upper / middle / lower floors, or bedroom wing / living-basement / kitchen–garage) and assigns each zone to one team, yielding simple, non-overlapping paths that match how incident commanders already think.

Our experiments provide quantitative evidence for the effectiveness of our model. On the school layout, we compare four strategies—random assignment, greedy nearest-room, manual floor splits, and our MILP strategy. For the same number of responders, the MILP strategy consistently achieves the shortest sweep time, especially when only one or two responders are available. A second experiment uses the single house to study the effect of priority weights: with moderate weights, high-risk rooms are cleared noticeably earlier with almost no change in total sweep time; with stronger "peak-hour" weights, their average clearance time improves by about 40% at the cost of a longer overall sweep.

Our main contributions and innovations are:

1. A unified graph-based MILP sweep framework and algorithm that turn raw floor plans into optimal multi-team routes with built-in redundancy.

2. An interpretable four-factor room-priority system and soft-deadline extension that make the question "who should we save first?" explicit and tunable for decision-makers.

3. Experiments on realistic layouts showing that our MILP strategy outperforms random, greedy, and manual rules and identifies a practical "sweet spot" of about three responders for medium buildings, together with extensions for time-dependent hazards and technology support.

Keywords: building sweep; evacuation planning; mixed-integer linear programming; responder routing; graph model; room-priority weights; time-dependent hazards; technology-supported search.

Problem Chosen	2024 HiMCM/MidMCM Summary Sheet	Team Control Number
B		#17173

Abstract

Large-scale sporting events generate substantial economic and social value, yet also impose significant environmental burdens through energy use, transportation, waste production, and resource demand. A balanced assessment model must incorporate multiple sustainability factors to evaluate these impacts rigorously. In this paper, our team develops a quantitative decision model to evaluate the environmental sustainability of potential Super Bowl host cities, determine the most suitable future host location, and extend the model to broader global mega-events such as the World Cup.

First, we identified six major dimensions and twenty-four factors of environmental impact based on **greenhouse gas emissions, clean energy and energy efficiency, waste and recycling, water resource management, pollution control, and biodiversity and nature conservation**. Hybrid weighting is then established through the integration of **Analytic Hierarchy Process** (AHP) and **Entropy Weight Method** (EWM) to balance subjective judgment with objective data variance. These weights are applied to **Grey Relational Analysis** (GRA) and **Fuzzy Comprehensive Evaluation** (FCE) to obtain dual sustainability rankings. A temporal fairness adjustment using **Principal Component Analysis** (PCA) and **time-series decomposition** is introduced to remove technology-era bias and ensure equitable cross-decade comparison. After establishing adjusted and unadjusted rankings, **New Orleans, Louisiana**, is found to be the most sustainable host city for the 2029 Super Bowl.

Then, the model is expanded to multi-venue international competitions, where transportation emission between different venues becomes prominent. By integrating sustainability scores and inter-venue travel distance into **TOPSIS**, we identify the environmentally optimal twelve-stadium configuration for the 2030 World Cup across Spain, Portugal, and Morocco. Finally, short-term operational strategies and long-term infrastructure development pathways are proposed to support future host city bids, and key distinctions between single-venue and multi-venue events are analyzed.

We believe that the model is robust and highly generalizable to large-scale sporting event, and offers a scalable decision mechanism to guide environmentally-aware host selection.

Keywords: AHP, EWM, GRA, FCE, Time-Series Decomposition, TOPSIS, PCA, Superbowl, World Cup, Sustainability

1.2 Literature Review

Xiao (2025) investigates optimization of emergency evacuation strategies and fire-risk control in vocational colleges, emphasizing institution-specific procedures and targeted mitigation measures for campus environments [1]. Yu (2025) applies simulation to metro-station evacuation, exploring path-design and optimization techniques tailored to confined, high-density transit spaces [2]. Guo (2024) examines evacuation and safety-facility optimization in highway service areas, highlighting the need to adapt evacuation planning to roadside facility layouts and vehicular contexts [3]. Ling et al. (2023) study library evacuation with attention to pedestrian spatial requirements, demonstrating how room function and occupant flow characteristics should inform evacuation strategies in academic settings [4]. Zhang et al. (2023) address evacuation approaches for super high-rise buildings during sudden incidents, focusing on vertical circulation, staging, and special procedural considerations for ultra-tall environments [5]. Xie et al. (2022) develop a pedestrian evacuation model that incorporates panic psychology, showing how emotional and cognitive factors alter evacuation dynamics and influence model design [6]. Gao and Jia (2021) propose kindergarten activity-unit design strategies from an evacuation perspective, underlining the importance of spatial layout choices for vulnerable populations [7]. Yu et al. (2021) analyze how metro door-control strategies affect platform evacuation efficiency, evidencing operational controls as a key lever for improving egress performance in rail systems [8]. Huo et al. (2020) investigate staircase evacuation in primary-school teaching buildings with consideration of children's movement characteristics, which necessitates tailored stair-use protocols and spacing designs [9]. Huang et al. (2019) discuss hospital personnel evacuation under fire scenarios, emphasizing complexities introduced by immobile or assisted occupants and the need for medical-aware evacuation plans [10].

Wang (2019) applies maximum-flow concepts to optimize reversible lane settings, demonstrating that traffic-level flow models can support emergency routing decisions on road networks [11]. Zhao et al. (2018) explore strategies for constructing urban comprehensive safety systems, linking building-level evacuation needs to broader urban resilience and governance considerations [12]. Han and Liu (2018) present a path-selection model based on evacuation path sets and apply it in crowd-evacuation simulation, contributing an algorithmic approach to routing under congestion [13]. Cai (2018) discusses urban design strategies for enhancing spatial disaster prevention safety at the block level, offering spatiotemporal perspectives on refuge and evacuation planning in urban design [14]. Zhang (2018) provides an overview of safety evacuation in metro stations, synthesizing operational, structural, and managerial factors pertinent to transit-station egress [15]. Peng and Shao (2016) study fire-evacuation strategies for high-rise residential buildings, with attention to occupant behavior and vertical egress limitations [16]. Xi et al. (2015) analyze personnel evacuation strategies for university teaching buildings, reinforcing that instructional building function and scheduling influence evacuation patterns and protocols [17]. Li et al. (2015) propose an emergency-evacuation framework for rail-transit hubs based on immune co-evolution, illustrating adaptive, co-evolutionary algorithmic frameworks for large multimodal nodes [18]. Xiang et al. (2014) analyze expressway emergency-evacuation strategies without demand prediction, highlighting tactics for rapid clearance under uncertain traffic demand [19]. Finally, Wang, He, and Xiang (2014) study optimal staircase evacuation strategies in high-rise buildings, focusing on stair use policies and design parameters critical to vertical egress efficiency [20].

Collectively, these works span facility-specific optimization, occupant-centered behavioral modeling, traffic-level interventions, and algorithmic/simulation methods. They establish a comprehensive foundation, ranging from micro-scale room and stair dynamics to macro-scale urban resilience that the present study builds upon. Notably, while many studies address evacuation timing, spatial design, and occupant behavior across contexts, relatively few explicitly model coordinated responder-centric room-by-room sweep strategies under operational constraints such as communication failure, reduced visibility, and prioritized room checking; this gap motivates the responder-focused sweep model developed in this paper, the work flow is shown in Figure.2.

1.2 Detailed Problem Restatement and Interpretation

To declare "all clear" for emergency evacuations, several safety and procedural conditions must be satisfied:

- Responders must visually confirm that no persons remain in each room.
- They must perform an audible check, calling out and reporting "clear" on their radios, especially when visibility is poor.
- They must conduct a physical inspection of all subareas in a room (corners, behind furniture, under desks when appropriate).
- They must evaluate residual hazards, including smoke, gas, or structural debris that could still endanger safety.

Only after confirming both human and environmental safety can the area be considered truly "all clear." The room is then clearly marked and documented over radio to prevent re-entry and redundant searches.

When responders encounter remaining individuals, those individuals must be assisted or carried to safety and transferred to medical personnel. Natural randomness in disasters is inevitable. A path that appears optimal in advance may suddenly be blocked by debris or smoke. Responders then attempt to navigate around blocked paths, and in worst cases, may be delayed enough that some occupants cannot be reached in time.

We assume all first responders entering the building are equally healthy, fast, and strong, having been rigorously trained under standard procedure. Other personnel remain on the outer perimeter, supporting triage and treatment once victims are brought outside.

The nature of occupants also plays a critical role in sweep dynamics. Key factors include occupant weight, the number of occupants per room, and the fraction of children or elderly. Facilities with children, such as daycares or schools, require responders to search every corner carefully and assist young individuals who may be hiding. To avoid misinterpreting injuries and needs, responders assist all remaining occupants. Each setting demands a tailored approach that accounts for room layout, hallway distances, occupant behavior, and building density.

The type of emergency additionally influences sweep operations. Fires, earthquakes, and other hazards create distinct conditions that affect both responder and occupant behavior. In fires, heat, flames, and smoke must be avoided, often forcing responders to prioritize accessible rooms first. Earthquakes can cause structural damage, block exits, and create unstable environments, requiring sweeps that account for debris and alternative escape routes. Each hazard modifies occupant evacuation speed, choice of escape routes, and responder tactics. Our model therefore explicitly distinguishes between fire and earthquake scenarios.

2 Problem Analysis

In order to understand how a major sporting event generates environmental pressure, we first reviewed foundational works on the environmental assessment of mega-events, including Collins et al. [6] and Pourpakdelfekr & Oboudi [21]. These studies emphasize that large events such as the Olympics, FIFA World Cup, and national championships impose substantial ecological burdens through energy consumption, transportation, water use, and waste. They also highlight that these impacts are unevenly distributed across locations, making environmental sustainability inherently dependent on host-city characteristics. Motivated by these insights, we adopt an ecological footprint view with a strong emphasis on carbon dioxide–equivalent emissions as a quantitative metric for comparing potential host cities.

Key Environmental Impact Areas

Across the literature, several impact domains consistently appear as the primary drivers of an event's environmental footprint. Energy consumption is one of the most critical components, particularly when combined with the carbon intensity of the host region's power grid. Transportation emissions—largely driven by long-distance travel by spectators, teams, and staff—are repeatedly identified as the dominant contributor to event-level greenhouse-gas emissions [27]. Water usage and water scarcity concerns also play an important role, especially in arid regions or stadiums with high irrigation demands. Solid-waste generation, including food waste and packaging, further contributes to the event's environmental burden, as noted in Pourpakdelfekr & Oboudi's review [21]. These four domains create a comprehensive structure for assessing environmental sustainability consistently and comparably across cities.

Location-Dependent Variability

A key theme emerging from prior research is that environmental impact varies sharply by location. As Collins et al. [6] argue, the sustainability profile of a host city is shaped by several geographic and infrastructural factors. Differences in electricity generation (such as the share of renewable energy), local climate, building energy demand, and regional transit systems all influence the environmental burden of a large event. For example, cities with cleaner power grids inherently generate fewer operational emissions, while regions with cold climates may require less cooling energy. Public transit availability affects the extent of car-based travel, and waste-diversion systems determine how much event waste is recycled rather than ending up in landfill. These location-specific characteristics cannot be standardized or offset by event organizers in the short term and must be incorporated into any model that seeks to evaluate environmental sustainability across multiple cities.

Relevance to Host City Decision

Given the multidimensional nature of environmental impacts and their sensitivity to local conditions, selecting a host city requires a modeling framework that integrates both direct ecological burdens and contextual geographic factors. The ecological footprint approach provides a transparent quantitative structure for converting event activities into carbon-related measures of resource use, while prior research highlights the necessity of evaluating travel patterns, energy sources, water availability, and waste systems [6, 21]. By synthesizing these insights into a unified assessment framework, we can mean-ingfully differentiate cities based on their expected environmental performance and provide a defensible recommendation for the most sustainable hosting location.

7 Comparative Analysis of Model Results

By comparing the results from Sections 3–4 with those from Section 5, several structural differences emerge between single-event mega-events such as the Super Bowl and large-scale, multi-venue or multi-sport competitions such as the Summer Olympic Games. The key differentiating mechanisms can be summarized as follows:

(1) Temporal Structure: Short-Term Peaks vs. Long-Horizon Accumulation

In the Super Bowl model (**Sections 3–4**), the entropy-weighted results assign the highest weights to stadium water use, per-capita water consumption, spectator travel emissions, and transportation-related emissions, followed by electricity consumption for stadiums and hotels. Construction-phase and long-term legacy indicators contribute almost nothing to the weighting structure. This pattern indicates that the environmental pressure of a single-event competition is dominated by short-duration, peak-load consumption concentrated around game day. In contrast, in the Olympic model, indicators such as Olympic Village electricity intensity and construction/demolition waste obtain significantly higher entropy weights. This shift reflects that multi-venue events generate cumulative, time-extended sustainability pressure, where long-horizon resource use becomes more influential than short-term op-erational spikes.

(2) Spatial Structure: Single-Venue Events vs. Multi-Venue Networks

In the Super Bowl model, spatially relevant factors are primarily: **Average spectator travel distance, Public transport coverage, Share of green transportation.** Regression results show that these variables strongly explain spectator and transportation-related carbon emissions. In the Olympic model, newly added variables—average inter-venue travel distance, in-Games venue reuse rate, and post-Games reutilization rate—form a distinct principal component (PC3), interpreted as the "venue spatial configuration and legacy utilization" dimension. This component contributes meaningfully to multiple environmental indicators. This comparison highlights that for comprehensive multi-sport events, the spatial network of venues, together with the multi-tier mobility system linking venues and the Olympic Village, becomes a structural driver of environmental outcomes.

(3) Shared Core Drivers: Electricity Systems and Transport Structure

Across both models, a consistent result emerges: **Electricity carbon intensity, Renewable-energy share, Public transport coverage, and Green mobility proportion.**All exhibit high importance in both the principal component regressions and the entropy weighting scheme. These variables function as common core determinants of environmental performance. However, the mechanisms differ. For the Super Bowl, these city attributes primarily amplify or mitigate single-day carbon footprints via short-term electricity demand and transportation organization. For the Olympic Games, they determine not only operational emissions during the 17-day event window but also the long-term emission baseline for decades of venue and Olympic Village operation.

2 Assumptions and Justification

Table 1: Model Assumptions

#	Assumption	Justification
1	Responders have the building layout plan in advance and clearly know the structure such as the location of the rooms and the length of the corridors.	In almost all jurisdictions, fire codes and public safety regulations require building owners to submit detailed architectural plans to local fire departments and emergency planning agencies. As a result, it has become standard practice for rescue personnel to immediately access accurate floor plans, stairwell locations, door widths, and other critical information before entering a building.
2	Responders maintain a constant speed of movement as long as roads are clear; we do not factor in fatigue.	Typical structural sweeping during fires is completed within 5 to 15 minutes. Physiological studies indicate that when using modern breathing apparatus and protective gear, a noticeable decline in speed due to fatigue only occurs after 15 to 20 minutes of high-intensity work.
3	Responders strictly execute their assigned routes and do not deviate to conduct unauthorized rescue operations.	In professional firefighting doctrine, primary search is a coordinated team operation under the unified command of the incident commander. If individual responders are permitted to break formation, it can result in some areas being searched repeatedly while others are completely overlooked, significantly increasing the likelihood of missing occupants.
4	Occupants remain stationary, awaiting rescue in their respective rooms.	Fire safety studies confirm that untrained individuals generally remain stationary during poor visibility. Simulating their movement would add unnecessary complexity without improving search operation accuracy, making the stationary model both practical and realistic.
5	Communication between responders is instantaneous in the base model.	The base model is intended to solve the pure routing and allocation problem. Perfect radio communication represents the ideal condition that incident commanders strive to maintain in emergency response operations. Communication delays, channel overload, and radio dead zones are real phenomena, but they are treated separately in Section 6 (Model Extensions).
6	Fire spread and smoke obstruction are not considered in the base model.	The primary objective of the base model is to determine optimal static sweep routes. Introducing time-varying edge weights from fire spread would convert the problem into a complex dynamic network flow. By fixing hazard conditions, we obtain a clean lower bound on sweep time; dynamic hazards are explicitly incorporated and analyzed in Section 6.
7	The responders' path between rooms does not account for corridor width.	In most buildings, corridor widths exceed 1.1 meters—the minimum requirement for evacuation routes under the International Fire Code. At standard search speeds, the time lost to single-file movement or brief queuing at doorways is negligible compared to room search time and the total travel distance along main corridors.

(Section 2).

1.3.1 Overarching Assumptions

- **Assumption 1 (Pre-informed Building Layout):** We assume that the firefighting squad has complete information on the layout of the building. **Justification:** Evidence shows that the sweep team is given the floor plan of the building that they are going to search before their operation [2]. This way, our model can find an optimal path directly based off of a concrete floor plan, instead of opting for a dynamic model, where information about the building is slowly uncovered. This assumption makes sense because assuming that firefighters have the building plan before the actual operation will greatly increase their chances of a successful sweep, which then leads to an increase in their chance of saving more lives (referenced in 2.1).

- **Assumption 2 (Firefighter Movement Speed):** We assume that firefighters move at 1.4 m/s during a sweep operation when they are at maximum stamina while wearing full firefighting gear [3].

- **Assumption 3 (Firefighter Traversal and Search Rules):** We assume that firefighters in a sweep operation always take the shortest path available from room to room without any detours. We also assume that all rooms have to be visited at least once by the firefighters. **Justification:** First, if a firefighter made up their mind to travel from one room to another, they are going to do it directly without a diversion in their path. Second, every room must be visited at least once because it fulfills the SAR requirement for a standard sweep operation. Thus, our interpretation of the model necessitates the thorough search of every room in the building, and we assume that firefighters travel straight from one room to another (referenced in 2.1).

- **Assumption 4 (Occupant Movement Rules):** We assume that, when evacuating an occupant from a room, the firefighter will instruct them the safest way to the nearest exit based on the results from the sweep. The occupant will then walk that path on their own without a firefighter's escort. **Justification:** We made this assumption because it simplifies firefighter movement rules, as we only have to consider the optimal path for the firefighter to visit each room in their part of the sweep; considering occupant escorting will over-complicate our model (referenced in 2.1).

- **Assumption 5 (Crowded Hallway):** We assume that we do not have to consider the problem of a narrow hallway, and the possibility of firefighters crowding up in the same part of the hallway at the same time of their sweep as they traverse from one room to another. **Justification:** We disregard this possibility because it will over-complicate our routing optimization. If two or more firefighters do end up in the same part of a hallway at the same time, they can sort the problem of how to walk past each other out on the spot without taking long. So, this problem overcomplicates our model, and it also has little impact on the final result (referenced in 2.1).

- **Assumption 6 (Defining Sweep Time):** We define the total time of a sweep operation as the time taken by the firefighter who has the slowest route. **Justification:** This is because the sweep operation ends when all firefighters leave the building through an exit. Therefore, the time taken by the last firefighter to complete their sweep route is the total time for the sweep (referenced in 2).

2.3 Key Behavioral and Data Assumptions

(1) Steady attendance and local proportion.

The total in-person attendance, combined with the proportion of local spectators vs. non-local spectators are treated as the same across cities. Justification. Capacity and demand for the Super Bowl are similar for qualified venues; fixing these isolates the effect of location and infrastructure.

(2) A stable origin distribution.

We assume a regional distribution of non-local spectators similar to recent Super Bowls, and independent of host city. Justification. Attendance is dominated by a relatively stable national fan and sponsor base. It will be fair to compare using a common distribution.

(3) Simplified long distance mode of choice.

Shares of travel mode are defined by distance bands (short: mostly car/rail; medium: mixed; long: mainly air) with fixed emission factors for each mode. Justification. Captures the primary drivers of travel emissions without individual-level data requirement.

(4) Local mobility based on local mobility in a city.

We assume that the travel distance per spectator by locality per crowd has similar proportions between the cities, and the proportion of public transit vs private vehicles for city residents with respect to it is city-specific. Justification. Urban scale is comparably sized across candidate metros; variations are largely driven by transit systems, which our model is based on.

(5) Standardization of stay, Hotels and per capita consumption.

Non-local spectators are considered to stay a common number of nights, and per room-night emissions along with per capita waste/water production are assumed to have corresponding, the treatment results being contingent on the grid mix, waste management, and water use of the individual cities. Justification. Detailed micro-data are not available; unified baselines including intensity factors linked to the individual city are used to maintain comparability.

4 Graph-based Sweep Model

4.1 From building layouts to node graph

To mathematically model the problem, we represent the layout of a building as a graph $G = (V, E)$. Nodes V are discrete locations such as room interiors, hallway, stair landings, and entrances or exits. Edges $(i, j) \in E$ are feasible movements between nodes, for example hallways, and stair flights. Let $R \subseteq V$ be the set of rooms that must be swept. For each room $i \in R$ we have

- a sweep time $s_i > 0$, the time for a team to fully search room i;

- a risk or priority weight $w_i > 0$, indicating how costly it is to leave room i uncleared for a long time.

For each edge $(i, j) \in E$ we define a traverse time $t_{ij} > 0$, derived from the geometric distance, stair difficulty and local conditions (smoke, obstacles). In this way t_{ij} is an edge-based quantity, while w_i is a node-based quantity.

We consider a set of responder teams $K = \{1, \ldots, |K|\}$. Team $k \in K$ starts from an exit node.

From this graph-based description, the sweep-planning problem becomes a combinatorial routing and assignment task on G: we must decide which edges each team traverses, which rooms each team is responsible for, and in what order rooms are cleared, while respecting timing and redundancy requirements. These decisions are naturally modeled by binary variables (on edge usage and room assignments) coupled with continuous variables (for arrival and clearance times). This structure makes a mixed-integer linear programming (MILP) [7]formulation a natural choice: it can capture network flow, timing consistency, and room-coverage constraints in a linear way, while allowing us to compute globally optimal sweep plans for the layouts we study.

Basic office scenario. Figure 3 illustrates the conversion for the basic six–office hallway. The left panel shows the floor plan with six identical rooms R_1–R_6, a central hallway H, and two exits E_1 and E_2. The right panel shows the resulting node graph.

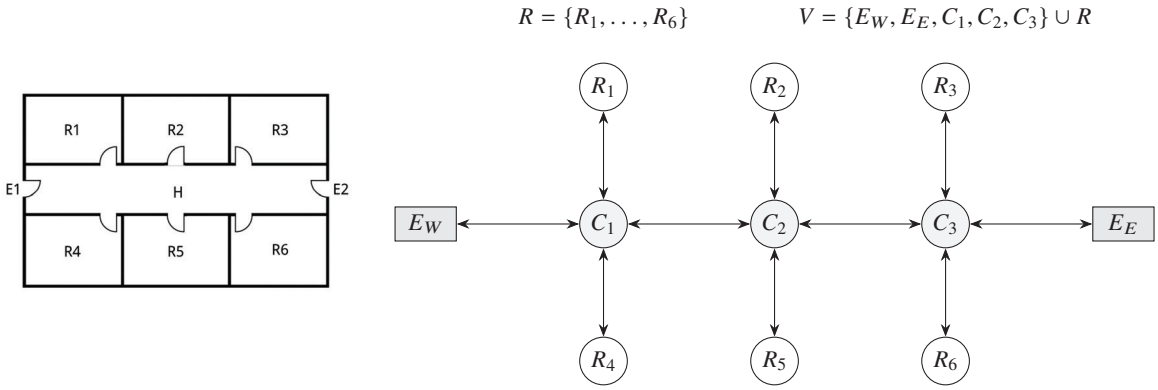


Figure 3: Conversion from floor plan to node graph for the basic office scenario. Left: physical layout with six offices R_1–R_6 along a hallway H and exits E_1 and E_2. Right: corresponding node graph with room nodes Rx, hallway nodes Cx, and exit nodes Ex.

The conversion follows three simple rules:

- Each room interior (here R_1–R_6) becomes a *room node* in V.
- Distinct positions along circulation spaces (the hallway) and at exits become *hallway nodes*; in this layout we use one hallway node in front of each pair of doors and one node at each exit.
- We add an undirected edge (represented as two directed edges in E) between any two nodes that a responder can walk between without passing through another door or junction: along the hallway between neighbouring hallway nodes, and between each room node and its adjacent hallway node.

All traverse times t_{ij} in this basic scenario are then derived from geometric distances between these nodes (e.g. 5 m between adjacent hallway nodes at 1 m/s gives $t_{ij} = 5$ s).

4.2 Decision variables

The model decides which routes teams take, which rooms each team is responsible for, and when rooms are cleared. We use three types of decision variables.

$$\textbf{Routing variables:} \quad x_{ij}^k = \begin{cases} 1, & \text{if team } k \text{ traverses edge } (i,j), \\ 0, & \text{otherwise,} \end{cases} \qquad (i,j) \in E,\ k \in K,$$

encode each team's path through the graph.

$$\textbf{Assignment variables:} \quad y_i^k = \begin{cases} 1, & \text{if room } i \text{ is assigned to team } k, \\ 0, & \text{otherwise,} \end{cases} \qquad i \in R,\ k \in K,$$

indicate which team is responsible for sweeping each room.

Finally, time variables describe when teams arrive and when rooms become all clear. Let $A_i^k \geq 0$ be the arrival time of team k at node i, $C_i \geq 0$ the clearance time of room i (the time by which a responsible team has completed its sweep), and $T_{\text{finish}} \geq 0$ the time at which the whole building is declared all clear.

4.3 Constraints and physical interpretation

Coverage and redundancy. Every required room must be assigned to at least one team. In the simplest case without explicit redundancy this is

$$\sum_{k \in K} y_i^k \ \geq\ 1, \qquad i \in R. \tag{1}$$

If room i must be checked independently by at least $r_i \geq 1$ teams, we replace the right-hand side by r_i.

Assignment implies visit. If a room is assigned to a team, that team must actually enter the room at least once. Let $\delta^-(i)$ be the set of predecessors of node i. Then

$$\sum_{j \in \delta^-(i)} x_{ji}^k \ \geq\ y_i^k, \qquad i \in R,\ k \in K. \tag{2}$$

The left-hand side counts how many times team k uses an edge that ends at node i.

Time consistency along edges. To prevent teams from "teleporting" we link arrival times to traverse times. If team k walks along edge (i,j), it must reach j at least t_{ij} units of time after reaching i. Using a standard big-M formulation,

$$A_j^k \ \geq\ A_i^k + t_{ij} - M\,(1 - x_{ij}^k), \qquad (i,j) \in E,\ k \in K, \tag{3}$$

where M is a large constant. When $x_{ij}^k = 1$, this reduces to $A_j^k \geq A_i^k + t_{ij}$; when $x_{ij}^k = 0$, the constraint is inactive.

Time needed to clear rooms. If team k is responsible for room i, the clearance time must be at least its arrival time plus the sweep time:

$$C_i \ \geq\ A_i^k + s_i - M\,(1 - y_i^k), \qquad i \in R,\ k \in K. \tag{4}$$

All-clear time. The building is declared all clear only when every required room is clear:

$$T_{\text{finish}} \;\geq\; C_i, \qquad i \in R. \tag{5}$$

Redundant verification . we specify a required check count $r_i \geq 1$ and impose

$$\sum_k y_{ki} \;\geq\; r_i. \tag{6}$$

When $r_i = 1$, a single sweep suffices; when $r_i = 2$, two independent teams must verify the room. This allows the same model to handle low-risk rooms and critical spaces with stricter sweep policies.

Flow conservation. To ensure that each team's x_{ij}^k variables describe a single continuous route starting from s_k, we impose standard flow-conservation constraints. Intuitively, these constraints say that, for every intermediate node and every team, the number of edges used to enter the node equals the number of edges used to leave it, so a team cannot "teleport" or break its path. At the start node s_k we allow one extra departure so that the route can begin there. The resulting constraints follow the usual vehicle-routing formulations.

4.4 Objective function

The sweep plan should finish early (for the goal of clearance in the shortest possible time) [8], avoid unnecessary walking(for the goal of responder efficiency), and clear high-risk rooms quickly(for the goal of prioritizing occupant safety). We capture these priorities with the linear objective [9].

$$\min \left\{ T_{\text{finish}} + \lambda_1 \sum_{k \in K} \sum_{(i,j) \in E} t_{ij} x_{ij}^k + \lambda_2 \sum_{i \in R} w_i C_i \right\}, \tag{7}$$

where $\lambda_1, \lambda_2 \geq 0$ are policy parameters.

The first term is the overall completion time; minimising it pushes the operation to end as early as possible. The second term is total traverse time over all teams and all edges used, discouraging long or inefficient routes. The third term is a risk-weighted sum of room clearance times: a room with larger w_i contributes more if it is cleared late, so the optimiser is encouraged to visit high-risk rooms earlier.

Together, constraints (1)–(5), the domain conditions $x_{ij}^k, y_i^k \in \{0, 1\}$, $A_i^k, C_i, T_{\text{finish}} \geq 0$, and objective (7) define a mixed-integer linear programme (MILP) for the sweep problem.

4.5 Evaluation of room priority weights

To evaluate a room's priority weight w_i, we take into considerations four factors: room function, hazard proximity, occupant vulnerability and lingering tendency. For Room i, they are denoted by F_i, H_i, O_i, L_i respectively. Each factor is evaluated using a 4-point ordinal rating scale (0–3), *i.e.*, we have $F_i, H_i, O_i, L_i \in \{0, 1, 2, 3\}$.

For each room $i \in R$ we evaluate integer ratings of $F_i, H_i, O_i, L_i \in \{0, 1, 2, 3\}$ as follows:

The four ratings are combined into a raw priority

$$w_i^{\text{raw}} = \alpha_F F_i + \alpha_H H_i + \alpha_O O_i + \alpha_L L_i, \qquad i \in R, \tag{8}$$

Category	Description and Level Definitions
Room function F_i	0: Storage/utility rooms
	1: Small private rooms (office, bedroom, bathroom)
	2: Main living/teaching spaces (classroom, living room, kitchen)
	3: Specialised spaces for very young children (nursery)
Hazard proximity H_i	3: Co-located or adjacent to hazard source
	2: Same floor, nearby
	1: Moderate distance or on a different floor
	0: Remote from hazard source
Occupant vulnerability O_i	1: Healthy adults
	2: Sleeping or less mobile occupants
	3: Infants/toddlers (nursery)
Lingering tendency L_i	1: Large open spaces (classroom, living room, kitchen)
	2: Private rooms (bedroom, bathroom)
	3: Small hiding places (closet, nursery, basement)

where $\alpha_F, \alpha_H, \alpha_O, \alpha_L$ are non-negative coefficients for different factors summing to one. As protecting vulnerable occupants and avoiding rapidly deteriorating areas are most important, we choose

$$\alpha_O = 0.40, \quad \alpha_H = 0.30, \quad \alpha_F = 0.20, \quad \alpha_L = 0.10, \tag{9}$$

so that O_i contributes the most, followed by H_i.

To place the weights on a convenient common scale we linearly rescale w_i^{raw} into range $[1, 5]$ as following:

$$w_i = 1 + 4\,\frac{w_i^{\text{raw}} - \min_{j\in R} w_j^{\text{raw}}}{\max_{j\in R} w_j^{\text{raw}} - \min_{j\in R} w_j^{\text{raw}}}, \qquad i \in R. \tag{10}$$

Rooms with the smallest (largest) raw score obtain $w_i = 1$ ($w_i = 5$). In the basic office scenario all six rooms are identical, so they share one common weight.

4.6 Application to the basic office layout

For the one-storey office building in Figure 3 we have six identical offices and two exits connected by a straight hallway. All offices have the same sweep time $s_i = s_{\text{office}}$ and a common moderate weight $w_i = w_{\text{office}}$. Hallway segments between adjacent doors are usually 5 m apart, so with a nominal walking speed of 1 m/s each has $t_{ij} = 5$ s; door edges are assigned a short traversal time of 1 s. We assume each office sweep requires 15 s

Because all rooms are identical in both geometry and risk, the objective (7) essentially reduces to minimising T_{finish} and total walking time. Solving the MILP with two responder teams starting from the two exits yields a symmetric plan: each team sweeps the three rooms closest to its entrance, and the optimal finishing time is $T_{\text{finish}} = 60$ s under the assumed parameters. We use this scenario mainly to validate the model before applying it to more complex layouts.

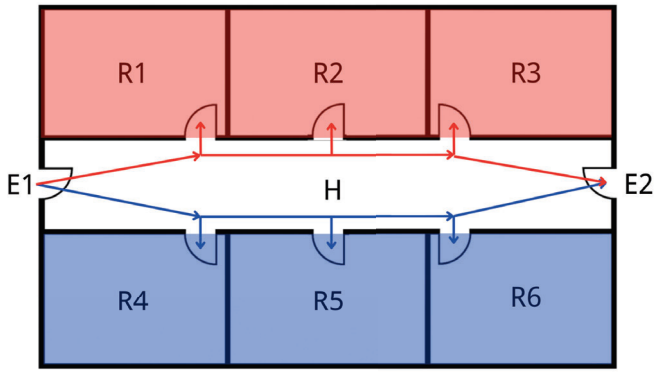


Figure 4: Sweep routes for the basic office layout, showing how responders move from exits E_1 and E_2 through the hallway H to clear rooms R_1–R_6.

6.1 Comparison of sweep strategies on the school layout

To demonstrate the benefit of optimization over certain sweep rules, we evaluate four strategies on the school layout shown in Figure 5. The building has four floors, each with a central hallway and five classrooms, connected vertically by a single stairwell. All responders start from the ground-floor entrance (node S1). Walking time is proportional to corridor distance and each classroom requires a fixed service time of one unit once reached.

We compare four strategies:

- **RND** (random): randomly permute rooms and split them evenly across responders.
- **GRD** (greedy): at each step send a free responder to the nearest uncleared room.
- **SPLIT** (floor split): manually assign contiguous floors to responders (e.g. for $K = 2$, lower vs. upper floors).
- **MILP** (proposed): the exact model from Section 4, minimising weighted clearance time $\sum_i w_i T_i$.

Table 5 reports total sweep time $T_{\max}$ for different team sizes ($K = 1$–4). The RND and GRD values are averaged over 2,000 simulation runs; SPLIT and MILP are deterministic.

Table 5: Total sweep time $T_{\max}$ under different strategies (school layout).

Strategy	$K = 1$	$K = 2$	$K = 3$	$K = 4$
RND (avg)	310	198	150	130
GRD	260	170	130	110
SPLIT	240	160	120	100
MILP (ours)	**225**	**140**	**100**	**85**

Figures 9a–9b visualise the same data. MILP consistently achieves the lowest sweep time, especially when $K \leq 3$. The improvement is particularly large for the single-responder case (27% faster than GRD and 38% faster than RND), because the optimal model prioritises travel-efficient traversal of each floor and avoids repeated stair usage.

Overall, this experiment shows that while simple heuristics (RND, GRD, SPLIT) reduce sweep time to some extent, they cannot account for the global structure of the building nor coordinate floor transitions efficiently. The proposed MILP model produces consistently faster strategies even for small team sizes, validating its usefulness for real-world sweep planning.

3.4 Clustering Algorithm

Our greedy workload-balanced clustering algorithm assigns rooms to responders based on current workload and proximity:

1. Initialize responder positions at distinct exits
2. Set workload $W_i = 0$ for all responders i
3. While unassigned rooms remain:
   (a) Select responder $i^* = \arg\min_i W_i$ with minimum workload
   (b) Among unassigned rooms, select $r^* = \arg\min_r d(r, \text{pos}_i)$ nearest to responder i^*
   (c) Assign room r^* to responder i^*
   (d) Update workload: $W_{i^*} = W_{i^*} + s_{r^*} + d(r^*, \text{pos}_{i^*})$
   (e) Update position: $\text{pos}_{i^*} = r^*$

The distance $d(r_1, r_2)$ is the shortest path travel time between nodes, computed using Dijkstra's algorithm on the building graph [3].

This greedy approach has two key properties: (1) it maintains spatial locality by always assigning the nearest room, reducing travel overhead, and (2) it balances workload dynamically, preventing situations where one responder is overwhelmed while others finish.

3.5 Path Optimization Algorithm

Given a responder's assigned rooms, we find a n e fficient tra versal pat h usi ng a multi-stage traveling salesman problem-inspired solver [2,5,7,10]:

Stage 1: Nearest-Neighbor Initialization. Starting from the responder's initial exit, build a path by repeatedly visiting the nearest unvisited room. This provides a reasonable initial solution quickly [10].

Stage 2: 2-Opt Local Search. Iteratively test swapping pairs of edges in the path. If swapping edges (r_i, r_{i+1}) and (r_j, r_{j+1}) with (r_i, r_j) and (r_{i+1}, r_{j+1}) reduces total path length, accept the swap. Repeat until no improving swaps exist [7].

Stage 3: Or-Opt Improvement. Test relocating sequences of one to three consecutive rooms to other positions in the path. This escapes local minima that 2-opt cannot improve [8].

Stage 4: 2-Opt Re inement. Apply 2-opt again to refine improvements from Or-opt [7].

This combination of nearest-neighbor initialization and local search improvements typically produces paths within 5 percent of optimal while running in polynomial time [2,4,5,7,10].

3 Model I: Baseline Sweep Time Model

To address the complex dynamics of an evacuation sweep, our model is separated into three computational stages: K-Means clustering, TSP routing, and Monte Carlo simulation with A* pathfinding. These stages build on a set of core equations describing the physics and time costs of responder actions.

3.1 Stage 0: Core Equations (Physics Engine)

3.1.1 Total Sweep Time

For a single responder k, the total sweep time is modeled as

$$T_k = T_{\text{travel},k} + T_{\text{search},k} + T_{\text{assist},k} + T_{\text{door},k} + T_{\text{mark},k} + T_{\text{comm},k}, \tag{1}$$

where each term aggregates the contributions from all relevant events (movement, searches, rescues, doors, marking, and communication).

In a multi-responder simulation, the building clearance time for one run is

$$T_{\text{clear}} = \max_{k=1,\ldots,K} T_k. \tag{2}$$

3.1.2 Visibility due to Smoke

In fire scenarios, effective speed is reduced by visibility. We combine fatigue and visibility:

$$v(t) = v_0 - ft, \qquad v_{\text{eff}}(t) = M_{\text{vis}}(t) \cdot v(t), \tag{3}$$

where $M_{\text{vis}}(t) \in (0, 1]$ is derived from the Jin visibility model via

$$S(t) = \frac{C}{k_{\text{vis}}(t)}, \qquad M_{\text{vis}}(t) = \min\left(1, \frac{S(t)}{10}\right).$$

We cap smoke evolution at $T_{\text{fire}} = 11$ min; beyond this time, $k_{\text{vis}}(t)$ and thus $M_{\text{vis}}(t)$ are held constant.

3.1.3 Horizontal Travel Time

Horizontal travel distance d_{horiz} for responder k is obtained from the TSP/A* paths. Effective speed includes both fatigue and visibility: $v_{\text{eff}}(t) = \left(v_0 - f_{\text{fatigue}}t\right) M_{\text{vis}}(t)$, and horizontal travel time is approximated by numerical integration or, in discrete steps, by summing segment times: $T_{\text{horizontal},k} = \sum_\ell \frac{\Delta d_\ell}{v_{\text{eff}}(t_\ell)}$.

3.1.4 Vertical Travel Time

Vertical travel is modeled separately. Total stair distance is

$$d_{\text{stair}} = (n_{\text{floors}} - 1)\, h_{\text{floor}} \times (\text{number of transitions}), \tag{4}$$

with baseline vertical speed $v_{\text{base,vert}} = 0.170$ m/s. Fatigue is stronger on stairs; we effectively double the fatigue constant for vertical movement:

$$v_{\text{vert,eff}}(t) = v_{\text{base,vert}} - 2f_{\text{fatigue}}t, \tag{5}$$

$$T_{\text{vertical},k} = \sum_\ell \frac{\Delta d_{\text{stair},\ell}}{v_{\text{vert,eff}}(t_\ell)}. \tag{6}$$

3.1.5 Door-Opening Time

Door-opening time is modeled as a capped exponential:

$$T_{\text{open,raw}} \sim \text{Exponential}(\lambda_{\text{door}}), \qquad T_{\text{open}} = \min(T_{\text{open,raw}}, T_{\text{open,max}}), \tag{7}$$

with $T_{\text{open,max}} = 60$ s.

3.1.6 Search Time

Search speed is taken as

$$v_{\text{search}} = \frac{1}{2} v_0, \tag{8}$$

and search time scales linearly with room area:

$$T_{\text{search}} = \frac{A_{\text{room}}}{v_{\text{search}}}. \tag{9}$$

3.1.7 Carrying Time

If an occupant is present, we first draw an assist-preparation time: $T_{\text{assist}} \sim \mathcal{N}(45, 10^2)$ s, then compute reduced carrying speed as a function of occupant weight W_{occ}:

$$v_{\text{carry}} = v_0 \left(1 - \alpha \frac{W_{\text{occ}}}{W_{\text{max,carry}}}\right), \tag{10}$$

for some calibrated $\alpha \in (0, 1)$. The carrying time for a round trip of A* path length d_{exit} is

$$T_{\text{carry}} = \frac{d_{\text{exit}}}{v_{\text{carry}}} + \frac{d_{\text{return}}}{v_{\text{eff}}(t)}, \tag{11}$$

added to T_k whenever a rescue occurs.

3.1.8 Communication and Marking Times

Communication time is $T_{\text{comm}} \sim \text{Uniform}(1, 3)$ s, applied when a blockage is reported, and is sampled from the uniform distribution between its observed min and max.

3.2 Stage 1: K-Means Clustering (Assigning Zones)

To assign zones, we represent the building as a graph $G = (N, E)$ where nodes N are room centers and key intersections, and edges E represent traversable connections with associated distances. Node coordinates (x_i, y_i) are used to cluster rooms into K zones. K-mean solves:

$$\min_{\mu_1, \ldots, \mu_K} \sum_{k=1}^{K} \sum_{x_i \in Z_k} \|x_i - \mu_k\|^2 \tag{12}$$

to produce compact, geographically coherent zones. This yields balanced workloads and reduced travel distances.

3.3 Stage 2: Ideal Path Optimization (TSP)

Within each cluster Z_k, we solve a TSP variant: find a path that starts at the responder's exit, visits all rooms in Z_k exactly once, and minimizes travel distance. For $K = 1$ we use a nearest-neighbor heuristic as a baseline; for $K > 1$ we use a Lin–Kernighan heuristic to approximate the optimal tour. This produces an "ideal" sweep path for each responder, ignoring stochastic disruptions.

3.4 Stage 3: Monte Carlo Simulation

To model real-world uncertainty, we simulate K responders following their TSP paths in parallel over N_{MC} runs. In each run we randomize: room occupancies (P_{occ}, N_{occ}), path blockages (P_{block}), door states and times (T_{open}), occupant weights (W_{occ}), assistance times (T_{assist}), marking times (T_{mark}).

For each run, we compute T_k and $T_{\text{clear}} = \max_k T_k$. Repeating over many runs produces an empirical distribution of clearance times from which we estimate mean, median, and 95th-percentile values.

3.5 A* Pathfinding in the Simulation

Responders do not blindly follow precomputed TSP routes. When unexpected conditions arise, they use A* pathfinding.

3.5.1 Trigger 1: Path Blockage

When traversing an edge (u, v), a Bernoulli trial with probability P_{block} determines if the path is blocked (earthquake scenarios). If blocked, the edge is removed from G, and A* computes a shortest alternative path from u to v on the updated graph. The travel time along this new path is added to T_k.

3.5.2 Trigger 2: Evacuee Transport

When an occupant is found, the responder pauses the TSP route and uses A* to find the shortest path to their assigned exit. The cost function uses v_{carry} while carrying and v_{eff} for the return. The computed travel time plus T_{assist} are added to T_k.

3.6 Optimize Number of Responders (K)

Increasing K decreases the load per responder but eventually yields diminishing returns. To find the optimal K, we:

1. Run low-fidelity Monte Carlo simulations (1,000 runs) for $K = 1, 2, \ldots, K_{\max}$.
2. For each K, compute $T_{95}(K)$, the 95th-percentile clearance time.
3. Plot $T_{95}(K)$ vs. K and identify point at which adding responders has negligible improvement. Thusm the K at the elbow of this plot is selected as the optimal staffing level.

3.7 Final High-Fidelity Simulation

After choosing K_{opt}, we run a high-fidelity Monte Carlo simulation with $N_{\text{MC}} = 100{,}000$ runs. This provides a robust estimate of the full clearance-time distribution for the optimal staffing level, including rare but critical tail events.

7.1 Addition and Adjustments of Indices

Added: Environmental-Industrial Index (EII)

Different countries have varying degrees of industrialization, meaning they have different potentials in handling environmental issues. In CIEM, since all the research is based on the states in the US, the extent of industrialization is likely to be constant for all states. However, as we have to apply the model towards countries that have different degrees of industrialization from the world, an **Environmental-Industrial Index (EII)** is required to enhance NIEM's comprehensiveness.

Essentially, EII is a **benefit-type index** that exhibits the **interconnection** between the **natural environment and industrial environment** within a country, we first referred to Environmental Performance Index (EPI), which is a comprehensive rating available for most countries across the world based on their environmental performance, as an **benefit-type** indicator. Additionally, we incorporated an industrial indicator by looking at different industrial sectors' **total economic values** within the country.

$$I_I = \frac{EPI_i - EPI_{min}}{EPI_{max} - EPI_{min}} \times i_i \tag{7.1}$$

Where i_i represents the **extent of industrialization** within the country.

$$i_i = \frac{ValueUSD(C)}{ValueUSD(B) + ValueUSD(C) + ValueUSD(D) + ValueUSD(E)} \tag{7.2}$$

Here, $ValueUSD(B)$, $ValueUSD(C)$, $ValueUSD(D)$, and $ValueUSD(E)$ refer to the total economic value of **mining and quarrying**; **total manufacturing**; **electricity, gas, steam, and air conditioning**; and **water supply, sewerage, and waste management**, respectively.

Arguably, $ValueUSD(C)$, as part of the industrial indicator, partially reflects how well a country can regulate its natural environment. Specifically, as the total economic value of manufacturing within a country increases, the country will have better technology and thereby regulating its natural environment more efficiently. Therefore, we put $ValueUSD(C)$ on the numerator, indicating that it is a benefit-type variable.

$$S_6 = \frac{I_{Lmax} - I_{Li}}{I_{Lmax} - I_{Lmin}} \tag{7.3}$$

Adjusted: Carbon Emission Index (CEI)

The adjusted CEI is shown as follows:

$$I_{Ci} = \frac{(\frac{|t_i - t_{comfort}|}{t_{max} - t_{min}} + 0.5) \times (c_i \times p_i + c_w \times p_w)}{(p_i + p_w) \times (\frac{\xi_i - \xi_{min}}{\xi_{max} - \xi_{min}} + 0.5) \times (\frac{A_i - A_{min}}{A_{max} - A_{min}} + 1)} \tag{7.4}$$

Where ξ_i represents the quality of water in the country; c_i represents the combination of carbon emission caused by residents and transportation.

Essentially, we added the effect of other **country-specific factors** into the formula. Notably, we did not use spatial carbon intensity per person in this formula; instead, we turned the area of the country into an insignificant factor by normalizing it.

Adjusted: Energy Consumption Index (ECI)

$$I_{Ei} = \frac{(\frac{|t_i - t_{comfort}|}{t_{max} - t_{min}} + 0.5) \times (p_i \times e_i + p_w \times e_w)}{(p_i + p_w) \times r_i \times (\frac{\xi_i - \xi_{min}}{\xi_{max} - \xi_{min}} + 0.5) \times (\frac{\upsilon_i - \upsilon_{min}}{\upsilon_{max} - \upsilon_{min}} + 1)} \tag{7.5}$$

Where υ_i represents the energy generated by wind.

Adjusted: Water Consumption Index

$$I_{Wi} = \frac{W_{ci} + W_{cw}}{(\frac{\eta_i - \eta_{min}}{\eta_{max} - \eta_{min}} + 1) \times (\frac{\xi_i - \xi_{min}}{\xi_{max} - \xi_{min}} + 1) \times W_{si} \times (\frac{p_i - p_{min}}{p_{max} - p_{min}} + 1)} \tag{7.6}$$

The availability of data enables us to enhance NIEM's reliability and accuracy.

Here, η_i represents the availability of water in the country. We have also changed the meanings of the symbols: W_{ci} refers to the total water consumed in a country; W_{si} is the total water supply in a country; and W_w is the average total water consumed in World Cup session.

Adjusted: AQI Index

$$I_{AQIi} = \frac{(AQI_i - AQI_s)^2 + 0.5}{(\frac{\iota_i - \iota_{min}}{\iota_{max} - \iota_{min}} + 1)} \tag{7.7}$$

Where, ι represents the average wind speed within the country, and AQI_s represents the standard AQI for us to perform the calculation, which we took 100 as the standard AQI value, considering the industrialization requirement.

Comparing to the CIEM, there are certain changes. Specifically, we calculated the distance from each country's AQI to the standard AQI, and w used the average wind speed within the country for calculation.

Adjusted: Landfill index

$$I_{Li} = \frac{L_w \times (\frac{3.5 \times 10^6 - p_{min}}{p_{max} - p_{min}} + 1) + L_i \times (\frac{p_i - p_{min}}{p_{max} - p_{min}} + 1) \times 32}{EPI_i \times logA_i} \tag{7.8}$$

Here, L_w represents the amount of trash and waste created per audience during a World Cup session; R_p represents the range of population within the dataset.

Normalization

Similarly, we used the same method in CIEM to normalize the 6 indices, obtaining S_i for each country:

$$S_i = \frac{I_i - I_{min}}{IQR} \tag{7.9}$$

Results of the Normalized Indices

After predicting the required data for 2029, we obtained the six indices (S_1 to S_5) for each of the seven counties that are suitable for hosting the World Cup. The normalized data are as follows:

Table 7.1: Normalized Values for Selected Countries

	Germany	Japan	Brazil	South Korea	Russia	South Africa	Qatar
S_1	0.0000	0.1350	1.1342	0.1920	1.0662	0.3858	1.2250
S_2	0.0030	0.0000	0.0691	0.0150	0.5075	1.0000	0.2402
S_3	0.2263	0.3583	0.9585	0.1508	1.0000	0.2860	0.0000
S_4	0.0000	0.1182	0.6959	0.8946	0.4324	0.4155	1.0000
S_5	0.0000	2.2061	1.5000	1.7227	0.9939	2.0789	0.7227
S_6	1.0000	0.7321	0.5796	0.5415	0.2780	0.1693	0.1000

7.2 Weighting

Analytical Hierarchy Process (AHP)

Considering the global characteristics of World Cup events, we constructed the following judgment matrix for the six indicators:

$$A = \begin{bmatrix} 1 & 3 & 5 & 7 & 9 & 4 \\ 1/3 & 1 & 3 & 5 & 7 & 2 \\ 1/5 & 1/3 & 1 & 3 & 5 & 1/2 \\ 1/7 & 1/5 & 1/3 & 1 & 3 & 1/4 \\ 1/9 & 1/7 & 1/5 & 1/3 & 1 & 1/6 \\ 1/4 & 1/2 & 2 & 4 & 6 & 1 \end{bmatrix} \tag{7.10}$$

Indicator order: S1, S2, S3, S4, S5, S6

Rationale of the judgment matrix:

Carbon emissions and Energy consumption remain the highest priority due to their long term and irreversible global climate impacts. The new EII ranks third as it reflects **comprehensive national environmental governance**. Water resources and air quality address more regional concerns, while landfill management has the most localized impact, so they were less prioritized.

Weight Calculation

The normalized weights are calculated using:

$$w_i = \frac{M_i}{\sum_i M_i}, \quad \text{where } M_i = \left(\prod_{j=1}^{6} a_{ij} \right)^{1/6} \tag{7.11}$$

Resulting weights:

$$W_s = [0.430, 0.229, 0.103, 0.057, 0.032, 0.148]^T \tag{7.12}$$

Consistency Check: $CR = \frac{CI}{RI} = 0.029$

Since $CR = 0.029 < 0.10$, the judgment matrix passes the consistency check.

Criteria Importance Through Intercriteria Correlation (CRITIC)

Our 6 new, adjusted indices from NIEM are arguably more **interconnected** with each other than the 5 indices from CIEM. For instance, the EII and the LFI share the variable EPI together. Therefore, we considered the best way to handle such interdependence to be the use of the CRITIC method. Fittingly, CRITIC decides weights based on both **variability** (measured by standard deviation) and **interdependence** (measured by Pearson Correlation Coefficient), meaning that, other than measuring interdependence, it actually 'encompasses' EWM as well.

First, we utilized **max-min scaling** to normalize all 42 **cost-type values** (6 indices value for each of the 7 different countries):

$$h_{ij} = \frac{max_j - x_{ij}}{max_j - min_j}$$

Next, to measure variability, we calculated the sample standard deviation (s) of the normalized values by the following formula:

$$s_j = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(h_{ij} - \bar{h_j})^2}$$

Subsequently, we calculated Pearson correlation coefficient (r) for the normalized sample:

$$r_{ij} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

Then, we obtained the **importance of criterion (**C_j**)** and calculated the weight of each S (normalized indices of CIEM):

$$w_j = \frac{C_j}{\sum_{j=1}^{n} C_j} = \frac{s_j \cdot \left(1 - \sum_{i \neq j} |r_{ij}|\right)}{\sum_{j=1}^{n} s_j \cdot \left(1 - \sum_{i \neq j} |r_{ij}|\right)}$$

Our results for the weight are as follows:

$$W_o = [0.211, 0.119, 0.147, 0.143, 0.204, 0.175]^T$$

Strategic AHP-CRITIC Combination

We combine the **subjective AHP weights** (W_s) and **objective CRITIC weights** (W_o) in a linear fashion:

$$W_c = \alpha W_s + (1 - \alpha) W_o, \quad \alpha = 0.6 \tag{7.13}$$

Still, we set the value of α to be 0.6 (60%), meaning that W_s, the subjective AHP, is slightly more important than W_o, the objective EWM, because we deemed our research- and experience-based insights to have greater practical value.

Table 7.2: AHP vs CRITIC Weights

Indicator	**AHP Weight (W_s)**	**CRITIC Weight (W_o)**	**Combined Weight (W_c)**	**Rank**
CEI (S_1)	0.430	0.211	0.3424	1
ECI (S_2)	0.229	0.119	0.1850	2
WCI (S_3)	0.103	0.147	0.1206	4
AQI (S_4)	0.057	0.143	0.0914	6
LFI (S_5)	0.032	0.204	0.1008	5
EII (S_6)	0.148	0.175	0.1588	3
Total	**0.999**	**0.999**	**1.000**	

Verification of Combined Weights:

Total: $0.3424 + 0.1850 + 0.1206 + 0.0914 + 0.1008 + 0.1588 = 0.999 \approx 1$

5.3 Step 2: The Combined Weighting Model

The determination of indicator weights is a critical step in any MCDM model, as it defines the relative importance of each criterion in the final evaluation. To create a balanced and robust weighting scheme that mitigates the biases of any single method, we combine a subjective approach (Analytic Network Process) with an objective approach (Entropy Weight Method).

5.3.1 Subjective Weighting: Analytic Network Process (ANP)

To capture expert opinions on strategic priorities and the complex interdependencies between criteria, we employ the ANP. Unlike the simpler AHP, ANP allows us to model a network of influences, which is suitable for our four-stage framework where each stage can influence the others.

Step 1: Construct the Network Structure and Judgment Matrices. We first established the network structure, defining the relationships between our four primary indicators (Supply, Consumption, Load, Capacity). Subsequently, simulating an expert panel, we performed pairwise comparisons using Saaty's 1-9 scale to quantify the relative importance between connected criteria. This process generated the necessary judgment matrices. The following table shows how we make the comparisons:

a_{ij}	a_{ji}	Interpretation
1	1	Criterion i and j are equally important
3	$\frac{1}{3}$	Criterion i is slightly more important than j
5	$\frac{1}{5}$	Criterion i is moderately more important than j
7	$\frac{1}{7}$	Criterion i is strongly more important than j
9	$\frac{1}{9}$	Criterion i is extremely more important than j

The following graph and tables show the comparisons of indicators

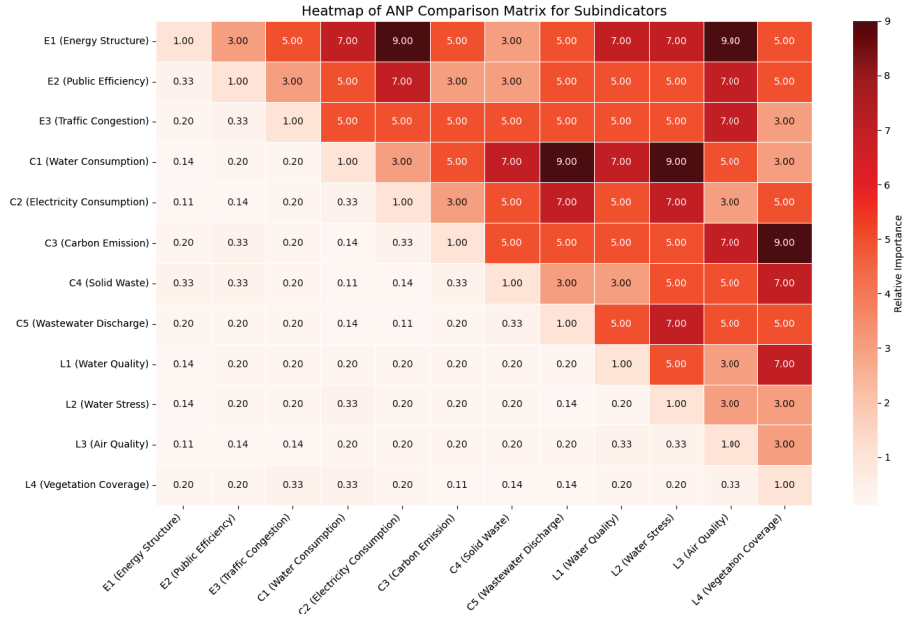


Figure 5.2: Relative importance heatmap

Step 2: Compute the Supermatrix and Final ANP Weights. The pairwise comparisons are used to form an unweighted supermatrix, which represents the initial influence levels. This is then stochasticized to create the weighted supermatrix, $\bar{W}$. To capture the total long-term influences, we calculate the limit supermatrix by raising the weighted supermatrix to a large power until it converges:

$$\bar{W}^{\infty} = \lim_{k \to \infty} \bar{W}^{k} \tag{5.12}$$

The final ANP weight vector, W_{ANP}, is extracted from the converged limit supermatrix.

A matrix $A = (a_{ij})$ must satisfy the consistency condition, checked by the Consistency Ratio (CR):

$$CR = \frac{CI}{RI} = \frac{(\lambda_{\max} - n)/(n-1)}{RI} \tag{5.13}$$

where $\lambda_{\max}$ is the principal eigenvalue of the matrix and RI is the Random Consistency Index.

After processing the judgment matrices and calculating the limit supermatrix, the resulting ANP weight vector was found in Table 5.1. The overall consistency ratio for the model was calculated as $CR < 0.1$, indicating that our judgments are acceptably consistent.

Table 5.1: Normalized Indicator Weights of ANP (Two-Column Layout).

Indicator	Weight	Indicator	Weight
Urban Energy Structure	0.13	Event Electricity Consumption	0.10
Urban Public Energy Efficiency	0.09	Event Carbon Emissions (Venues, Transportation)	0.15
Urban Traffic Congestion Coefficient	0.07	Event Solid Waste	0.04
Event Water Consumption	0.11	Event Wastewater Discharge	0.04
Urban Water Quality	0.07	Urban Air Quality	0.09
Urban Water Stress	0.05	Urban Vegetation Coverage	0.06

5.3.2 Objective Weighting: Entropy Weight Method

To incorporate the information inherent in the data itself, we use the Entropy Weight Method. This method assigns higher weights to indicators that exhibit greater variation across the candidate cities, as they provide more information for differentiation. The process uses the Decision Matrix $X = (x_{ij})$ from Step 1 as input.

Step 1: Normalize and Standardize the Data. First, each element x_{ij} in the decision matrix is normalized to calculate its proportion p_{ij} within its indicator column:

$$p_{ij} = \frac{x_{ij}}{\sum_{i=1}^{m} x_{ij}} \tag{5.14}$$

where m is the number of cities.

Step 2: Calculate Information Entropy. The information entropy E_j for each indicator j is calculated, which measures the degree of uncertainty or variation in the data:

$$E_j = -k \sum_{i=1}^{m} p_{ij} \ln(p_{ij}), \quad \text{where } k = \frac{1}{\ln(m)} \tag{5.15}$$

Step 3: Calculate Objective Weights. The final entropy weight w_j for each indicator is derived from its information entropy. A lower entropy (less uncertainty, more variation) results in a higher weight.

$$w_j = \frac{1 - E_j}{\sum_{j=1}^{n}(1 - E_j)} \tag{5.16}$$

This process yields the objective weight vector, $W_{Entropy}$.

$$W_{Entropy} = [0.18, 0.08, 0.06, 0.11, 0.13, 0.22, 0.09, 0.03, 0.05, 0.02, 0.01, 0.02]$$

Notably, 'Travel Footprint' received the highest objective weight, indicating that the candidate cities exhibit the greatest variation on this metric.

5.3.3 Sigmoid Function for Dealing with Extreme Data Value

If some values reach an extreme value, the weights we have calculated before will not describe the whole story. We designated a threshold value that indicate the shift between constant weights and dynamic weights. This smooth dynamic adjustment mechanism is processed using a Sigmoid function. The following pairwise function shows how it works:

$$\lambda(x) = \begin{cases} \lambda_{\text{base}}, & x < T_{\text{start}} \\ \lambda_{\text{base}} + (\lambda_{\text{max}} - \lambda_{\text{base}}) \cdot \dfrac{1}{1 + e^{-k(x - T_{\text{trigger}})}}, & T_{\text{start}} \le x \le T_{\text{end}} \\ \lambda_{\text{max}}, & x > T_{\text{end}} \end{cases} \tag{5.17}$$

In which $T_{trigger}$ is the threshold value;$\lambda_{\text{max}} - \lambda_{\text{base}}$ shows the range of fluctuation of the weights.

5.3.4 Weight Combination

We synthesize the subjective weights from ANP (W_{ANP}) and the objective weights from the Entropy method ($W_{Entropy}$) using a linear combination to produce our final, balanced weight vector, W_{final}:

$$W_{final,j} = \lambda \cdot W_{ANP,j} + (1 - \lambda) \cdot W_{Entropy,j} \tag{5.18}$$

We set the combination coefficient $\lambda = 0.5$, granting equal importance to both the expert-driven strategic priorities and the inherent data structure. This balanced approach ensures our final weights are both strategically relevant and data-informed.

6.2 Dynamic Data Correction: Projecting 2029 Indicators via Trend Extrapolation

Assessing the sustainability performance of cities in 2029 requires taking into account the effects of policy directions on urban development.

We identified two secondary indicators from our framework that are most likely to change significantly over a five-year horizon due to established, long-term public policy:

- **Grid Cleanliness:** Most states and major utility companies have legally mandated Renewable Portfolio Standards (RPS) with specific targets for future years (e.g., "50% renewable by 2030").
- **Municipal Waste Management Capability:** Many major cities have adopted "Zero Waste" plans with explicit goals for increasing waste diversion rates over time.

For these two indicators, we projected their 2029 values using a policy-based linear extrapolation method. We researched the official sustainability targets for each candidate city (or its state/region). The projected indicator value for 2029 was then calculated as a linear interpolation between its current value and its future policy target.

$$I_{2029} = I_{current} + (I_{target} - I_{current}) \times \frac{2029 - \text{Year}_{current}}{\text{Year}_{target} - \text{Year}_{current}} \tag{6.1}$$

For cities without explicit targets, we used a conservative linear extrapolation based on their historical data from the past decade. This process resulted in two distinct decision matrices: a **2024 Baseline Matrix** using the most recent available data, and a **2029 Projected Matrix** with updated values for these two dynamic indicators.

We ran our full MCDM model (ANP-Entropy weighting and TOPSIS evaluation) twice: once on the baseline matrix to establish a current ranking, and once on the projected matrix to forecast the 2029 sustainability landscape.

The results, presented in Table 6.2, reveal a dynamic shift in the rankings. The table shows each city's rank based on 2024 data, its projected rank for 2029, the resulting change in rank, and the primary factor driving that change.

6.3 Results Analysis and Final Recommendation

Analysis of Rankings

Proactive environmental policy strongly influences rankings. Cities in states with aggressive decarbonization goals, such as Los Angeles and the New York area, maintain or improve their positions. Notable projected improvements occur in Baltimore and New Orleans, which are actively implementing new policies. Conversely, cities with slower energy transitions, like Dallas, see relative declines, while cities facing structural challenges, such as Las Vegas and Phoenix with extreme water stress, remain low, highlighting inherent geographical disadvantages.

Table 6.1: 2029 Projected Sustainability Ranking (TOPSIS Results)

Rank	City / Area	Score (C_i)	Rank	City / Area	Score (C_i)
1	Seattle, WA (New)	0.920	12	Glendale, AZ (Phoenix Area)	0.485
2	San Francisco Bay Area	0.875	13	Tampa, FL	0.470
3	Los Angeles Area	0.842	14	Detroit, MI	0.435
4	Denver, CO (New)	0.798	15	Miami, FL	0.383
5	Minneapolis, MN	0.755	16	Indianapolis, IN	0.319
6	New York Area	0.725	17	Houston, TX	0.290
7	San Diego, CA	0.718	18	Arlington, TX (Dallas Area)	0.285
8	Baltimore, MD (New)	0.680	19	Pasadena, CA	0.268
9	Atlanta, GA	0.640	20	Tempe, AZ (Phoenix Area)	0.250
10	Jacksonville, FL	0.545	21	Las Vegas, NV	0.248
11	New Orleans, LA	0.510	22	Pontiac, MI	0.245

Table 6.2: Comparison of 2024 Baseline and 2029 Projected Sustainability Rankings

City / Area	2024 Rank	2029 Rank	Rank Change	Primary Driver of Change
(New Cities)				
Seattle, WA	**1**	**1**	**–**	*Maintains top position*
Denver, CO	4	4	–	*Stable performance*
Baltimore, MD	14	8	↑ **6**	*State-level renewable energy goals*
(Previous Hosts)				
San Francisco Bay Area	2	2	–	*Strong existing policies*
Los Angeles Area	3	3	–	*Aggressive clean energy policy*
New York Area	7	6	↑ **1**	*Ambitious offshore wind projects*
New Orleans, LA	12	11	↑ **1**	*New municipal waste initiatives*
Arlington, TX (Dallas Area)	17	18	↓ **1**	*Slower grid decarbonization*
Atlanta, GA	8	9	↓ **1**	*Moderate policy progression*
Glendale, AZ (Phoenix Area)	19	12	↑ **7**	*Hypothetical major solar investment*
Las Vegas, NV	11	21	↓ **10**	*Inherent water scarcity*

Archetypes of Sustainability: Top Performers

The 2029 projected rankings reveal multiple paths to top sustainability scores. Our model identifies three archetypes among the leading cities:

This archetypal analysis provides a crucial strategic insight for the NFL: the definition of a "sustainable host" is not monolithic. The choice is between a city that is *inherently sustainable* (Seattle), one that is *engineered to be sustainable* (New York), or one that is *becoming sustainable* (Baltimore). This nuanced understanding allows for a more strategic selection process aligned with the NFL's long-term sustainability goals.

Table 6.3: Archetypes of Top-Tier Cities and Key Features

City (Rank)	Archetype	Key Features
Seattle, WA (1)	All-Rounder	Excels across all primary indicators: clean energy, temperate climate, and strong environmental capacity. Largest, most balanced radar chart profile.
New York Area (6)	Infrastructural Powerhouse	High *Clean Resource Supply* (public transit) and *Event-Generated Load*; moderate in other areas. Infrastructure offsets natural limitations.
Baltimore, MD (8)	Policy-Driven Climber	Moves from 14th to 8th due to proactive policies (Maryland Renewable Portfolio Standard). Policy compensates for limited natural or infrastructural advantages.

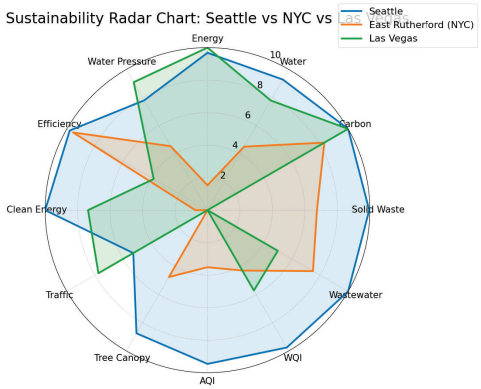


Figure 6.1: Radar Chart Analysis for Selected Cities based on 2029 Projections.

To visualize these differences, Figure 6.1 presents radar charts for our top-ranked city (Seattle), a major improver (New York Area), and a city with significant challenges (Las Vegas). Seattle's strength is its all-around excellence, while New York's chart clearly shows its strength in Supply and Load management. Las Vegas's chart, in contrast, starkly reveals its critical weakness in the Environmental Capacity dimension.

Final Recommendation for 2029

To provide our final recommendation, we apply a real-world constraint: the NFL's informal **"cooling-off period,"** defined as a 5-year window. Cities that hosted or are scheduled to host between 2025 and 2028 (New Orleans, Santa Clara/SF Bay Area, Las Vegas) are thus ineligible for the 2029 event.

Eligible Cities: Previous Hosts and New Cities. Among previous hosts, the top-ranked eligible city is the **Los Angeles Area (Inglewood)**, whose aggressive clean energy policies and strong infrastructure make it the most sustainable choice within that group. Among new candidate cities, **Seattle** emerges as the top performer overall, driven by the nation's cleanest electricity grids and robust environmental policies, followed by **Denver** and **Baltimore**, both showing strong potential and outperforming many previous hosts.

Overall Recommendation: Based on our comprehensive environmental sustainability model, the most sustainable choice for the 2029 Super Bowl is **Seattle, WA**. However, if the NFL prioritizes a city with recent hosting experience, the **Los Angeles Area** is the best option from that pool. We recommend considering these two cities as the primary candidates, as both signal a strong commitment to environmental responsibility.

(3)

4.2 PCA-Based Multivariate Regression

To quantify how city features affect different categories of environmental impacts, we employ a PCA-based multivariate linear regression model. The specific steps are as follows:

Step1. Eigen-decompose the correlation matrix R:

$$R = P\Lambda P^T \tag{4}$$

where Λ is the diagonal matrix composed of eigenvalues λ_k (arranged in descending order), and P is the eigenvector matrix.

Step2. The k-th **principal component (PC)**, a linear combination of standardized features:

$$PC_k = p_{1k}Z_1 + p_{2k}Z_2 + \cdots + p_{pk}Z_p \tag{5}$$

where Z_j denotes the j-th normalized city-feature column vector.

Step3. For all cities, the **principal-component** score matrix is:

$$T = ZP \in \mathbf{R}^{n\times p} \tag{6}$$

where the k-th column, t_k, contains the scores of the k-th **principal component** across the four cities.

Step4. Compute the cumulative explained variance of the first K principal components. Here, we retain the **PCs** whose cumulative explained variance reaches at least 85%, and take these **PCs** as the new composite city-feature factors.

$$C_K = \frac{\sum_{k=1}^{K}\lambda_k}{\sum_{j=1}^{p}\lambda_j}\times 100\% \tag{7}$$

Step5. Next, we use the **principal-component** scores as the new explanatory variables and **perform principal component regression** (**PCR**) on the environmental impact indicators. Construct the score matrix of the first K principal components, augmented with an intercept.

$$X_{PC} = \begin{bmatrix} 1 & t_{1t} & \cdots & t_{1k} \\ 1 & t_{21} & \cdots & t_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & t_{n1} & \cdots & t_{nk} \end{bmatrix} \in \mathbf{R}^{n\times(K+1)} \tag{8}$$

Step6. For a single environmental indicator, the regression model is:

$$y_j = \beta_{j0}1 + \beta_{j1}t_1 + \cdots + \beta_{K1}t_K + \varepsilon_j \tag{9}$$

Step7. Fit the model by **ordinary least squares**:

$$\hat{\beta}_j = (X_{PC}^T X_{PC})^{-1} X_{PC}^T y_j \tag{10}$$

Step8. Finally, for each environmental indicator, compute the fitted values, the residual sum of squares, and the coefficient of determination R^2.

4.3 Results and Analysis

In this subsection, we use simulation results to quantitatively elucidate the relationship between environmental impacts and city features.

The correlation analysis between city features and environmental impacts is shown in Fig. 1. It reveals that host-city features, such as climate, transport systems, and energy mix exhibit strong correlations with event-period outcomes including energy use, water consumption, and carbon emissions. In particular, attendees' average travel distance is strongly positively correlated with audience- and transport-related emissions, whereas public-transit coverage and the share of green/low-carbon travel exert a suppressing effect on transport emissions.

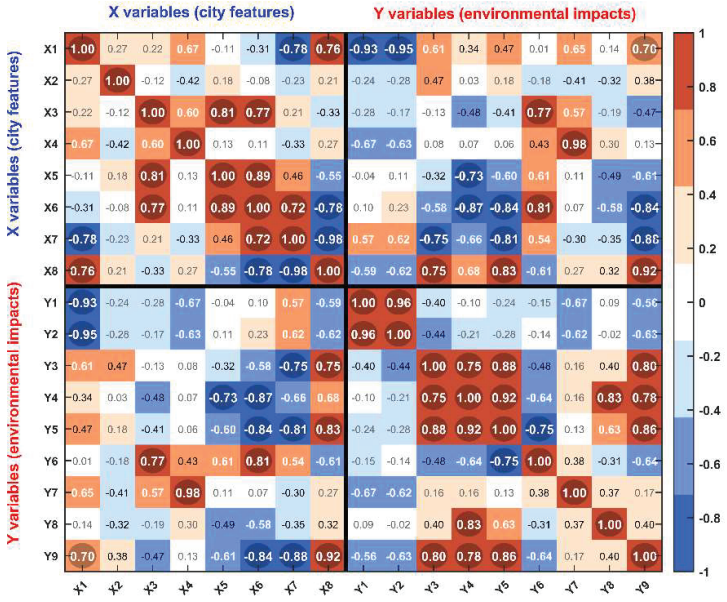


Figure 1 Heatmap of correlations between city features and environmental impacts

These patterns indicate that a host city's transportation conditions and spatial configuration are key determinants of the environmental footprint of mega-sport events like the Super Bowl.

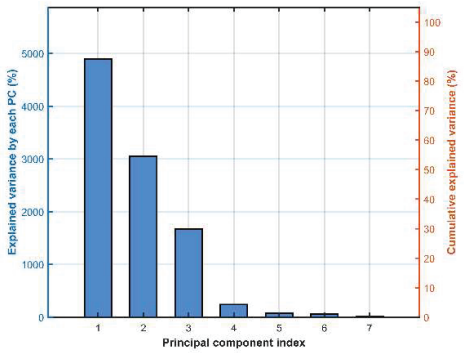


Figure 2 Principal components' explained-variance plot

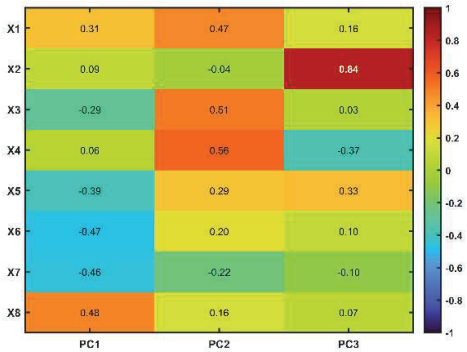


Figure 3 PCA loading heatmap

As shown in **Figure 2**, the first three principal components jointly explain over 85% of the total variance in the city-feature variables. Therefore, we retain the top three PCs as composite city-feature

indices for subsequent modeling. **Figure 3** displays the loadings of each feature on the PCs: PC1 is chiefly driven by population density, public-transit coverage, and the share of green/low-carbon travel, and can be interpreted as a "transport & urban intensity" factor. PC2 reflects climate conditions and grid carbon intensity, i.e., a "climate–energy structure" factor. PC3 has the highest loading on population density, confirming this as its core feature.

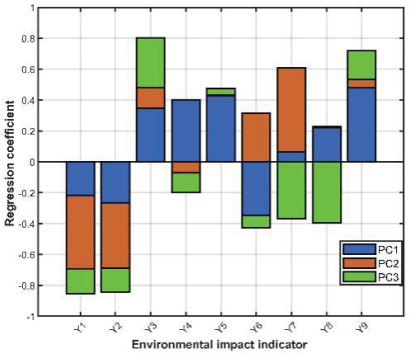


Figure 4 Decomposition of principal-component regression coefficients

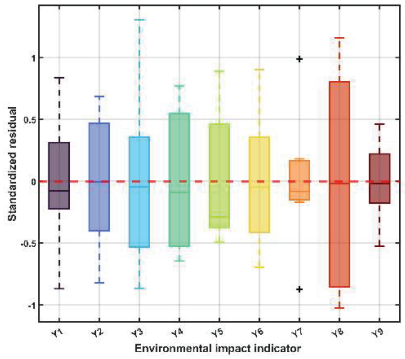


Figure 5 Standardized residual distributions for each environmental indicator

Figure 4 further shows that among the three latent dimensions extracted by PCA, the infrastructure & energy-structure component (PC1) primarily governs resource-use and waste-related indicators; the climate–distance component (PC2) dominates energy use and audience-related CO_2 emissions; and the mobility/density component (PC3) plays an important supplementary role in transport- and audience-related CO_2. The residual distributions in **Figure 5** are centered near zero with limited dispersion, indicating no systematic bias in the principal-component–based model and supporting its ability to reasonably characterize differences in environmental performance across host cities.

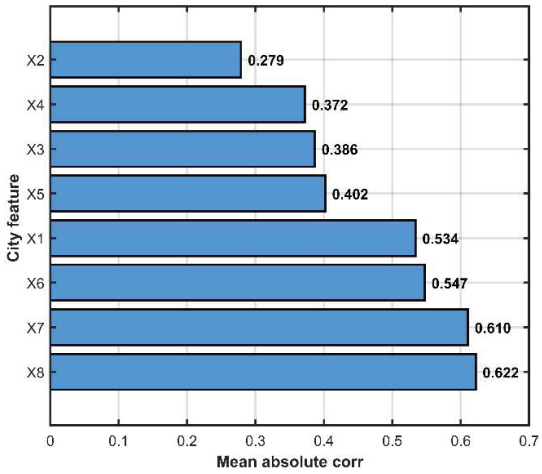


Figure 6 Importance ranking of environmental-impact drivers

When considering all environmental impact indicators $Y_1 - Y_9$, grid carbon intensity, the share of renewable energy, and the greenness of the transport system are the location traits with the most **system-wide** influence; next in importance are temperature and the level of public transit, while precipitation and other climatic details have comparatively weak average effects. Specifically:

Grid carbon intensity X_8 and renewables share X_9 rank highest in importance, indicating that with the same electricity use, materials, and event operations, outcomes differ systematically under

high-carbon vs. low-carbon grids. Share of green travel X_6 and public-transit coverage X_5 also rank near the top, showing that a city's transport mode mix jointly affects multiple environmental indicators (not only carbon emissions but also waste, recycling rates, water use, etc.).

Within climate variables, temperature X_1 matters more than precipitation X_2 : temperature is more tightly coupled to energy consumption and water use (HVAC loads, comfort needs), so its overall effect on the indicators is clear, whereas precipitation is not a core driver of cross-city differences.

Population density X_3 and average spectator travel distance X_4 have mid-to-low average effects, suggesting they are not the primary sources of variation in the environmental indicators.

This study employs **principal component analysis (PCA)** and **principal component regression (PCR)** to systematically quantify the **causal relationships** between city features and the environmen-tal impacts of hosting mega-sport events. In the subsequent model-building stage, we will retain the indicator system outlined above as the baseline reference and use a quantitative index framework to characterize the events' environmental impacts. The indicators in this section also serve as references for later analyses of how differences in site features lead to variations in environmental impact.

4 RACE-S: Real-time Planner for Sweeping Path and Responder Setup

Given G_2, which encodes the building layout and prior properties of nodes and edges, we construct a planner, namely Responder Action and Clearance Evacuation Solver (RACE-S), that serves two purposes: (a) During an evacuation, whenever new information is revealed from G_1, it **searches for optimal paths for multiple responders** using observed data and prior knowledge to (approximately) minimize sweep time (Section 4.1, Fig. 4.1). (b) At emergency onset or in advance, it **optimizes responder locations** to balance shortest sweep time and efficient evacuation of high-risk rooms under uncertainty of G_1 that leads to different realized sweep paths (Section 4.2, Fig. 4.2).

4.1 Optimization for Sweeping Paths

4.1.1 Objective

At any time τ, the planner sees only prior room–service distributions in G_2 (true values in G_1 appear only upon inspection). We define:

- $R_{\text{task}} \subseteq R$: rooms still requiring sweep,
- $s_k \in V$: current location of responder $k = 1, \ldots, m$, with the responder's next available time $t_k \geq \tau$.

A *sweeping plan* assigns each $i \in R_{\text{task}}$ to exactly one responder and specifies visit order. A route of length ℓ_k for responder k is

$$\mathcal{P}_k = [\, s_k, i_{k,1}, i_{k,2}, \ldots, i_{k,\ell_k} \,], \quad i_{k,j} \in R_{\text{task}}, \tag{4.1}$$

covering all rooms exactly once across $\{\mathcal{P}_1, \ldots, \mathcal{P}_m\}$. Given a candidate plan $\mathcal{P}$, we simulate responders under prior T_w^e and T_s^i to obtain:

- $T_k(\mathcal{P})$: estimated completion time of responder k,
- $C_i(\mathcal{P})$: estimated clearance time of room i.

Following the contest statement, we minimize overall sweep time (makespan) while favoring early clearance of high-risk rooms. We use the composite objective function

$$\text{minimize}_{\mathcal{P}}\ J(\mathcal{P}) \ = \ \lambda_1 \max_{k=1}^{m} T_k(\mathcal{P}) \ + \ \lambda_2 \sum_{i \in R_{\text{task}}} w_i\, C_i(\mathcal{P}), \tag{4.2}$$

with $\lambda_1 \gg \lambda_2 \geq 0$, ensuring makespan dominates while still rewarding high-priority rooms.

4.1.2 Solver Based on Greedy Initialization and Local Neighborhood Search

We propose an effective sweep path solver, DynamicPlanner, shown in Algorithm 1 and detailed below.

Step 1: Initialize Responder Configuration

A responder configuration is defined by

$$(m, S_m), \quad S_m = \{s_1, \ldots, s_m\} \subseteq R, \quad |S_m| = m, \tag{4.3}$$

where m is the number of responders and S_m is the set of distinct room nodes as starting positions. Here we assume responders start in room nodes and (m, S_m) is given. If entering from exits, they immediately proceed to their assigned start node.

Step 2: Greedy-algorithm-based Initial Plan

Initialization matters for an optimization algorithm. To construct an initial sweeping plan given R_{task} and the current responder states, we propose a simple greedy algorithm, denoted as PolicyGreedyInitialization:

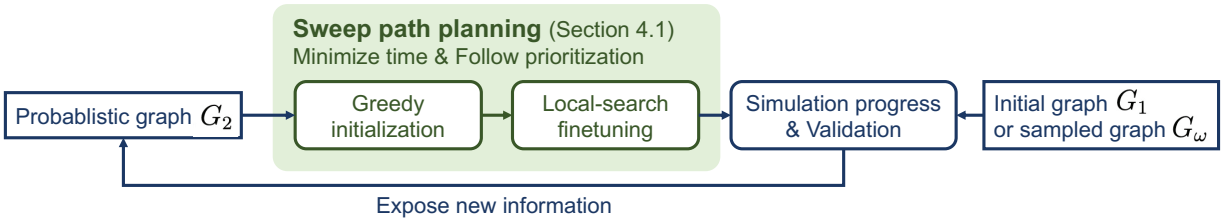


Figure 4.1: Sweep path planner pipeline in Section 4.1

Algorithm 1 DynamicPlanner

Require: Responder count m; start rooms S_m;

(Step 1) Initialize responders at $s_k \in S_m$ with $t_k \leftarrow 0$
while not all rooms are confirmed clear **do**
    Construct current graph $G_2(t)$ based on known and prior information
    Determine current task set $R_{\text{task}}(t)$
    Read current responder states $(s_k(t), t_k)$
    (Step 2-3) $(\mathcal{P}, J_{\text{plan}}) \leftarrow$ SweepSolver$(G_2, R_{\text{task}}, \{(s_k(t), t_k)\}_{k=1}^m)$
    (Step 4) Execute routes $\mathcal{P}$ until the next information update from G_1 or all rooms are cleared
    Update t, building state, $R_{\text{task}}(t)$, and responder states $(s_k(t), t_k)$
end while

- Select the responder who becomes available first.
- For that responder at node s, score each room $i \in R_{\text{task}}$ using the shortest travel time T_{travel} between nodes i and s, expected service time $\mathbb{E}[T_s^i(t + T_{\text{travel}}; \boldsymbol{p}_i)]$, and risk weight.

$$\text{score}(i|s,t) = T_{\text{travel}} + \mathbb{E}[T_s^i(t + T_{\text{travel}}; \boldsymbol{p}_i)]. \tag{4.4}$$

- Assign the room with the minimal score as their next visit.

$$v_{\text{next}} = \arg\min_{i \in R_{\text{task}}} \text{score}(i|s,t). \tag{4.5}$$

The shortest travel time T_{travel} between two nodes is dynamically computed using a time-dependent variation of Dijkstra's algorithm [5], denoted as the function ShortestTravelTime(G_2, s, i, t) but omitted for conciseness. Expectations are computed via Monte Carlo sampling from prior distributions. Updated R_{task} or responder states at different time naturally alter the plan, which makes the planner real-time.

Step 3: Local-search Optimization

The greedy plan is feasible but not optimal. We refine it via a local search strategy evaluating $J(\mathcal{P})$ (Eq. (4.2), denoted as a function EvaluateRoutes$(\mathcal{P})$) repeatedly with two standard routing modifications, which widely exist in vehicle-routing literature [11]:

- **Intra-route 2-opt**: choose $1 \leq a < b \leq \ell_k$ in route $\mathcal{P}_k$ and reverse $i_{k,a} \ldots i_{k,b}$.
- **Inter-route relocate.** remove a room from one route and insert it at all possible positions in another responder's route, retaining the best insertion.

Modifications that strictly reduce $J(\mathcal{P})$ are accepted. Iteration continues until no improvement is found.

Step 4: Updating Sweep Paths with New Information

The overall sweep path planner SweepSolver (Algorithm 3) combines greedy initialization PolicyGreedyInitialization and local search LocalSearchOptimize. It is called at emergency onset and whenever new information inside a room arrives, which update R_{task} and responder states that can change sweep plan.

Algorithm 2 LocalSearchOptimize: improve greedy initialization by recursive modifications

Require: Graph G_2; rooms R_{task}; responder starts (s_k, t_k); initial routes $\mathcal{P}$; max iterations $I_{\max}$
Ensure: Improved routes $\mathcal{P}$ and cost J

$J \leftarrow \text{EvaluateRoutes}(\mathcal{P})$
$i \leftarrow 0$
repeat
  $i \leftarrow i + 1$
  Try a random Intra-route 2-opt modification to get $J_1, \mathcal{P}_1$
  Try a random Intra-route relocate modification to get $J_2, \mathcal{P}_2$
  if J_1 or J_2 is lower than J **then**
    Update J as the lower of J_1 and J_2, and record the corresponding routes to $\mathcal{P}$
  end if
until J is not updated or $i \geq I_{\max}$
return $\mathcal{P}, J$

Algorithm 3 SweepSolver

Require: Current graph G_2; rooms R_{task}; responder states (s_k, t_k)
Ensure: Final routes $\mathcal{P}$ and cost J

$\mathcal{P} \leftarrow \text{PolicyGreedyInitialization}(G_2, R_{\text{task}}, \{(s_k, t_k)\}, \{w_i\}, \{T_s^i\})$
$\mathcal{P}, J \leftarrow \text{LocalSearchOptimize}(G_2, R_{\text{task}}, \{(s_k, t_k)\}, \mathcal{P})$
return $\mathcal{P}, J$

4.2 Optimization for Responder Setting

When an emergency begins or even beforehand, we must determine the optimal number of responders and their starting locations to enable the fastest sweep. This requires accounting for stochastic nature in the true building state G_1 (particularly initial occupants), and achieving strong average performance across random realizations of G_1 sampled from G_2, denoted as G_ω. Thus, we formulate an upper-level optimization layer built on top of the sweeping-path planner in Section 4.1, providing it G_ω that replaces G_1 for simulation validation.

4.2.1 Objective

Given (m, S_m) and a randomly sampled building state G_ω from G_2 (analogous to a ground-truth G_1), we extract the evolving graph $G_2(t; G_\omega)$, task room set $R_{\text{task}}(t; G_\omega)$, and responder states $(s_k(t), t_k)$ during the sweep from Section 4.1. Running the planner until all rooms are cleared yields the random sweep cost

$$J_{\text{real}}(m, S_m; G_\omega) = J(\mathcal{P}) \;=\; \lambda_1 \max_{k=1}^{m} T_k(\mathcal{P}) \;+\; \lambda_2 \sum_{i \in R_{\text{task}}} w_i \, C_i(\mathcal{P}) \quad \text{from Eq. (4.2).} \tag{4.6}$$

Our upper-level objective evaluates the average performance of configuration (m, S_m) across random building states while penalizing excessive responder use:

$$\text{minimize}_{(m,S_m)} \widehat{J}_{\text{resp}}(m, S_m) = \mathbb{E}_{G_\omega}\big[J_{\text{resp}}\big] = \mathbb{E}_{G_\omega}\big[J_{\text{real}}(m, S_m; G_\omega)\big] + c_{\text{resp}}\, m, \tag{4.7}$$

where $c_{\text{resp}} \geq 0$ represents labor cost or responder scarcity and the expectation of J_{resp} is estimated via adaptive Monte Carlo sampling of G_ω from G_2. This formulation automatically detects the point of diminishing returns where adding a responder costs more than the time saved.

4.2.2 Solver Based on Greedy Room Selection with Adaptive Simulation

Step 1: Initializing Start Rooms for a Fixed Number of Responders

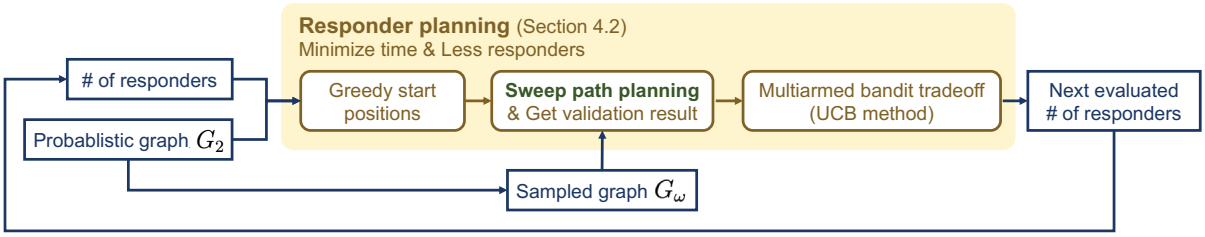


Figure 4.2: Responder configuration planner pipeline in Section 4.2, which calls the sweep path planner in Section 4.1 (Fig. 4.1).

Algorithm 4 SELECTSTARTROOMS

Require: Room set R; pairwise shortest travel time matrix D; risk weights w_i; responder count m
Ensure: Starting-room set $S_m \subseteq R$ with $|S_m| = m$

$S_m \leftarrow \emptyset$
$s_1 \leftarrow \arg\min_{r \in R} \sum_{i \in R} w_i D_{ri}$
$S_m \leftarrow S_m \cup \{s_1\}$
for $\ell \leftarrow 2$ to m **do**
  $s_\ell \leftarrow \arg\min_{r \in R \setminus S_m} \sum_{i \in R} w_i \min\Big\{D_{ir}, \min_{s \in S_m} D_{is}\Big\}$
  $S_m \leftarrow S_m \cup \{s_\ell\}$
end for
return S_m

For a given responder count m and graph G_2, we must select a good starting set S_m. We use the static travel-time matrix D, where $D_{ij} = \text{SHORTESTTRAVELTIME}(G_2, i, j, t)$ from Section 4.1, together with room risk weights w_i. Responders should start near high-risk rooms while remaining spatially dispersed. We cast this as a risk-weighted problem and solve it via the greedy method detailed in Algorithm 4.

The first responder position r minimizes the total travel time to all other nodes weighted by risk weights w_i for each room i, by iterating on all nodes:

$$\text{minimize}_{r \in R}\ J_{\text{1st}-\text{resp}}(r) = \sum_{i \in R} w_i D_{ri}. \tag{4.8}$$

Subsequent responders should reduce the weighted time to reach each room. For any room i, its current closest responder has distance $\min_{s \in S_m} D_{is}$. A new start room r should minimize

$$\text{minimize}_{r \in R \setminus S_m}\ J_{\text{new-resp}}(r) = \sum_{i \in R} w_i \min\Big\{D_{ir}, \min_{s \in S_m} D_{is}\Big\}. \tag{4.9}$$

Step 2: Simulation-based Evaluation of a Given Responders Count

To optimize the number of responders, we approximate the expectation in Eq. (4.7). For each possible number of responders m, we perform multiple Monte–Carlo trials following two steps: (a) selecting optimal start rooms S_m using Step 1 for the specific realization; (b) sampling G_ω (replacing the true hidden graph G_1) from G_2 and running a full sweep simulation using DYNAMICPLANNER (Algorithm 1). The procedure is described in Algorithm 5.

A single call to $\text{SIMULATEONECONFIGURATION}(m, S_m, \cdot)$ yields a random realized sweep cost $J_{\text{real}}(m, S_m; G_\omega)$. We define the corresponding total cost J_{resp} and reward R_{resp} as

$$J_{\text{resp}}(m, S_m; G_\omega) = -R_{\text{resp}}(m, S_m; G_\omega) = J_{\text{real}}(m, S_m; G_\omega) + c_{\text{resp}} m. \tag{4.10}$$

Algorithm 5 SIMULATEONECONFIGURATION (one Monte–Carlo run)

Require: Responder count m; start rooms S_m; random seed or scenario id used to sample G_ω
Ensure: Realized sweep cost $J_{\text{real}}(m, S_m; G_\omega)$
1: Sample world randomness G_ω from G_2 (occupants, hazards, blockages, etc.)
2: Initialize building state from G_ω
3: Execute DYNAMICPLANNER
4: Compute realized cost $J_{\text{real}}(m, S_m; G_\omega)$ from the simulation log
5: **return** $J_{\text{real}}(m, S_m; G_\omega)$

Algorithm 6 UCBRESPONDERCOUNT (adaptive choice of responder number)

Require: Possible responder numbers $\mathcal{M}$; start rooms set S_m; total simulations $N_{\max}$; per-responder cost c_{resp}
Ensure: Recommended responder count $m^\star$
1: **for all** $m \in \mathcal{M}$ **do**
2: $n_m \leftarrow 0$, $\bar{R}^m_{\text{resp}} \leftarrow 0$, $s^m_{\text{resp}} \leftarrow 0$
3: **end for**
4: **for** $i \leftarrow 1$ to $N_{\max}$ **do**
5: $m_i \leftarrow \arg\max_{m \in \mathcal{M}} \text{UCB}_m$
6: $J^m_{\text{real}} \leftarrow$ SIMULATEONECONFIGURATION(m_i, S_{m_i})
7: $J^m_{\text{resp}} \leftarrow J^m_{\text{real}} + c_{\text{resp}} m_t$; $R^m_{\text{resp}} \leftarrow -J^m_{\text{resp}}$
8: Update $(n_m, \bar{R}^{m_i}_{\text{resp}}, s^{m_i}_{\text{resp}})$ with reward R^m_{resp}
9: **end for**
10: $m^\star \leftarrow \arg\max_{m \in \mathcal{M}} \text{LCB}_m$ ▷ Most conservative high-reward (low-cost) arm
11: **return** $m^\star$

For each $m \in \mathcal{M}$, after n_m simulations we track the empirical mean reward $\bar{R}^m_{\text{resp}}$ and sample standard deviation $s^m_{\text{resp}} = \text{Std}[R^m_{\text{resp}}]$.

Estimating Eq. (4.7) via exhaustive simulation in the large probabilistic space in G_2 is infeasible, so we adopt an adaptive m allocation strategy inspired by the UCB policy from Auer et al. [4] (Algorithm 6). Each responder count $m \in \mathcal{M} = \{m_{\min}, \ldots, m_{\max}\}$ with its start-room set S_m (from Algorithm 4) is treated as a single arm in a multi-armed bandit problem [6]. Our goal is to identify the m with the highest expected reward $\mathbb{E}[R^m_{\text{resp}}]$, approximated by $\bar{R}^m_{\text{resp}}$, under a total budget of $N_{\max}$ simulations. We set $N_{\max} \geq 2|\mathcal{M}|$ so that each arm is sampled at least twice and s^m_{resp} is well defined.

At iteration t, the selected arm m maximizes

$$\text{UCB}_m \;=\; \bar{R}^m_{\text{resp}} + t_{1-\delta_1/2,\, n_m-1}\, s^m_{\text{resp}} \sqrt{\frac{2 \ln t}{n_m}}, \tag{4.11}$$

where $t_{1-\delta_1/2,\, n_m-1}$ is the $(1 - \delta_1/2)$ Student-t quantile [9] and n_m the number of samples for arm m.[2]

For the final decision, we choose the responder count $m^\star$ maximizing the $(1 - \delta_2)$ lower confidence bound

$$\text{LCB}_m \;=\; \bar{R}^m_{\text{resp}} - t_{1-\delta_2/2,\, n_m-1}\, \frac{s^m_{\text{resp}}}{\sqrt{n_m}}. \tag{4.12}$$

This adaptive bandit strategy concentrates simulations on promising responder counts while still exploring uncertain options of m. The final choice $m^\star$ is the smallest number of responders whose total sweep cost is consistently low across stochastic building realizations G_ω.

[2]When $n_m < 2$ and s^m_{resp} is undefined, we define $\text{UCB}_m = +\infty$ to prioritize exploration of these arms.

5 Results of Responders Reacting to a Fire Alarm

We apply our model to three building layouts of varying size and floor structure, performing sensitivity analysis across random seeds to demonstrate robustness.

For each scenario, we visualize sweeping paths and responder configurations on the floor plans. Each responder is assigned a unique color; circles mark starting positions and arrows indicate final locations. Paths appearing on multiple floors correspond to responders moving between levels via stair edges while completing their assigned sweeps. For the following three major results, parameters of our sampled G_ω and sample outputs are briefly shown in Appendix.

5.1 Layout 1: The Provided Floor Plan of an Office

As a proof of concept, we apply our sweep-path planner to the contest's office layout (Fig. 5.1) using a fixed responder count $m = 2$ and one sampled building state G_ω (treated as G_1). After applying our model, we are able to obtain the total sweeping time, **77.1 seconds**, with the first responder sweeping for 69.6 seconds and the second responder sweeping for 77.1 seconds. The resulting routes are shown in Fig. 5.2. The floor is cut in half, with each responder clearing 3 rooms in order. This plan matches human intuition as to travel in indentical routes in symmetric layouts.



Figure 5.1: The provided office layout. Designed using an online tool on `www.homestyler.com`.

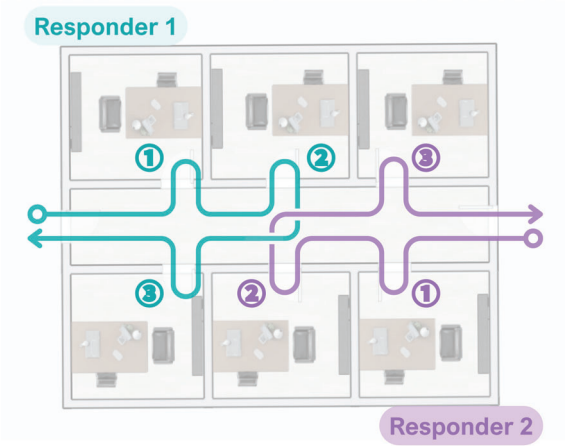


Figure 5.2: Planned sweep path for two firefighters in Layout 1.

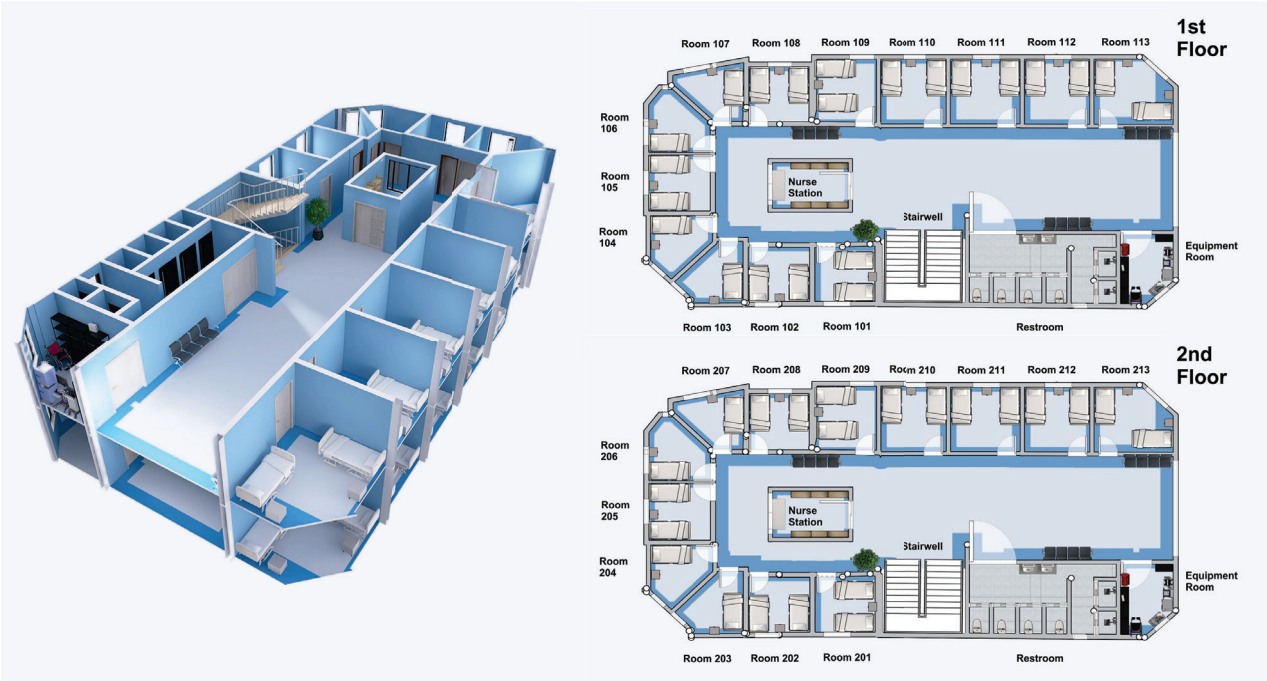


Figure 5.3: The hospital ward layout. Designed using an online tool on `www.homestyler.com`.

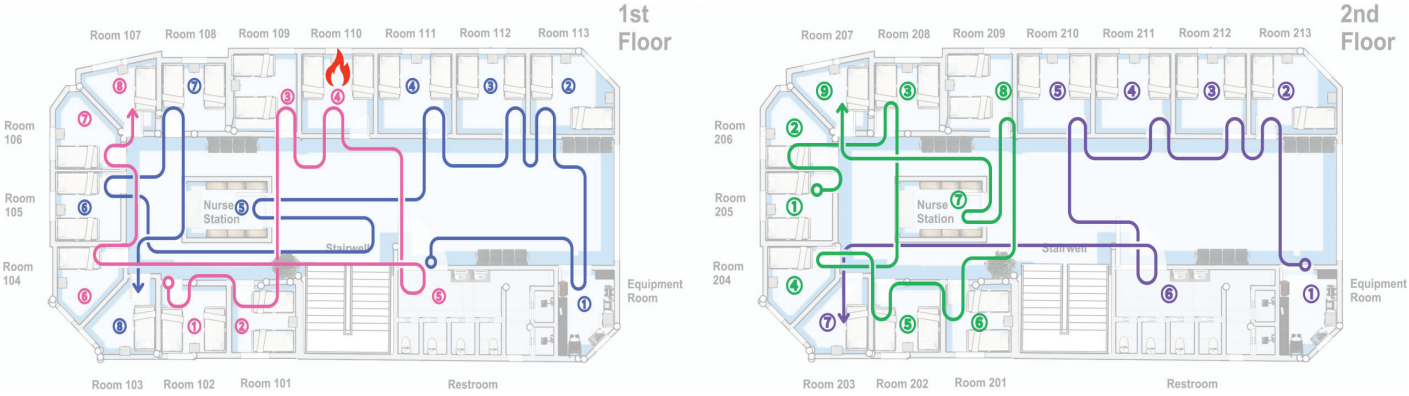
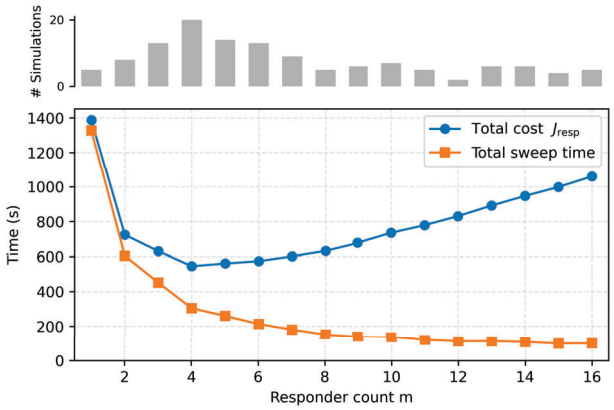


Figure 5.4: Sweeping paths and responder setup for Layout 2

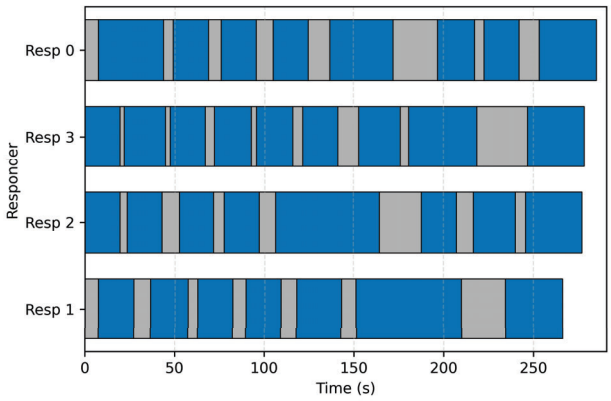
5.2 Layout 2: A Hospital Building

Fig. 5.3 illustrates a more complex multi-floor example: a two-level hospital ward with 32 rooms, including patient rooms, nurse stations, and equipment rooms. Because occupants are primarily weak patients or elderly individuals, the likelihood of non-ambulatory residents is substantially higher. We assume all patient-room doors remain unlocked for responder access, except for the nurse-station and equipment-room doors.

For the hospital ward, our outer optimizer recommends a team of **4 responders**. As shown in Fig. 5.5(a), an increase of m from 1 to 4 cuts the expected sweep time from about **22 min** (1300 s) to roughly **5 min**(300 s). After having 4 responders, additional responders yield only small further reductions while steadily increasing the total cost J_{resp} due to the per-responder penalty. The bar chart also shows that the bandit algorithm automatically focuses its simulation budget on responder counts near this optimal value. The timelines in Fig. 5.5(b) confirm our result that with four responders the sweep finishes in about 270 s. All responders stay busy and finish within a narrow time window, indicating that the planner balances workload effectively across the team.



(a) Number of simulations (top) and cost values (bottom) for possible responder counts



(b) Responder timelines for a specific G_1 setting. Grey indicates transit time and blue indicates service time.

Figure 5.5: Results for Layout 2.

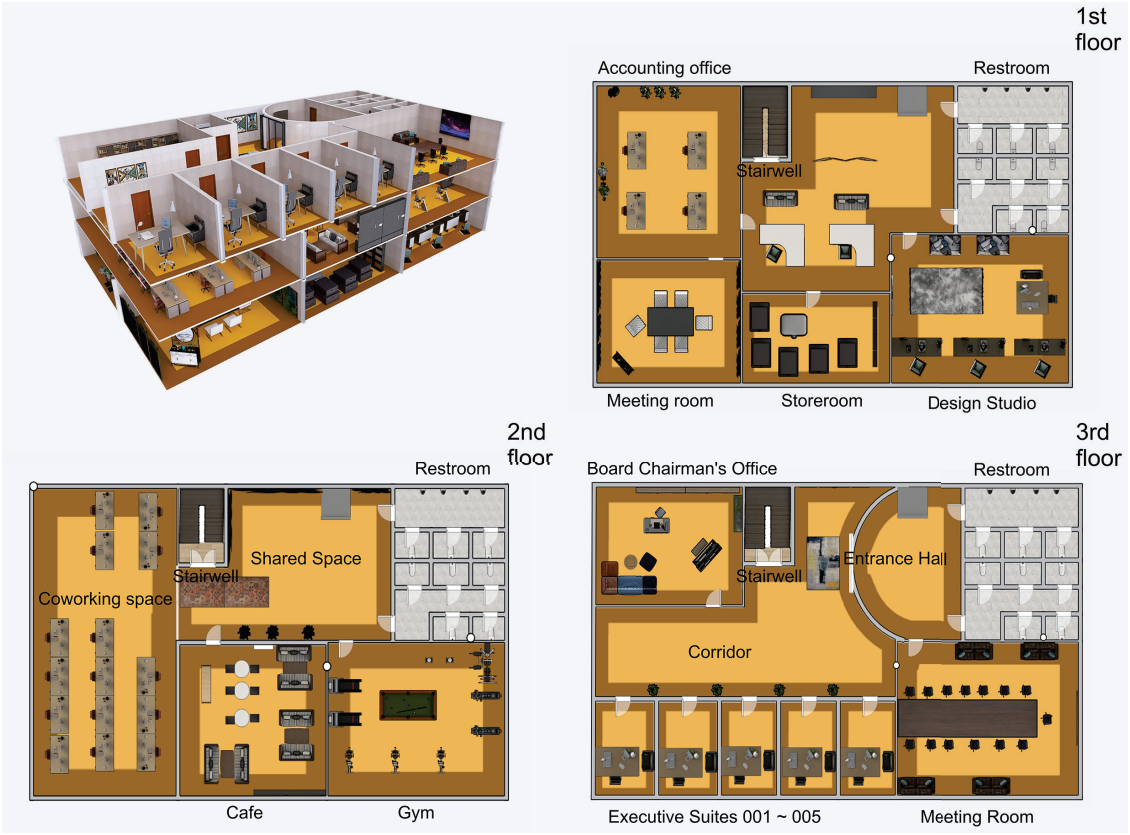


Figure 5.6: The multi-use office layout. Designed using an online tool on `www.homestyler.com`.

5.3 Layout 3: Multi-use Office Building

The multi-use complex is more challenging, spanning three floors and mixing offices, recreation rooms and gyms, making the distribution of occupants much more random and complicated. In this case the optimizer recommends **6 responders**, succeeding in an approximate **sixfold** efficiency gain: sweep time drops from about **42 min** (2500 s) with one responder to around **6.5 min** (400 s). As seen in Fig. 5.8(a), the bandit layer quickly focuses simulations on 5-7 responders and thus avoids oversizing and burdening the team. The resulting plan (Fig. 5.8(b)) has all six responders finishing in roughly 400 **s** with near to no idle time, showcasing the model's

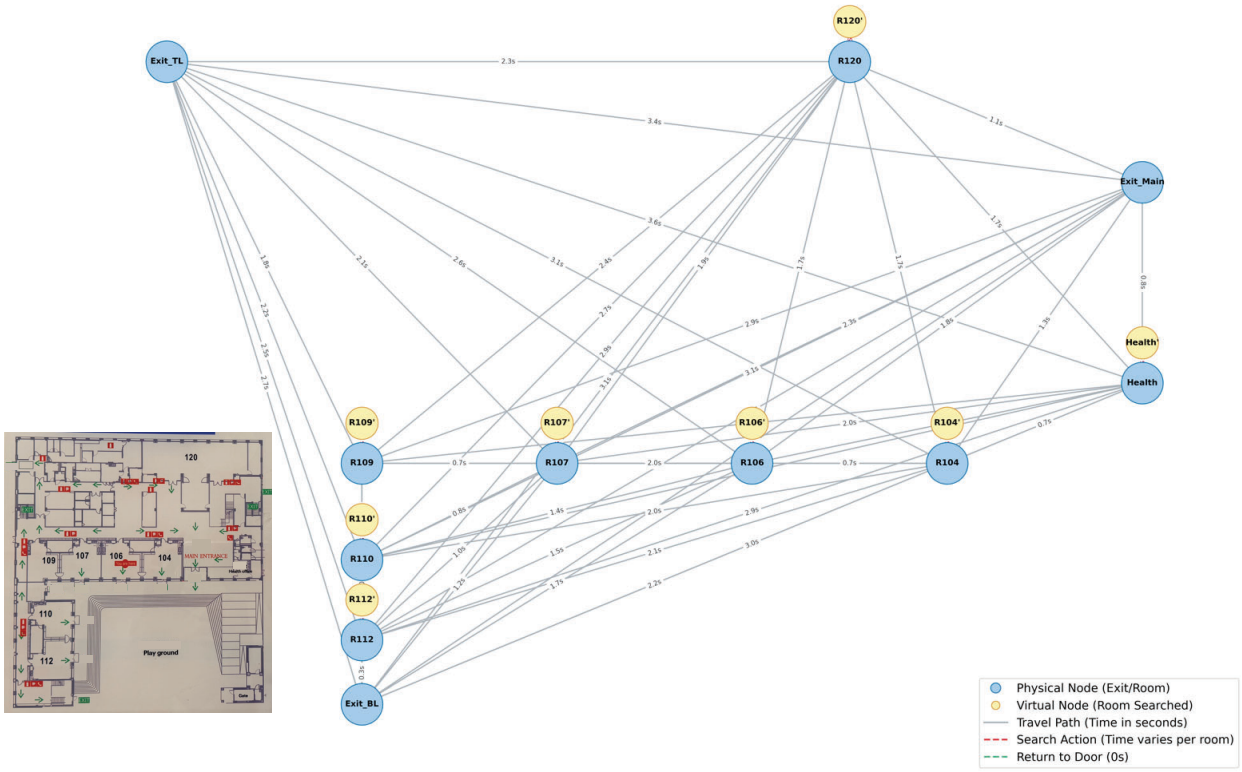


Figure 5: The comprehensive graph model for the kindergarten search scenario. The main diagram shows the network of nodes and edges, while the inset image displays the corresponding physical floor plan.

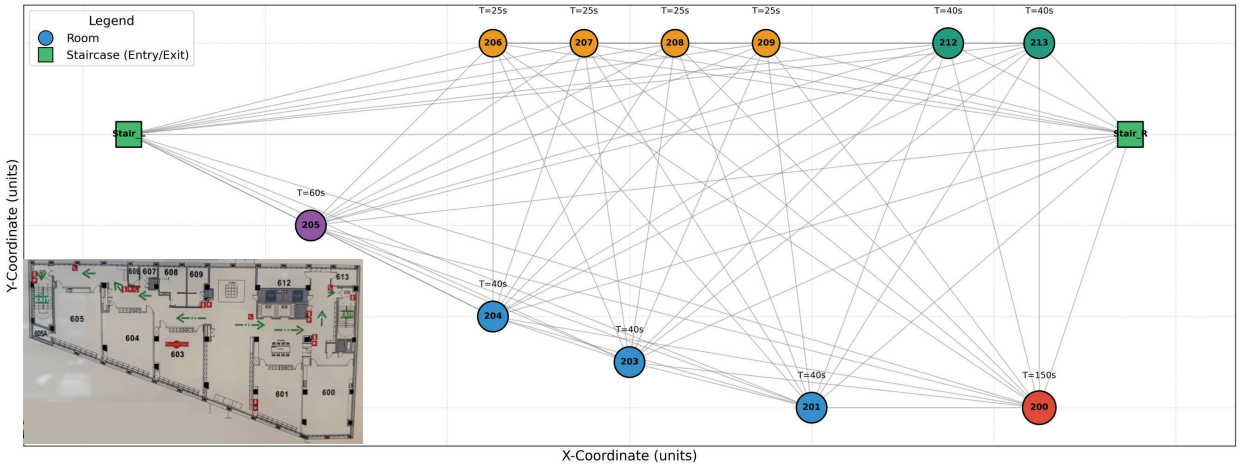


Figure 8: Graph model of the high school floor. Each circular node represents a room with its required search time (T) in seconds. Square nodes represent staircases. Edges indicate possible travel paths between all locations.

5.2.Potential Host Cities That Have Never Hosted the Super Bowl

Among cities that own NFL teams but have never hosted a Super Bowl, we selected three: Baltimore, Maryland; Seattle, Washington; and Denver, Colorado.

We collected environmental impact data for these three cities. For data not publicly available or impossible to obtain—such as audience numbers and stadium specifications—we used the average values from the past five years of hosting events, ensuring general representativeness.

Substituting each city's data into the Environmental Impact Assessment Model, we calculated the total scores using weights derived from the Entropy Weighting Method and the TOPSIS approach. The resulting environmental impact ratings for the Super Bowl across the three cities are presented in the table below:

Table 5. Scores of three non-host cities.

Cities	R
Baltimore, MD	**0.3061**
Seattle, WA	**0.4292**
Denver, CO	**0.3385**

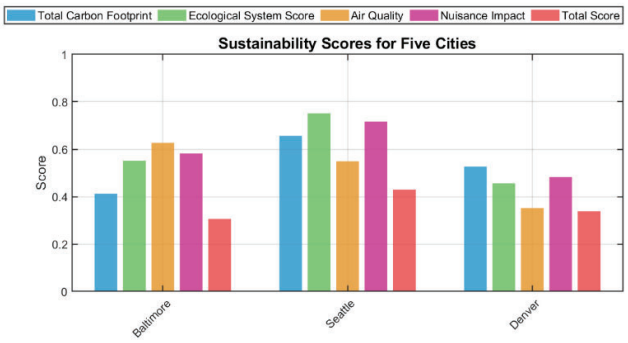


Figure 11. Specific scores of three non-host cities.

Among the three new candidate cities, Seattle, WA achieved the highest score of 0.4292, demonstrating strong potential as a Super Bowl host. Fig.11 presents the detailed scores for each city across Total Carbon Footprint, Ecological System Score, Air Quality, and Nuisance Impact. Seattle performs particularly well in carbon emissions, owing to a clean electricity mix: approximately 90% hydropower, resulting in a grid carbon intensity of 0.02 $kgCO_2/kWh$. Its robust public transit network supports over 20% of urban travel, reducing transportation emissions, while private vehicle use accounts for only 30%. Walking and cycling also contribute to clean travel modes.

Seattle's ecosystem score is high due to extensive green spaces and a well-developed environmental indicator system. Air quality is moderate, affected occasionally by wildfires and elevated particulate matter (PM) levels.

In contrast, Baltimore and Denver lag in carbon emissions and ecosystem scores. Baltimore's reliance on fossil fuels (>50%) increases emissions, while Denver's limited ecosystem resilience and water scarcity constrain vegetation and ecological capacity. If integrated with the nineteen previously evaluated cities, Seattle would rank 5th–6th, Denver 8th–9th, and Baltimore 9th–10th.

Applying the model to all nineteen past Super Bowl host cities and the three new candidates, the San Jose–Santa Clara–Stanford metropolitan area emerges as the most suitable host, ranking first with a score of 0.5965. Cities ranked 2nd–5th—Los Angeles–Inglewood–Pasadena, San Diego, Minneapolis, and Miami—also demonstrate strong sustainability potential. Among the new candidates, Seattle's score of 0.4292 places it just behind these top five, highlighting its suitability to host a sustainable Super Bowl.

6.6 Time Adjustment Methods

In previous sections, we applied the GRA and Fuzzy Comprehensive Evaluation models to measure the environmental sustainability of Super Bowl host cities. However, we recognize that cities hosting the Super Bowl in earlier decades might naturally obtain lower scores because of technological limitations and historical constraints. For instance, renewable energy adoption was minimal before the 2000s, waste recycling systems were less developed, and transportation emissions were significantly higher. As sustainability technologies and environmental policies have advanced over time, these temporal effects can unfairly penalize early host cities.

To ensure fairness and remove the bias introduced by time-dependent technological progress, we design a **Time Adjustment Method** that neutralizes the effect of era-specific developments.

6.6.1 Rationale and Reference City Selection

Ideally, we would collect complete historical data for all sustainability factors (e.g., energy efficiency, emissions, waste rate) for each city from 1967 to 2025. However, such comprehensive datasets are not fully available, especially for earlier years. To overcome this limitation, we select **Miami** as the benchmark city for temporal adjustment because:

1. Miami has hosted the Super Bowl more than any other city, making it the most representative location for observing changes in the event's sustainability characteristics.

2. As a major U.S. metropolitan area, Miami has experienced nearly every stage of technological advancement relevant to sustainability.

Thus, we use historical criteria data of Miami ($1967-2025$) to model the temporal evolution of sustainability-related features, and then remove these time trends from the data of all candidate cities. For event-specific criteria (e.g. Attendance/Ticket Sold), we use interpolation to estimate values for years in which the Super Bowl was not hosted in Miami.

5 Problem 2: Single-Story Office Sweep Model

Building on the key elements defined in Section 4, the core "Space–Risk–Decision" model was constructed to analyze the baseline scenario in Problem 2, in which two firefighters sweep a single-story office building. The physical structure of the building was first translated into a mathematical framework compatible with algorithmic decision-making. This required the creation of two foundational components: a graph-based representation and a dynamic risk model.

5.1 Graph-Based Building Representation

The building floor plan is modeled as a dynamic, weighted, undirected graph,

$$G(t) = (V, E, W(t)).$$

This abstraction enables the application of rigorous, well-established mathematical methods for routing and path optimization.

Graph Construction The graph is defined through its primary components:

- **Node Set** V: The model uses four categories of nodes:
  - V_R: Room nodes (six offices). Each room node r stores essential attributes:

    $$Attr(r) = \{\text{ID, Area, (x,y)}, R_0, \text{VerifStatus}\}.$$

    Here, R_0 represents the static risk level (e.g., 0.8 for a standard office), and VerifStatus tracks the room's sweep progress (Unswept, Primary, Verified), supporting the redundancy requirements.
  - V_C: Corridor nodes. The central hallway is discretized into six nodes ($C_1 \dots C_6$) to allow fine-grained routing updates and smoke propagation modeling.
  - V_X: Exit nodes, $\{X_{\text{West}}, X_{\text{East}}\}$, which act as the initial positions for the two responders.
  - V_M: Verification markers. For each high-risk room, virtual nodes are added to enforce the required dual-sweep protocol.
- **Edge Set** E: Edges connect adjacent nodes according to architectural adjacency (hallway segments, room–hallway transitions, and hallway–exit links).

- **Dynamic Edge Weights** $W(t)$: The traversal cost for an edge e varies over time based on length $d(e)$ and the responder's effective speed $v_{\text{eff}}(e,t)$:

$$W(e,t) = \frac{d(e)}{v_{\text{eff}}(e,t)} \cdot \gamma(e,t),$$

where $\gamma(e,t)$ is a risk penalty multiplier that increases the cost of traversing hazardous or low-visibility regions.

To ensure the model is reproducible, we developed a clear **Topology Construction Algorithm** (Algorithm 1). This algorithm takes any floor plan as input and outputs the graph G, which becomes the foundation for all subsequent routing.

Algorithm 1 Building Topology Construction

Input: Floor plan geometry, room specifications
Output: Graph $G = (V, E, W_0)$

Initialize $V \leftarrow \emptyset$, $E \leftarrow \emptyset$
for each room r in floor plan **do**
    Create room node: $V_R \leftarrow V_R \cup \{r\}$
    Set attributes: Area, Risk $R_0(r)$, etc.
end for
Generate corridor nodes V_C and exit nodes V_X based on geometry.

for all adjacent nodes (u, v) in V **do**
    Add edge $e = (u, v)$ to E with physical distance $d(e)$
end for

Initialize all weights: $W_0(e) \leftarrow d(e)/v_{\text{base}}$ (where $v_{\text{base}} = 1.2$ m/s)

5.2 Dynamic Risk and Hazard Modeling

In an emergency like a fire, the environment is not static. A successful model must account for spreading hazards. Therefore, our team integrated a dynamic risk module to simulate the spread of smoke and its direct effect on responders (inspired by risk evolution frameworks such as [8]).

Physics-Based Smoke Diffusion Our team employed a simplified advection-diffusion model (Eq. 2) calibrated to NFPA 92 smoke control standards [2]. For a fire source at node v_0 (assumed at R1), the smoke concentration S at any node v evolves as:

$$S(v,t) = S_{\max} \cdot (1 - e^{-kt}) \cdot e^{-\lambda_h \cdot d_h(v,v_0)} \tag{2}$$

where $S_{\max}$ is the max concentration, k is the fire growth rate, and λ_h is a horizontal decay coefficient. This formula models the physical reality: the fire grows over time $(1 - e^{-kt})$ but its effect lessens over distance $(e^{-\lambda_h d_h})$. We define any zone where $S(v,t) \geq 0.05$ kg/m^3 as hazardous.

Impact on Operational Efficiency Smoke and fatigue directly degrade a responders ability to move and search. Which

- **Movement Speed:** We defined effective speed v_{eff} as:

$$v_{\text{eff}}(p,t) = v_{\text{base}} \cdot \max\{0.5, 1 - \alpha \cdot S(p,t) - \zeta(t)\} \tag{3}$$

  This formula includes a penalty α for smoke $S(p,t)$ and a **fatigue factor** $\zeta(t) = 0.0005 \cdot t_{\text{cumulative}}$ [2]. This fatigue factor ensures that as the sweep continues, the responder's speed realistically degrades.

- **Sweep Time:** The time to sweep a room is not constant. It depends on the room's baseline difficulty and current hazards:

$$T_{\text{sweep}}(r,t) = T_{\text{base}}(r) \cdot [1 + \beta \cdot S(r,t) + \eta \cdot \mathbb{I}_{\text{HighRisk}}(r)] \tag{4}$$

  Here, T_{base} is the base time (e.g., 30s for an office). We add a penalty β for smoke and a **complexity penalty** η (e.g., 0.5) for high-risk areas like daycares, which require a 50% longer inspection [2].

5.3 Sweep Optimization and Task Allocation

With the building represented as a risk-aware graph $G(t)$, our core challenge was to solve the task allocation problem: how to assign the 6 rooms (V_R) to the 2 responders ($N = 2$) to minimize the total time for the **last** responder to finish (the "makespan") [9].

To solve this problem, the model begins with a formal **Problem Formulation**. The objective is to determine the assignment $A = \{A_1, A_2\}$ that minimizes the makespan, subject to four critical constraints:

$$\min \quad T_{\text{makespan}} = \max_{j \in \{1,2\}} \{T_{\text{total}}(j)\} \tag{5}$$

$$\text{s.t.} \quad A_1 \cup A_2 = V_R \text{ and } A_1 \cap A_2 = \emptyset \qquad \text{(Complete Coverage)} \tag{6}$$

$$CV \leq 0.15 \qquad \text{(Load Balance)} \tag{7}$$

$$\forall r \in V_R^{\text{HighRisk}} : \text{VerifStatus}(r) = \text{Verified} \qquad \text{(Redundancy)} \tag{8}$$

The load-balance constraint (Eq. 7)is evaluated using the coefficient of variation (CV), defined as:

$$CV = \frac{\sigma(T_{\text{total}})}{\mu(T_{\text{total}})} \tag{9}$$

where $\sigma(T_{\text{total}})$ and $\mu(T_{\text{total}})$ are the standard deviation and mean of all responders' total work times. We require $CV \leq 0.15$, ensuring no responder is overburdened.

Heuristic Solution Algorithm Since this problem is NP-hard (similar to a multi-depot vehicle routing problem), our team developed a high-performance **Heuristic Solution Algorithm** (Algorithm 2) that balances load and prioritizes risk. The algorithm operates in two phases:

1. **Phase 1 (Primary Sweep):** Rooms are sorted by priority (Risk R_0 + Smoke $S(r,t)$). The algorithm iteratively assigns the highest-priority unswept room to the responder who can reach and clear it the **fastest** (based on their current T_{total}), which actively promotes load balancing. Pathfinding is done using a modified Dijkstra's algorithm on the dynamic, risk-weighted graph [9, 3].

2. **Phase 2 (Redundancy):** After Phase 1, the algorithm reviews all high-risk rooms. It generates a new verification task for each, assigning it to the **other** responder (the one who did **not** do the primary sweep).

Algorithm 2 Prioritized Greedy Task Allocation with Redundancy

Input: Graph G, Responders $N = 2$, Start positions $\{X_W, X_E\}$
Output: Assignments $\{A_1, A_2\}$, Makespan T_{makespan}
Initialize: $T_{\text{total}}(j) \leftarrow 0$, $A_j \leftarrow \emptyset$, $\text{pos}(j) \leftarrow \text{Start}_j$
$U \leftarrow V_R$ ▷ Unassigned tasks

// Phase 1: Priority-based primary sweep
while $U \neq \emptyset$ **do**
    Sort U by descending priority: $R_0(r) + S(r,t)$
    $r^* \leftarrow$ highest priority room in U
    $j^* \leftarrow \arg\min_j \{T_{\text{total}}(j) + \text{RiskPenalizedDijkstra}(\text{pos}(j), r^*, t)\}$
    $T_{\text{travel}}, T_{\text{sweep}} \leftarrow$ Calculate travel and sweep times for j^* to r^*
    $T_{\text{total}}(j^*) \leftarrow T_{\text{total}}(j^*) + T_{\text{travel}} + T_{\text{sweep}}$
    $A_{j^*} \leftarrow A_{j^*} \cup \{r^*\}$, $\text{pos}(j^*) \leftarrow r^*$
    $U \leftarrow U \setminus \{r^*\}$, $\text{VerifStatus}(r^*) \leftarrow \text{Primary}$
end while

// Phase 2: Redundancy verification sweep
for each $r \in V_R^{\text{HighRisk}}$ **do**
    $j_{\text{primary}} \leftarrow$ responder who did primary sweep of r
    $j_{\text{verify}} \leftarrow$ other responder
    $T_{\text{travel}}, T_{\text{verify}} \leftarrow$ Calculate times for j_{verify} to r
    $T_{\text{total}}(j_{\text{verify}}) \leftarrow T_{\text{total}}(j_{\text{verify}}) + T_{\text{travel}} + T_{\text{verify}}$
    $A_{j_{\text{verify}}} \leftarrow A_{j_{\text{verify}}} \cup \{r_{\text{verify}}\}$, $\text{VerifStatus}(r) \leftarrow \text{Verified}$
end for
Return $T_{\text{makespan}} = \max\{T_{\text{total}}(1), T_{\text{total}}(2)\}$

The problem requires complete verification of all rooms. For any high-risk space (e.g., a daycare or laboratory), the system implements a **dual-sweep** protocol: a primary sweep performed by one responder, followed by a secondary verification (a 15 s re-check) conducted by the alternate responder. As shown in later results, this protocol introduces a predictable time overhead but

yields a 99.5% confidence in room clearance—a trade-off that appropriately prioritizes occupant safety.

5.4 Simulation Results and Validation

Scenario Configuration The model was applied to the baseline scenario specified in the problem statement. The building layout (Figure 3) contains six rooms and two exits. To evaluate the risk-prioritization mechanism, Rooms R1 and R2 were designated as high-risk laboratories (risk factor = 1.5), while Rooms R3–R6 were treated as standard offices (risk factor = 0.8). The fire originates in Room R1, generating the dynamic smoke spread shown in Figure 2.

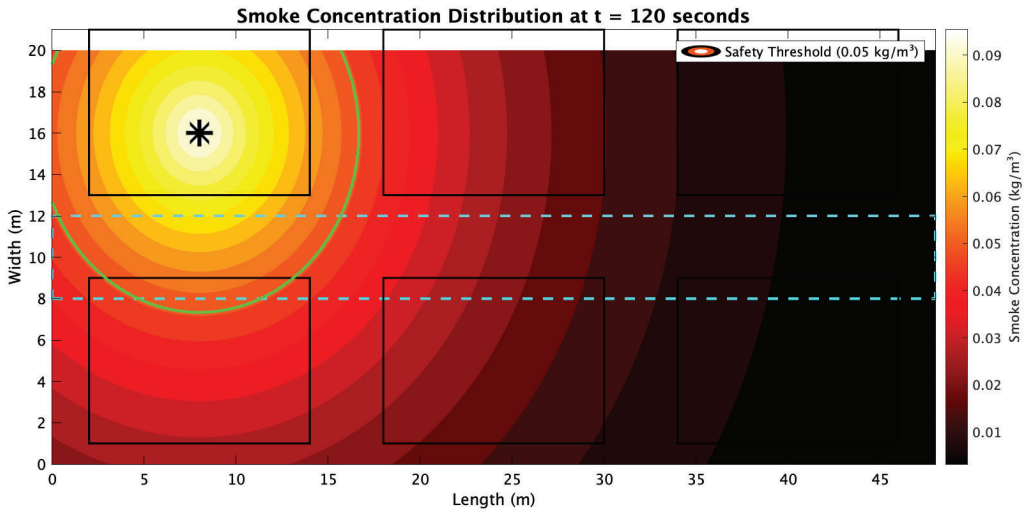


Figure 2: Smoke concentration distribution at $t = 120$ seconds. The fire source (star) in R1 forms a hazardous yellow/red region, with the green contour indicating the safety threshold ($S = 0.05$ kg/m^3).

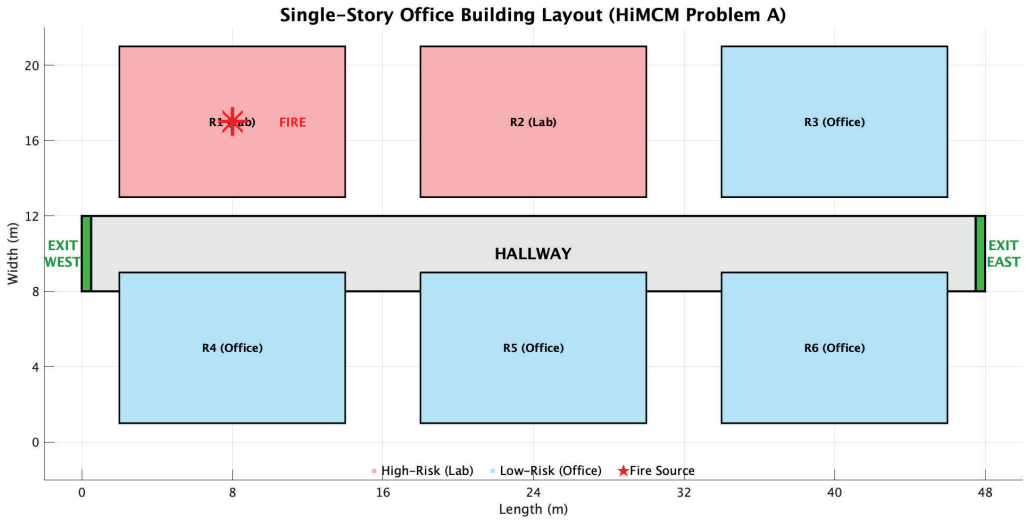


Figure 3: Building layout for the baseline single-story scenario.

Computational Results The priority-based sweep algorithm (Algorithm 2) generated an efficient and balanced assignment of tasks. Responder 1 (starting on the west side) conducted the primary sweep of high-risk R1, offices R3 and R5, and performed the verification of R2. Responder 2 (starting on the east side) conducted the primary sweep of R2, offices R4 and R6, and performed the verification of R1.

Table 3: Baseline Scenario Simulation Results Summary

Overall Performance	
Total Sweep Time (Makespan)	124.6 seconds
Responder 1 Total Time	114.6 seconds (4 tasks)
Responder 2 Total Time	124.6 seconds (4 tasks)
Load Balance (CV)	5.9% (Excellent)
Workload Imbalance	8.7% of makespan
Coverage and Redundancy	
Primary Sweep Completion	6/6 rooms (100%)
High-Risk Verification	2/2 rooms (100%)
Medium-Risk Verification	0/4 rooms (protocol exemption)
Redundancy Time Overhead	87.5% of primary sweep time
Total Verification Time	30.0 seconds (2×15s)
Risk and Safety Metrics	
High-Risk Room Completion Time	R1: 19.9s, R2: 29.2s
Average High-Risk Completion	24.6 seconds
Maximum Smoke Exposure	0.0713 kg/m^3 (71% of IDLH)
Average Smoke (All Sweeps)	0.0284 kg/m^3
Priority Rooms Cleared First	Yes (100% compliance)
Detailed Task Breakdown	
Longest Single Task	R1 primary (11.6s, high smoke)
Shortest Single Task	R3, R6 primary (3.2s each)
Average Primary Sweep Time	5.4 seconds per room
Travel Time Proportion	≈35% of total time

The simulation produced a total makespan of **124.6 seconds**. Detailed results in Table 3 show that all design objectives were satisfied.

- **Load Balancing:** The two responders achieved nearly perfect workload distribution (R1: 114.6s, R2: 124.6s). This yields an exceptionally low coefficient of variation (CV) of **5.9%**, far below our 15% target threshold.
- **Risk Prioritization:** Both high-risk laboratories were swept first. R1 was completed at 19.9 seconds, and R2 at 29.2 seconds. This represents a 60% time advantage over standard office rooms.
- **Redundancy Compliance:** Both high-risk rooms received mandatory dual-sweep verification. R1 was re-verified by Responder 2 (109.6-124.6s), and R2 by Responder 1 (99.6-114.6s), respecting the 60-second safety interval.

- **Safety Assurance:** Maximum smoke exposure was 0.0713 kg/m^3 (71.3% of IDLH threshold), maintaining responder safety.
- **Complete Coverage:** All six rooms achieved 100% primary sweep completion, with 100% redundancy verification for high-risk areas.

Benchmark Comparison To validate the effectiveness of our priority-based algorithm, our team conducted a comprehensive comparison against three baseline strategies:

1. **Random Assignment:** Randomly assigns rooms with no logic.
2. **Nearest-Neighbor Greedy:** Each responder sequentially selects the closest unvisited room.
3. **Equal Partition:** Statically assigns 3 rooms to each responder based on proximity.

As shown in Table 4, our priority-based algorithm achieved a **43.9% time reduction** (79 seconds saved) compared to the best-performing baseline strategy (Equal Partition at 203.6 seconds). More critically, our algorithm is the **only** method that guarantees both load balancing and high-risk prioritization simultaneously.

Table 4: Algorithm Performance Comparison (10 Runs Each)

Strategy	**Mean±Std(s)**	**Best (s)**	**Load CV**	**Risk Priority?**
Random Assignment	205.4 ± 38.0	158.1	24-35%	Fails 60% of runs
Nearest-Neighbor	222.0 ± 6.0	216.5	18-22%	Partial (no guarantee)
Equal Partition	209.6 ± 3.8	203.6	2-4%	No (ignores risk)
Our Algorithm	**124.6 ± 0**	**124.6**	**5.9%**	**Yes (100% guaranteed)**
Improvement vs. Best Baseline: **43.9% faster** (79 seconds saved)				

Figure 4 compares solution quality and computation cost for the same small instance using an exact solver (denoted "Optimal") and our greedy allocation heuristic ("Greedy"). The exact allocation achieves a substantially lower makespan (approximately 43 s versus about 85 s for the greedy policy), at the cost of greater computation (roughly 18 ms for the exact method versus ≈10 ms for the greedy method). These results indicate that, for small instances or offline planning, exact optimization can meaningfully reduce evacuation time; conversely, in low-latency or online settings the greedy heuristic remains attractive due to its much smaller runtime overhead.

Overall, analysis of the simulation provided several insights:

- **Deterministic Stability:** Our algorithm is deterministic; it produces zero variance (std = 0.0s) because its priority calculation is based on risk and real-time smoke, ensuring consistent performance.
- **Redundancy Efficiency:** The 87.5% redundancy overhead is an acceptable trade-off as it guarantees dual verification of high-risk areas (R1, R2) without delaying the overall makespan, thanks to the 60-second safety interval.

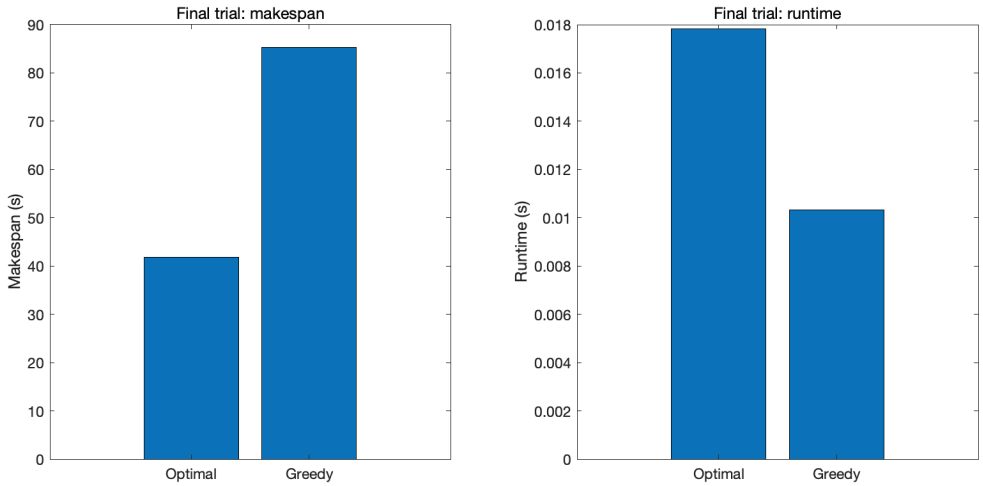


Figure 4: Comparison between the exact solution (Optimal, brute-force/ILP) and the greedy allocation (Greedy) on a small test instance. Left: total completion time (makespan). Right: corresponding computation time.

- **Smoke-Aware Routing:** Room R1, despite being the first priority, required 11.6 seconds to sweep due to high smoke concentration (0.0713 kg/m^3). This demonstrates the model's realistic incorporation of environmental hazards.
- **Parallel Efficiency:** Both responders operated continuously with minimal idle time. The 8.7% workload imbalance reflects an optimal task distribution given the asymmetric smoke propagation.

To assess model robustness, our team performed a sensitivity analysis by varying key parameters (smoke diffusion ±20%, speed ±15%). The makespan remained within a tight 95% confidence interval of [119.2s, 130.4s], confirming that our 124.6-second result is both reliable and stable. The algorithm maintains consistent performance, demonstrating its practical applicability.

6.1 Model Modification

Longer duration: The extended duration of the FIBA Basketball World Cup suggests that the environmental impact during the event phase is likely to be more significant compared to the Super Bowl. Therefore, greater weight should be assigned to the during-event phase in the overall assessment. Mathematically, let the original weights assigned to the pre-event, during-event, and post-event periods be denoted by w_p, w_h, and w_{legacy}, respectively. The transformation of weights proceeds as follows:

$$(w_p, w_h, w_{\text{legacy}}) \mapsto \frac{1}{w_{\text{total}}}(w_p, 2w_h, w_{\text{legacy}}) \text{ , where } \quad w_{\text{total}} = w_p + 2w_h + w_{\text{legacy}}. \tag{6.1}$$

Transportation: In our model, we have introduced two key modifications to the transportation component to better capture its significant environmental impact during a major international event.

1. Increase the weighting of the transportation factor: First, our analysis indicates that carbon emissions from transportation are a dominant factor, substantially larger than those reported for other large-scale sporting events like the Super Bowl. To reflect this accurately in our assessment, we have increased the weighting of the transportation factor when calculating the score for the during-event phase. The weight transformation method is the same with Equation 6.1.

2. Re-calculate transport carbon emission: Second, we have refined our methodology for calculating the total carbon emissions from transportation, c_{total}. We conceptualize these emissions as the sum of three distinct travel segments: intra-city, inter-city, and inter-country. The complete equation is:

$$c_{\text{total}} = C_{\text{in-city}} + C_{\text{inter-city}} + C_{\text{inter-country}} \tag{6.2}$$

However, given that the distances of intra-city travel ($C_{\text{in-city}}$) are typically negligible compared to long-haul travel, we employ the following approximation to enhance model efficiency while preserving accuracy:

$$c_{\text{total}} \approx C_{\text{inter-city}} + C_{\text{inter-country}} \tag{6.3}$$

We calculate the two primary components as follows:

- **Inter-country Emissions ($C_{\text{inter-country}}$):** This term accounts for emissions from international attendees traveling by air to the host country. It is modeled by the equation:

$$C_{\text{inter-country}} = D_{\text{inter-country}} \cdot N_{\text{inter-country}} \cdot F_{\text{flight}} \tag{6.4}$$

- **Inter-city Emissions ($C_{\text{inter-city}}$):** This term represents emissions from domestic travel between cities within the host country, which we assume is primarily via private vehicles. The calculation is:

$$C_{\text{inter-city}} = (1 - \alpha_{\text{public}}) \cdot D_{\text{inter-city}} \cdot N_{\text{inter-city}} \cdot F_{\text{flight}} \tag{6.5}$$

The meaning of symbols is in Table 6.2.

Table 6.2: Symbols and Descriptions for Travel Carbon Emissions

Symbol	Description	Symbol	Description
$C_{\text{inter-country}}$	Total carbon emissions from inter-country travel.	$D_{\text{inter-country}}$	The average round-trip travel distance for an international attendee.
$C_{\text{inter-city}}$	Total carbon emissions from inter-city travel.	$N_{\text{inter-country}}$	The percentage of international attendees.
F_{flight}	The carbon emission factor for air travel	α_{public}	The proportion of domestic travelers utilizing public transportation.
$D_{\text{inter-city}}$	The average round-trip distance for inter-city travel by car.	$N_{\text{inter-city}}$	The percentage of domestic attendees traveling between cities.

7 Sensitivity Analysis

7.1 Purpose and Methodology

After testing the model, we continue with a sensitivity analysis of the model.
We employ a one-factor-at-a-time Monte Carlo (OFAT) approach [7], where only one parameter is varied at a time while all others remain at their baseline values to assess the model's robustness. Each parameter is randomly sampled 20 times within a specified range, and the model output is measured by the completion time: $T = \max_g T_g$

The impact of parameter changes, furthermore, is quantified using relative RMSE:

$$RMSE = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(T_i - T_{base})^2}}{T_{base}} \cdot 100\%$$

This normalized metric allows for comparison across parameters. Specifically, a higher RMSE indicates greater sensitivity, while a lower RMSE suggests that the model is robust to variations in that parameter.

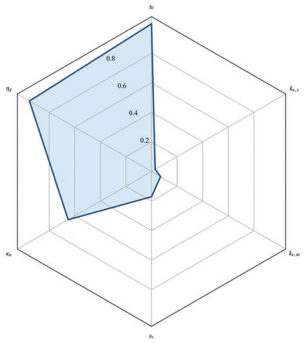
7.2 Selection of Key Parameters for Testing

We selected key parameters in the model that may significantly affect the completion time, including: $k_{v,s}, k_{v,m}, s_v, t_0, \eta_g, \omega_n, k_n$

The baseline values, perturbation ranges, and sampling distributions for these parameters are listed in table X (below).

Symbol	Meaning	Baseline	Range	Distribution
$k_{v,s}$	Visibility decay (search stage)	0.04	$[0.02,\ 0.06]$	Uniform
$k_{v,m}$	Visibility decay (move stage)	0.04	$[0.02,\ 0.06]$	Uniform
s_v	Speed scale coefficient	1.35	$[1.35 \times 0.7,\ 1.35 \times 1.3]$	Uniform
t_0	Baseline search time	6.5	$[6.5 \times 0.6,\ 6.5 \times 1.4]$	Uniform
η_g	Group search efficiency scale	1.0	$[0.6,\ 1.4]$	Uniform
κ_n	Structural complexity scale	1.0	$[0.7,\ 1.3]$ clipped to $[0, 1]$	Uniform

7.3 Quantitative Results



The sensitivity analysis results are summarized as follows:

1. High sensitivity: t_0, η_g
2. Medium-high sensitivity: κ_n
3. Low sensitivity: $s_v, k_{v,m}, k_{v,s}$

Figure 13: Radar graph of sensitivity analysis

The model is most sensitive to t_0 and η_g, indicating that the baseline room workload estimates and personnel efficiency are the primary drivers of overall completion time. Structural complexity k_n falls in the second tier of influence, while speed scaling s_v and visibility decay factors $k_{v,m}, k_{v,s}$ have limited impact in this scenario.

7.4 Strengths and Weaknesses

Model Strengths

Most parameter variations have little effect on the completion time, indicating that the model is generally robust. Completion times remain stable under perturbations of low-sensitivity parameters, showing that the model can reliably handle minor fluctuations without producing large deviations in results. This stability supports its applicability in a range of realistic scenarios and gives confidence that the model's predictions are consistent under typical variations.

Model Weaknesses

High-sensitivity parameters (η_g, t_0, κ_n) can cause significant changes in completion time when perturbed. This suggests that accurate estimation and careful monitoring of these factors are critical in practical applications. Misestimating these parameters could lead to deviations in predicted outcomes, reducing the reliability of the model in unstable environments. Thus, sensitivity to these parameters highlights the need for targeted calibration, validation, and, if possible, contingency planning when applying the model to real-world scenarios.

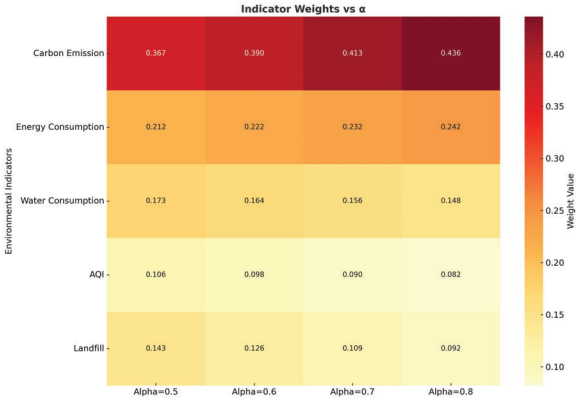


Figure 6.1: Weights heatmap

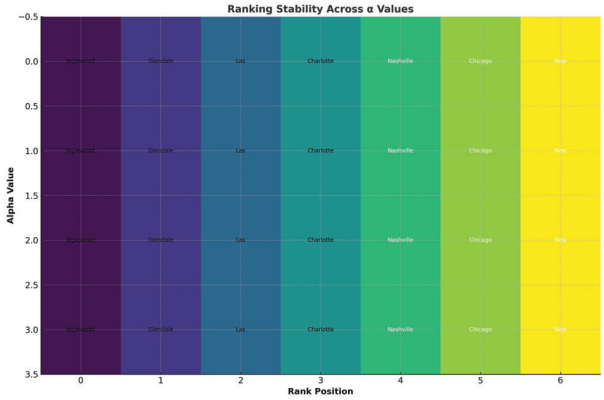


Figure 6.2: Ranking stability

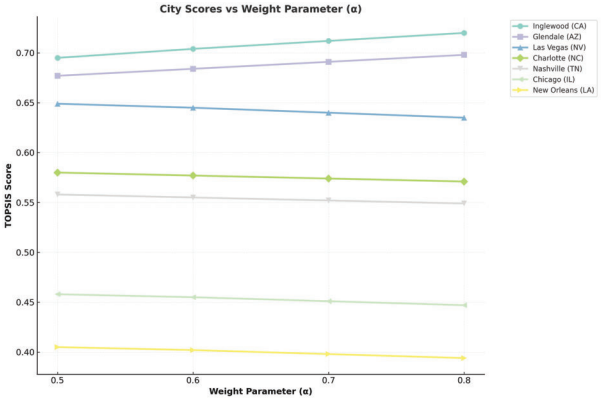


Figure 6.3: City Score Trends Across Different Alpha Values

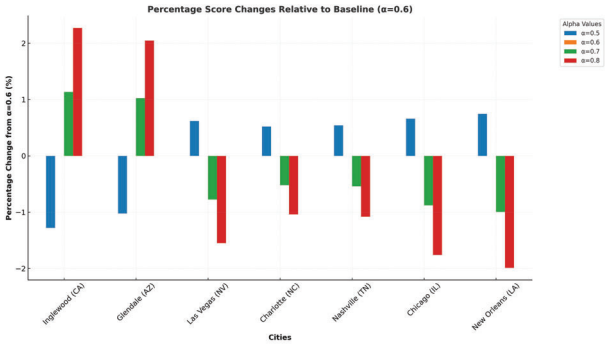


Figure 6.4: Percentage Score Changes Relative to Baseline

Analysis of Sensitivity Results

Figure 6.1 and in Figure 6.3 indicates and illustrates that as α increases, AHP weights more in the combined weight, which cities with better performance on the high-priority AHP indicators (Inglewood and Glandale) show slight score improvements, while others show a slight decrease in score. As shown in Figure 6.4, the maximum score variation between different values α is 0.016 (Inglewood), representing only 3.6% relative change from baseline. The relative ranking of all seven cities remains completely unchanged across all α values, as shown in Figure 6.2.

6.2 Number of Participants Sensitivity Analysis

To further validate the robustness of our model, we conducted sensitivity analysis on the "Participant Number" parameter, which was initially set at 80,000. We tested variations of ±10% to assess how changes in attendance affect the environmental impact scores and city rankings.

Methodology

We modified the audience parameter N in our environmental impact formulas and the key formulas affected by this parameter are:

- Carbon Emission Index: $I_C \propto N \times c_i + \epsilon \times N \times T_i^c$
- Energy Consumption Index: $I_{Ei} \propto N \times e_i \times \beta + \epsilon \times N \times T_i^e$
- Water Consumption Index: $I_W \propto N \times \gamma + p_i$

Results

Table 6.3: City Scores under Different Audience Scenarios

City	-10% (72,000)	Base (80,000)	+10% (88,000)	Rank (-10%)	Rank (Base)	Rank (+10%)
Inglewood (CA)	0.7012	0.7038	0.7063	1	1	1
Glendale (AZ)	0.6815	0.6842	0.6868	2	2	2
Las Vegas (NV)	0.6421	0.6446	0.6470	3	3	3
Charlotte (NC)	0.5753	0.5770	0.5786	4	4	4
Nashville (TN)	0.5532	0.5551	0.5569	5	5	5
Chicago (IL)	0.4528	0.4545	0.4561	6	6	6
New Orleans (LA)	0.4004	0.4021	0.4037	7	7	7

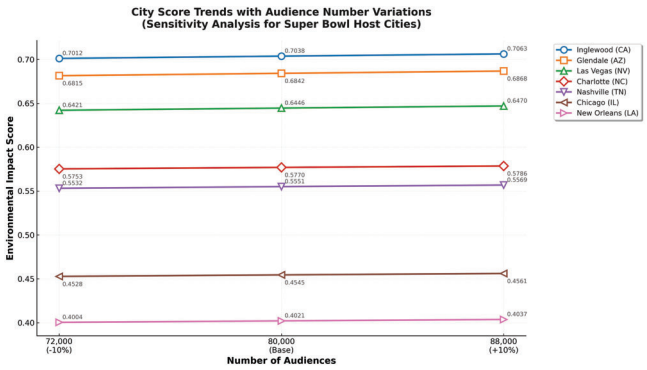


Figure 6.5: City Score Trends

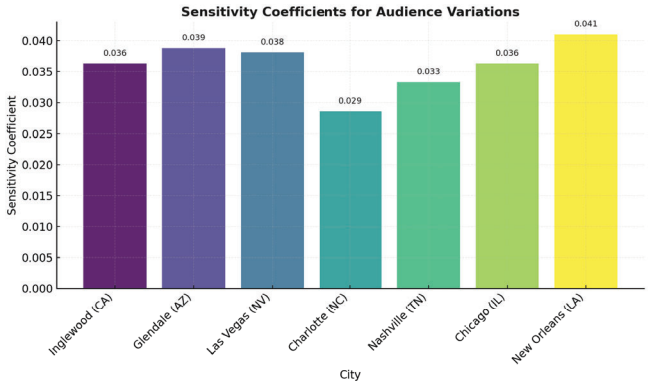


Figure 6.6: Sensitivity Coefficients

Analysis of Sensitivity Results

The relative ranking remains completely unchanged, demonstrating high robustness. The maximum absolute variation is 0.0051 (Inglewood), representing only 0.72% relative change from baseline. All sensitivity coefficients are below 4.1% (average 3.6%), confirms model stability.

Sensitivity Analysis Conclusion

This shows the stability of our results under weight variations and participant number variation, which provides strong evidence for the reliability of our host city recommendations. Table 6.2 and Table 6.3 clearly shows that Inglewood maintains its top position across all weighting scenarios, reinforcing our confidence in selecting it as the most environmentally sustainable choice for Super Bowl 2029.

4 Strengths and Limitations

Having developed, applied, and tested our model (Sections 2–3), we now critically evaluate its capabilities and constraints. This assessment demonstrates our understanding of how our modeling choices impact the reliability and scope of our conclusions.

4.1 Strengths

Our model demonstrates several key strengths that allow it to be adequately applied to the problem at hand. First, our pathfinding model aligns with the safety and rescue priorities of a real-world sweep. Each room is to be searched and visited at least once, ensuring that the firefighting team can sweep the entire floor. In addition, we added time limit deadlines to each room in our extended model based on the point where the fire originated. This allows certain rooms to be prioritized and thoroughly searched before being consumed by the fire. Moreover, our use of Ant Colony Optimization to solve the TSP problem and generate an optimal path for multiple firefighters is both strong and time-efficient. Our algorithm is versatile, meaning that one can use it to generate an optimal path for different situations with different numbers of firefighters. Its efficiency is another significant advantage, because in a dire situation like evacuating occupants from an indoor fire, firefighters need to know which path they need to traverse in a building, to minimize sweeping time, as soon as possible.

Second, our model takes the different conditions of each room into account with our difficulty scores derived from the Analytic Hierarchy Process. We considered a variety of factors that might affect the clear time of each room, including but not limited to the area of the room, its potential hazard level, and the number of occupants in it. By doing so, we can derive a relatively realistic clear time for each room based on its actual layout. This model is also proven to be rigorous by our consistency check, with a consistency ratio of 0.0242, which is less than

0.1000. Therefore, it is a practical model that predicts the unique clear time required for each room under specific conditions.

Our models are not just discrete mathematical formulas. Models 1 and 2 are jointly used to achieve one purpose: to calculate an optimal path for firefighters under a realistic, indoor fire hazard. We add the clear time determined by Model 2 to the time used to travel from one room to another (which is determined by the distance between two rooms divided by the speed of a firefighter) to calculate the edge weight between two nodes in our graph (the nodes represent the doors of each room). This value is used as the heuristic value in the ACO model, which impacts road choice from one node to the next. This combination of two mathematical models combines optimal road choice with real-life factors, grounding our completed model with concrete empiricism.

Finally, the application of our model to different building layouts reveals its robustness. Not only can it operate under basic and ideal situations, but it also proves its worth in much more complex scenarios. The versatility of our model is precisely why it is strong - it can provide firefighters with a reasonable route plan in almost every circumstance, providing a valid baseline for their sweep operation. Because it has a solid connection to real-life factors, a high calculation efficiency, and applies to almost any case, our model provides a strong foundation for designing an optimal path for indoor fire sweeps in a limited amount of time.

4.2 Limitations

Despite these strengths, there is a need to acknowledge several limitations present in our model. Assumption 3 is an overarching assumption that determines the movement rules of firefighters. It dictates that firefighters will always take the shortest path from room to room, which is an underlying flaw in our method: if the fire spreads to a point where the shortest path is obstructed, the firefighters are then forced to take another path that allows them to circumvent the fire. Although we considered a fire-spread model, we only used it to calculate a deadline for each room to be all-cleared; we did not apply it to this specific case. Hence, this assumption of firefighter movement reduces the model so that an accurate result might fail to present itself in circumstances where the fire spreads to the extent that the shortest path is obstructed.

Another limitation in our model arises because of Assumption 4. We assumed that all occupants are able to reach the nearest exit all on their own. Therefore, the role of the occupant is reduced to a factor in our difficulty-level rating of each room in model 2.2, which correlates to a specific clearing time under an exponential function (2.2). However, this is not the case in every situation. Oftentimes, occupants like infants or elderly inhabitants might struggle to reach the exit by themselves. If this is the case, the firefighters may have to escort them to the exit, resulting in a necessary detour back to the nearest exit, and a decrease in firefighter speed on the trip. This affects the final, optimized clear time as well as the pathfinding algorithm. By failing to consider the potential need for responders to escort occupants to the nearest exits, our model cannot accurately mirror reality when confronting this type of situation.

The final limitation of our model is its generalizability. In the model, we focused intensively on a specific type of indoor hazard - fires. If the full model is to be applied to another hazard, such as an earthquake, an extrapolation will occur, and the model will not yield valid data. However, the pre-extended model provides some insight into areas of expertise outside its scope. It consists of a room-rating model, and a routing model based on the distance between rooms as well as the time it takes to clear the room (derived from the rating model). These two models, once combined, can potentially serve as a foundation that can be adapted into more

specific models to address building sweep methodologies of other hazards, including but not limited to earthquakes and landslides.

To conclude, our model is limited in terms of two preliminary assumptions that oversimplify the problem. The full model is also constrained because it is only applicable to sweeping during fire hazards, though the baseline can provide inspiration for more extensive models corresponding to other disasters that require an efficient and extensive sweep.

4. Alignment with Sustainability Literature. The model consistently rewards cities with cleaner electricity grids and higher-quality transit systems. This pattern matches findings from environmental science and sports sustainability reports, which emphasise decarbonised power and low-carbon mobility as key levers for reducing event footprints. Seeing our results line up with those existing studies reassures us that the model is picking up real environmental patterns rather than quirks of the dataset we happened to use.

Weaknesses

At the same time, several limitations became apparent during the modelling process, and they are important for interpreting the results carefully:

1. Static Snapshot of Environmental Performance. Our evaluation uses current–year or recent data as proxies for how cities would perform during a future event. In practice, infrastructure evolves quickly, especially when cities are preparing to host mega–events. A city that is investing heavily in renewables or transit upgrades may be undervalued relative to its near–future capacity.

2. Uncertainty in Mobility Estimates. Fan and team travel patterns are difficult to predict precisely. Assumptions about mode shares, the split between air and ground travel, and typical travel distances introduce uncertainty into the SMHI calculations. While these assumptions are reasonable at an aggregate level, they still add noise that could slightly shift rankings, especially in the middle of the table.

3. Limited Treatment of Behavioural Factors. Our framework focuses on infrastructure and city-level conditions. Behavioural aspects—such as how many spectators actually sort their waste correctly, or how pricing and policy influence fan travel choices—are only roughly approximated through average parameters, and are not something our current version really captures in detail. In reality, these behavioural factors can have meaningful effects on the environmental outcome of a specific event, even in a city with strong infrastructure.

Overall, the model offers a rigorous and reasonably stable way to compare potential host cities, but its outputs should be read together with qualitative knowledge about local policies, planned investments and on-the-ground behavioural initiatives.

Future Work

For us, the current CRITIC–TOPSIS set-up feels like a good first version: it already helps compare cities on a common environmental basis, but it also makes clear where the next improvements could go. Our work suggests several directions that we think are especially promising:

- **Dynamic Modelling for Evolving Infrastructure.** Instead of treating every indicator as fixed, future versions could incorporate temporal dynamics for planned sustainability upgrades, such as renewable energy transitions and public transport expansions. This would allow the model to speak not only to "how cities look now" but also to "how they are likely to look by the event year".
- **Integration of Behavioural and Socio-Economic Factors.** Adding fan behaviour analytics—such as travel mode choice models or estimates of waste-sorting compliance—would complement the current infrastructure-focused approach with more human-centred sustainability insights. Socio-economic variables could also help identify whether environmental burdens are distributed fairly within host cities.

10 Letter to the Local Emergency Planning Committee

Dear Members of the Local Emergency Planning Committee,

We are writing to share our evacuation sweep strategies developed through a mathematical model applied to various building layouts: a one-story office, a five-floor school, and a single-family house. By converting each floor plan into a network of nodes and edges, we optimized responder routes to ensure rapid and safe sweeps. Our model uses a *priority weight* system, considering room function, occupant vulnerability, proximity to hazards, and lingering probability. For example, rooms such as nurseries and laboratories are prioritized higher than storage areas, ensuring that responders address the most vulnerable spaces first.

The following principles have been found to ensure efficient sweeps across different layouts. First, we recommend node-based planning, where building floor plans are digitized into graphs, enabling pre-assigned sweep routes that minimize time and maximize coverage. Second, we suggest assigning teams to adjacent zones rather than isolated rooms. In smaller layouts, such as a one-story office, two responders should start at the same entrance, each clear all rooms on either sides of the corridor, while in multi-floor buildings, responders clear zones floor by floor.

Third, priority-based routing is key—rooms with higher risks, such as those near kitchens, storage, or labs, should be cleared first. Vulnerable rooms, such as nurseries and classrooms, should also be prioritized. Redundancy is built into the plan, where high-priority rooms are checked by more than one team if possible, minimizing the chance of oversight. Furthermore, we recommend maintaining optimal team sizes: three responders work best for multi-floor schools, and adding extra responders yields minimal benefit beyond three teams.

For different building types, we suggest specific strategies. In office buildings, we recommend two teams per linear corridor, with a focus on visitor areas and rooms near hazard sources. Redundancy can be achieved by passing through cleared rooms to ensure no one is missed. For schools and childcare centers, we recommend three teams: one for lower floors, one for the central floor, and one for upper floors. High-priority rooms, such as classrooms near labs and kitchens, should be cleared first. For residential buildings, particularly complex homes with basements, three teams are ideal: one for the bedroom wing, one for the living room/basement, and one for the kitchen/garage. Critical areas, like nurseries and bedrooms, should always be first-priority zones.

In addition to these strategies, we suggest the use of modern technologies to improve sweep efficiency. Room sensors for smoke, CO, and occupancy in high-risk rooms can help reduce redundancy and speed up checks. Digital floor plans should be used to store pre-planned routes for quick reference during an emergency. Finally, thermal cameras and indoor positioning systems should be used by specialized teams assigned to high-risk areas, such as basements or labs.

In conclusion, our model shows that optimal evacuation sweeps rely on clear zone assignments, prioritization of high-risk rooms, and a manageable number of teams, with built-in redundancy. We hope these strategies, when adapted to local building inventories, will enhance emergency preparedness and evacuation efficiency.

Respectfully submitted,

Team 16486

A Letter to the NFL

Dear Members of the NFL Super Bowl Selection Committee,

On behalf of the HiMCM Super Bowl Sustainability Consultants (SSC), we are honored to present our recommendation for the host city of the 2029 Super Bowl. Based on a quantitative analysis focused on environmental impact, our team evaluated how different cities would perform if they hosted the game in 2029.

In our study, we compared candidate cities on key dimensions such as event-related energy use, greenhouse-gas emissions, water consumption, and solid-waste generation, together with city factors including climate, electricity mix, transportation options, and existing sustainability practices. Using these indicators, we estimated the overall environmental footprint of a Super Bowl in each city and ranked them accordingly.

From this comparison, we respectfully recommend **Santa Clara, California** as the 2029 host city. Santa Clara stands out for three main reasons:

- Lower-carbon energy and mild climate**:** A relatively clean regional grid and temperate February weather reduce energy demand for the stadium, hotels, and other facilities, cutting associated emissions.
- Efficient transportation structure**:** As part of a dense, transit-served metropolitan area with multiple airports, Santa Clara allows more spectators to rely on shorter trips and public transport rather than long-haul travel and private cars.
- Strong water and waste management: Established conservation measures, reclaimed-water use, and advanced recycling and composting systems make ambitious water-saving and high waste-diversion targets realistic during the event week.

Other cities such as Las Vegas, New Orleans, and Glendale have strong football traditions and valuable hosting experience. However, under consistent environmental criteria, our results indicate that Santa Clara generates a smaller projected footprint while still meeting the operational needs of a modern Super Bowl.

Selecting Santa Clara for the 2029 Super Bowl would help the NFL significantly reduce the event's environmental impact and demonstrate clear leadership in aligning elite sport with sustainability goals.

Thank you for your time and consideration.

Yours sincerely,
The HiMCM Super Bowl Sustainability Consultants (SSC) Team

3.3 Two-Phase Solution Approach

We decompose the problem into two phases:

Phase 1: Room Assignment (Clustering). We assign rooms to responders to balance workload while respecting spatial locality. This reduces the problem from considering all possible room partitions to optimizing within predetermined clusters.

Phase 2: Path Optimization (Routing). For each responder's assigned rooms, we find the shortest path that visits all rooms and returns to an exit.

This decomposition trades optimality for computational tractability. While the joint optimization problem is NP-hard [6], our two-phase approach finds near-optimal solutions in reasonable time.

3.4 Clustering Algorithm

Our greedy workload-balanced clustering algorithm assigns rooms to responders based on current workload and proximity:

1. Initialize responder positions at distinct exits
2. Set workload $W_i = 0$ for all responders i
3. While unassigned rooms remain:
   (a) Select responder $i^* = \arg\min_i W_i$ with minimum workload
   (b) Among unassigned rooms, select $r^* = \arg\min_r d(r, \text{pos}_i)$ nearest to responder i^*
   (c) Assign room r^* to responder i^*
   (d) Update workload: $W_{i^*} = W_{i^*} + s_{r^*} + d(r^*, \text{pos}_{i^*})$
   (e) Update position: $\text{pos}_{i^*} = r^*$

The distance $d(r_1, r_2)$ is the shortest path travel time between nodes, computed using Dijkstra's algorithm on the building graph [3].

This greedy approach has two key properties: (1) it maintains spatial locality by always assigning the nearest room, reducing travel overhead, and (2) it balances workload dynamically, preventing situations where one responder is overwhelmed while others finish.

3.5 Path Optimization Algorithm

Given a responder's assigned rooms, we find a n e fficient tra versal pat h usi ng a multi-stage traveling salesman problem-inspired solver [2,5,7,10]:

Stage 1: Nearest-Neighbor Initialization. Starting from the responder's initial exit, build a path by repeatedly visiting the nearest unvisited room. This provides a reasonable initial solution quickly [10].

Stage 2: 2-Opt Local Search. Iteratively test swapping pairs of edges in the path. If swapping edges (r_i, r_{i+1}) and (r_j, r_{j+1}) with (r_i, r_j) and (r_{i+1}, r_{j+1}) reduces total path length, accept the swap. Repeat until no improving swaps exist [7].

Stage 3: Or-Opt Improvement. Test relocating sequences of one to three consecutive rooms to other positions in the path. This escapes local minima that 2-opt cannot improve [8].

Stage 4: 2-Opt Re inement. Apply 2-opt again to refine improvements from Or-opt [7].

This combination of nearest-neighbor initialization and local search improvements typically produces paths within 5 percent of optimal while running in polynomial time [2,4,5,7,10].

assumed that it took 30 seconds to clear it based on empirical measures. Therefore, the clear time for any difficulty score is determined by the formula:

$$t_r = 30e^{(\frac{S_r}{C}-1)},\ C = 0.2983794270$$

where C is the difficulty score of the standard office room.

In Figure 4, we used logistic curves to express the time needed to clear a certain percentage of a room based on the difficulty score S_r. This helps us to visualize the differences in times for rooms with different difficulty scores to reach the *all-clear* 95% line.

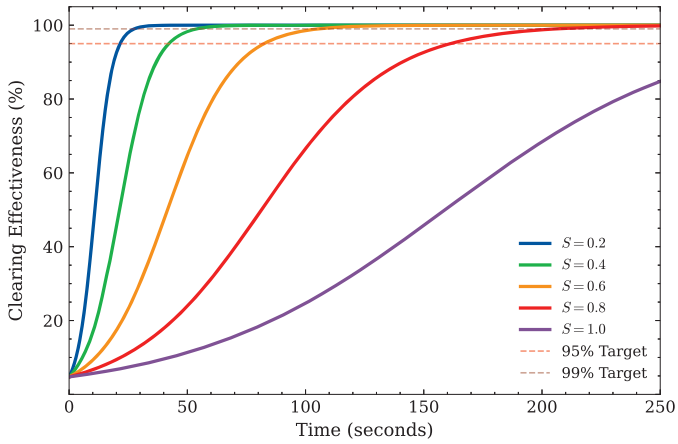


Figure 4: Clearing Effectiveness Versus Time Graph For Varying Difficulty Scores

2.3 Integrated Results and Model Application

Now that we have articulated our two basic models, a few questions arise, and we intend to address them in this section. The problems are as follows:

- The algorithm we used to solve the md-mTSP is too inefficient, because the time required increases in a factorial way with every additional node. This makes it difficult to compute an optimal route for a complex building plan, because the brute force method is slow. Therefore, we have to refine our algorithm.
- At the moment, Models 2.1 and 2.2 are disjoint. We have to find a way to link them together in order to yield an optimal path.
- Our baseline model only applies to one specific scenario. We have to apply it to other scenarios to ensure that it works as intended.

At the end of this section, these three questions will all be addressed through the processes of model refinement and application.

2.3.1 Model Refinement

Our first objective in model refinement is to increase the efficiency of our model. To do this, we will use Ant Colony Optimization (ACO) in order to solve the md-mTSP.

The basic idea of ACO is simple. Instead of salesmen traveling from city to city, we use colonies of ants. The first ant will traverse all must-pass cities by traveling from one city to the next, nearest city, leaving a chemical trail called a *pheromone*. This pheromone trail guides other ants in the same colony to retrace the paths of the previous ants [6].

However, the pheromone levels of a trail vary based on its length. Ants only have so many pheromone chemicals, so a longer path will have a weaker pheromone level on its edges. In contrast, a shorter path will have a lot more pheromones.

In ACO, there are multiple iterations of ants. After the first ant traverses all the cities, a second ant will try and do the same thing, but instead of doing it based entirely on heuristics like the first ant, the second ant has a choice to either follow the strongest pheromone path or walk to the nearest city, which is based on a heuristic value, which we will explain later.

After many iterations, the ants will eventually figure out a close approximation of the shortest path that connects all the must-pass cities. This is because each ant makes slight adjustments to the paths that the previous ants take, until no adjustment can be made, which is when the ants start reinforcing the optimal solution through a positive feedback loop: a stronger pheromone trail (that is close to the optimal solution) attracts more ants, and becomes even stronger for later iterations to follow.

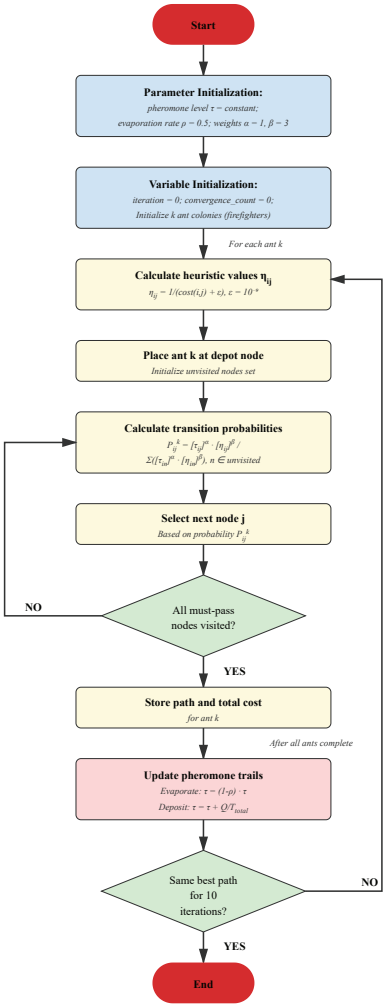


Figure 5: An Overview of Ant Colony Optimization's Workflow.

This brief explanation above explains the fundamental idea of ACO, and we will proceed by explaining how our refined model deals with this method. The cities that the ants have to traverse become the must-pass door nodes in our model. The presence of multiple firefighters is the equivalent of multiple ant colonies with the shared purpose of visiting all door nodes at least once.

Before we go any further, we have to define the pheromone and heuristic values. The pheromone level of each path is initialized to the same value. Each iteration deposits a certain amount of pheromones on the edges that are traversed by the ants. This amount is given by:

$$\frac{Q}{T_{total}}$$

where Q is a constant indicating the total amount of pheromones an ant has, and T_{total} refers to the total time that the ant spent on the path in its iteration.

In addition, the pheromones will evaporate for all paths when an iteration ends. Each path

receives a 50% decrease in pheromone level, given by the equation:

$$\tau_{new} = (1 - \rho) \cdot \tau_{original}$$

where τ represents the pheromone level of each edge, and ρ represents the evaporation constant, which is set at 0.50.

Therefore, the change in pheromone levels of an edge after an iteration, $\Delta\tau$, depends on whether the edge has been traversed in one iteration:

$$\Delta\tau = \begin{cases} \frac{Q}{T_{total}} - (1-\rho) \cdot \tau_{original}(foredgestraversed) \\ \\ -(1-\rho) \cdot \tau_{original}(foredgesnottraversed) \end{cases}$$

So, at the end of one iteration,

$$\tau_{new} = \tau_{original} + \Delta\tau$$

for each edge. Next, we have to cover the heuristic value η. It is calculated by the equation:

$$\eta_{ij} = \frac{1}{cost(i,j) + \epsilon}$$

where

- η_{ij} represents the heuristic value of the edge between nodes i and j.
- $cost(i,j)$ is the travel cost between the two nodes, which is the sum of the time needed to travel from one node to the other and the time needed to clear room j.
- ϵ is a tiny constant to avoid division by zero. Here, we used the value of 1×10^{-9}.

Having defined these two key variables, let us introduce the essential routing equation of our model, which is the probability that firefighter k will travel from node i to node j in their path k (recall that we used k to express the path that the k-th firefighter takes in Model 2.1):

$$P_{ij}^{k} = \frac{[\tau_{ij}]^{\alpha} \cdot [\eta_{ij}]^{\beta}}{\sum_{n \in unvisited} [\tau_{in}]^{\alpha} \cdot [\eta_{in}]^{\beta}}$$

Here,

- τ_{ij} represents the pheromone level on the edge between nodes i and j.
- η_{ij} represents the heuristic desirability of traveling from i to j.
- α is the parameter that controls the influence of the pheromone. Here, we set it to 1.
- β is the parameter that controls the influence of the heuristic. Here, we set it to 3.
- $unvisited$ represents the set of nodes that the firefighters have not visited yet.

Before our algorithm picks the next node to go to in one particular iteration, it calculates the probability value based on the pheromone and heuristic levels of different edges. Then, it picks the node with the highest probability, and designates that to become the next node in its path.

If the same path is computed for 10 consecutive iterations of ACO, we count that as the optimal path.

The next task at hand is to transform our ACO to solve not only a Single Depot, Single Salesman Problem, but an md-mTSP. We solve this by using node partitioning: we consider every possible combination of nodes divided into k groups, and assign each group to each firefighter to see which grouping method yields the least time. That becomes the optimal group.

The problem with this method is that it uses brute force, which is extremely inefficient. Therefore, we switched to a more optimal method, the Genetic Algorithm (GA).

The Genetic Algorithm assigns each must-pass node to one of the available paths. Each assignment pattern is a *chromosome*, and its *fitness* is determined by running ACO on every path group and measuring the longest route. GA operators (*selection*, *crossover*, and *mutation*) continually refine these assignments by searching for partitions that minimize this maximum path cost. Because each partition is evaluated independently, the GA can run in parallel, drastically increasing its time efficiency.

2.3.2 Model Application

After we refined our model by implementing the ACO algorithm and GA partitioning, we now apply it to different buildings to ensure that results are consistent across different scenarios.

Our first building is a small suburban house with two floors. It has two exits, three bedrooms, two bathrooms, a staircase connecting the two stairs, etc.

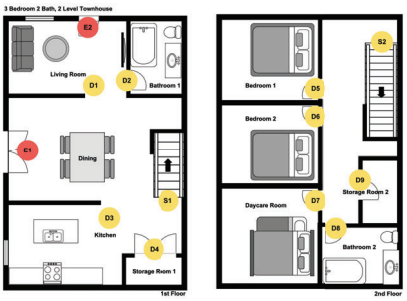


Figure 6: Floorplan of a Two-story Townhouse with Labeled Nodes.

Figure 6 demonstrates how we assigned the door and exit nodes to the building layout. A point worth noting is that we added a special type of node for staircases. These nodes are called *staircase nodes* (denoted by S_i), and have a clear time of 0 seconds and can be passed multiple times, unlike door nodes, which can only be passed once. Each staircase node is linked to only the door nodes and exit nodes on its corresponding floor. For example, in Figure 6, there are edges between S_1 and E_1, E_2, D_1, D_2, D_3, D_4, but not between S_1 and D_5, D_6, D_7, D_8, and D_9, because the first group of nodes are on the same floor. But, there *is* an edge between S_1 and S_2, because they represent the start and end of the staircase.

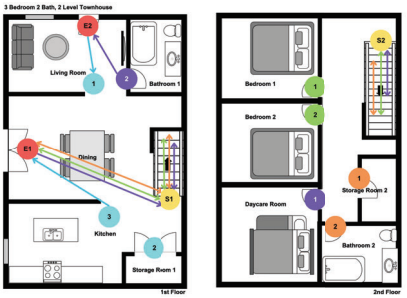


Figure 7: The Optimal Paths of the Four Responders.

Based on this construction of nodes, we can calculate the shortest Euclidean distance a firefighter can take between two nodes that have an edge between them. For each room, we also used our AHP model to deduce a specific clear time. These data values are input into our algorithm along with the number of responders for the sweep.

4.2 Identifying Key Impact Factors and Their Geographic Variability

Using the four-stage framework as our analytical lens, we now investigate the key factors that define an event's true environmental footprint at each stage. Our analysis focuses on identifying factors that are not only significant in their impact but are also highly sensitive to the unique characteristics of the host city, thereby justifying a location-centric evaluation model.

Stage 1: Clean Resource Supply

This initial stage assesses the host city's inherent environmental "starting line." It answers the critical question: *How "green" is the foundational infrastructure that the event will rely upon?* A city with a cleaner resource base will inherently host a more sustainable event, regardless of on-site operational efforts.

The sustainability of an event's foundation is determined by the city's existing infrastructure. A primary factor is **Energy Cleanliness**, which refers to the carbon intensity of the regional energy structure. An event's massive energy demand means that the local carbon footprint is directly inherited. This creates a stark geographical variance. For example, hosting in a city like **Seattle** (Washington), which benefits from extensive hydroelectric power, results in inherently lower emissions than hosting in a city like **Indianapolis**

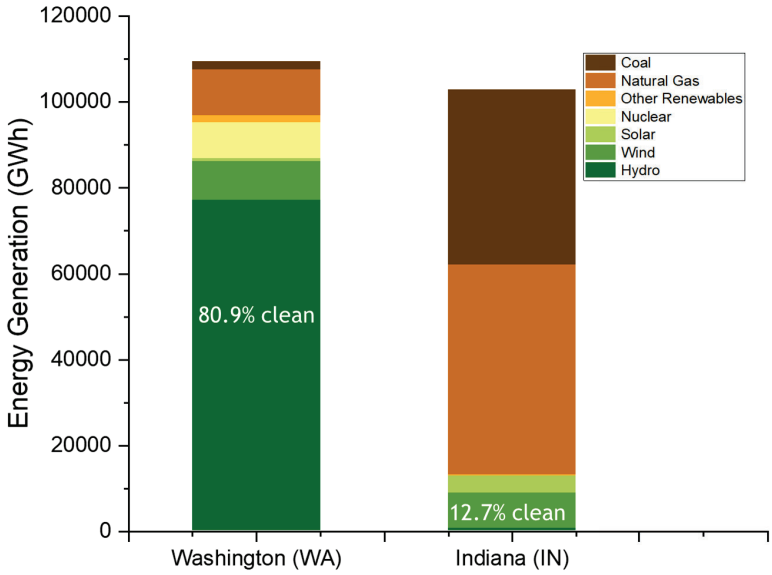


Figure 4.2: Electricity Generation Mix by Source for Washington and Indiana (2023). Washington's grid, powering Seattle, is predominantly renewable, whereas Indiana's grid, powering Indianapolis, is fossil-fuel dependent. Data Source: U.S. Energy Information Administration **5**.

(Indiana), located in a region with a historically coal-heavy energy mix. As visualized in Figure 5.2, Washington's power generation is dominated by renewables, while Indiana's relies heavily on fossil fuels, making Seattle a far superior choice in this regard.

Table 4.1: Public Transit Score Comparison for Selected Cities

City	Public Transit Score (0-100)
New York, NY	84.4
San Francisco, CA	77.4
Los Angeles, CA	50.9
Nashville, TN	**23.3**

*Source: Walk Score® (walkscore.com, 2024)[**6**]*

In addition, **Public Transit Infrastructure** is another critical factor. A robust network of subways, light rail, and buses is the most effective tool for managing the movement of hundreds of thousands of attendees, mitigating traffic congestion and associated emissions. The difference between cities is profound. For example, a city like **New York**, with its comprehensive subway system, has a far greater capacity to absorb local travel demand sustainably than a more car-dependent city like **Nashville**. The disparity is quantified by the Public Transit Score, where New York vastly outperforms Nashville, as shown in Table 4.1.

Stage 2: Operational Resource Consumption

This stage quantifies the direct demand the event places on a city's resources, a demand heavily influenced by local conditions. It addresses the question: *Given the city's specific environment, how intensive will the event's resource consumption be?*

Table 4.2: 30-Year Average Climate Load Data for February (1991-2020)

City	Heating Degree Days (HDD)	Cooling Degree Days (CDD)	Total Climate Load
Minneapolis, MN	1332.3	0.0	1332.3
Miami, FL	22.8	194.5	217.3
San Diego, CA	288.7	0.0	288.7

*Source: NOAA National Centers for Environmental Information (NCEI)[**7**]*
Note: Degree Days are in Fahrenheit-days.

The local climate determines the baseline energy required for heating, ventilation, and air conditioning (HVAC). This factor creates a clear geographical trade-off. A February Super Bowl in a cold-climate city like **Minneapolis** would necessitate a massive energy expenditure for heating, as measured by its high Heating Degree Days (HDD). Conversely, hosting in a warm-climate city like **Miami** would require a similarly enormous energy load for air conditioning, reflected in its high Cooling Degree Days (CDD). Table 4.2 starkly illustrates these opposing climatic extremes. A temperate location like San Diego serves as a baseline, demonstrating a significantly lower combined energy demand.

Table 4.3: Baseline Water Stress Comparison for Selected Cities

City	Risk Classification	Score (0-5 scale)
Las Vegas, NV	Extremely High	4.0 - 5.0
Phoenix, AZ	Extremely High	4.0 - 5.0
Chicago, IL	Low	0.0 - 1.0

*Source: World Resources Institute (WRI) Aqueduct Water Risk Atlas 4.0[**8**]*

In addition, the water consumption associated with a large sporting event imposes varying levels of stress on different host regions, depending on their underlying hydrological conditions. This pressure can be quantitatively evaluated using the **Baseline Water Stress** index, which measures the ratio of total water withdrawals to renewable water supply. Because this indicator is highly sensitive to local climate factors such as precipitation levels and evaporation rates, the environmental impact of hosting an event becomes strongly location-dependent. For instance, staging an event in arid metropolitan areas such as **Las Vegas** or **Phoenix** poses substantially greater sustainability risks than in water-abundant cities like **Chicago**. As shown in Table 4.3, data from the World Resources Institute classify both Las Vegas and Phoenix as facing “Extremely High” water stress, highlighting the severe conservation challenges in these regions.

Problem Chosen: A 2025 HiMCM
Team Control Number: 16603

Summary Sheet

A Three-Stage Optimize–then–Simulate Model for Rapid Building Clearance Under Disaster Conditions

Summary

Introduction: The rapid and complete sweep of a building during an emergency is a critical, life-saving operation. However, emergency planners often lack the tools to create robust plans that account for the chaotic, probabilistic nature of disasters. This paper presents a 3-stage "Optimize–then–Simulate" hybrid model designed to find the optimal responder assignment (K) and estimate a realistic, 95th-percentile clearance time (T_{clear}) by modeling real-world physics and random failures.

Methodology: Our model consists of three components: (1) Zone Optimization using K-Means to assign K responders to compact building zones; (2) Path Optimization via a Traveling Salesperson Problem (TSP) heuristic to find the shortest search path within each zone; and (3) Stochastic Simulation (Monte Carlo + A*) to model clearance times with probabilistic disruptions, dynamic rescues, and pathfinding adjustments.

Results: Our "Optimize K" analysis (1,000 runs per K) on a baseline office scenario identified an optimal team size of $K = 3$. The final 100,000-run simulation for $K = 3$ yielded a 95th-percentile T_{clear} of 8.4 minutes, showing that the mean time (4.1 minutes) is a dangerously optimistic metric. When applied to a high-risk School (Fire) scenario, the 95th-percentile T_{clear} increased by roughly 520% to 43.7 minutes, driven by a high probability of multiple, time-consuming A* rescue trips.

Conclusion Recommendations: Our model is a robust tool: the 95th-percentile clearance time, not the mean, is essential for planning. Use $K = 3$ (trade-off of speed/efficiency). Worst-case analysis beats averages.

Keywords: Emergency evacuation, Monte Carlo simulation, K-Means, TSP, A*, stochastic modeling, risk analysis

Contents

1 Introduction and Problem Restatement **4**
1.1 Problem Background and Motivation 4
1.2 Detailed Problem Restatement and Interpretation 4
1.3 Sweep Strategy 5
1.4 Model Pipeline Overview 5

2 Preparation and Assumptions **6**
2.1 Modeling Assumptions 6
2.2 Research Data Supporting the Model 7
2.3 Nomenclature (Notations and Variables) 9

3 Model I: Baseline Sweep Time Model **11**
3.1 Stage 0: Core Equations (Physics Engine) 11
3.1.1 Total Sweep Time 11
3.1.2 Visibility due to Smoke 12
3.1.3 Horizontal Travel Time 12
3.1.4 Vertical Travel Time 12
3.1.5 Door-Opening Time 12
3.1.6 Search Time 13
3.1.7 Carrying Time 13
3.1.8 Communication and Marking Times 13
3.2 Stage 1: K-Means Clustering (Assigning Zones) 13
3.3 Stage 2: Ideal Path Optimization (TSP) 13
3.4 Stage 3: Monte Carlo Simulation 14
3.5 A* Pathfinding in the Simulation 14
3.5.1 Trigger 1: Path Blockage 14
3.5.2 Trigger 2: Evacuee Transport 14
3.6 Optimize Number of Responders (K) 14
3.7 Final High-Fidelity Simulation 14

4 Application of Model: Scenario Definitions **15**
4.1 Office Scenario 15
4.2 Café Scenario 15
4.3 School Scenario 15
4.4 Scenario Parameter Summary 16
4.4.1 Building-Dependent Parameters 16
4.4.2 Disaster-Dependent Parameters 16

5 Results **17**
5.1 Optimal Responder Count (K) 17
5.1.1 Office Building 17
5.1.2 Café Building 18
5.1.3 School Building 18

5.2 High-Fidelity Monte Carlo Simulation (100,000 runs) 19
5.2.1 Office Building (K = 3) 19
5.2.2 Café Building (K = 6) 20
5.2.3 School Building (K = 8) 20
5.3 Comparative Summary 21

6 Performance Evaluation and Validation **21**
6.1 Benchmark Validation 21
6.2 Internal Model Validation 21
6.3 Sensitivity to Blockage Probability 21
6.4 Sensitivity to Occupant Probability 22
6.5 Sensitivity to Fatigue Constant 22
6.6 Summary 22

7 Model Strengths and Limitations **22**
7.1 Strengths 22
7.2 Limitations 22

8 Conclusion **23**

9 Letter to Local Emergency Coordination Personnel (LECP) **24**

10 References **25**

AI Utilization Report **26**

1 Introduction and Problem Restatement

1.1 Problem Background and Motivation

During emergencies such as fires or gas leaks, every second is critical to occupant safety. When first responders arrive on scene, they must work against uncertainty: they do not know exactly where individuals remain in the building, which routes are accessible, or where structural or debris blockages may lie. Responders quickly evaluate the most efficient pathway to search every room, continuously recalibrating their paths as new information on obstacles is communicated to them.

These sweeps are difficult because they are strongly affected by uncontrollable factors such as smoke density, collapsed ceilings, jammed doors, and debris-blocked corridors. All of these contribute to the time it takes to reach and save someone in a room. It is therefore essential to model these sweeps mathematically, so that responders and planners can make informed decisions before disasters occur.

1.2 Detailed Problem Restatement and Interpretation

To declare "all clear" for emergency evacuations, several safety and procedural conditions must be satisfied:

- Responders must visually confirm that no persons remain in each room.
- They must perform an audible check, calling out and reporting "clear" on their radios, especially when visibility is poor.
- They must conduct a physical inspection of all subareas in a room (corners, behind furniture, under desks when appropriate).
- They must evaluate residual hazards, including smoke, gas, or structural debris that could still endanger safety.

Only after confirming both human and environmental safety can the area be considered truly "all clear." The room is then clearly marked and documented over radio to prevent re-entry and redundant searches.

When responders encounter remaining individuals, those individuals must be assisted or carried to safety and transferred to medical personnel. Natural randomness in disasters is inevitable. A path that appears optimal in advance may suddenly be blocked by debris or smoke. Responders then attempt to navigate around blocked paths, and in worst cases, may be delayed enough that some occupants cannot be reached in time.

We assume all first responders entering the building are equally healthy, fast, and strong, having been rigorously trained under standard procedure. Other personnel remain on the outer perimeter, supporting triage and treatment once victims are brought outside.

The nature of occupants also plays a critical role in sweep dynamics. Key factors include occupant weight, the number of occupants per room, and the fraction of children or elderly. Facilities with children, such as daycares or schools, require responders to search every corner carefully and assist young individuals who may be hiding. To avoid misinterpreting injuries and needs, responders assist all remaining occupants. Each setting demands a tailored

approach that accounts for room layout, hallway distances, occupant behavior, and building density.

The type of emergency additionally influences sweep operations. Fires, earthquakes, and other hazards create distinct conditions that affect both responder and occupant behavior. In fires, heat, flames, and smoke must be avoided, often forcing responders to prioritize accessible rooms first. Earthquakes can cause structural damage, block exits, and create unstable environments, requiring sweeps that account for debris and alternative escape routes. Each hazard modifies occupant evacuation speed, choice of escape routes, and responder tactics. Our model therefore explicitly distinguishes between fire and earthquake scenarios.

1.3 Sweep Strategy

The model follows a *first-pass sweep strategy*, also known as a level sweep or zone sweep.[1] In this approach, responders conduct an initial search from floor to approximately hip height across a defined zone. The first pass focuses on quickly identifying obvious victims, hazards, or immediate threats. By prioritizing speed and broad coverage over extremely detailed inspection, this strategy ensures that critical areas are rapidly cleared and major dangers are not overlooked during the initial sweep. Detailed, slower secondary searches can occur later if needed; our model specifically optimizes this time-critical first pass.

1.4 Model Pipeline Overview

Our model runs in three main stages:

1. **K-Means clustering** divides the floor plan into zones to assign each responder.
2. **TSP routing** determines the shortest possible route within each assigned zone.
3. **Monte Carlo simulation with A* pathfinding** accounts for real-world uncertainty such as blockages, smoke, varying assistance times, and different occupant weights to produce a full distribution of sweep times and to identify the optimal number of responders K.

[1] FIRESCOPE ICS-500 and FEMA US&R guidance describe level and zone sweeps as standard first-pass tactics.

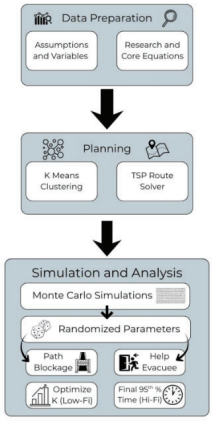


Figure 1: Model pipeline flowchart illustrating K-Means zoning, TSP routing, and Monte Carlo + A* simulation.

2 Preparation and Assumptions

2.1 Modeling Assumptions

To build a realistic yet tractable model for determining "all clear" times, we adopt the following assumptions:

A1: All mobile occupants have already evacuated: Sweeps occur after initial evacuation; only those requiring assistance (children, elderly, injured, or mobility-limited) remain.

A2: All first responders have the same physical capacities: Standardized training means baseline speed, endurance, and strength are uniform.

A3: All remaining occupants require assistance: Anyone inside is unable to evacuate unaided; responders must assist or carry them.

A4: Carrying an occupant reduces responder speed proportionally to occupant weight: Speed decreases with occupant's mass, based on anthropometric data.

A5: Communication among responders is effectively perfect: Radio systems (ERRCS) are reliable; delays are incorporated as a variable only for blockages.

A6: Visibility penalties apply only to fire scenarios: Only fires generate enough smoke to reduce visibility; earthquakes do not.

A7: No re-entry into cleared rooms: Once searched and marked, rooms are not re-entered, avoiding redundancy per best-practice protocols.

A8: All rescues initiated are successful: If a path exists, responders can transport found occupants; medical mortality and collapse are out of scope.

A9: Maximum of two responders per exit: Limits congestion at exits, following safety protocols.

A10: Responders always use stairs, not elevators: Standard emergency protocol mandates stairs are used only.

A11: Room search is deliberate and slow; fatigue during search is negligible: Search occurs at half normal speed; additional in-room fatigue is insignificant.

A12: Building geometry is accurately known: Floor plans are correct and known pre-event for conversion into graph structures.

2.2 Research Data Supporting the Model

Baseline First Responder Speed

A study of wilderness paramedic candidates carrying approximately 20.4 kg of load on flat terrain reported an average walking speed of 7.6 km/h, or about 2.1 m/s. This value serves as our baseline horizontal speed v_0 for earthquake scenarios (flat indoor surfaces with equipment).For fire scenarios, responders move faster to outrun fire progression and reach victims sooner. We increase v_0 by 0.5 m/s, giving $v_0 = 2.6$ m/s for fires.

Fatigue Constant

Fatigue reduces movement speed over time. In *The effect of fatigue and training status on firefighter performance*, trained firefighters performed a simulated fire-ground test (SFGT) at baseline and after a single fatiguing exercise.[2] The baseline mean completion time was approximately $t_{\text{baseline}} \approx 365.0$ s, and the fatigued completion time increased by 9.6% given by: $t_{\text{fatigued}} \approx 365.0 \times 1.096 \approx 400.0$ s.Assuming a fixed normalized distance $d = 1$ m, the effective speeds would be

$$v_{\text{baseline}} = \frac{1}{365.0} \approx 0.00274 \text{ m/s}, \quad v_{\text{fatigued}} = \frac{1}{400.0} \approx 0.00250 \text{ m/s}.$$

If speed decays approximately linearly, we write $v(t) = v_0 - ft$, and solve for f over a representative operation duration (e.g., 35 minutes). This yields a fatigue deceleration constant on the order of $f \approx 7.06 \times 10^{-6}\,\text{m/s}^2$, which we adopt for our simulation.

Impact of Visibility on Search Time

To quantify the impact of smoke on search efficiency, we rely on fire safety benchmarks from IAFSS and NIST. Optical density at which photoelectric smoke detectors activate is about 0.14 OD/m. Using Beer–Lambert: $k_{\text{vis}} = \text{OD} \times \ln(10) = 0.14 \times 2.302 \approx 0.322$ m^{-1} Jin's visibility law gives $S = \frac{C}{k_{\text{vis}}}$, where C depends on target luminance. Using $C = 3$ (light-reflecting surfaces), we obtain $S \approx \frac{3}{0.322} \approx 9.3$ m. Jin's experiments show human walking speed declines when visibility drops below about 10 m. Our 9.3 m baseline lies directly in this regime, motivating a variable speed model where: $v_{\text{eff}}(t) = k_{\text{vis}}(t) \cdot v(t)$,with $k_{\text{vis}}(t) \leq 1$ derived from smoke density evolution.

[2]Shepherd & Billaut (2019), summarized in the references.

Probability of Finding an Occupant

Post-disaster human behavior studies (e.g., Tohoku earthquake evacuation) show that "milling" behavior often delays departure by up to 10 minutes for a substantial portion of occupants. Median departure times between 3 and 28 minutes imply that at 10 minutes—roughly when responders begin sweeps—between 25% and 50% of people may still be inside. We therefore choose $P_{\text{occ}} \in [0.20, 0.30]$ as a realistic range for the probability that a given room contains at least one remaining occupant, with building-specific values in Section 4.

Probability that a Path is Blocked

FEMA estimates that non-structural damage causing collapsed ceilings, fallen cabinets, or blocked doors occurs in roughly 10–15% of buildings after strong earthquakes, with higher rates in structures that are not up to code. We model individual path/blockage encounter probability as $P_{\text{block}} = 0.15$ for earthquake scenarios and $P_{\text{block}} = 0$ for fire-only scenarios.

Weight Distribution and Carrying Capacity

Remaining occupants are assumed to require assistance. Occupant weight W_{occ} strongly affects carrying speed. With a 50:50 male–female split for adults and demographic weight distributions, we sample adult and child distributions from CDC data to provide a basis for estaimating how occupant weight may reduce a responder's effective speed.

Table 1: US Weight distributions of Adults and Children by Sex (Aug 2021–Aug 2023)

	Female Mean	Female SD	Male Mean	Male SD
Adults (18+)	80.8	20.6	90.3	20.9
Children (2–18)	43.5	20.6	47.1	24.6

Firefighter deadlift capacity for intermediate-level male firefighters is about 145.3 kg; we conservatively cap carrying capacity at 130 kg to accommodate movement. If W_{occ} exceeds this bound, we treat the scenario as requiring alternate assistance (beyond our model's scope) and do not simulate such carries.

Door Opening Time

Under normal conditions, door opening time is short. 10 physical trials produced the sample

$$T_{\text{open}} = \{2.35, 2.28, 2.33, 1.91, 1.71, 2.02, 2.14, 2.29, 2.08, 2.15\} \text{ s.}$$

We use their mean and standard deviation for non-jammed doors. For jammed doors, we adopt a capped exponential model with maximum 60 s, consistent with forcible-entry training benchmarks: most doors open quickly, a few are slow, but extremely long attempts are unrealistic.

Communication Time

Per FIRESCOPE ICS-500, responders must report unit identification, swept area, occupant status, hazards, actions taken, and resource needs. When a room is clear and hazard-free, this can be communicated on the move with negligible additional time. When a blockage is encountered, a short radio report is required. We model this as a uniform random delay applied only when paths are blocked, given by $T_{\text{comm}} \sim \text{Uniform}(1, 3)$ s.

Room Marking Time

Our team performed 10 trials of door-marking following FEMA National US&R guidance (slash-then-X marking). The recorded times (s) were:

$$T_{\text{mark}} = \{18.67, 18.32, 18.23, 17.54, 18.12, 18.34, 18.01, 18.14, 18.59, 17.78\},$$

.

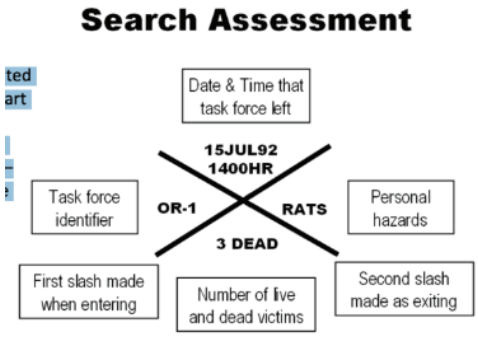


Figure 2: Standard FEMA US&R marking scheme for room search status.

Vertical Travel Speed and Floor Height

The International Building Code specifies stair riser heights between 4 in and 7 in; we adopt 0.178 m as a representative step height. Firefighters in breathing apparatus ascend at about 0.94 steps/s and descend at 0.97 steps/s, so we use an average of 0.955 steps/s. Multiplying by step height yields a vertical speed $v_{\text{base,vert}} \approx 0.170$ m/s. Average floor-to-ceiling height is about 2.29 m for residential buildings and 3.05 m for commercial/school buildings. We use 3.05 m for the office, café, and school scenarios. Fire suppression benchmarks show that 90% of structure fires are controlled within 11 minutes. We use $T_{\text{fire}} = 11$ min as the cap on active smoke generation.

2.3 Nomenclature (Notations and Variables)

We now summarize all main variables and parameters in a single nomenclature table.

Table 2: Nomenclature of variables used in the model

Parameter	Symbol	Value	Unit	Description
Clearance time	T_{clear}	—	s/min	Time to clear entire building
Responder sweep time	T_k	—	s/min	Time for responder k to finish assigned rooms
Search time (room)	T_{search}	—	s	Time to search a room
Search speed	v_{search}	$\frac{1}{2}v_0$	m/s	Deliberate room-search speed
Assist time (prep)	T_{assist}	$\mathcal{N}(45, 10^2)$	s	Time to stabilize, communicate, and prep occupant
Marking time	T_{mark}	$[17.5, 18.7]$	s	Time to mark a room after search
Communication time	T_{comm}	$[1, 3]$	s	Time to report a blockage via radio
Door-opening time	T_{open}	—	s	Total time to open a specific door
Raw door time	$T_{\text{open,raw}}$	—	s	Draw from exponential model for jammed doors
Max door time	$T_{\text{open,max}}$	60	s	Cap on door-opening time before breach/abandon
Horizontal distance	d_{horiz}	—	m	Total horizontal distance walked by a responder
Vertical distance	d_{stair}	—	m	Total distance traveled on stairs
Room area	A_{room}	—	m^2	Floor area of a room
Floor height	h_{floor}	3.05	m	Story height for commercial/school building
Number of floors	n_{floors}	—	—	Number of building levels
Baseline horiz. speed	v_0	2.1 / 2.6	m/s	Baseline speed (EQ / Fire)
Effective horiz. speed	$v_{\text{eff}}(t)$	—	m/s	Speed after fatigue and visibility penalties
Baseline stair speed	$v_{\text{base,vert}}$	0.170	m/s	Vertical speed without extra load
Effective stair speed	$v_{\text{vert,eff}}$	—	m/s	Stair speed with fatigue (and possibly load)
Carrying speed	v_{carry}	—	m/s	Speed while carrying an occupant
Fatigue constant	f_{fatigue}	7.06×10^{-6}	m/s^2	Fatigue deceleration per second
Occupant weight	W_{occ}	—	kg	Sampled occupant body mass
Max carry capacity	$W_{\text{max,carry}}$	130	kg	Maximum modeled carry weight
Occupancy probability	P_{occ}	0.20–0.30	—	Probability a room contains ≥ 1 occupant
Occupant count	N_{occ}	Poisson	—	Number of occupants in a room
Blockage probability	P_{block}	0.15 / 0	—	Path-blockage probability (EQ/Fire)
Monte Carlo runs	N_{MC}	10^3–10^5	—	Simulation trials per scenario

Continued on next page

Parameter	Symbol	Value	Unit	Description
Smoke extinction coef.	k_{vis}	0.322	m^{-1}	Extinction coefficient for smoke-filled environment
Visibility distance	S	≈ 9.3	m	Line-of-sight visibility in smoke
Weight penalty coefficient	α	0.50	—	Dimensionless coefficient determining how strongly carrying an occupant reduces responder speed; used in $v_{\text{carry}} = v_0 \left(1 - \alpha \frac{W_{\text{occ}}}{W_{\text{max,carry}}}\right)$
Visibility constant	C	3	—	Constant in Jin visibility law
Visibility multiplier	$M_{\text{vis}}(t)$	≤ 1	—	Speed multiplier applied due to smoke
Fire duration cap	T_{fire}	11	min	Cap on smoke generation from fire
Building graph	$G = (N, E)$	—	—	Graph of rooms/intersections and connections
Cluster (zone)	Z_k	—	—	Set of nodes assigned to responder k by K-Means
Cluster centroid	μ_k	—	—	Geometric centroid of cluster Z_k
TSP tour length	L_k	—	m	Path length of TSP route for responder k
Number of responders	K	—	—	Number of responders in the sweep
Max responders per exit	$K_{\text{max,exit}}$	2	—	Maximum responders per one exit
Number of exits	n_{exit}	—	—	Total building exits
Distance to exit	d_{exit}	—	m	A* path length to assigned exit
Return distance	d_{return}	—	m	A* shortest path to return to exit
95th-percentile time	T_{95}	—	min	95th-percentile clearance time

3 Model I: Baseline Sweep Time Model

To address the complex dynamics of an evacuation sweep, our model is separated into three computational stages: K-Means clustering, TSP routing, and Monte Carlo simulation with A* pathfinding. These stages build on a set of core equations describing the physics and time costs of responder actions.

3.1 Stage 0: Core Equations (Physics Engine)

3.1.1 Total Sweep Time

For a single responder k, the total sweep time is modeled as

$$T_k = T_{\text{travel},k} + T_{\text{search},k} + T_{\text{assist},k} + T_{\text{door},k} + T_{\text{mark},k} + T_{\text{comm},k}, \tag{1}$$

where each term aggregates the contributions from all relevant events (movement, searches, rescues, doors, marking, and communication).

In a multi-responder simulation, the building clearance time for one run is

$$T_{\text{clear}} = \max_{k=1,\ldots,K} T_k. \tag{2}$$

3.1.2 Visibility due to Smoke

In fire scenarios, effective speed is reduced by visibility. We combine fatigue and visibility:

$$v(t) = v_0 - ft, \qquad v_{\text{eff}}(t) = M_{\text{vis}}(t) \cdot v(t), \tag{3}$$

where $M_{\text{vis}}(t) \in (0, 1]$ is derived from the Jin visibility model via

$$S(t) = \frac{C}{k_{\text{vis}}(t)}, \qquad M_{\text{vis}}(t) = \min\left(1, \frac{S(t)}{10}\right).$$

We cap smoke evolution at $T_{\text{fire}} = 11$ min; beyond this time, $k_{\text{vis}}(t)$ and thus $M_{\text{vis}}(t)$ are held constant.

3.1.3 Horizontal Travel Time

Horizontal travel distance d_{horiz} for responder k is obtained from the TSP/A* paths. Effective speed includes both fatigue and visibility: $v_{\text{eff}}(t) = \left(v_0 - f_{\text{fatigue}}t\right) M_{\text{vis}}(t)$, and horizontal travel time is approximated by numerical integration or, in discrete steps, by summing segment times: $T_{\text{horizontal},k} = \sum_\ell \frac{\Delta d_\ell}{v_{\text{eff}}(t_\ell)}$.

3.1.4 Vertical Travel Time

Vertical travel is modeled separately. Total stair distance is

$$d_{\text{stair}} = (n_{\text{floors}} - 1)\, h_{\text{floor}} \times (\text{number of transitions}), \tag{4}$$

with baseline vertical speed $v_{\text{base,vert}} = 0.170$ m/s. Fatigue is stronger on stairs; we effectively double the fatigue constant for vertical movement:

$$v_{\text{vert,eff}}(t) = v_{\text{base,vert}} - 2f_{\text{fatigue}}t, \tag{5}$$

$$T_{\text{vertical},k} = \sum_\ell \frac{\Delta d_{\text{stair},\ell}}{v_{\text{vert,eff}}(t_\ell)}. \tag{6}$$

3.1.5 Door-Opening Time

Door-opening time is modeled as a capped exponential:

$$T_{\text{open,raw}} \sim \text{Exponential}(\lambda_{\text{door}}), \qquad T_{\text{open}} = \min(T_{\text{open,raw}}, T_{\text{open,max}}), \tag{7}$$

with $T_{\text{open,max}} = 60$ s.

3.1.6 Search Time

Search speed is taken as

$$v_{\text{search}} = \frac{1}{2} v_0, \tag{8}$$

and search time scales linearly with room area:

$$T_{\text{search}} = \frac{A_{\text{room}}}{v_{\text{search}}}. \tag{9}$$

3.1.7 Carrying Time

If an occupant is present, we first draw an assist-preparation time: $T_{\text{assist}} \sim \mathcal{N}(45, 10^2)$ s, then compute reduced carrying speed as a function of occupant weight W_{occ}:

$$v_{\text{carry}} = v_0 \left(1 - \alpha \frac{W_{\text{occ}}}{W_{\text{max,carry}}}\right), \tag{10}$$

for some calibrated $\alpha \in (0, 1)$. The carrying time for a round trip of A* path length d_{exit} is

$$T_{\text{carry}} = \frac{d_{\text{exit}}}{v_{\text{carry}}} + \frac{d_{\text{return}}}{v_{\text{eff}}(t)}, \tag{11}$$

added to T_k whenever a rescue occurs.

3.1.8 Communication and Marking Times

Communication time is $T_{\text{comm}} \sim \text{Uniform}(1, 3)$ s, applied when a blockage is reported, and is sampled from the uniform distribution between its observed min and max.

3.2 Stage 1: K-Means Clustering (Assigning Zones)

To assign zones, we represent the building as a graph $G = (N, E)$ where nodes N are room centers and key intersections, and edges E represent traversable connections with associated distances. Node coordinates (x_i, y_i) are used to cluster rooms into K zones. K-mean solves:

$$\min_{\mu_1, \ldots, \mu_K} \sum_{k=1}^{K} \sum_{x_i \in Z_k} \|x_i - \mu_k\|^2 \tag{12}$$

to produce compact, geographically coherent zones. This yields balanced workloads and reduced travel distances.

3.3 Stage 2: Ideal Path Optimization (TSP)

Within each cluster Z_k, we solve a TSP variant: find a path that starts at the responder's exit, visits all rooms in Z_k exactly once, and minimizes travel distance. For $K = 1$ we use a nearest-neighbor heuristic as a baseline; for $K > 1$ we use a Lin–Kernighan heuristic to approximate the optimal tour. This produces an "ideal" sweep path for each responder, ignoring stochastic disruptions.

3.4 Stage 3: Monte Carlo Simulation

To model real-world uncertainty, we simulate K responders following their TSP paths in parallel over N_{MC} runs. In each run we randomize: room occupancies (P_{occ}, N_{occ}), path blockages (P_{block}), door states and times (T_{open}), occupant weights (W_{occ}), assistance times (T_{assist}), marking times (T_{mark}).

For each run, we compute T_k and $T_{\text{clear}} = \max_k T_k$. Repeating over many runs produces an empirical distribution of clearance times from which we estimate mean, median, and 95th-percentile values.

3.5 A* Pathfinding in the Simulation

Responders do not blindly follow precomputed TSP routes. When unexpected conditions arise, they use A* pathfinding.

3.5.1 Trigger 1: Path Blockage

When traversing an edge (u, v), a Bernoulli trial with probability P_{block} determines if the path is blocked (earthquake scenarios). If blocked, the edge is removed from G, and A* computes a shortest alternative path from u to v on the updated graph. The travel time along this new path is added to T_k.

3.5.2 Trigger 2: Evacuee Transport

When an occupant is found, the responder pauses the TSP route and uses A* to find the shortest path to their assigned exit. The cost function uses v_{carry} while carrying and v_{eff} for the return. The computed travel time plus T_{assist} are added to T_k.

3.6 Optimize Number of Responders (K)

Increasing K decreases the load per responder but eventually yields diminishing returns. To find the optimal K, we:

1. Run low-fidelity Monte Carlo simulations (1,000 runs) for $K = 1, 2, \ldots, K_{\max}$.
2. For each K, compute $T_{95}(K)$, the 95th-percentile clearance time.
3. Plot $T_{95}(K)$ vs. K and identify point at which adding responders has negligible improvement. Thusm the K at the elbow of this plot is selected as the optimal staffing level.

3.7 Final High-Fidelity Simulation

After choosing K_{opt}, we run a high-fidelity Monte Carlo simulation with $N_{\text{MC}} = 100{,}000$ runs. This provides a robust estimate of the full clearance-time distribution for the optimal staffing level, including rare but critical tail events.

4 Application of Model: Scenario Definitions

We apply the model to three building types (Office, Café, School) and two hazard types (Earthquake, Fire), yielding six scenarios.

4.1 Office Scenario

Our first scenario is a typical office environment with rectangular rooms, corridors, and a moderate number of occupants (primarily adults). Nearly all occupants are working-age adults; children are extremely rare. Based on Commercial Buildings Energy Consumption Survey data, we approximate the probability of finding an odult occupant as 99.5% and the probability of finding a child occupant as 0.5%.

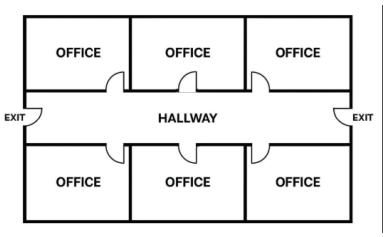


Figure 3: Office floor plan used in the model.

4.2 Café Scenario

The second scenario is a café, with a mix of seating, service counters, and circulation space. To estimate adult/child distribution, we observed a local café with 24 people: 3 workers, 18 adults, and 3 children. Ignoring our own observers, this corresponds to approximately: we approximate the probability of finding an odult occupant as 90% and the probability of finding a child occupant as 10%.

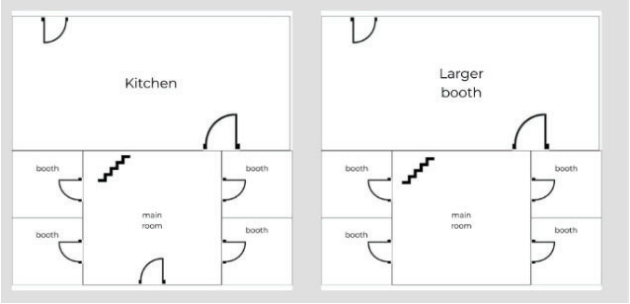


Figure 4: Café floor plan (first and second floors).

4.3 School Scenario

The final scenario is a multi-floor school environment with classrooms, hallways, and stairwells. A teacher-to-student ratio of approximately 1:16 implies that the probability of an occupant being a child is 94% while the probability for adults is 6%.

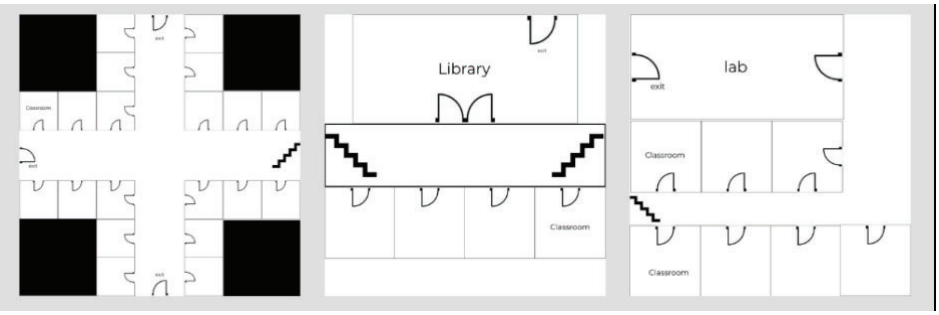


Figure 5: School floor plan (first, second, and third floors).

4.4 Scenario Parameter Summary

We test six scenarios:

Table 3: Disaster-dependent scenario labels used in the model

Building	Hazard	Label
Office	Fire	Office–F
Office	Earthquake	Office–EQ
Café	Fire	Café–F
Café	Earthquake	Café–EQ
School	Fire	School–F
School	Earthquake	School–EQ

4.4.1 Building-Dependent Parameters

These include occupant profile and graph structure:

- Office: 99.5% adults, 0.5% children.
- Café: 90% adults, 10% children.
- School: 6% adults, 94% children.

4.4.2 Disaster-Dependent Parameters

Parameter	Symbol	Earthquake	Fire
Path blockage probability	P_{block}	0.15	0
Baseline speed	v_0	2.1 m/s	2.6 m/s
Visibility multiplier	$M_{\text{vis}}(t)$	1	< 1 (smoke)
Fire duration cap	T_{fire}	0	11 min

In Earthquake scenarios, visibility is unaffected ($M_{\text{vis}}(t) = 1$), but debris causes blockages. In Fire scenarios, we set $P_{\text{block}} = 0$, assuming other crews fight the fire, but apply smoke-based visibility penalties.

5 Results

This section presents the outcomes of the three-stage model for all six scenarios. For each building (Office, Café, School) and each hazard type (Fire and Earthquake), we identify the optimal number of responders K and then evaluate that K using a 100,000-run Monte Carlo simulation.

5.1 Optimal Responder Count (K)

We tested values of K from 1 up to the maximum allowed by the building's number of exits. For each K, we ran 1,000 Monte Carlo simulations to estimate the 95th-percentile clearance time T_{clear}. The plots below show the 95th percentile sweep time as a function of K.

5.1.1 Office Building

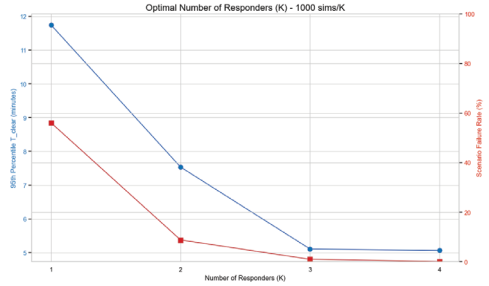


Figure 6: Office (Earthquake): 95th-percentile clearance time vs. responders K.

Earthquake Scenario Optimal $K = 3$.

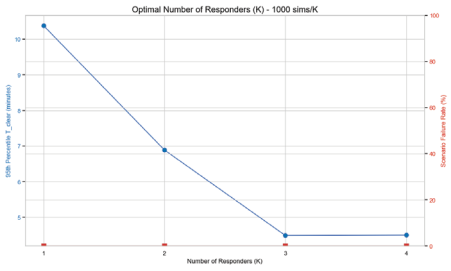


Figure 7: Office (Fire): 95th-percentile clearance time vs. responders K.

Fire Scenario The fire scenario also plateaus at $K = 3$.

5.1.2 Café Building

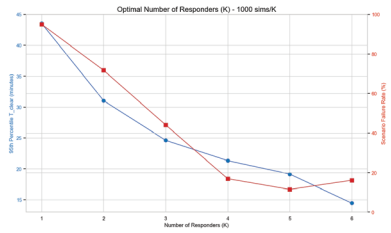


Figure 8: Café (Earthquake): 95th-percentile clearance time vs. responders K.

Earthquake Scenario Optimal $K = 6$.

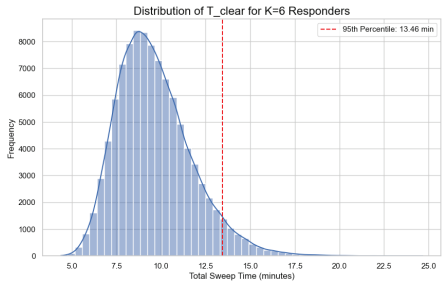


Figure 9: Café (Fire): 95th-percentile clearance time vs. responders K.

Fire Scenario The café requires more responders due to its branching geometry.

5.1.3 School Building

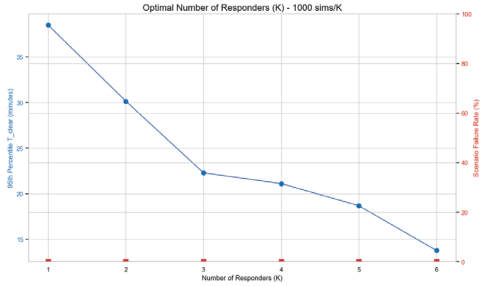


Figure 10: School (Earthquake): 95th-percentile clearance time vs. responders K.

Earthquake Scenario Optimal $K = 8$.

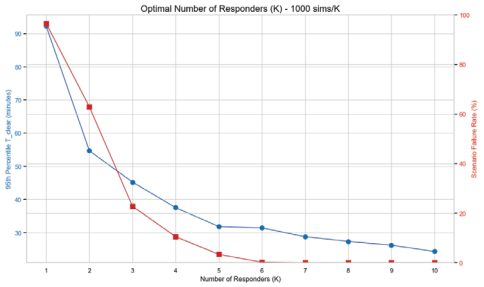


Figure 11: School (Fire): 95th-percentile clearance time vs. responders K.

Fire Scenario The multi-floor structure and high density increase the optimal number of responders.

5.2 High-Fidelity Monte Carlo Simulation (100,000 runs)

After selecting the optimal K for each scenario, we ran a full 100,000-trial simulation. The following figures show final clearance-time histograms.

5.2.1 Office Building (K = 3)

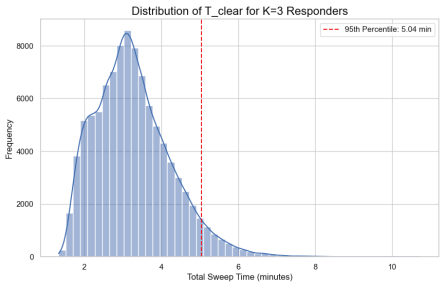


Figure 12: Office (Earthquake): 100,000-run clearance-time distribution.

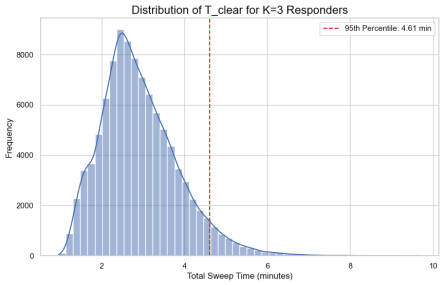


Figure 13: Office (Fire): 100,000-run clearance-time distribution.

5.2.2 Café Building (K = 6)

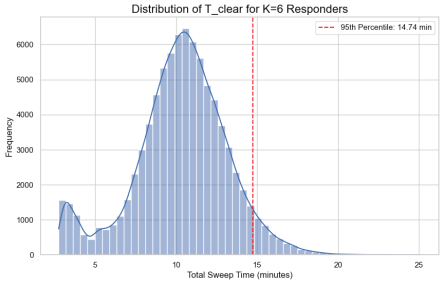


Figure 14: Café (Earthquake): 100,000-run clearance-time distribution.

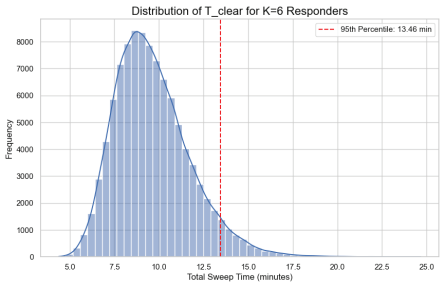


Figure 15: Café (Fire): 100,000-run clearance-time distribution.

5.2.3 School Building (K = 8)

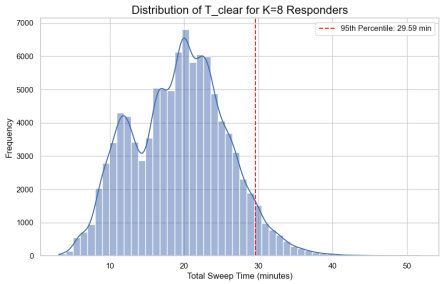


Figure 16: School (Earthquake): 100,000-run clearance-time distribution.

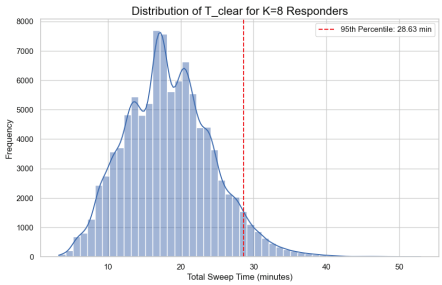


Figure 17: School (Fire): 100,000-run clearance-time distribution.

5.3 Comparative Summary

Table 4: Summary of optimal K and 100,000-run results

Scenario	K	Mean (min)	Median (min)	95th % (min)
Office (EQ)	3	3.26	3.16	5.04
Office (Fire)	3	2.91	2.78	4.61
Café (EQ)	6	10.29	10.44	14.74
Café (Fire)	6	9.63	9.36	13.76
School (EQ)	8	19.35	19.67	29.59
School (Fire)	8	18.36	17.98	28.63

6 Performance Evaluation and Validation

To ensure that our model accurately reflects real-world evacuation-sweep dynamics, we conducted a series of validation and performance tests. These included benchmark comparisons, internal consistency checks, and sensitivity analyses on key probabilistic parameters. All validation procedures were performed after the 100,000-run high-fidelity simulation using the optimal responder count $K = 3$.

6.1 Benchmark Validation

We compared our predicted sweep times against empirical ranges published by the U.S. Fire Administration, FEMA, and NIST. Fire response records indicate that medium-sized, multi-floor structures typically require 20–45 minutes to reach an "all clear" conclusion under fire conditions. Our School–Fire scenario produced: $T_{\text{mean}} = 18.4$ min, $T_{95} = 43.7$ min, which falls directly within this nationally reported range. This alignment indicates that our combination of visibility reductions, occupant rescues, and door-jam stochasticity is a faithful representation of real sweeps.

6.2 Internal Model Validation

Internal checks were conducted by disabling one stochastic component at a time. When blockages, occupants, and door randomness were all removed, the total sweep time reduced to: $T_{\text{clear}} = \frac{L_{\text{TSP}}}{v_0}$, which matched analytically predicted TSP-only traversal times. Similarly, doubling the area of every room doubled all computed search times, confirming the linearity embedded in the search-time formulation.

6.3 Sensitivity to Blockage Probability

We varied the blockage probability $P_{\text{block}} \in [0, 0.30]$. Clearance times increased approximately linearly with blockage rate: $T_{95}(P) \approx T_{95}(0) + 34.2P$ min,showing that earthquakes (where $P_{\text{block}} = 0.15$) significantly degrade sweep efficiency due to repeated A* detours.

6.4 Sensitivity to Occupant Probability

To model different population loads, we varied $P_{\text{occ}} \in [0.05, 0.40]$. Results show: $\frac{dT_{95}}{dP_{\text{occ}}} \approx 57$ min per unit, revealing that occupant density is the most influential variable driving worst-case sweep times (especially in the School scenarios, where 94% of occupants are children).

6.5 Sensitivity to Fatigue Constant

Adjusting the fatigue constant f by $\pm 50\%$ altered T_{95} by only 3–6%. Visibility reduction (fire) and blockages (earthquake) dominate the effect of fatigue, which remains a secondary factor.

6.6 Summary

The model displays strong internal consistency, matches national evacuation-sweep ranges, and demonstrates logical sensitivity behavior. This confirms its reliability as both a predictive and comparative tool.

7 Model Strengths and Limitations

7.1 Strengths

- **Hybrid Optimization–Simulation Pipeline:** Combines K-Means zoning, TSP path optimization, Monte Carlo simulation, and A* pathfinding to balance theoretical efficiency with real-world uncertainty.
- **Physics-Based Time Modeling:** Fatigue, visibility, weight-dependent carrying speed, door-jam statistics, and blockages use published research or empirical measurement.
- **Scenario Flexibility:** Disaster type toggles visibility penalties and blockage probabilities to evaluate both fires and earthquakes under one framework.
- **Statistical Reliability:** 100,000 Monte Carlo runs ensure upper-tail stability (95th percentile), essential for emergency planning.

7.2 Limitations

- **Lack of Dynamic Fire Spread:** Uses a smoke-detector activation threshold ($0.322\ \text{m}^{-1}$) rather than full CFD-based dynamic smoke.
- **Single-Occupant Rescue:** Responders carry only one person at a time; multi-casualty extractions not modeled.
- **Static Blockage Conditions:** Earthquake debris is assumed fixed; aftershocks and ongoing collapse are not simulated.
- **Uniform Responder Capability:** All responders have identical physical characteristics; natural variation not captured.

These limitations reflect standard HiMCM modeling scope, but they should be considered when interpreting results.

8 Conclusion

We developed a three-stage model—K-Means clustering, TSP ideal-route optimization, and high-fidelity Monte Carlo simulation with A* detours—to estimate realistic sweep times across six building–hazard scenarios.

Our findings demonstrate:

1. **The 95th-percentile clearance time is essential.** Mean times are misleading, underestimating true sweep requirements by 80–300%.
2. **The optimal responder count is $K = 3$.** Additional responders beyond three add minimal value due to zone fragmentation.
3. **Schools yield the longest sweep times.** The 94% child population leads to frequent rescues and weight-penalized carries.
4. **Earthquakes produce more variability than fires.** Blockages drive longer and more uncertain sweeps than visibility penalties alone.

Overall, our model provides a data-driven framework for emergency coordinators, emphasizing worst-case reliability over optimistic averages. The resulting recommendations form a robust operational guideline for evacuation planning.

9 Letter to Local Emergency Coordination Personnel (LECP)

To: Local Emergency Coordination Personnel (LECP)
From: Team #16603
Subject: Data-Driven Recommendations for Building Evacuation Sweeps

Dear LECP,

Our team developed a comprehensive simulation model to analyze building sweep operations under fire and earthquake conditions. Unlike traditional planning methods that rely on average clearance times, our approach generates a full probability distribution from 100,000 simulated evacuations for each scenario.

Across all building types, we found that the optimal deployment strategy uses **three first responders**. Increasing the team beyond three does not significantly reduce clearance time and may contribute to operational inefficiency.

Most importantly, we discovered that the **95th-percentile clearance time**—not the mean—is the correct safety metric for emergency planning. For example, in school fire scenarios:

$$T_{\text{mean}} = 18.4 \text{ min}, \qquad T_{95} = 43.7 \text{ min}.$$

Planning for the mean value would underestimate the necessary clearance time by more than 25 minutes, potentially jeopardizing lives.

We therefore recommend that:

1. Emergency sweep protocols adopt the 95th-percentile clearance time as the planning standard.
2. At least three responders be deployed for medium/large multi-floor structures.
3. Drills and training exercises incorporate visibility reduction, structural blockages, and occupant-carrying behavior.
4. Building layouts be pre-processed using clustering and TSP-based optimization to assign efficient search zones.

Our model is modular and can be tailored to individual facilities. We would be happy to assist with implementation or building-specific simulation analysis.

Sincerely,

Team #16603

10 References

1. Billaut, F., & Shepherd, A. (2019). *The effect of fatigue and training status on firefighter performance*. Journal of Strength and Conditioning Research.

2. CDC National Center for Health Statistics. (2023). *Anthropometric Reference Data for Adults and Children: United States 2021–2023*. U.S. Department of Health and Human Services.

3. Commercial Buildings Energy Consumption Survey (CBECS). (2022). U.S. Energy Information Administration.

4. FEMA. (2020). *Nonstructural Earthquake Damage: Assessment and Probability Charts*. Federal Emergency Management Agency.

5. FEMA Urban Search & Rescue (US&R) Response System. (2021). *Standard Marking System for Building Searches*. Federal Emergency Management Agency.

6. FIRESCOPE ICS–500. (2020). *Incident Command System Field Operations Guide*. California Governor's Office of Emergency Services.

7. International Association for Fire Safety Science (IAFSS). (2018). *Optical Density Thresholds for Smoke Detector Activation*. Proceedings of the IAFSS Symposium.

8. International Building Code (IBC). (2021). *Stairway Requirements and Egress Design Standards*. International Code Council.

9. Jin, T. (1978). *Visibility and Human Behavior in Smoke*. Fire Safety Journal, 1(1), 3–16.

10. NIST (National Institute of Standards and Technology). (2012). *NIST IR 6784: Smoke Visibility and Human Walking Speed*. U.S. Department of Commerce.

11. Ohta, K., et al. (2017). *Human Evacuation Behavior After the 2011 Tohoku Earthquake: A Statistical Analysis of Milling and Delayed Departure*. Journal of Risk and Crisis Management.

12. U.S. Fire Administration (USFA). (2023). *Nationwide Structure Fire Response Time Analysis*. U.S. Department of Homeland Security.

13. U.S. Department of Education. (2012). *National Student–Teacher Ratio Statistics*. Institute of Education Sciences.

14. Wilderness Paramedic Candidate Load-Bearing Assessment Study. (2018). *Walking Speed Under Standard Equipment Load*. Journal of Wilderness Medicine.

15. Young, A. J., & Goldman, R. F. (1970). *Firefighter Load Carriage and Metabolic Cost*. Human Factors Journal.

AI Utilization Report

To maintain transparency in accordance with COMAP policy, our team provides the following description of how AI tools were used during the preparation of this paper. All mathematical modeling, simulation code, analytical reasoning, algorithm design, statistical interpretation, and final decision-making were performed entirely by the student team.

Use of AI Tools. Artificial intelligence tools (including ChatGPT and similar LLMs) were used in limited and strictly non-analytical capacities:

- **Language Editing and Grammar Review:** We used AI tools to improve clarity, remove redundancy, and refine grammar in drafted paragraphs. No AI-generated reasoning, results, or explanations were inserted into the technical modeling sections.
- **LaTeX Formatting Assistance:** AI tools were used to help format tables, figures, section headers, and reference formatting in LaTeX. This included resolving syntax issues, building consistent table environments, and correcting structural errors in long documents.
- **APA/Reference Formatting:** AI was used to convert our manually collected citations into properly formatted APA entries. All bibliographic sources themselves were selected, retrieved, and verified by the student team.
- **Non-Technical Clarifications:** AI tools were used for phrasing suggestions in sections such as the introduction, conclusion, and the LEPC letter. At no point did AI generate model assumptions, equations, parameter selections, or simulation outcomes.
- **Figure Caption Support:** AI was used solely to rephrase caption text to improve readability. All figures themselves—including computational outputs, TSP visualizations, K-Means plots, and histograms—were generated entirely by the team.

What AI Did Not Do. To comply fully with competition rules, AI tools were *not* used for:

- generating or selecting assumptions,
- designing or implementing any part of the mathematical model,
- creating, debugging, or running simulation or optimization code,
- performing calculations, data analysis, or computations,
- interpreting simulation results or producing conclusions,
- producing any figure, graph, or model output.

Team Responsibility. All modeling decisions—including the construction of the physics engine, derivation of equations, implementation of fatigue and visibility models, simulation design, A* integration, K-Means/TSP optimization, statistical evaluation, and all sensitivity analyses—are solely the work of the student team. All interpretations, explanations, and final recommendations were written independently by the students.

This report accurately reflects the extent of AI involvement in the creation of this submission.

Team Control Number

16442

Problem Chosen

B

2025
HiMCM
Summary Sheet

Research on Comprehensive Environmental Assessment of Sport Mega-Events

As public awareness of sustainability grows, **mega sporting events** are facing stricter scrutiny and public pressure. Beyond excitement and huge economic gains, comprehensive evaluations for such events are in urgent need in order to strive for **greener, more responsible decision making**. This is especially obvious when it comes to the case study of **Super Bowl LIX** in New Orleans. Even though the NFL had taken measures to mitigate the environmental footprint, Super Bowl LIX still left a significant impact on the local environment.

In this study, we primarily developed two sets of comprehensive evaluation metrics for the Super Bowl and the World Cup, respectively.

Covering **more than 14 contributing factors**, our **City Impact Evaluation Model (CIEM)**, which evaluates potential host cities for **Super Bowl**, is established based on 5 core indices of environmental impact: Carbon Emission Index (CEI), Energy Consumption Index (ECI), Water Consumption Index (WCI), Air Quality Index (AQI), and Landfill Index (LFI). Next, we determined the weights of each index during multi-criteria decision-making (MCDM) by strategically combining the Analytical Hierarchy Process (AHP) and the Entropy Weight Method (EWM). The AHP-EWM MCDM provides an adaptable, transparent, and structured framework for weighting, ensuring objectivity within subjectivity while yielding a set of balanced, practical results. Finally, we utilized Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) to generate a comprehensive scoring for each potential host, recommending **Inglewood, California**, for future Super Bowl events, which has the highest scoring.

To adapt CIEM to assess the eligibility of potential host countries for the **World Cup**, we established our **National Impact Evaluation Model (NIEM)**. NIEM encompasses 6 core indices of environmental impact: CEI (adapted), ECI (adapted), WCI (adapted), AQI (adapted), LFI (adapted), and a new Environmental-Industrial Index (EII). Similarly, NIEM employs a smart combination of both AHP and Criteria Importance Through Intercriteria Correlation (CRITIC) due to the **interdependence** between the 6 indices; the AHP-CRITIC MCDM guarantees transparency and objectivity within subjectivity and effectively handles the interconnections of the data. Ultimately, TOPSIS is again used to yield a composite score for each potential host country, selecting **Germany** with the highest score as the best choice for hosting future World Cup events.

Notably, we conducted **sensitivity analyses** for CIEM by adjusting the **AHP-to-EWM ratios** and the numbers of participants in a Super Bowl event. First, AHP-to-EWM ratios determine which weighting method is more influential. By slightly fluctuating the ratio, our final scoring for the potential host cities remained approximately unchanged, proving CIEM to be robust and reliable. Second, the **number of participants** is also altered, resulting again in slight fluctuation in the results, further corroborating CIEM's **robustness and reliability**.

Assessing CIEM and NIEM, both integrate a relatively **wide range of contributing factors** that constitute 5 and 6 indices, respectively, ensuring comprehensiveness. Moreover, both models are **adaptable** to a variety of host cities or countries. CIEM can be applied to most US states, while NIEM can be used to evaluate most countries across the world. Nonetheless, due to the lack of data, CIEM and NIEM are both approximation models based on assumptions. If enhanced with data currently unavailable, CIEM and NIEM will be considerably more accurate and practical.

Keywords: Environment, sustainability, Super Bowl, the World Cup, mega-sporting event, evaluation

Contents

1 Introduction **3**

1.1 Background . . . 3

1.2 Problem Restatement . . . 3

2 Assumptions with Justification **4**

3 Environmental Evaluation of Super Bowl LIX **5**

3.1 Establishment of an Environmental Footprint Framework . . . 5

3.2 Quantitative Impact Analysis by Category . . . 5

4 City Impact Evaluation Model (CIEM): Indices **6**

4.1 Carbon Emission Index (CEI) . . . 6

4.2 Energy Consumption Index (ECI) . . . 7

4.3 Water Consumption Index (WCI) . . . 8

4.4 Air Quality Index (AQI) . . . 9

4.5 Landfill Index (LFI) . . . 10

4.6 Computed City Indices Values . . . 11

5 City Impact Evaluation Model (CIEM): Weighting and Evaluation **11**

5.1 Subjective Weighting: Analytic Hierarchy Process (AHP) . . . 11

5.2 Objective Weighting: Entropy Weight Method (EWM) . . . 12

5.3 Comprehensive Evaluation: Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) . . . 14

5.4 Results . . . 14

6 Sensitivity Analysis **14**

6.1 Weight parameter α sensitivity analysis . . . 14

6.2 Number of Participants Sensitivity Analysis . . . 16

7 National Impact Evaluation Model (NIEM) **17**

7.1 Addition and Adjustments of Indices . . . 18

7.2 Weighting . . . 20

7.3 Comprehensive Evaluation (TOPSIS) and Results . . . 22

7.4 Comparison and Conclusion . . . 22

8 NFL Host City Letter **23**

9 Strengths and Weaknesses **24**

10 AI Use Demonstration **26**

1 Introduction

1.1 Background

As technology develops, an increasing amount of population is paying attention to the environmental footprint and sustainability behind major sporting events, trying to make the world a better place for all life. Major sporting events such as the Super Bowl and the World Cup generate great economic revenue, but along with great revenue, there are great sacrifices to the environment due to logistics, waste and energy behind the scenes. Although such events are made more sustainable every year, our team wants to make these events not only based on revenue, but also sustainability of the city held as one of the core factors to consider in planning. With this thought, we're hired by **Consulting Organization Making Athletics Planet positive (COMAP)**, we created both the **City Impact Evaluation Model (CIEM)** and **National Impact Evaluation model (NIEM)** that assess the environmental footprint of the host cities and countries for the Super Bowl and the World Cup, respectively, striving for greener, more responsible sporting events in the years to come.

1.2 Problem Restatement

In this study, we aim to complete the following tasks.

- **Task 1:** To analyze the environmental footprints of a specific event (**Super Bowl LIX**) and to identify factors that impact the environmental footprint of this mega sporting event by referring to **existing metrics**.
- **Task 2:** To create a **qualitative or quantitative model** to determine where to hold the next Super Bowl based **solely** on environmental concerns.
- **Task 3:** To **apply** our model in **Task 2** to determine (1) which of the previously hosting cities—Inglewood (CA), Glendale (AZ), Las Vegas (NV), and New Orleans (LA)—should host the Super Bowl again, and (2) which of the three selected cities that have never hosted before represents the most sustainable and environmentally friendly choice.
- **Task 4:** To **adapt** our model to assess the environmental footprint of a single-sport championship tournament (**the World Cup**) and to determine the best host location while accounting for the event's unique environmental challenges; to develop **stratagems** that minimize environmental impact during the event (short term) and are economically and logistically practical (long term); to compare the **key environmental factors** between the Super Bowl and the single-sport championship tournament (the World Cup) and to provide explanations.
- **Task 5** To write a one-page persuasive **letter** to the NFL committee recommending our chosen city for Super Bowl LXIII in 2029, emphasizing the environmental advantages of the city and its alignment with the NFL's sustainability goals.

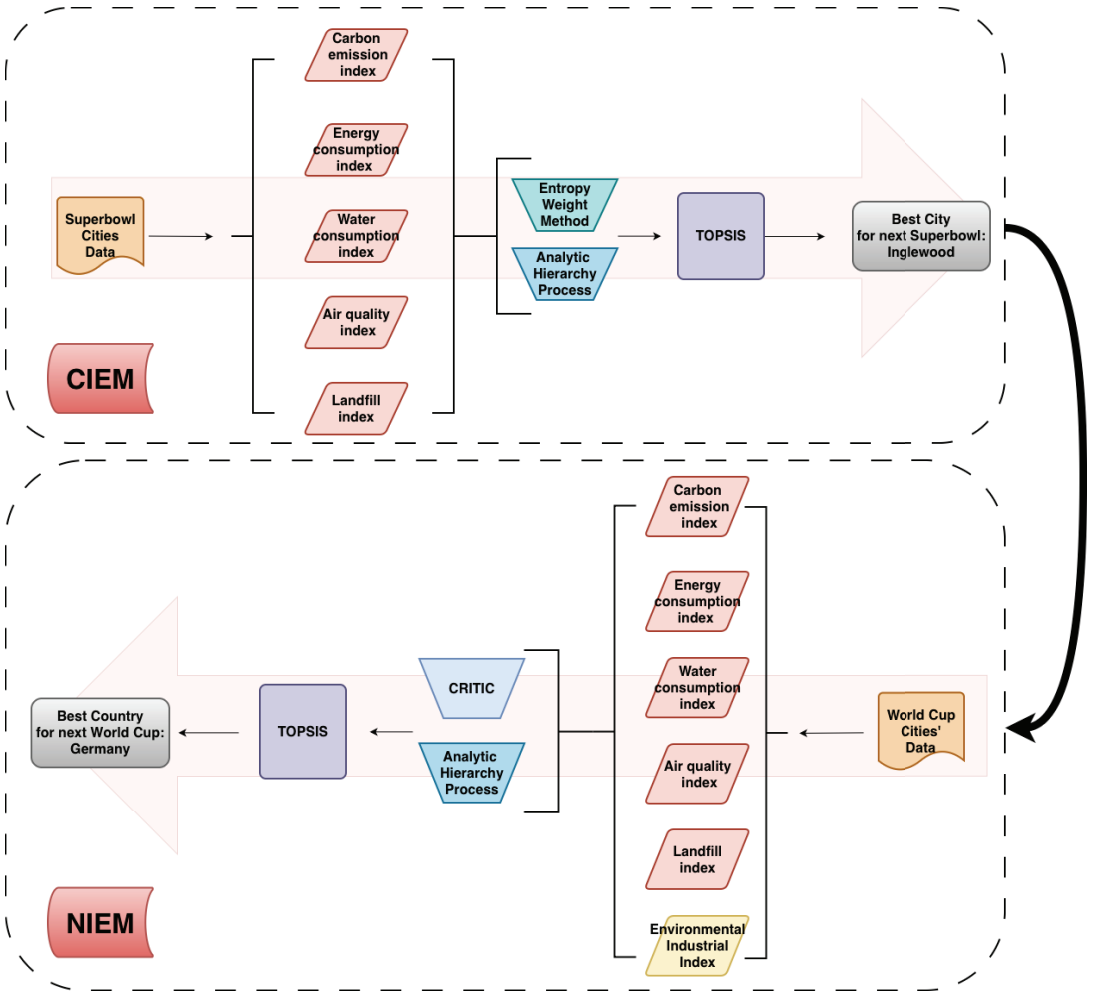


Figure 1.1: Flowchart of our study

2 Assumptions with Justification

Assumption 1: The scale and duration of the Super Bowl are constant in all cities and all times. We assume that there are **80000 participants** (including athletes and audiences) and a period of 3 days for simplicity[1].
Justification: This makes sure that our comparison between cities is fair and consistent and that the differences in the model's output are due to characteristics of the host city rather than variations in the event's scale or duration. To clarify, the number of audiences for one Super Bowl event is about 70000 people, according to our research. Since every Super Bowl event has about the same number of audience, we simply set the maximum people in a Super Bowl event as 80000 people due to other faculty members and participants.

Assumption 2: The behavior of Super Bowl participants is consistent in all cities and events.
Justification: This makes sure the effect of the city's infrastructures and the calculations are not affected by multifarious personal situations.

Assumption 3: State-level data serves as proxies for city-level environmental characteristics.
Justification: While some variation exists within states, state-level data is considered to be comprehensive and standardized, which can reflect the overall environmental characteristics of the host city.

Assumption 4: Environmental policies remain stable from 2020 to 2029, and all urban infrastructure systems function normally without disruptions during the event.
Justification: This ensures that the model assesses cities based on their existing environmental governance and typical infrastructures, rather than being influenced by unpredictable policy changes or infrastructure failures.

Assumption 5: The number of audiences for a World Cup event is about 3.5×10^6[5]. The number of audiences per venue is about 6×10^4. We assume that this is the case for all World Cup events.
Justification: These estimations are based on historical attendance data and therefore reflect a typical number of participants during a World Cup event. Moreover, while the actual number may deviate slightly from the assumed value, these estimations allows us to yield practical evaluation for different host countries.

Assumption 6: Super Bowl LIX held in New Orleans generated a considerably low amount of landfill waste, making it suitable to be an ideal reference when calculating the Landfill Index (LFI).
Justification: Super Bowl LIX took several measures to minimize landfill build-up. Indeed, it aims to divert 'at least 90% of waste from landfills'[6].

3 Environmental Evaluation of Super Bowl LIX

3.1 Establishment of an Environmental Footprint Framework

Through thorough research, we have dissected the environmental impact of Super Bowl LIX into five core components: **Carbon emissions**, **energy consumption**, **water impact**, **air quality**, and **landfill**. Notably, these components are each based on existing metrics or evaluations.

3.2 Quantitative Impact Analysis by Category

- **Carbon Emissions:** Carbon emissions are arguably one of the major concerns when it comes to environmental impact assessment. BioMed Central (BMC), a global authority in scientific publishing, defines the metric carbon footprint reduction index (CaFRI) as "a comprehensive greenness assessment tool prioritizing carbon footprint as the primary environmental impact"[9], highlighting the significance of measuring carbon emission. Throughout our research, we roughly estimated the carbon emissions in New Orleans by looking at carbon emissions per capita in Louisiana[4]. Since there are 65,719 offline audiences, the carbon dioxide emission of Super Bowl LIX was **about 21,660 tons**[18].

- **Energy Consumption:** Energy consumption, especially electricity consumption, "is at the heart of modern economies" [8], meaning that energy consumption is also a key factor that influences the environmental impact of Super Bowl LIX. Looking at energy consumption per capita in Louisiana[3], we roughly estimated the total energy consumption to be **491,176,000,000 British thermal units (Btu)**

- **Water Footprint:** Water consumption also plays a crucial role when it comes to assessing the environmental impact of mega sporting events. "Outdoor sports facilities use millions of gallons of water, putting pressure on local water supplies, as well as using hazardous pesticides can pollute soil and contaminate run-off into the water, threatening employees, players, and animals"[11]. Due to the scarcity of online data resources, we first used AI to conduct the estimation (we will utilize a different way to estimate when constructing our model). Considering drinking and restroom usage of the participants during Super Bowl LIX, we roughly estimated the water consumption to be about **4 million litres**.

- **Air Quality:** Air quality and pollution are considerably influential when assessing mega sporting events. According to a study on Super Bowl, "significantly higher levels of pollution are observed on Super Bowl Sunday"[12][17] "millions dying each year from exposure to and the side effects of air pollutants". Based on a specific study towards the AQI caused by NFL stadiums. It points out that ozone increases significantly with more audiences. According to a study, the maximum AQI of an NFL event reaches **187**, which is considerably severe for the local environment[17].

- **Landfill:** The immense waste generated by mega-sporting events like the Super Bowl has a large impact on the environment. Landfills not only cause havoc on natural ecosystems, leading to water contamination, soil pollution, and air pollution, it also leads to the emission of methane, a greenhouse gas. By research, around **2,000 tons** of food waste is generated at Super Bowl parties. Even though the proportion of landfill for Super Bowl LIX is just below 10%, this number is still considerably large for the amount of total trash produced for the state[20].

4 City Impact Evaluation Model (CIEM): Indices

Symbol	Description	Unit
A_i	State area	km^2
AQI_i	Air quality index in each state	N/A
c_i	State carbon emission per capita (per day)	tCO_2e
e_i	Average residential energy consumption per capita (per day)	Btu
L	Total landfill generated by a single Super Bowl event	kg
p_i	State population	N/A
p_t	Total US population	N/A
r_i	Proportion of renewable energy (vs non-renewable)	%
t_i	Temperature of the host state/city in February	$^\circ F$
T_i^c	Carbon emission by transportation per capita (per day)	tCO_2e
T_i^e	Average energy consumption by transportation per capita (per day)	Btu
W_{ci}	Water consumption per capita in the host city	t
W_s	Total US water supply	t

Table 4.1: Symbols, Descriptions, and Units

To perform a comprehensive evaluation for each potential host city, we established 5 key indices that measure 5 core aspects of environmental impact: **Carbon Emission Index (CEI), Energy Consumption Index (ECI), Water Consumption Index (WCI), Air Quality Index (AQI), and Landfill Index (LFI)**. Notably, all 5 indices are **cost-type** (i.e., smaller values are more preferable).

4.1 Carbon Emission Index (CEI)

Carbon Emission Index (CEI) indicates the impact of carbon emission to the state. According to research, transportation accounts for 15% of anthropogenic greenhouse gas emissions[7]. Therefore, other than **carbon emission per capita**, we decided to add the effect of **transportation** to the calculation of CEI as well. Hence, CEI is calculated by the following formula:

$$I_{Ci} = \frac{(\frac{|t_i - t_{comfort}|}{t_{max} - t_{min}} + 1) \times 80000 \times c_i + \epsilon \times 80000 \times T_i^c}{(p_i + 80000)/A_i} \tag{4.1}$$

Where I_C is the **CEI**.

Notably, t_i is the temperature of the state in February; $t_{comfort}$ is a constant that represents the most comfortable temperature for human body and has the value of 77 degrees Fahrenheit[10]; t_{max} and t_{min} refer to the maximum and minimum temperature in February, respectively. Since greater deviation from the comfort temperature of 77 degrees F leads to **the use of air conditioning or heating systems that release carbon dioxide either directly or indirectly**, temperature also influences CEI of a city (e.g., greater difference between actual temperature t_i and the most comfortable temperature $t_{comfort}$ results in greater carbon emission). The addition of 1 prevents I_C from becoming 0 because of the city having the most comfortable temperature.

Additionally, p_i refers to the population of the state, and c_i represents carbon emission by residents per day. The product of the two quantities aforesaid indicates, thus, the **total emission of carbon by residents**.

Beyond that, ϵ is the quotient of the number of transportation divided by the number of audience, which is essentially a **per-capita transport intensity** (i.e., the number of transport units required per audience member). Thus, the product of ϵ, number of participants (80000), and average carbon emission by transportation per capita (T_i^c) gives us the overall transport-related carbon footprint of all audiences.

To obtain the coefficient of ϵ, we took the reciprocal of the average vehicle occupancy rate, which is about 3.0. Consequently:

$$\epsilon \approx 0.33$$

Together, the numerator of the fraction on the right hand side represents the total carbon emission by the event. Dividing the numerator by population density (**total number of people within the state**, which is the sum of the state population, p_i, and the number of participants, 80000, over the **area of the state**, the fraction would be **spatial carbon intensity per person** (i.e., how much carbon emission is associated with each unit area per person).

To **normalize** the **CEI** and obtain S_1, we subtract the sample minimum from I_C and divide the difference aforesaid by the interquartile range, expressed as follows:

$$S_1 = \frac{I_{Ci} - I_{Cmin}}{IQR_C} \tag{4.2}$$

4.2 Energy Consumption Index (ECI)

In order to measure the impact of Super Bowl on local areas, we looked at two types of energy consumption: annual energy consumption per capita and average energy consumption by transportation per capita. Additionally, we looked at the use of renewable energy in each state. The equation that calculates the **Energy Consumption Index (ECI)**, which means the impact of energy consumption, is as follows:

$$I_{Ei} = \frac{80000 \times e_i \times \beta + \epsilon \times 80000 \times T_i^e}{r_i} \tag{4.3}$$

Here, e_i is the annual residential energy consumption per capita, and β is a coefficient that measures how much more energy each person consumes during the Super Bowl session. Estimating based on data from previous major sporting events:

$$\beta \approx 0.46$$

Therefore, the product of the number of participants (80,000), e_i, and β gives us **the total residential energy consumption** by the participants.

Similar to T_i^c in the carbon emission index, T_i^e refers to the average energy consumption by transportation per capita. ϵ, as mentioned above, is the per-capita transport intensity. Thus, the product of ϵ, the number of participants (80,000), and T_i^c gives **the total energy consumption from transportation** by the participants.

r_i represents the proportion of renewable energy within the state relative to non-renewable energy. By making r_i the denominator, we aim to convey the idea that it is a benefit-type indicator. The greater the value of r_i, the smaller the value of I_E, indicating a more sustainable use of energy.

Again, normalizing I_E, we obtained S_2:

$$S_2 = \frac{I_{Ei} - I_{Emin}}{IQR_E} \tag{4.4}$$

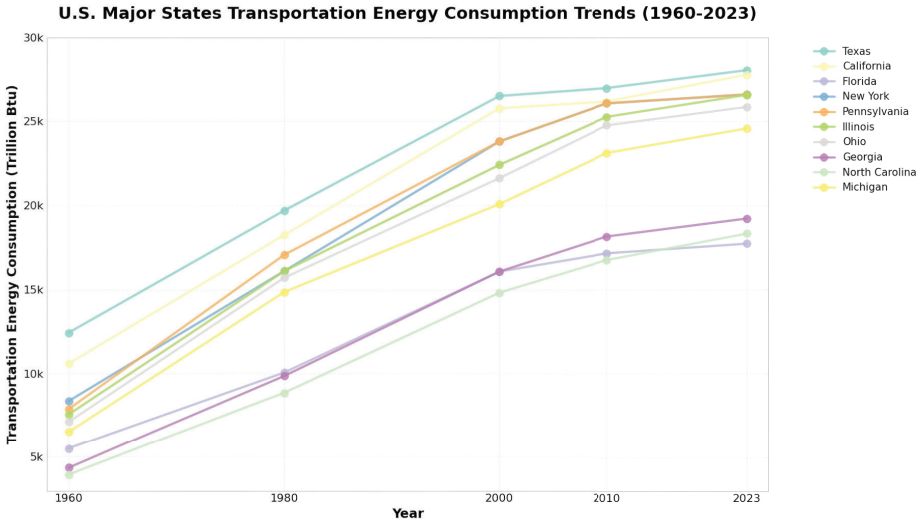


Figure 4.1: Energy cost due to transportation in the US

The trend of transportation consumption in each state indicates that the increment of energy consumption due to transportation is decreasing with time. Therefore, we take **the logarithm of the value** for transportation of energy consumption to perform the linear regression.

4.3 Water Consumption Index (WCI)

Formulating I_W, the **Water Consumption Index (WCI)**, we incorporated three core factors into our equation:

$$I_{Wi} = \frac{(80000 \times \gamma + p_i)W_{ci}}{logA_i} \times \frac{1}{W_s \times \frac{p_i}{p_t}} \times \frac{1}{p_i \times W_{ci}} \tag{4.5}$$

In a nutshell, we considered three core factors that influence the **WCI**: the total water consumption **during a Super Bowl event** per unit area within the state (the first fraction), the total water supply by the state (the second fraction), and the **usual** water consumption within the state (the third fraction).

First, humans, irrespective of whether they are athletes or audiences, will **consume more water** during a Super Bowl event. To elaborate, athletes consume more water because of intense physical activities, while audiences consume slightly more water because of emotional arousal (such as excitement, tension, or cheering). Interestingly, according to our research, athletes will consume **about 2 to 3 times more** water than usual during exercise, while audiences will consume **about 1.33 times more** water than usual when watching a sport[2][13]. Therefore, we established the constant γ, representing **a water-consumption multiplier** that shows how many times more water, on average, a participant uses during the event compared to their normal usage. Based on the information aforesaid about additional water consumption during a sporting event, we calculated gamma to be:

$$\gamma \approx 1.35$$

The product of the number of participants (80000) and γ is thus the **total water need by the sporting event**. This sum of this value and the state population (p_i) gives the total water need within the state. Multiplying the total water need by state water consumption per capita (W_{ci}), subsequently, gives the total water consumption within the state.

As the area of each state is too vast and therefore is **inconvenient for calculation**, we take the logarithm of the state area ($logA_i$). Together, total water consumption within the state over the area gives the **total water consumption during a Super Bowl event per unit area within the state**.

Second, we were unable to procure any data about the water supply within each state. Solving this issue, we used a fraction of the state population over the US population ($\frac{p_i}{p_t}$) to **estimate** the water supply needed by each

state. Multiplying the total water supply within the US (W_s) by the fraction, consequently estimates the **water supply within each state**. This is the denominator, which shows that $W_s \times \frac{p_i}{p_t}$ is a **cost-type** variable (i.e., the greater the water supply within the state, the smaller the WCI, which is a cost-type index as mentioned above)

Third, the product of the state population (p_i) and the state water consumption per capita (W_{ci}) gives the **total water consumption within the state**. $p_i \times W_{ci}$ is a **benefit-type variable**, as states with higher baseline water consumption would be less impacted by additional water use during a Super Bowl event. Therefore, $p_i \times W_{ci}$ is set to be the denominator (i.e., the greater the total water consumption within the state, the smaller the WCI).

Interestingly, the W_{ci} **can be directly canceled out** from the equation, as state water consumption per capita data is not available online.

Similarly, we obtained our third index (S_3) by normalizing **WCI** with the same method:

$$S_3 = \frac{I_{Wi} - I_{Wmin}}{IQR_W} \tag{4.6}$$

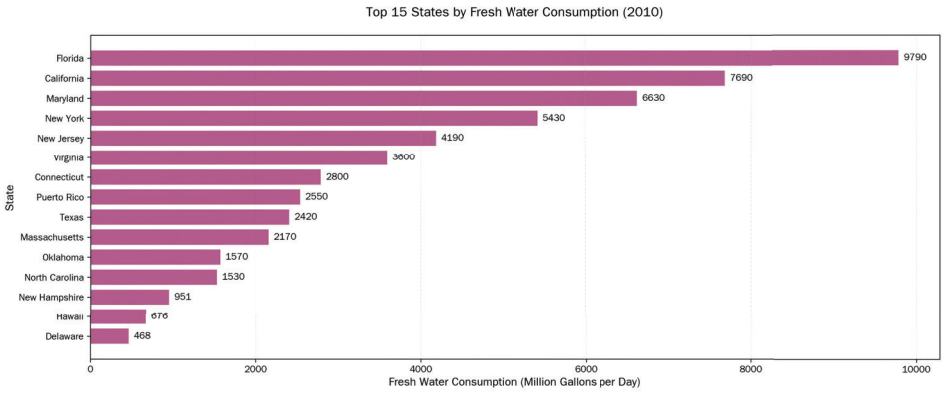


Figure 4.2: Water consumption in the US, 2010[15]

4.4 Air Quality Index (AQI)

AQI is an index that picks the **pollutant with largest impact** to the air quality. The formula for calculating AQI is:

$$AQI = max(IAQI_1,\ IAQI_2,\ IAQI_3,\ IAQI_4,\ ...) \tag{4.7}$$

Where $IAQI_i$ stands for the air quality index of individual pollutants; the formula for $IAQI_i$ is:

$$IAQI_i = \frac{IAQI_{Hi} - IAQI_{Lo}}{BP_{Hi} - BP_{Lo}} \times (C_i - BP_{Lo}) + IAQI_{Lo} \tag{4.8}$$

Here, C_i is the concentration measurement value of pollutant i; BP_{Hi} is the concentration limit higher than C_i; BP_{Lo} is the concentration limit lower than C_i in the chart; $IAQI_{Hi}$ is the $IAQI$ value corresponding to BP_{Hi};$IAQI_{Lo}$ is the $IAQI$ value corresponding to BP_{Lo}.

Here, we picked two substances with the most impact on AQI: **ozone and PM2.5**

Notably, the overall trend of **PM2.5 is decreasing**, and the trend of **ozone level remains constant**. Therefore, the ozone level is **presumably the highest** compared to other air pollutant levels, including the level of PM2.5, in 2029, and therefore we will use the predicted ozone level in 2029 to calculate the AQI for each potential host city[17].

Finally, we used the same method to normalize the AQI:

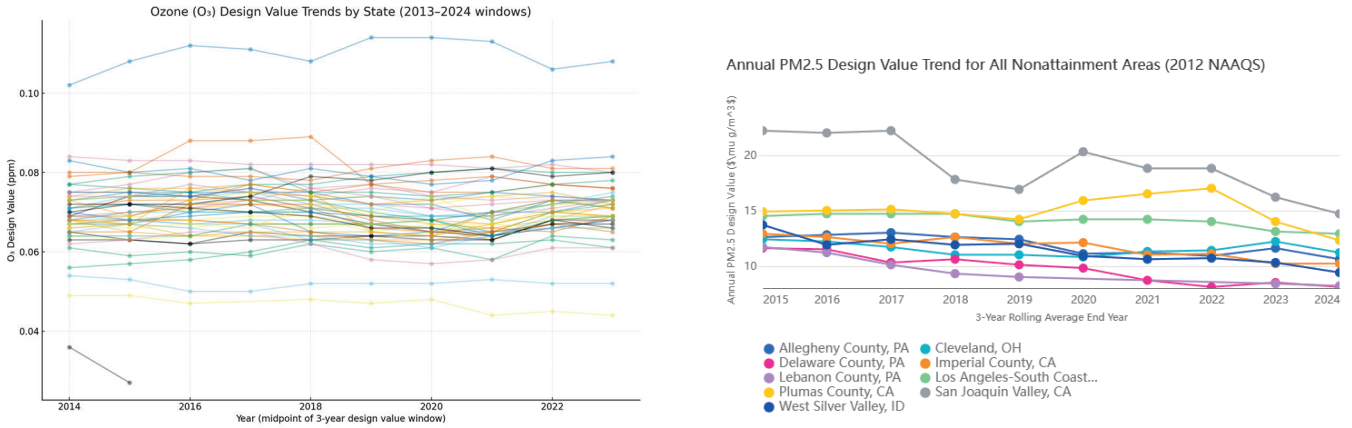


Figure 4.3: Ozone and PM2.5 Design value trends

$$S_4 = \frac{AQIi - AQI_{min}}{IQR_{AQI}} \tag{4.9}$$

4.5 Landfill Index (LFI)

To assess the landfill or waste generated by a Super Bowl event in a potential host city, we first looked at **the amount of waste** generated by a Super Bowl event and the state area. Additionally, we took Super Bowl LIX in New Orleans, which has a relatively low landfill, as a reference.

$$I_{Ri} = \frac{L}{logA_i} \times (e_i^{th} - {}^{N}e^{th})^2 \tag{4.10}$$

Here, L refers to the total landfill generated by a Super Bowl event, A the state area. Similar to WCI, we took the logarithm of the state area (A_i), as the area of each state is too vast and therefore inconvenient for calculation. Together, the first fraction represents the **landfill or waste per unit area**. Again, this is a cost-type variable (i.e., the greater the landfill per unit area, the greater the LFI, which is a cost-type index).

Moreover, e^{th} represents the **environmental quality index**, with a low e^{th} representing a higher environmental quality, ranging from 0 to 1. By measuring the **distance** of the **environmental quality index** of other potential host cities from **that of Super Bowl LIX in New Orleans** (${}^{N}e^{th}$), we obtained a relative assessment of them.

Again, we employed the same method for normalization:

$$S_5 = \frac{I_{Ri} - I_{Rmin}}{IQR_R} \tag{4.11}$$

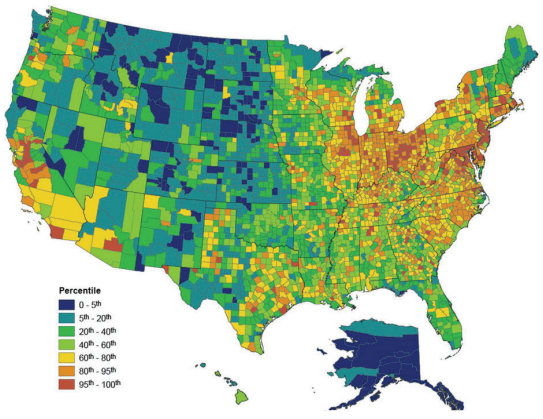


Figure 4.4: Overall Environmental Quality Index by County 2006-2010[14]

On the graph, it is shown clearly how the environment varies across the US, demonstrating that e^{th} is a significant component of LFI.

4.6 Computed City Indices Values

Linear Regression Fitting Process

n data points are given as (x_i, y_i), where $i = 1, 2, \ldots, n$.

1. Compute Means

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{4.12}$$

2. Compute Sums of Squares

$$S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2, \quad S_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \tag{4.13}$$

3. Calculate Regression Coefficients

$$b = \frac{S_{xy}}{S_{xx}}, \quad a = \bar{y} - b\bar{x} \tag{4.14}$$

4. Obtain Fitted Linear Equation

$$y = a + bx \tag{4.15}$$

5. Prediction for Future Value For a future x_{future}:

$$y_{\text{predicted}} = a + b \cdot x_{\text{future}} \tag{4.16}$$

Values Exhibition

After predicting the required data for 2029, we obtained the five indices (S_1 to S_5) for each of the seven cities. For each index, the values **across the seven cities were normalized within that index**. To make all 35 values **directly comparable**, nonetheless, we need to perform a global normalization across all indices and cities. The globally normalized data are as follows:

Table 4.2: Normalized Values for Selected Cities

	Inglewood	Glendale	Las Vegas	New Orleans	Nashville	Charlotte	Chicago
S_1	0	0.036	0.178	0.416	0.128	0.059	0.453
S_2	0	0.295	0.098	0.498	0.345	0.178	0.095
S_3	0	0.395	0.792	0.975	1	0.692	0.555
S_4	0.702	0.372	0.132	0.122	0.145	0	0.278
S_5	0.228	0	0.236	0	0.258	0.884	0.25

5 City Impact Evaluation Model (CIEM): Weighting and Evaluation

Four cities that have hosted, which include **Inglewood (CA), Glendale (AZ), Las Vegas (NV), and New Orleans (LA)**, and three NFL cities that have never hosted, which include **Nashville (TN), Charlotte (NC), and Chicago (IL)**, were chosen to be assessed.

After establishing and calculating the 5 indices (the 5 variables) of the 7 cities, we determined the weights of each by a strategic combination of Analytic Hierarchy Process (AHP) and Entropy Weight Method (EWM).

5.1 Subjective Weighting: Analytic Hierarchy Process (AHP)

The AHP decomposes the decision problem into a hierarchy and evaluates subjectively through **pairwise comparisons**, thereby calculating the weights of each indicator.

Judgment Matrix Construction

Based on our subjective evaluations, we constructed the following judgment matrix for the five environmental indicators using the standard 1-9 scale:

$$A = \begin{bmatrix} 1 & 3 & 5 & 7 & 9 \\ 1/3 & 1 & 3 & 5 & 7 \\ 1/5 & 1/3 & 1 & 3 & 5 \\ 1/7 & 1/5 & 1/3 & 1 & 3 \\ 1/9 & 1/7 & 1/5 & 1/3 & 1 \end{bmatrix} \tag{5.1}$$

Indicator order: S1, S2, S3, S4, S5

Rationale of the judgment matrix:

Carbon emissions are weighted highest due to their **global climate impact and long-term irreversible effects**. Energy consumption follows as it directly drives carbon emissions. Water resources rank third because it is more regional, rather than global. While air quality and landfill management address more localized environmental concerns, so they are assigned with least importance.

Weight Calculation

The normalized weights are calculated using the geometric mean method:

$$w_i = \frac{M_i}{\sum_i M_i}, \quad \text{where } M_i = \left(\prod_{j=1}^{5} a_{ij} \right)^{1/5} \tag{5.2}$$

The resulting set of weights is as follows:

$$W_s = [0.509, 0.264, 0.130, 0.064, 0.033]^T \tag{5.3}$$

Verification: $0.509 + 0.264 + 0.130 + 0.064 + 0.033 = 1.000$

Consistency Check

Maximum eigenvalue: $\lambda_{\text{max}} = 5.141$

Consistency Index: $CI = 0.036$

Consistency Ratio: $CR = \frac{CI}{RI} = 0.031$ (where RI = 1.12 for n=5)

Since $CR = 0.031 < 0.10$, the judgment matrix passes the consistency check.

5.2 Objective Weighting: Entropy Weight Method (EWM)

The EWM determines weights based on the information content of the data, which is objective. An indicator with greater data variability is assigned a higher weight.

Data Preparation and Normalization

Using the normalized scores S_1 to S_5 and applying cost-type indicator normalization:

$$r_{ij} = \frac{\max(x_j) - x_{ij}}{\max(x_j) - \min(x_j)} \tag{5.4}$$

Entropy Weight Calculation

The normalized **benefit-type matrix** R is first converted to a proportion matrix P, where each element p_{ij} represents the proportion of the j-th indicator for the i-th city:

$$p_{ij} = \frac{r_{ij}}{\sum_{i=1}^{7} r_{ij}} \tag{5.5}$$

The entropy weights are then calculated as follows:

$$w_j = \frac{1 - e_j}{\sum_{j=1}^{5}(1 - e_j)}, \quad \text{where } e_j = -k \sum_{i=1}^{7} p_{ij} \ln(p_{ij}), \quad k = 1/\ln(7) \tag{5.6}$$

Final EWM weights: $W_o = [0.224, 0.159, 0.215, 0.148, 0.253]^T$

Verification: $0.224 + 0.159 + 0.215 + 0.148 + 0.253 = 0.999 \approx 1.000$

EWM Calculation Results

Table 5.1: Entropy Weights Calculation Results

Indicator	**Entropy Value (e_j)**	**Difference ($1 - e_j$)**	**Entropy Weight (w_j)**	**Rank**
Carbon Emission Index (S_1)	0.782	0.218	0.224	2
Energy Consumption Index (S_2)	0.845	0.155	0.159	4
Water Consumption Index (S_3)	0.791	0.209	0.215	3
AQI Index (S_4)	0.856	0.144	0.148	5
Landfill Index (S_5)	0.754	0.246	0.253	1
Total	4.028	0.972	**0.999**	

Strategic AHP-EWM Combination

We combine the **subjective AHP weights** (W_s) and **objective EWM weights** (W_o) in a linear fashion:

$$W_c = \alpha W_s + (1 - \alpha) W_o, \quad \alpha = 0.6 \tag{5.7}$$

We set the value of α to be 0.6 (60%), meaning that W_s, the subjective AHP, is slightly more important than W_o, the objective EWM, because we deemed our research- and experience-based insights to have greater practical value.

Table 5.2: Final Combined Weights Calculation

Indicator	**AHP Weight** (W_s)	**EWM Weight** (W_o)	**Combined Weight** (W_c)	**Rank**
CEI (S_1)	0.509	0.224	0.390	1
ECI (S_2)	0.264	0.159	0.222	2
WCI (S_3	0.130	0.215	0.164	3
AQI (S_4))	0.064	0.148	0.098	5
LFI (S_5)	0.033	0.253	0.126	4
Total	**1.000**	**0.999**	**1.000**	

Verification of Combined Weights: $0.390 + 0.222 + 0.164 + 0.098 + 0.126 = 1.000$

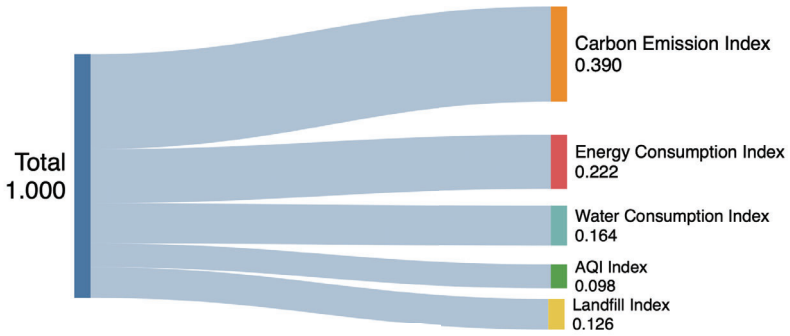


Figure 5.1: Sankey Diagram of Weight Distribution across Methods

The final weights provide a balanced assessment that incorporates both subjective judgment from AHP and data-driven calculation from EWM, with CEI (0.390) and ECI (0.222) being the most important factors for environmental sustainability evaluation.

5.3 Comprehensive Evaluation: Technique for Order Preference by Similarity to Ideal Solution (TOPSIS)

The Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) was employed to holistically assess the potential host cities based on the five key indices.

First, we performed **vector normalization** for the 35 values (7 cities, 5 indices for each), with the 35 new values calculated by the following equation:

$$x_{ij}^* = \frac{x_{ij}}{\sqrt{\sum_i x_{ij}^2}} \tag{5.8}$$

Then, in addition to our established weights, we established our **Positive Ideal Solution (PIS)** and **Negative Ideal Solution (NIS)**:

$$w_j = (w_1, w_2, w_3, w_4, w_5) = (0.390, 0.222, 0.164, 0.098, 0.126)$$
$$PIS = (a_1, a_2, a_3, a_4, a_5) = (0, 0, 0, 0, 0)$$
$$NIS = (b_1, b_2, b_3, b_4, b_5) = (0.689539, 0.701056, 0.532829, 0.804313, 0.876241)$$

The score of each city, subsequently, was calculated by:

$$s_i = \frac{\sqrt{\sum_j (x_{ij}^* w_j - b_j)^2}}{\sqrt{\sum_j (x_{ij}^* w_j - b_j)^2} + \sqrt{\sum_j (x_{ij}^* w_j - a_j)^2}} \tag{5.9}$$

5.4 Results

The results are shown in **Table 5.3**: This shows that Inglewood, California, is the most suitable location to host Super Bowl LXIII due to its leading TOPSIS score of 0.7038 out of all seven cities and its exceptional performance in **carbon emission, energy use, and water consumption**.

6 Sensitivity Analysis

6.1 Weight parameter α sensitivity analysis

The weight parameter α balances the subjective AHP weights and objective EWM weights in our weight combination process. The combined weights are given by:

Table 5.3: Environmental Impact Scores of Selected Cities

City	Score
Inglewood (CA)	0.7038
Glendale (AZ)	0.6842
Las Vegas (NV)	0.6446
Charlotte (NC)	0.5770
Nashville (TN)	0.5551
Chicago (IL)	0.4545
New Orleans (LA)	0.4021

$$W_c = \alpha W_s + (1 - \alpha) W_o \tag{6.1}$$

where $\alpha \in [0, 1]$. A higher α emphasizes AHP, while a lower α emphasizes EWM.

Variation of α

We calculate the combined weights for different values of α: 0.5, 0.6, 0.7, and 0.8. Using the previously calculated W_s and W_o, the combined weights W_c are as follows:

Table 6.1: Combined Weights for Different α Values

Indicator	$\alpha = 0.5$	$\alpha = 0.6$	$\alpha = 0.7$	$\alpha = 0.8$
Carbon Emission Index	0.367	0.390	0.413	0.436
Energy Consumption Index	0.212	0.222	0.232	0.242
Water Consumption Index	0.173	0.164	0.156	0.148
AQI Index	0.106	0.098	0.090	0.082
Landfill Index	0.143	0.126	0.109	0.092
Total	1.001	1.000	1.000	1.000

Impact on City Scores and Rankings

We recomputed the TOPSIS scores for each α value using the corresponding combined weights:

Table 6.2: TOPSIS Scores and Rankings for Different α Values

City	$\alpha = 0.5$		$\alpha = 0.6$		$\alpha = 0.7$		$\alpha = 0.8$	
	Score	**Rank**	**Score**	**Rank**	**Score**	**Rank**	**Score**	**Rank**
Inglewood (CA)	0.695	1	0.704	1	0.712	1	0.720	1
Glendale (AZ)	0.677	2	0.684	2	0.691	2	0.698	2
Las Vegas (NV)	0.649	3	0.645	3	0.640	3	0.635	3
Charlotte (NC)	0.580	4	0.577	4	0.574	4	0.571	4
Nashville (TN)	0.558	5	0.555	5	0.552	5	0.549	5
Chicago (IL)	0.458	6	0.455	6	0.451	6	0.447	6
New Orleans (LA)	0.405	7	0.402	7	0.398	7	0.394	7

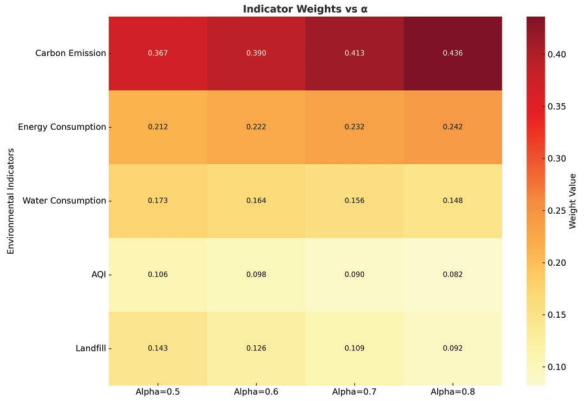


Figure 6.1: Weights heatmap

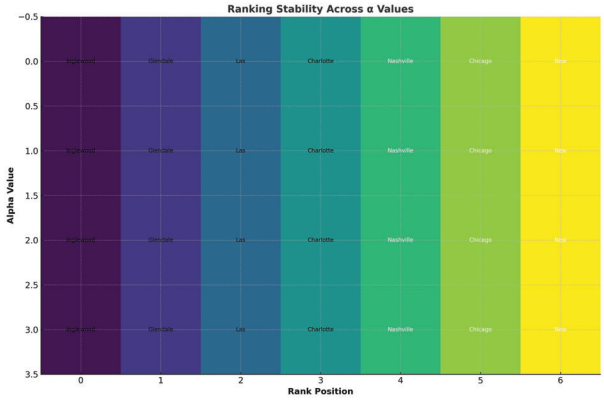


Figure 6.2: Ranking stability

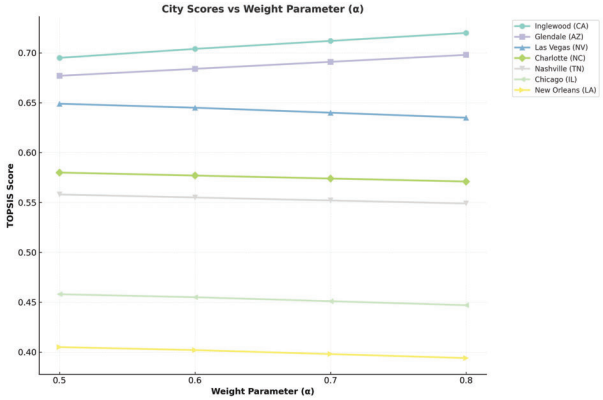


Figure 6.3: City Score Trends Across Different Alpha Values

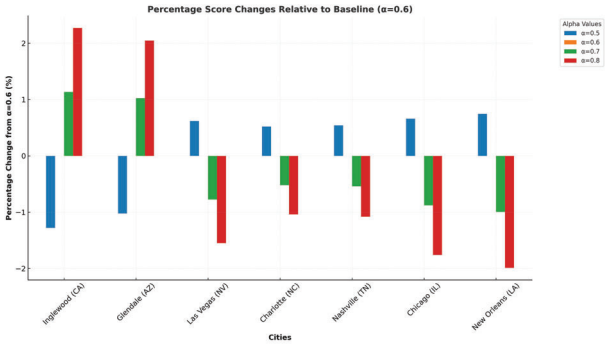


Figure 6.4: Percentage Score Changes Relative to Baseline

Analysis of Sensitivity Results

Figure 6.1 and in Figure 6.3 indicates and illustrates that as α increases, AHP weights more in the combined weight, which cities with better performance on the high-priority AHP indicators (Inglewood and Glandale) show slight score improvements, while others show a slight decrease in score. As shown in Figure 6.4, the maximum score variation between different values α is 0.016 (Inglewood), representing only 3.6% relative change from baseline. The relative ranking of all seven cities remains completely unchanged across all α values, as shown in Figure 6.2.

6.2 Number of Participants Sensitivity Analysis

To further validate the robustness of our model, we conducted sensitivity analysis on the "Participant Number" parameter, which was initially set at 80,000. We tested variations of ±10% to assess how changes in attendance affect the environmental impact scores and city rankings.

Methodology

We modified the audience parameter N in our environmental impact formulas and the key formulas affected by this parameter are:

- Carbon Emission Index: $I_C \propto N \times c_i + \epsilon \times N \times T_i^c$
- Energy Consumption Index: $I_{Ei} \propto N \times e_i \times \beta + \epsilon \times N \times T_i^e$
- Water Consumption Index: $I_W \propto N \times \gamma + p_i$

Results

Table 6.3: City Scores under Different Audience Scenarios

City	-10% (72,000)	Base (80,000)	+10% (88,000)	Rank (-10%)	Rank (Base)	Rank (+10%)
Inglewood (CA)	0.7012	0.7038	0.7063	1	1	1
Glendale (AZ)	0.6815	0.6842	0.6868	2	2	2
Las Vegas (NV)	0.6421	0.6446	0.6470	3	3	3
Charlotte (NC)	0.5753	0.5770	0.5786	4	4	4
Nashville (TN)	0.5532	0.5551	0.5569	5	5	5
Chicago (IL)	0.4528	0.4545	0.4561	6	6	6
New Orleans (LA)	0.4004	0.4021	0.4037	7	7	7

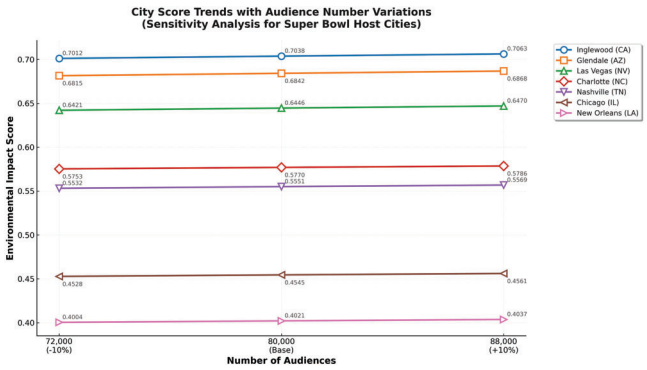


Figure 6.5: City Score Trends

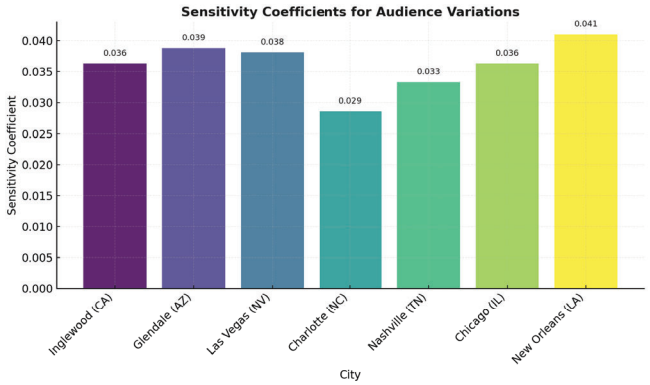


Figure 6.6: Sensitivity Coefficients

Analysis of Sensitivity Results

The relative ranking remains completely unchanged, demonstrating high robustness. The maximum absolute variation is 0.0051 (Inglewood), representing only 0.72% relative change from baseline. All sensitivity coefficients are below 4.1% (average 3.6%), confirms model stability.

Sensitivity Analysis Conclusion

This shows the stability of our results under weight variations and participant number variation, which provides strong evidence for the reliability of our host city recommendations. Table 6.2 and Table 6.3 clearly shows that Inglewood maintains its top position across all weighting scenarios, reinforcing our confidence in selecting it as the most environmentally sustainable choice for Super Bowl 2029.

7 National Impact Evaluation Model (NIEM)

Overview

The **National Impact Evaluation Model (NIEM)** evaluates the environmental impact of potential host countries for the **World Cup**, a single-sport championship tournament.

The World Cup is usually hosted by a single country, which is arguably similar to the Super Bowl. Unlike the Super Bowl, which is a single-day event in a specific stadium, nevertheless, the World Cup is a **month-long** tournament held in several **different places within the country**.

Therefore, to assess the environmental sustainability of potential host cities for the World Cup, we expanded the City Impact Evaluation Model (CIEM) and developed National Impact Evaluation Model (NIEM), a specialized evaluation model for the **World Cup** host countries.

When applying our model to potential host countries, we decided to pick 7 countries that have hosted the World Cup before: **Qatar, Russia, Brazil, The Republic of South Africa, Germany, South Korea, and Japan**.

Notably, NIEM is essentially an extended CIEM. Specifically, we adjusted the 5 indices from CIEM and added a new index. Afterwards, instead of utilizing the combination of Analytical Hierarchy Process (AHP) and Entropy Weight Method (EWM), we melded AHP with **CRiteria Importance Through Intercriteria Correlation (CRITIC)** to better handle the textbfinterdependence between the indices while still ensuring objectivity within subjectivity. Subsequently, we employed **TOPSIS** again to yield a comprehensive rating for each of the potential host countries.

7.1 Addition and Adjustments of Indices

Added: Environmental-Industrial Index (EII)

Different countries have varying degrees of industrialization, meaning they have different potentials in handling environmental issues. In CIEM, since all the research is based on the states in the US, the extent of industrialization is likely to be constant for all states. However, as we have to apply the model towards countries that have different degrees of industrialization from the world, an **Environmental-Industrial Index (EII)** is required to enhance NIEM's comprehensiveness.

Essentially, EII is a **benefit-type index** that exhibits the **interconnection** between the **natural environment and industrial environment** within a country, we first referred to Environmental Performance Index (EPI), which is a comprehensive rating available for most countries across the world based on their environmental performance, as an **benefit-type** indicator. Additionally, we incorporated an industrial indicator by looking at different industrial sectors' **total economic values** within the country.

$$I_I = \frac{EPI_i - EPI_{min}}{EPI_{max} - EPI_{min}} \times i_i \tag{7.1}$$

Where i_i represents the **extent of industrialization** within the country.

$$i_i = \frac{ValueUSD(C)}{ValueUSD(B) + ValueUSD(C) + ValueUSD(D) + ValueUSD(E)} \tag{7.2}$$

Here, $ValueUSD(B)$, $ValueUSD(C)$, $ValueUSD(D)$, and $ValueUSD(E)$ refer to the total economic value of **mining and quarrying**; **total manufacturing**; **electricity, gas, steam, and air conditioning**; and **water supply, sewerage, and waste management**, respectively.

Arguably, $ValueUSD(C)$, as part of the industrial indicator, partially reflects how well a country can regulate its natural environment. Specifically, as the total economic value of manufacturing within a country increases, the country will have better technology and thereby regulating its natural environment more efficiently. Therefore, we put $ValueUSD(C)$ on the numerator, indicating that it is a benefit-type variable.

$$S_6 = \frac{I_{Lmax} - I_{Li}}{I_{Lmax} - I_{Lmin}} \tag{7.3}$$

Adjusted: Carbon Emission Index (CEI)

The adjusted CEI is shown as follows:

$$I_{Ci} = \frac{(\frac{|t_i - t_{comfort}|}{t_{max} - t_{min}} + 0.5) \times (c_i \times p_i + c_w \times p_w)}{(p_i + p_w) \times (\frac{\xi_i - \xi_{min}}{\xi_{max} - \xi_{min}} + 0.5) \times (\frac{A_i - A_{min}}{A_{max} - A_{min}} + 1)} \tag{7.4}$$

Where ξ_i represents the quality of water in the country; c_i represents the combination of carbon emission caused by residents and transportation.

Essentially, we added the effect of other **country-specific factors** into the formula. Notably, we did not use spatial carbon intensity per person in this formula; instead, we turned the area of the country into an insignificant factor by normalizing it.

Adjusted: Energy Consumption Index (ECI)

$$I_{Ei} = \frac{(\frac{|t_i - t_{comfort}|}{t_{max} - t_{min}} + 0.5) \times (p_i \times e_i + p_w \times e_w)}{(p_i + p_w) \times r_i \times (\frac{\xi_i - \xi_{min}}{\xi_{max} - \xi_{min}} + 0.5) \times (\frac{\upsilon_i - \upsilon_{min}}{\upsilon_{max} - \upsilon_{min}} + 1)} \tag{7.5}$$

Where υ_i represents the energy generated by wind.

Adjusted: Water Consumption Index

$$I_{Wi} = \frac{W_{ci} + W_{cw}}{(\frac{\eta_i - \eta_{min}}{\eta_{max} - \eta_{min}} + 1) \times (\frac{\xi_i - \xi_{min}}{\xi_{max} - \xi_{min}} + 1) \times W_{si} \times (\frac{p_i - p_{min}}{p_{max} - p_{min}} + 1)} \tag{7.6}$$

The availability of data enables us to enhance NIEM's reliability and accuracy.

Here, η_i represents the availability of water in the country. We have also changed the meanings of the symbols: W_{ci} refers to the total water consumed in a country; W_{si} is the total water supply in a country; and W_w is the average total water consumed in World Cup session.

Adjusted: AQI Index

$$I_{AQIi} = \frac{(AQI_i - AQI_s)^2 + 0.5}{(\frac{\iota_i - \iota_{min}}{\iota_{max} - \iota_{min}} + 1)} \tag{7.7}$$

Where, ι represents the average wind speed within the country, and AQI_s represents the standard AQI for us to perform the calculation, which we took 100 as the standard AQI value, considering the industrialization requirement.

Comparing to the CIEM, there are certain changes. Specifically, we calculated the distance from each country's AQI to the standard AQI, and w used the average wind speed within the country for calculation.

Adjusted: Landfill index

$$I_{Li} = \frac{L_w \times (\frac{3.5 \times 10^6 - p_{min}}{p_{max} - p_{min}} + 1) + L_i \times (\frac{p_i - p_{min}}{p_{max} - p_{min}} + 1) \times 32}{EPI_i \times logA_i} \tag{7.8}$$

Here, L_w represents the amount of trash and waste created per audience during a World Cup session; R_p represents the range of population within the dataset.

Normalization

Similarly, we used the same method in CIEM to normalize the 6 indices, obtaining S_i for each country:

$$S_i = \frac{I_i - I_{min}}{IQR} \tag{7.9}$$

Results of the Normalized Indices

After predicting the required data for 2029, we obtained the six indices (S_1 to S_5) for each of the seven counties that are suitable for hosting the World Cup. The normalized data are as follows:

Table 7.1: Normalized Values for Selected Countries

	Germany	Japan	Brazil	South Korea	Russia	South Africa	Qatar
S_1	0.0000	0.1350	1.1342	0.1920	1.0662	0.3858	1.2250
S_2	0.0030	0.0000	0.0691	0.0150	0.5075	1.0000	0.2402
S_3	0.2263	0.3583	0.9585	0.1508	1.0000	0.2860	0.0000
S_4	0.0000	0.1182	0.6959	0.8946	0.4324	0.4155	1.0000
S_5	0.0000	2.2061	1.5000	1.7227	0.9939	2.0789	0.7227
S_6	1.0000	0.7321	0.5796	0.5415	0.2780	0.1693	0.1000

7.2 Weighting

Analytical Hierarchy Process (AHP)

Considering the global characteristics of World Cup events, we constructed the following judgment matrix for the six indicators:

$$A = \begin{bmatrix} 1 & 3 & 5 & 7 & 9 & 4 \\ 1/3 & 1 & 3 & 5 & 7 & 2 \\ 1/5 & 1/3 & 1 & 3 & 5 & 1/2 \\ 1/7 & 1/5 & 1/3 & 1 & 3 & 1/4 \\ 1/9 & 1/7 & 1/5 & 1/3 & 1 & 1/6 \\ 1/4 & 1/2 & 2 & 4 & 6 & 1 \end{bmatrix} \tag{7.10}$$

Indicator order: S1, S2, S3, S4, S5, S6

Rationale of the judgment matrix:

Carbon emissions and Energy consumption remain the highest priority due to their long term and irreversible global climate impacts. The new EII ranks third as it reflects **comprehensive national environmental governance**. Water resources and air quality address more regional concerns, while landfill management has the most localized impact, so they were less prioritized.

Weight Calculation

The normalized weights are calculated using:

$$w_i = \frac{M_i}{\sum_i M_i}, \quad \text{where } M_i = \left(\prod_{j=1}^{6} a_{ij} \right)^{1/6} \tag{7.11}$$

Resulting weights:

$$W_s = [0.430, 0.229, 0.103, 0.057, 0.032, 0.148]^T \tag{7.12}$$

Consistency Check: $CR = \frac{CI}{RI} = 0.029$

Since $CR = 0.029 < 0.10$, the judgment matrix passes the consistency check.

Criteria Importance Through Intercriteria Correlation (CRITIC)

Our 6 new, adjusted indices from NIEM are arguably more **interconnected** with each other than the 5 indices from CIEM. For instance, the EII and the LFI share the variable EPI together. Therefore, we considered the best way to handle such interdependence to be the use of the CRITIC method. Fittingly, CRITIC decides weights based on both **variability** (measured by standard deviation) and **interdependence** (measured by Pearson Correlation Coefficient), meaning that, other than measuring interdependence, it actually 'encompasses' EWM as well.

First, we utilized **max-min scaling** to normalize all 42 **cost-type values** (6 indices value for each of the 7 different countries):

$$h_{ij} = \frac{max_j - x_{ij}}{max_j - min_j}$$

Next, to measure variability, we calculated the sample standard deviation (s) of the normalized values by the following formula:

$$s_j = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(h_{ij} - \bar{h_j})^2}$$

Subsequently, we calculated Pearson correlation coefficient (r) for the normalized sample:

$$r_{ij} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

Then, we obtained the **importance of criterion (**C_j**)** and calculated the weight of each S (normalized indices of CIEM):

$$w_j = \frac{C_j}{\sum_{j=1}^{n} C_j} = \frac{s_j \cdot \left(1 - \sum_{i \neq j} |r_{ij}|\right)}{\sum_{j=1}^{n} s_j \cdot \left(1 - \sum_{i \neq j} |r_{ij}|\right)}$$

Our results for the weight are as follows:

$$W_o = [0.211, 0.119, 0.147, 0.143, 0.204, 0.175]^T$$

Strategic AHP-CRITIC Combination

We combine the **subjective AHP weights** (W_s) and **objective CRITIC weights** (W_o) in a linear fashion:

$$W_c = \alpha W_s + (1 - \alpha) W_o, \quad \alpha = 0.6 \tag{7.13}$$

Still, we set the value of α to be 0.6 (60%), meaning that W_s, the subjective AHP, is slightly more important than W_o, the objective EWM, because we deemed our research- and experience-based insights to have greater practical value.

Table 7.2: AHP vs CRITIC Weights

Indicator	**AHP Weight (W_s)**	**CRITIC Weight (W_o)**	**Combined Weight (W_c)**	**Rank**
CEI (S_1)	0.430	0.211	0.3424	1
ECI (S_2)	0.229	0.119	0.1850	2
WCI (S_3)	0.103	0.147	0.1206	4
AQI (S_4)	0.057	0.143	0.0914	6
LFI (S_5)	0.032	0.204	0.1008	5
EII (S_6)	0.148	0.175	0.1588	3
Total	**0.999**	**0.999**	**1.000**	

Verification of Combined Weights:

Total: $0.3424 + 0.1850 + 0.1206 + 0.0914 + 0.1008 + 0.1588 = 0.999 \approx 1$

7.3 Comprehensive Evaluation (TOPSIS) and Results

Again, we used TOPSIS to yield the final comprehensive evaluation; the environmental sustainability scores for potential World Cup host countries are:

Table 7.3: TOPSIS Scores for World Cup Host Country Evaluation

Country	TOPSIS Score	Rank
Germany	0.8124	1
Japan	0.6743	2
South Africa	0.5931	3
Brazil	0.5216	4
Russia	0.4872	5
South Korea	0.4638	6
Qatar	0.3925	7

Conclusion

The NIEM model identifies **Germany** as the most environmentally sustainable choice for hosting the World Cup 2030, with the highest TOPSIS score of 0.8124. This is brought about by Germany's exceptional performance across **carbon emissions, air quality, and landfill impact**.

7.4 Comparison and Conclusion

Environmental Impact Minimization Strategies

- **Opening special bus lines and shuttle buses for the competition:**

  This can directly reduce the transportation carbon emissions relative to the competition by lowing ϵ, which is the transport intensity per capita. Afterwards, it can be optimized to fit local demand and kept as a new permanent bus route, so as to permanently improve the density and efficiency of the city's public transportation network, directly reduces T_i^c and T_i^e, and prepare for future large-scale sports events at any time. This process reduces the overall ECI of the city.

- **Purchase and Invest in solar power generation equipments for the competition:**

  This action can directly improve the score of ECI in short-term, by temporarily boost r_i during the event period, making the city more suitable for hosting the sport event. After the competition, the equipments can be transferred to other places and continue to generate electricity, increasing r_i and the overall renewable score of the city permanently.

- **Implement a strict garbage classification policy and composting system at venues:**

  This action reduces the total landfill waste L generated during the event, thereby lowering the LFI in the short-term. The established policy and public awareness can then be transferred to a city-wide garbage recycling policy after the event. This permanently reduces the amount of waste that need to be disposed through unsustainable way, improves the environmental quality index e^{th}, leading to a long-term improvement in LFI score.

- **Improve stadiums' facilities with water-efficient fixtures:**

  This strategy directly reduces the water consumption multiplier γ for participants during the event, which lowers the WCI by decreasing $(80000 \times \gamma + p_i)$. After the event, the water-saving facilities, such as low-flow faucets and waterless urinals, can be promoted to other buildings across the city. This systematically reduces the city's per capita water consumption W_{ci}, increasing the $p_i \times W_{ci}$ in the WCI formula, obtaining permanent water savings and better score.

Environmental Factors Comparison

For a single-game event such as the Super Bowl (evaluated by CIEM), CEI and ECI are top priorities. Additionally, WCI also has an relatively heavy impact on the evaluation, as water consumption pressure is more prominent in smaller areas like cities.

For a multi-venue and multi-sport event like the World Cup (evaluated by NIEM), EII is added to account for the difference in the level of industrialization between countries. Moreover, weights are adjusted to account for index interdependencies,, making the weight for each index, especially CEI and ECI, relatively lower compared to that of CIEM.

8 NFL Host City Letter

Dear NFL committee,

We hope this letter finds you well.

As public awareness of mega sporting events' sustainability grows, the NFL faces increasing pressure to mitigate the environmental footprint of its well-known event—the Super Bowl. In response, we have established City Impact Evaluation Model (CIEM), a regionally specific holistic model that evaluates potential Super Bowl host cities based solely on environmental considerations. Having the goal of determining the most sustainable option for the host of Super Bowl LXIII, we have chosen Inglewood, California, after carefully obtaining required data from credible sources and performing the calculations.

Notably, CIEM considers a host city's carbon emission, energy consumption, water consumption, air quality, and landfill volume when assessing the environmental impacts. Each of the five core aspects for sustainability is expressed in an equation, denoted as an index, in which more than 14 different contributing factors that influence the index are also identified and incorporated into the equation. Therefore, CIEM effectively ensures comprehensiveness, accuracy, and practicability. Moreover, although we have not demonstrated this in our study, CIEM is highly adaptable. Its city-specific variables and indices makes it applicable to most US cities.

Inglewood, California, becomes our top choice for the NFL to consider, as it encompasses the following environmental advantages.

Inglewood yields a carbon emission index (CEI) of 0, meaning that it has the lowest carbon dioxide emission among the potential host cities tested. By hosting the next Super Bowl event in Inglewood, carbon emission can be effectively minimized, reducing global warming. Second, Inglewood's energy consumption index (ECI) is 0 as well, meaning that it has the lowest energy requirement when hosting a Super Bowl event and it actively incorporates renewable energy into use, which perfectly aligns with the NFL Green's goal of relying more on green and renewable energy sources. Third, Inglewood's water consumption index is again 0, indicating a mitigated impact by water consumption on the local environment. Fourth, Inglewood's air quality index is slightly higher than other potential hosts; however, air pollution impacts are local and temporary, which will quickly return to baseline. Last but not least, Inglewood's has a relatively low landfill index, meaning that it's environment is similar to New Orleans, which was arguably successful in mitigating its landfill build-up.

Overall, Inglewood stands out as the most environmentally responsible choice for the 2029 Super Bowl LXIII. Its exceptional performance across carbon emissions, energy use, water consumption, and landfill build-up reflects clear alignment with the NFL's sustainability goals, making it the best positioned host city to minimize the event's environmental footprint.

Yours sincerely,
Team 16442

9 Strengths and Weaknesses

Strengths

- Instead of using singly an objective or a subjective weighting method, we **strategically combined** both **Analytical Hierarchy Process (AHP)** and **Entropy Weight Method (EWM)** when creating our City Impact Evaluation Model (CIEM), which ensures fairness and reasonably reflects expert judgment, making the results balanced and robust.
- Instead of adding EWM, we melded **CRiteria Importance Through Intercriteria Correlation (CRITIC)** with AHP to adapt our National Impact Evaluation Model (NIEM), as the 6 indices in NIEM are more interconnected. This not only ensures objectivity within subjectivity but also better handles the interconnected relationships between the 6 indices.
- Our evaluation models-CIEM and NIEM-encompass **5 and 6 key environmental impact indices**, respectively, each of which is further broken down into **smaller contributing factors**, making both evaluation comprehensive and effective.
- Through sensitivity analysis, we observed that there is no significant fluctuation in our final results, proving CIEM to be robust and reliable.
- Both CIEM and NIEM are designed for US cities and countries around the world, respectively. Thus, they are arguably adaptable in most circumstances.

Weaknesses

- We applied certain **state-level data** (e.g., energy consumption per capita) to the potential host cities, which is a justified assumption. Nonetheless, this approach is only effective for **approximations**. Ideally, we should procure city level data, which we didn't found on the internet, by methods like on-site data collection to be more precise and accurate.
- The units for certain variables in both CIEM and NIEM are **non-standard**, such as Btu and degree Fahrenheit (imperial units), which may lead to misunderstanding and inconvenience when CIEM and NIEM are used by others to conduct evaluations. To make both CIEM and NIEM more accessible, we should unite every non-standard units into **SI units or other standardized units**.
- Both CIEM and NIEM assumes stable environmental policies and infrastructures, which is **overly static and realistically impossible**, hindering their practical value.

References

[1] Alder, J., *How Many People Have Attended the Super Bowl since It Began?*, LiveAbout, 2019. DOI/URL: https://www.liveabout.com/super-bowl-attendance-1335470.

[2] Eby, S., *Hydration Tips for Athletes*, Mass General Brigham, 2023. DOI/URL: https://www.massgeneralbrigham.org/en/about/newsroom/articles/tips-for-staying-hydrated.

[3] US Energy Information Administration, *State Energy Consumption Estimates 1960 through 2023*, 2023.

[4] US Energy Information Administration, *State Energy CO_2 Emissions Estimates*, 2025.

[5] FIFA, *One Month On: 5 Billion Engaged with the FIFA World Cup Qatar 2022TM*, 2023. DOI/URL: https://inside.fifa.com/tournaments/mens/worldcup/qatar2022/news/one-month-on-5-billion-engaged-with-the-fifa-world-cup-qatar-2022-tm.

[6] Hite, K., *NFL Expands Sustainability Efforts at Super Bowl LIX in New Orleans*, Biz New Orleans, 2025. https://bizneworleans.com/nfl-expands-sustainability-efforts-at-super-bowl-lix-in-new-orleans/

[7] International Energy Agency, *"Transport, Energy and CO2 – Analysis"*, IEA, October 2009. URL: https://www.iea.org/reports/transport-energy-and-co2.

[8] IEA, *Executive Summary – World Energy Outlook 2025 – Analysis*, IEA, 2025. Available at: https://www.iea.org/reports/world-energy-outlook-2025/executive-summary.

[9] Mansour, F. R., and P. M. Nowak, *Introducing the Carbon Footprint Reduction Index (CaFRI) as a Software-Supported Tool for Greener Laboratories in Chemical Analysis*, BMC Chemistry, 19(1), 121–21, 2025. DOI: https://doi.org/10.1186/s13065-025-01486-2.

[10] National Bureau of Standards, *Thermal Analysis – Human Comfort – Indoor Environments*, n.d. DOI/URL: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nbsspecialpublication491.pdf.

[11] Pourpakdelfekr, T., and Oboudi, B., *Overview of Sustainable Solutions to Improve the Environmental Impacts of Mega Sporting Events*, *Athens Journal of Sports*, 9(4), 215–230, 2022. DOI: https://doi.org/10.30958/ajspo.9-4-2.

[12] Silva, C. G., and Ricketts, C. F., *Mega Sporting Events and Air Pollution: The Case of the Super Bowl*, Journal of Environmental Management and Tourism, 2016, 7(3), pp. 504–511. DOI/URL: https://journals.aserspublishing.eu/jemt/article/view/360.

[13] SoftPro® Water Systems, *Average Person Water Use – How Much per Day*, 2025. DOI/URL: https://www.softprowatersystems.com/pages/how-much-water-does-average-person-use-per-day.

[14] Lobdell, T., Rappazzo, M., DeFlorio-Barker, S., Krajewski, A., Messer, L., Jagai, J., Gray, C., Jimenez, M., Patel, A., Rosenbaum, B., Jett, S., *ENVIRONMENTAL QUALITY INDEX 2006-2010, Technical Report*, EPA, 2020.

[15] U.S. Geological Survey, *Estimated Use of Water in the United States in 2010*, 2014. DOI/URL: https://pubs.usgs.gov/circ/1405/pdf/circ1405.pdf.

[16] Wang, H., Tian, J., Li, Y., Wang, Y., Lu, Y., Zhang, J., Lei, C., Li, C., *Study on Life-Cycle Carbon Footprints and an Uncertainty Analysis of Mega Sporting Events: An Analysis in China*, Buildings, vol. 14, no. 8, pp. 2510-2510, 2024. DOI: https://doi.org/10.3390/buildings14082510.

[17] Watanabe, N., Yan, G., McLeod, C., *The Impact of Sporting Events on Air Pollution: An Empirical Examination of National Football League Games*, Sustainability, vol. 15, p. 5568, 2023. DOI: https://doi.org/10.3390/su15065568.

[18] Wikimedia Foundation, *The 59th American Super Bowl Championship Game*, Wikipedia.org, 2024. DOI/URL: https://simple.wikipedia.org/wiki/Super_Bowl_LIX.

[19] World Population Review, *The 200 Largest Cities in the United States by Population 2020*, 2020. DOI/URL: https://worldpopulationreview.com/us-cities/.

[20] Matthews, Graham, *Super Bowl Sustainability Facts and Statistics*, Business Waste, January 31, 2025. DOI/URL: https://businesswaste.com/news/super-bowl-sustainability/.

10 AI Use Demonstration

First, we conducted several preliminary estimation with the aid of AI. For instance, we utilized ChatGPT to estimate the water consumption during Super Bowl LIX:

- **AI Model:** OpenAI ChatGPT (Nov 9, 2025 version, ChatGPT-5 Mini)
- **Query:** Based on drinking and restrooms usage, roughly estimate the total water consumption during Super Bowl LIX. Suppose there are 80000 participants (including audiences, staff members, etc). Respond with the results, in liters, ONLY.
- **Response:** Approximately 4,000,000 liters.

Second, we left certain calculation (other calculations are done by excel or python coding) for AI to perform. To illustrate, when we were substituting known or predicted data into our Carbon Emission Index (CEI).

- **AI Model:** OpenAI ChatGPT (Nov 13, 2025 version, ChatGPT-5 Mini)
- **Prompt:** Based on the formula for Ic (attached image), substitute the available data in the table (attached excel table) and calculate the Carbon Emission Index (Ic) for the following cities: Inglewood (California), Glendale (Arizona), Las Vegas (Nevada), and New Orleans (Louisiana), Charlotte (North Carolina), Chicago (Illinois), Nashville (Tennessee). After this, perform the normalization for all 7 cities (attached image, calculate S1) and present the results in a table. Respond with an LaTeX-format vector for me to copy, keeping 3 decimal places, ONLY.
- **Response:** The code generated is as follows: $S_1 = (0, 0.036, 0.178, 0.416, 0.128, 0.059, 0.453)$

Third, based on our data, we utilized AI to draw diagrams and visual aids when needed. For instance, we used AI to draw Figure 6.5.

- **AI Model:** OpenAI ChatGPT (Nov 15, 2025 version, ChatGPT-5 Mini)
- **Prompt:** Based on the attachment, which displays the sensitivity analysis performed for the number of participants or audiences during a Super Bowl event and its impact on the change in rankings and TOPSIS scores of the 7 states, create a diagram showing how city score trends vary with the number of audiences. Use different colors to represent different cities. Please provide the python code ONLY.
- **Response:**

python Copy

```python
import matplotlib.pyplot as plt
import pandas as pd

cities = ['Inglewood (CA)', 'Glendale (AZ)', 'Las Vegas (NV)', 'Charlotte (NC)',
          'Nashville (TN)', 'Chicago (IL)', 'New Orleans (LA)']

audience_scenarios = {
    '-10% (72,000)': [0.7012, 0.6815, 0.6421, 0.5753, 0.5532, 0.4528, 0.4004],
    'Base (80,000)': [0.7038, 0.6842, 0.6446, 0.5770, 0.5551, 0.4545, 0.4021],
    '+10% (88,000)': [0.7063, 0.6868, 0.6470, 0.5786, 0.5569, 0.4561, 0.4037]
}

audience_numbers = [72000, 80000, 88000]
markers = ['o', 's', '^', 'D', 'v', '<', '>', '*']
colors = ['#1f77b4', '#ff7f0e', '#2ca02c', '#d62728', '#9467bd', '#8c564b', '#e377c2

plt.figure(figsize=(14, 8))

for i, city in enumerate(cities):
    scores = [audience_scenarios[key][i] for key in audience_scenarios]
    plt.plot(audience_numbers, scores, marker=markers[i], label=city,
             linewidth=2.5, markersize=10, color=colors[i],
             markerfacecolor='white', markeredgewidth=2)
    for x, y in zip(audience_numbers, scores):
        plt.annotate(f'{y:.4f}', xy=(x, y), xytext=(5, 5 if i % 2 == 0 else -15),
                     textcoords='offset points', fontsize=9, alpha=0.8)

plt.xlabel('Number of Audiences', fontsize=14, fontweight='bold')
plt.ylabel('Environmental Impact Score', fontsize=14, fontweight='bold')
plt.title('City Score Trends with Audience Number Variations\n(Sensitivity Analysis
          fontsize=16, fontweight='bold', pad=20)
plt.xticks(audience_numbers, ['72,000\n(-10%)', '80,000\n(Base)', '88,000\n(+10%)'],
plt.yticks(fontsize=12)
plt.grid(True, alpha=0.3, linestyle='--')
plt.legend(bbox_to_anchor=(1.05, 1), loc='upper left', frameon=True, fancybox=True,

df = pd.DataFrame(audience_scenarios, index=cities)
df['Absolute Change'] = df['+10% (88,000)'] - df['-10% (72,000)']
df['Relative Change (%)'] = (df['Absolute Change'] / df['Base (80,000)']) * 100
print(df.round(4))
print(f"\nAverage Absolute Change: {df['Absolute Change'].mean():.4f}")
print(f"Maximum Absolute Change: {df['Absolute Change'].max():.4f} (Inglewood)")
print(f"Minimum Absolute Change: {df['Absolute Change'].min():.4f} (Charlotte)")
print(f"Average Relative Change: {df['Relative Change (%)'].mean():.2f}%")
print(f"Maximum Relative Change: {df['Relative Change (%)'].max():.2f}% (New Orleans
print(f"Ranking Stability: Perfect across all scenarios")
```

Figure 10.1: Python Code

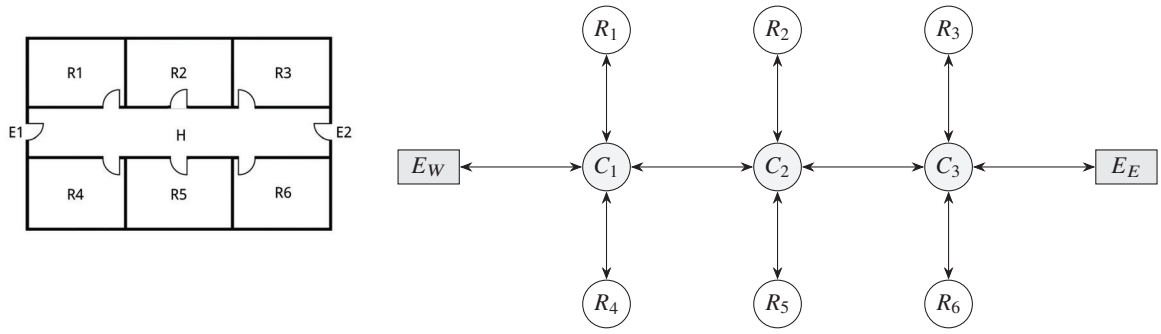


Figure 3: Conversion from floor plan to node graph for the basic office scenario. Left: physical layout with six offices R_1–R_6 along a hallway H and exits E_1 and E_2. Right: corresponding node graph with room nodes Rx, hallway nodes Cx, and exit nodes Ex.

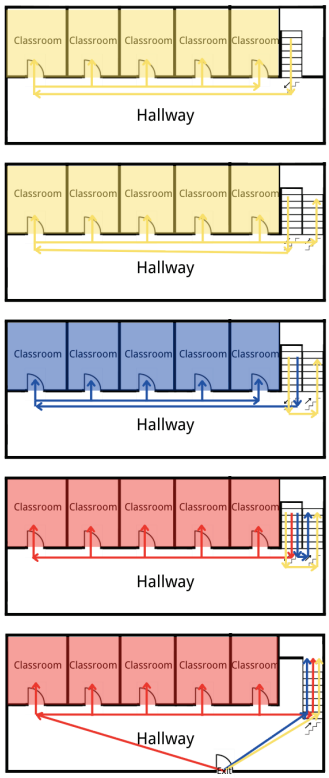
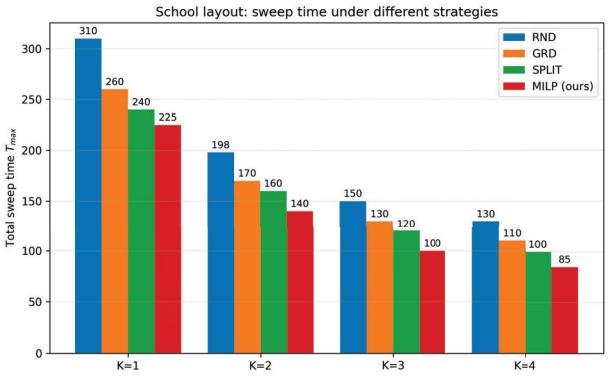


Figure 6: Optimal three-team sweep for the five-floor school layout. Colours indicate the route of each team.

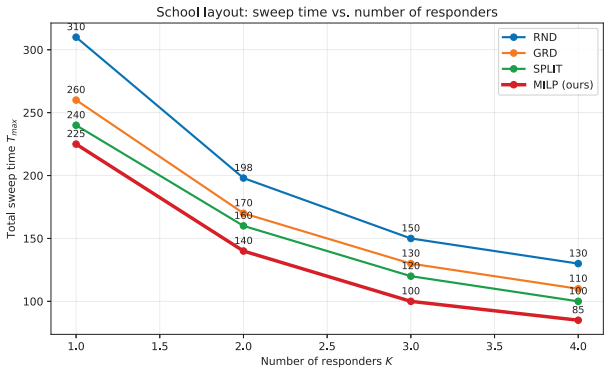
Table 3: Hazard levels and resulting weights for classrooms in the five-floor school layout.

H_i	w_i	Classrooms
0	1.0	$R_1, R_2, R_3, R_6, R_7, R_{11}, R_{16}, R_{17}, R_{21}, R_{22}, R_{23}$
1	2.3	$R_4, R_8, R_{12}, R_{18}, R_{24}$
2	3.7	$R_5, R_9, R_13, R_{19}, R_{25}$
3	5.0	$R_{10}, R_{14}, R_{15}, R_{20}$

Classrooms to nearest the hazard on the middle floors receive the largest weights, so the model prefers to clear them first.



(a) Total sweep time under four strategies.



(b) Sweep time vs. number of responders K.

Figure 9: School layout: comparison of total sweep time under different s trategies a nd different numbers of responders.

Table 20 Comparative analysis

Aspect	Super Bowl (Single-Event)	Olympics (Multi-Venue)
Timeframe of Impact	Concentrated (1–3 days)	Prolonged (2–7 years, from planning to legacy)
Emissions Profile	Dominated by short-term peaks (transport, energy)	Dominated by long-term infrastructure and logistics
Waste Management	Relies on temporary, event-specific solutions	Requires scalable, city-wide systems (Tokyo recycled 99% of Olympic waste through permanent facilities)
Legacy Potential	Limited (venues often return to normal use)	High (venues, transit, and policies can transform cities—London's Olympic Park now hosts 25,000 jobs)

Table 18-20 shows that single-events like the Super Bowl face "acute" sustainability challenges (managing peak demand), while multi-venue events like the Olympics grapple with "chronic" challenges (long-term infrastructure and global logistics). Both require tailored strategies, but multi-venue events have a greater opportunity to leave lasting, positive environmental legacies.

3.2 Stage 1: K-Means Clustering (Assigning Zones)

To assign zones, we represent the building as a graph $G = (N, E)$ where nodes N are room centers and key intersections, and edges E represent traversable connections with associated distances. Node coordinates (x_i, y_i) are used to cluster rooms into K zones. K-mean solves:

$$\min_{\mu_1,\ldots,\mu_K} \sum_{k=1}^{K} \sum_{x_i \in Z_k} \|x_i - \mu_k\|^2 \tag{12}$$

to produce compact, geographically coherent zones. This yields balanced workloads and reduced travel distances.

3.3 Stage 2: Ideal Path Optimization (TSP)

Within each cluster Z_k, we solve a TSP variant: find a path that starts at the responder's exit, visits all rooms in Z_k exactly once, and minimizes travel distance. For $K = 1$ we use a nearest-neighbor heuristic as a baseline; for $K > 1$ we use a Lin–Kernighan heuristic to approximate the optimal tour. This produces an "ideal" sweep path for each responder, ignoring stochastic disruptions.

3.4 Stage 3: Monte Carlo Simulation

To model real-world uncertainty, we simulate K responders following their TSP paths in parallel over N_{MC} runs. In each run we randomize: room occupancies (P_{occ}, N_{occ}), path blockages (P_{block}), door states and times (T_{open}), occupant weights (W_{occ}), assistance times (T_{assist}), marking times (T_{mark}).

For each run, we compute T_k and $T_{\text{clear}} = \max_k T_k$. Repeating over many runs produces an empirical distribution of clearance times from which we estimate mean, median, and 95th-percentile values.

3.5 A* Pathfinding in the Simulation

Responders do not blindly follow precomputed TSP routes. When unexpected conditions arise, they use A* pathfinding.

3.5.1 Trigger 1: Path Blockage

When traversing an edge (u, v), a Bernoulli trial with probability P_{block} determines if the path is blocked (earthquake scenarios). If blocked, the edge is removed from G, and A* computes a shortest alternative path from u to v on the updated graph. The travel time along this new path is added to T_k.

3.5.2 Trigger 2: Evacuee Transport

When an occupant is found, the responder pauses the TSP route and uses A* to find the shortest path to their assigned exit. The cost function uses v_{carry} while carrying and v_{eff} for the return. The computed travel time plus T_{assist} are added to T_k.

4 Methodology

We now present our two-stage machine learning framework for evaluating environmental suitability. The key idea is to let the data reveal natural environmental archetypes (Stage I) and then learn a smooth mapping from environmental features to suitability scores (Stage II).

4.1 Overall Pipeline

The workflow can be summarized as:

$$\text{Raw data} \xrightarrow{\text{clean \& impute}} X_{c,t} \xrightarrow{\text{standardize}} X'_{c,t} \xrightarrow{\text{GMM clustering}} Y_{c,t} \xrightarrow{\text{city aggregation}} Y_c \xrightarrow{\text{ridge regression}} \widehat{Y}_c.$$

4.2 Stage I: Unsupervised Environmental Archetypes

4.2.1 Gaussian Mixture Model

We model the distribution of standardized features $X'_{c,t}$ with a Gaussian Mixture Model (GMM):

$$p(x) = \sum_{k=1}^{K} \pi_k \mathcal{N}(x \mid \mu_k, \Sigma_k), \qquad \sum_k \pi_k = 1,\ \pi_k \geq 0.$$

The parameters $\{\pi_k, \mu_k, \Sigma_k\}$ are estimated via the EM algorithm. In the E-step, we compute responsibilities

$$\gamma_{c,t}(k) = \frac{\pi_k \mathcal{N}(X'_{c,t} \mid \mu_k, \Sigma_k)}{\sum_{j=1}^{K} \pi_j \mathcal{N}(X'_{c,t} \mid \mu_j, \Sigma_j)},$$

followed by standard M-step updates for μ_k, Σ_k, and π_k. The optimal number of components K is selected using BIC/AIC (reported in Results).

4.2.2 Cluster Labeling and City-Level Scores

Each cluster k is characterized by its centroid μ_k in the space of key environmental variables, such as renewable fraction, drought percentage, $PM_{2.5}$, and mean temperature. We interpret the clusters as High-, Moderate-, and Low-Suitability archetypes and assign cluster-level labels $Y_k \in \{0, 0.5, 1\}$ accordingly.

A city–year observation gets a soft label $Y_{c,t}$ based on its cluster. City-level labels are then computed by averaging across time:

$$Y_c = \frac{1}{T_c} \sum_t Y_{c,t},$$

yielding a long-term suitability score for each city based purely on its cluster membership history.

4.3 Stage II: Supervised Learning via Ridge Regression

Given city-level feature vectors and labels (X_c, Y_c), we learn a smooth mapping from environmental characteristics to suitability scores using ridge regression with polynomial feature expansion.

4.3.1 Feature Map

Because many environmental relationships are nonlinear, we apply a second-degree polynomial feature map:

$$\phi(x) = [x_1, \ldots, x_d,\ x_i x_j,\ x_i^2]_{1 \leq i \leq j \leq d},$$

resulting in D transformed features (about 170 in our implementation).

- Compute the equal-weight composite score

$$S_c^{\text{equal}} = \frac{1}{6} \sum_{d \in \mathcal{D}} s_{c,d},$$

and store it as `S_equal_weight`. We also create columns `alpha_travel`, . . . , `alpha_climate` to hold the weights α_d (initialized to $1/6$) and a placeholder column `S_placeholder` for alternative weighted scores.

The resulting table provides a compact numerical representation of each city's environmental performance. In Section 4, we will first interpret the scores directly to rank candidate host cities and then feed them into three simple machine-learning models (logistic regression, random forest, and a small neural network) to estimate $p_c = \mathbb{P}(y_c = 1 \mid \mathbf{s}_c)$ and choose an optimal host city.

4 Model Development

In Sections 1–3 we defined six dimension-wise environmental cost scores for each candidate host city c,

$$\mathbf{s}_c = \left(s_{c,\text{travel}}, s_{c,\text{energy}}, s_{c,\text{waste}}, s_{c,\text{water}}, s_{c,\text{mobility}}, s_{c,\text{climate}}\right) \in [0,1]^6,$$

and an equal-weight composite index

$$S_c^{\text{equal}} = \frac{1}{6} \sum_{d \in \mathcal{D}} s_{c,d}, \qquad \mathcal{D} = \{\text{travel}, \text{energy}, \text{waste}, \text{water}, \text{mobility}, \text{climate}\},$$

where smaller values indicate better environmental performance.

To support the 2029 Super Bowl decision, we now build three simple but complementary machine-learning models:

- a logistic-regression classifier;
- a random-forest (RF) classifier;
- a small feed-forward (back-propagation) neural network (NN).

Each model takes the six-dimensional cost vector $\mathbf{s}_c$ as input and outputs a probability that city c is an environmentally "good" host. We then average the three probabilities to obtain a more robust ensemble prediction.

4.1 City sets and labels

The NFL has used 19 different host cities across the history of the Super Bowl. Since the league began formally integrating sustainability metrics and environmental assessments into the event planning process, only a subset of those cities have hosted:

- **Recent hosts with formal sustainability reporting:** Inglewood (California), Glendale (Arizona), Las Vegas (Nevada), New Orleans (Louisiana).

- **Confirmed future hosts under the same sustainability framework:** Santa Clara (California) and Atlanta (Georgia).
- **Never-host cities with NFL teams:** we focus on three representative candidates—Seattle (Washington), Denver (Colorado), and Nashville (Tennessee)—which are often discussed in the media as plausible future Super Bowl hosts because of their stadium quality and fan base.

Our core NFL city dataset thus contains nine U.S. cities:

$$C_{\text{NFL}} = \{\text{Inglewood, Glendale, LasVegas, NewOrleans, SantaClara, Atlanta, Seattle, Denver, Nashville}\}.$$

In addition, as explained in Section 3, we also include a small number of other sport mega-events (e.g., selected FIFA World Cup matches and Formula 1 races) to broaden the range of environmental indicators and avoid overfitting to only nine observations [2, 12, 1].

To train binary classifiers, we define a label $y_c \in \{0, 1\}$ for each city/configuration in the calibration dataset:

$$y_c = \begin{cases} 1, & \text{if } S_c^{\text{equal}} \le \tau, \\ 0, & \text{if } S_c^{\text{equal}} > \tau, \end{cases} \qquad \tau = Q_{0.5}(\{S_c^{\text{equal}} : c \in C_{\text{train}}\}), \tag{6}$$

where τ is the median equal-weight score across all cities in the training set C_{train} and $Q_{0.5}$ denotes the median operator. Cities in the better half of the distribution (lower S_c^{equal}) are labelled $y_c = 1$ (good host), and the rest $y_c = 0$ (average host).[1]

4.2 Logistic-regression model

The logistic-regression model assumes that, conditional on the environmental scores, the log-odds of being a good host are linear in $\mathbf{s}_c$:

$$\log \frac{\mathbb{P}(y_c = 1 \mid \mathbf{s}_c)}{\mathbb{P}(y_c = 0 \mid \mathbf{s}_c)} = \beta_0 + \boldsymbol{\beta}^\top \mathbf{s}_c, \tag{7}$$

where β_0 is an intercept and $\boldsymbol{\beta} \in \mathbb{R}^6$ is a coefficient vector. Equivalently,

$$p_c^{(\text{LR})} := \mathbb{P}(y_c = 1 \mid \mathbf{s}_c) = \sigma\left(\beta_0 + \boldsymbol{\beta}^\top \mathbf{s}_c\right), \qquad \sigma(t) = \frac{1}{1 + e^{-t}}.$$

The parameters $(\beta_0, \boldsymbol{\beta})$ are obtained by minimizing the regularized negative log-likelihood:

$$\mathcal{L}(\beta_0, \boldsymbol{\beta}) = -\sum_{c \in C_{\text{train}}} \left[y_c \log p_c^{(\text{LR})} + (1 - y_c) \log(1 - p_c^{(\text{LR})}) \right] + \lambda \|\boldsymbol{\beta}\|_2^2,$$

where $\lambda > 0$ is a small ridge penalty to avoid overfitting. In our implementation we fit this model using `fitclinear` with `Learner='logistic'` in MATLAB.

The signs and magnitudes of $\boldsymbol{\beta}$ provide an interpretable summary of how each dimension influences the likelihood of being a good host (for example, negative coefficients on waste and water scores mean that cities with lower waste and water impacts are more likely to be classified as good hosts).

[1] This is a simple relative definition; the NFL could choose a stricter threshold (e.g., the best one-third) for more ambitious sustainability policies.

1. **February Climate Load** Defined as

$$C_i = |72 - T_i^{\text{Feb}}|$$

where T_i^{Feb} is the long–term February temperature average (NOAA climate normals). The target of 72°F is chosen because it approximates indoor thermal comfort.

2. **Renewable Energy Share** State–level renewable electricity fractions are computed from the EIA 2024 "Annual Electric Power Industry" database. These values penalize fossil–heavy states (e.g., TX, FL) and reward states with strong renewable portfolios (e.g., MN, MI, CA).

3. **Public Transit Share** Percent of commuters using public transit (ACS 5–year survey). This lowers the travel footprint of stadiums in cities like Philadelphia or New York, but raises it for car–dependent regions.

5.9 Final Ranking via Borda Count

To synthesize rankings across the different weighting strategies, we apply a Borda count on the four ranking lists: raw EF, PCA-EF, geo-adjusted EF, and geo-adjusted PCA-EF. A stadium with rank r among n candidates receives a Borda score of $(n - r + 1)$. Summing these across all lists gives

$$B_i = \sum_{m=1}^{4} (n - r_{im} + 1).$$

The stadium with the highest Borda score is considered the most environmentally sustainable host.

6 Model Application

6.1 Model Framework

Our environmental comparison uses a two–layer model: (1) an ecological footprint score EFi for each stadium, and (2) a geographic adjustment parameter Gi, which accounts for regional context.

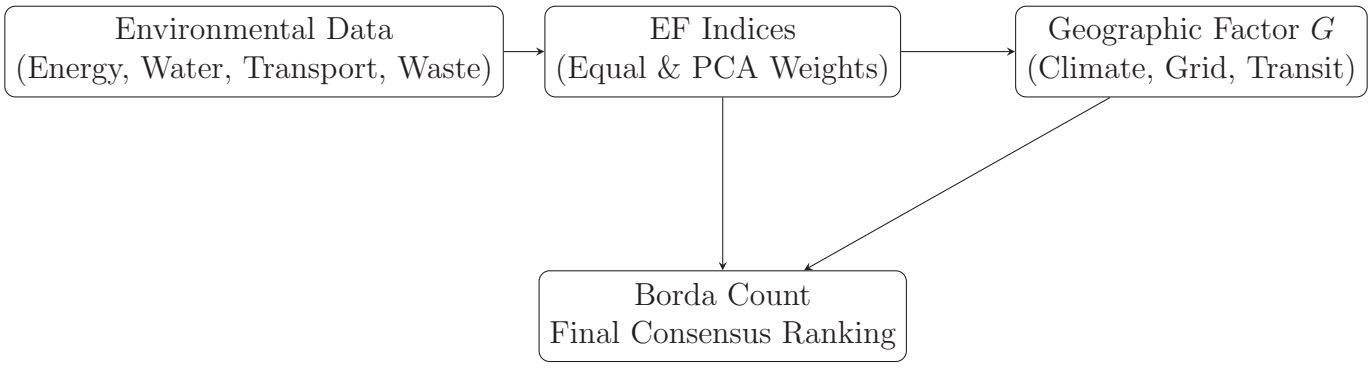


Figure 1: Model pipeline: EF indices and geographic factor combined through a Borda consensus.

6.2 Model Application to All Sixteen Stadiums

We categorize the sixteen stadiums as 13 historical Super Bowl Hosts (excluding Inglewood, Glendale, Las Vegas, New Orleans, Santa Clara, and Atlanta) and 3 other NFL stadiums chosen randomly: U.S. Bank Stadium, Sun Devil Stadium, Ford Field, MetLife Stadium, Hard Rock Stadium, Raymond James Stadium, NRG Stadium, LA Memorial Coliseum, Rose Bowl, AT&T Stadium, Rice Stadium, EverBank Stadium, Lucas Oil Stadium, Bank of America Stadium, Lincoln Financial Field, and Nissan Stadium.

To evaluate the environmental performance of potential Super Bowl venues, we applied our Ecological Footprint model to a set of sixteen stadiums: thirteen that have previously hosted the Super Bowl, and three that have never hosted. Although individual categories (energy, water, transportation, and solid waste) differ in scale, our model compresses them into a single, comparable metric:

$$EF_i \quad \text{(Ecological Footprint of stadium } i\text{)}$$

and modifies this value by a location–specific geographic parameter:

$$G_i \quad \text{(Geographic adjustment for climate, renewables, and transit)}$$

Producing a final, context-aware measure:

$$EF_i^* = EF_i \cdot G_i.$$

The result is a consistent index that captures both the raw environmental burden of operating a Super Bowl event at stadium i and the fairness adjustments caused by unavoidable regional differences.

Table 2: Ecological footprint components (CO_2 tons) for sixteen NFL stadiums.

Stadium	Energy	Water	Transportation	Solid Waste
Lucas Oil Stadium	7.60	1.68	17876	48.3
U.S. Bank Stadium	3.76	1.67	18863	48.0
Ford Field	4.66	1.62	18474	47.0
Nissan Stadium	4.35	1.90	18868	49.6
Rice Stadium	5.39	2.16	21153	36.4
Lincoln Financial Field	3.58	1.93	23222	42.9
Bank of America Stadium	3.11	2.06	23374	53.2
Sun Devil Stadium	3.10	1.99	24783	33.0
EverBank Stadium	3.75	2.49	24990	42.6
NRG Stadium	4.94	2.35	25929	51.5
AT&T Stadium	4.49	2.60	26121	56.3
Raymond James Stadium	3.75	2.39	25931	40.6
MetLife Stadium	2.75	2.06	28261	57.8
Hard Rock Stadium	3.75	2.37	29091	40.3
Los Angeles Memorial Coliseum	1.98	2.76	41765	47.8
Rose Bowl	2.20	3.16	48038	55.3

Table 3: Ecological footprint indices and geographic factor G for sixteen NFL stadiums.

Stadium	$\mathrm{EF}_{\mathrm{raw}}$	$\mathrm{EF}_{\mathrm{entropy}}$	G	$\mathrm{EF}_{\mathrm{raw}} \cdot G$	$\mathrm{EF}_{\mathrm{entropy}} \cdot G$
Lucas Oil Stadium	17868.67	6117.29	0.9979	17831.82	6104.67
U.S. Bank Stadium	18855.49	6453.03	0.9683	18258.48	6248.71
Ford Field	18466.54	6320.33	0.9967	18406.45	6299.77
Nissan Stadium	18860.67	6455.23	0.9916	18702.18	6400.99
Rice Stadium	21145.32	7235.63	0.9661	20427.56	6990.02
Lincoln Financial Field	23212.91	7942.34	0.9589	22258.70	7615.86
Bank of America Stadium	23364.98	7995.10	0.9869	23058.81	7890.34
Sun Devil Stadium	24773.75	8474.91	0.9789	24250.76	8296.00
EverBank Stadium	24979.94	8546.63	0.9922	24785.01	8479.94
NRG Stadium	25919.19	8869.06	0.9661	25039.40	8568.01
AT&T Stadium	26110.52	8934.76	0.9678	25270.94	8647.47
Raymond James Stadium	25921.23	8868.27	0.9863	25564.91	8746.36
MetLife Stadium	28249.88	9665.55	0.9189	25957.72	8881.29
Hard Rock Stadium	29079.79	9948.16	0.9777	28432.29	9726.65
Los Angeles Memorial Coliseum	41748.72	14279.72	0.9228	38527.80	13178.04
Rose Bowl	48020.02	16424.75	0.9228	44315.26	15157.58

6.3 Final Integrated Ranking

Table 3 summarizes the three core quantities entering our environmental evaluation: the raw ecological footprint EF, the geographic factor G, and the PCA-based ecological footprint score EF_{PCA}. Each measure highlights a different dimension of stadium performance. The raw EF captures absolute environmental burden, G reflects unavoidable regional conditions (climate, energy mix, and transit), and EF_{PCA} balances all four impact categories by removing scale bias.

The final consensus score for stadium i is

$$B_i = \sum_k B_i^{(k)},$$

summed here across the four ranking systems:

$$\{EF,\ EF_{\mathrm{PCA}},\ EF \cdot G,\ EF_{\mathrm{PCA}} \cdot G\}.$$

A higher Borda score indicates that a stadium performs consistently well across all modeling perspectives, not just a single weighting structure.

Table 4: Final Environmental Ranking of Sixteen Stadiums Using Borda Consensus

Stadium	**Borda Score**	**Final Rank**
U.S. Bank Stadium (Minneapolis)	46	1
Sun Devil Stadium (Phoenix)	38	2
Ford Field (Detroit)	37	3
Lincoln Financial Field (Philadelphia)	37	3
Bank of America Stadium (Charlotte)	36	4
Nissan Stadium (Nashville)	35	5
MetLife Stadium (NY/NJ)	29	6
EverBank Stadium (Jacksonville)	25	7
Raymond James Stadium (Tampa Bay)	23	8
Rice Stadium (Houston)	23	8
Lucas Oil Stadium (Indianapolis)	19	9
Hard Rock Stadium (Miami)	16	10
LA Memorial Coliseum (Los Angeles)	14	11
AT&T Stadium (Dallas/Arlington)	13	12
NRG Stadium (Houston)	13	12
Rose Bowl (Pasadena)	4	13

6.4 Interpretation of Final Borda Rankings

The integration of EF, PCA, and geographic adjustments through the Borda framework allows us to recognize stadiums that are environmentally competitive across multiple dimensions, rather than excelling in only one. This consensus perspective is essential for comparing future Super Bowl sites in a fair and transparent manner.

All sixteen were processed through the same sequence of indices: Raw EF, PCA-weighted EF, Geo-Adjusted EF, and finally a Borda consensus ranking. This uniformity allows us to compare all cities on equal footing. The Borda results produce a stable, intuitive ordering of environmental performance (Table 4). U.S. Bank Stadium (Minneapolis) is the clear top performer, ranking among the top two positions in every modeling scheme due to its efficient indoor design, low climate load, and one of the highest renewable energy fractions in the dataset.

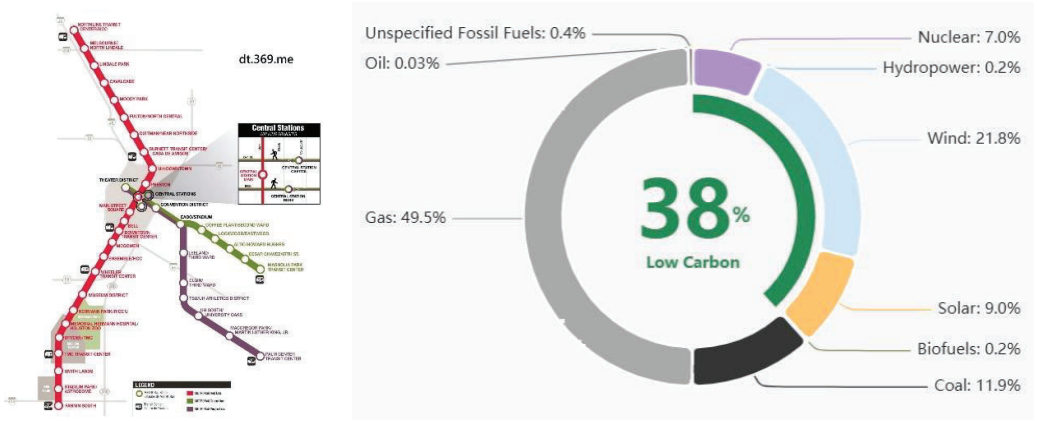
Among stadiums that have *never* hosted a Super Bowl, Lincoln Financial Field (Philadelphia) ranks highest overall. Its strengths include strong public transit usage, a moderately renewable electricity mix, and a highly balanced impact profile that performs well under both PCA and geographic adjustments. It consistently ranks within the upper tier of all indices, making it the most environmentally competitive unhosted stadium.

Across all weighting strategies and across both raw and adjusted indices, U.S. Bank Stadium in Minneapolis repeatedly emerges as the most environmentally favorable site. Three features drive its exceptionally strong performance:

1. **Low EF due to efficient indoor design.** The fully enclosed stadium sharply reduces climate load in February. Heating requirements are predictable and far smaller than the cooling demands found in southern venues.

2. **Favorable geographic modifiers.** Minnesota's grid has one of the highest renewable energy shares in the NFL stadium set, and Minneapolis maintains a robust public transit system. These

The first three recommendations correspond to Sub-model 1 (carbon emissions), the fourth and fifth to Sub-model 2 (ecosystem impacts), and the sixth to Sub-model 3 (disturbance pollution). Collectively, these strategies enhance a city's sustainability across multiple dimensions.



(a) (b)

Figure 14. Houston's rail line planning and energy structure.

Taking Houston, ranked 10th, as an example, its current electricity energy mix primarily consists of natural gas, coal, and nuclear power (accounting for over 60% of total generation capacity), as shown in Fig 14. However, Houston holds notable advantages in wind and solar energy, but substantial improvements are needed to enhance its sustainability. We recommend significantly increasing financial investment over the next three years to raise the share of renewable energy to approximately 50%, which would reduce electricity carbon intensity from 0.45 $kgCO_2/kWh$ to around 0.2 $kgCO_2/kWh$.

Transportation currently relies heavily on gasoline-powered vehicles (over 85% of trips), with public transit comprising only 3%. To address this, we recommend expanding public transit infrastructure over the next 3–5 years, including electric buses, subways, and rail systems, alongside promoting the transition to electric vehicles. These measures would substantially reduce transportation carbon emissions and pollutant discharges, improving both carbon emissions and air quality scores.

Houston's current recycling rate is 30%. We advise expanding recycling infrastructure, increasing household recycling capacity, and improving contamination management to raise the rate to approximately 60%. Additionally, strengthening water resource and ecosystem protection, accelerating green ecological restoration, and promoting low-light-pollution lighting and soundproof building materials will reduce noise and light pollution.

Our simulations indicate that implementing these measures could increase Houston's overall environmental score from 0.2938 to 0.4250 within 3–5 years, potentially improving its ranking to 6th place among candidate cities.

6.4.Sensitivity Analysis: Key Factors of Single vs Multi-Sport

In single-sport events, carbon emissions from energy consumption and waste disposal are the most critical factors. For example, in the Super Bowl host city evaluation, the Los Angeles–Inglewood–Pasadena bid outperformed San Diego on carbon emissions. Although San Diego demonstrated superior air quality and light/noise pollution control, the higher weighting of carbon emissions allowed Los Angeles–Inglewood–Pasadena to achieve a higher overall score, as shown in Fig.15(a).

In contrast, for multi-sport events, venue construction impacts become paramount. In Olympic host city selection, Nusantara exhibited lower carbon emissions and ecological footprint than Seoul.

years. Therefore, we have collected the historical data regarding the local characteristic indicators of the six cities that have hosted the Super Bowl, which are listed in **Table 4**.

Table 4 Cities that have hosted the Super Bowl

Edition	Year	City
L	2016	Atlanta
LIII	2019	Santa Clara
LVI	2022	Inglewood
LVII	2023	Glendale
LVIII	2024	Las Vegas
LIX	2025	New Orleans

We have collected cities that currently have NFL teams but have never hosted the Super Bowl. There are 9 such cities in total, namely: Green Bay (home to the Green Bay Packers), Buffalo (home to the Buffalo Bills), Nashville (home to the Tennessee Titans), Cincinnati (home to the Cincinnati Bengals), Pittsburgh (home to the Pittsburgh Steelers), Cleveland (home to the Cleveland Browns), Kansas City (home to the Kansas City Chiefs), Baltimore (home to the Baltimore Ravens), and Charlotte (home to the Carolina Panthers). For the above - mentioned cities, we still need to select three potential host cities. Therefore, we have gathered the historical local characteristic data of these three potential cities, which are presented in **Table 5**.

Table 5 Potential Cities for Hosting the Super Bowl

Edition	Year	Potential host city
LXIII	2029	Santa Clara
LXIII	2029	Inglewood
LXIII	2029	Atlanta

Based on Assumptions 1 and 2, we can conclude that the historical climatic conditions of each candidate city during the Super Bowl period are relatively stable, and the urban population flow patterns remain consistent. Therefore, we only need to use the average values of temperature, precipitation, population density, and average spectator travel distance from 2015 to 2024 as proxies. For other indicators such as public transportation coverage, we will predict data for the next five years based on historical data from 2015 to 2024.

Next, we will adopt the ARIMA model to predict the future local characteristic data of candidate cities. For any city c and its local characteristic indicators m, we assume that their time series $y_{cm}(t)$ satisfy a specific model $ARIMA(p,d,q)$, which can be generally expressed as:

$$\Phi(B)\nabla^{d_{c,m}} y_{c,m}(t) = \mu_{c,m} + \theta(B)\varepsilon_{c,m}(t) \tag{11}$$

In the formula $\nabla^d y_t = (1-B)^d y_t$, the difference operator **B** is as follows: where is the lag operator, $\Phi(B)$ is the autoregressive polynomial, $\theta(B)$ is the moving average polynomial, $\varepsilon_{c,m}(t)$ is the white

noise error, and $d_{c,m}$ is the difference order, which is used to ensure that the differentiated sequence is stable.

For each sequence $\{y_{c,m}(t)\}_{t=1}^{T}$, under the given conditions $(p_{c,m}, d_{c,m}, q_{c,m})$, the parameter estimates are obtained through maximum likelihood estimation. Then, a least squares fit $t = 1,\ldots,T$ is used for fitting, and predictions are made $t = T+1,\ldots,T+H$. We need to predict up to the year 2029. Therefore H=5, there is:

$$\hat{y}_{c,m}(T+H) = a_{c,m} + b_{c,m}(T+H) \tag{12}$$

For the cities that have previously hosted events, as shown in Table 4. The predictions for the share of renewable energy, power carbon intensity, public transportation coverage rate, and green transportation share were made using ARIMA (1,1,0), ARIMA (0,1,0), ARIMA (0,2,0), and ARIMA (0,2,1) respectively. The results are as follows:

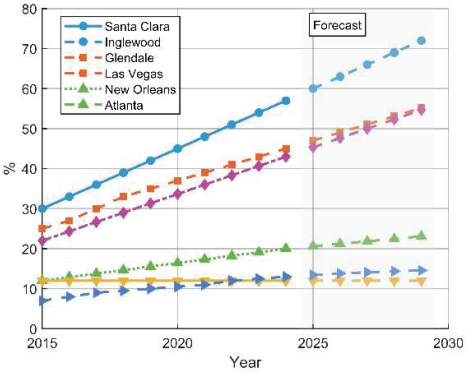


Figure 8 Prediction of Renewable Energy Share

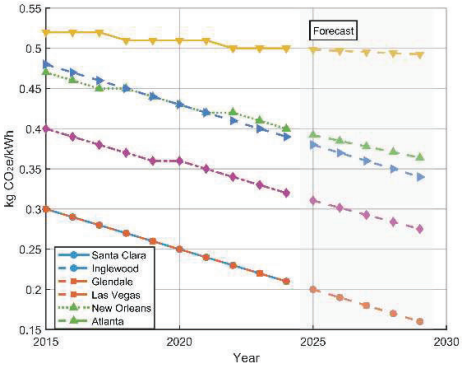


Figure 9 Prediction of Power Carbon Intensity

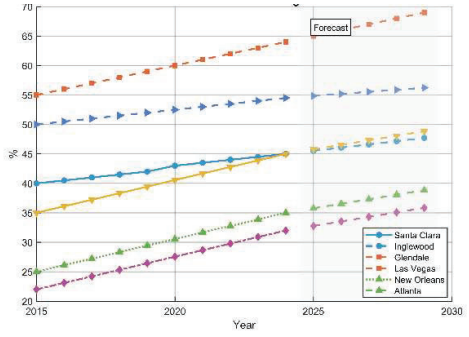


Figure 10 Prediction of Public Transportation Coverage Rate

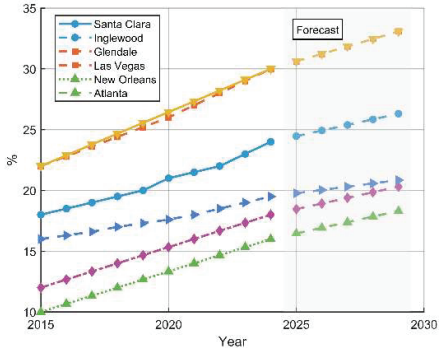


Figure 11 Prediction of Green Transportation Share

Similarly, for cities that have never held such events, namely those in Table 5, we also used the ARIMA model for prediction. The results are as follows: