

Article

Not peer-reviewed version

Multi-Objective Optimization of Itinerary Planning Based on Weather Changes and User Preferences

[Shuang Li](#) and [Ka-Cheng Choi](#)*

Posted Date: 20 April 2026

doi: 10.20944/preprints202604.1364.v1

Keywords: itinerary planning; weather; multi-objective optimization; genetic algorithm



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Multi-Objective Optimization of Itinerary Planning Based on Weather Changes and User Preferences

Shuang Li and Ka-Cheng Choi 

Faculty of Applied Sciences, Macao Polytechnic University, Macao 999078, China

* Correspondence: rebeccachoi@mpu.edu.mo

Abstract

Existing travel planning systems lack user participation in itinerary scoring and apply coarse, binary weather treatment that risks excluding high-quality outdoor attractions under mild precipitation. This paper presents a multi-objective genetic algorithm (GA)-based itinerary planning system that addresses both limitations. The system incorporates a weather-adaptive POI scoring framework mapping nine weather conditions to three strategies, applying intensity-proportional rain penalties and a geographic flexibility bonus scaled by local indoor alternative density. User preferences are encoded via three integer sliders whose normalised values directly set the GA fitness weights for POI quality, traveling efficiency, and preference satisfaction. The system is evaluated on 144 attractions in Macao. Results show that outdoor POI representation decreases proportionally with precipitation intensity across all nine weather conditions and is substantially suppressed under official extreme weather alerts, while itinerary quality is preserved through the flexibility bonus. Slider adjustment experiments confirm that amplifying each weight produces statistically consistent, direction-correct improvements in its target sub-objective without degrading the others. These findings validate the functional independence of the three-objective fitness formulation and demonstrate that graduated weather treatment and direct user weight control together yield a more responsive and robust itinerary planning system.

Keywords: itinerary planning; weather; multi-objective optimization; genetic algorithm

1. Introduction

The tourism industry has long served as a significant driver of global economic development [1], and its post-pandemic recovery bolstered by supportive policy frameworks [2] which has accelerated demand for more sophisticated and personalized travel planning tools [3]. At the core of a satisfying travel experience lies effective itinerary planning: the task of selecting and sequencing a set of points of interest (POIs) that simultaneously maximize attraction quality, minimize travel overhead, respect operational constraints, and align with individual traveler preferences. As the combinatorial complexity of this problem grows with city scale and the diversity of user-defined requirements, automated planning systems have become an important research priority.

The itinerary planning (IP) problem has been formally demonstrated to be nondeterministic polynomial-time hard (NP-hard) [4], rendering exact optimization intractable for practical-scale instances. Genetic algorithms (GAs), as population-based metaheuristics, are well-suited to multi-objective itinerary optimization owing to their global search capability and inherent support for constraint handling [5–8]. Existing GA-based itinerary systems, however, share two structural limitations. First, their scoring and evaluation criteria are predefined by system developers, affording users no participation in establishing what constitutes a high-quality itinerary. Yochum et al. [4] propose an adaptive crossover and mutation strategy for GA-based IP, yet their objective weights remain statically configured rather than user-driven. Huang et al. [8] extend GA planning to the multiple-itinerary planning (MIP) setting through a niching genetic algorithm (NGA), offering multi-day and multi-solution generation, but similarly without user-facing evaluation customization. Beyond GAs,

data-driven approaches such as the Transformer- and graph convolutional network (GCN)-based deep learning-based itinerary recommendation (DLIR) model of Halder et al. [9] capture complex spatiotemporal user preference dynamics, yet depend on large-scale historical trajectory corpora and do not support real-time user-directed evaluation adjustment. Second, the integration of weather into itinerary scoring remains coarse. Zhang et al. [10] account for weather in route generation through an encoder-decoder model, but apply a broadly categorical weather treatment that risks undue exclusion of high-quality outdoor POIs during mild precipitation and provides no mechanism for user-declared weather sensitivity.

To address these limitations, this paper proposes a multi-objective GA-based itinerary planning system that integrates user-participatory preference weights and a graduated weather-adaptive scoring mechanism. A weather-adaptive POI scoring framework, built around nine weather conditions mapped to three distinct scoring strategies, calibrates outdoor attraction competitiveness according to precipitation intensity and the density of nearby indoor alternatives. User preferences are captured through an interactive weight interface that directly governs the GA fitness function, enabling diverse traveler profiles to produce correspondingly distinct optimized itineraries. The system is developed and evaluated on a real-world POI dataset covering 144 attractions in Macao, a high-profile tourist destination [11] characterized by high annual rainfall [12] and well-documented tourist sensitivity to precipitation [13,14]. Weather data is retrieved via the OpenWeatherMap application programming interface (API) both during itinerary generation and as users progress through their trips, enabling real-time adaptive re-routing when conditions deteriorate.

The principal contributions of this paper are as follows:

1. A weather-adaptive POI scoring framework that maps nine distinct weather conditions (clear, drizzle, heavy drizzle, light shower rain, light rain, shower rain, moderate rain, extreme rain, and extreme weather) to three qualitatively differentiated scoring strategies, all modulated by the local density of indoor alternatives. This resolves the binary exclusion limitation of existing weather-aware systems while preserving qualitatively superior outdoor attractions within generated itineraries under mild adverse conditions.
2. A user-participatory evaluation architecture in which three integer sliders produce normalized weight coefficients $\alpha, \beta, \gamma \in [0.1, 1.0]$ (with $\alpha + \beta + \gamma = 1$) that directly govern the GA fitness function across three sub-objectives: POI quality, traveling efficiency, and preference satisfaction. This mechanism minimizes the information burden on users while generating itineraries that transparently reflect individual traveler priorities.

2. Related Work

Research on automated itinerary planning has advanced along three principal directions: GA and metaheuristic-based optimization, learning-based and advanced optimization methods, and context-aware planning incorporating real-world environmental factors and emerging large language model (LLM)-based tools.

GAs and related metaheuristics have established themselves as a dominant approach to IP given their ability to navigate large combinatorial search spaces under multi-objective constraints without requiring problem-specific mathematical tractability. Yochum et al. [4] formulate IP as a multi-objective optimization (MOO) problem and propose the adaptive genetic algorithm with crossover and mutation (AGAM), in which crossover and mutation probabilities are dynamically scaled in response to population fitness statistics; POI popularity is quantified through geotagged photo frequency, and the model simultaneously optimizes mandatory POI coverage, visit count, visit duration, popularity, entrance cost, and rating across six real-world city datasets. Şehab and Turan [5] design a custom crossover operator (COX) and a dynamic mutation strategy within a constrained GA framework that handles budget limits, time windows, and category preferences through a parameterized fitness function, achieving competitive performance on multi-constraint itinerary instances. Anshari and Baizal [6] extend GA-based planning to multi-day scenarios by coupling GA route generation with

multi-attribute utility theory (MAUT), modeling the problem as a traveling salesman problem (TSP) variant and evaluating routes across attraction ratings, travel costs, and sightseeing time with user-defined weight distributions. Cena et al. [7] decouple the planning task into POI recommendation and temporal scheduling phases, demonstrating that any ranked POI list can supply a GA-based scheduler that accommodates opening hours, travel time, and user preferences for trip intensity and crowd avoidance. Huang et al. [8] address the MIP problem through an NGA, applying biological niche theory to limit intra-population fitness competition and thereby enable simultaneous convergence toward multiple high-quality, diverse D-day itineraries, overcoming the single-solution limitation of most prior GA approaches. Tenemaza et al. [15] propose a two-stage pipeline combining k -means geographic clustering with a parameterized GA, augmented by real-time web data, to support dynamic, scalable multi-day recommendation. Despite their methodological diversity, these works share a common limitation: scoring criteria and objective weights are universally developer-defined, providing no mechanism for users to directly participate in formulating evaluation logic. Choi et al. [16] develop a multi-objective GA for multi-day tourism route planning in Macao, explicitly incorporating real-world temporal constraints including attraction operation time windows and recommended stay durations into the chromosome encoding and fitness evaluation; a local fine-tuning operator is further applied post-evolution to improve schedule feasibility and user experience. While these work shares both the GA framework, their fitness function parameters are developer-fixed, attraction scores are not modulated by weather conditions, and no mechanism is provided for users to adjust objective weights at runtime. Our system addresses this gap by directly exposing fitness sub-objective weights to user configuration.

Parallel developments in deep learning and advanced optimization have introduced more expressive preference models and specialized algorithmic machinery to the IP domain. The DLIR model proposed by Halder et al. [9] integrates a multi-head attention-based Transformer to capture users' periodic and recent preference dynamics with a co-visiting patterns-based GCN to model non-linear spatial dependencies; itineraries are assembled via a greedy policy minimizing queuing time within budget, outperforming Monte Carlo Tree Search (MCTS)-based baselines by 7.79%–26.28% in F1-score across multiple datasets. Chen et al. [17] propose a triple-agent deep reinforcement learning (RL) framework in which three specialized agents independently optimize POI exploration direction, distance constraints, and selection strategy, with inter-agent dependencies modeled via Bayesian inference, yielding more exploratory and diverse itineraries than traditional multi-objective methods across multiple urban datasets. Noguchi et al. [18] apply hybrid subQUBO annealing, a quantum annealing (QA) technique, to multi-day intermodal trip planning, decomposing large-scale quadratic unconstrained binary optimization (QUBO) problems into hardware-compatible subproblems with a correction mechanism for constraint satisfaction, demonstrating competitive results on a Tokyo dataset. Lim et al. [19] enrich the orienteering problem (OP) formulation with dynamic queuing time modeling by inferring time-dependent visitor arrival distributions from geo-tagged photos, compensating for the static cost assumptions of prior orienteering-based approaches. Choachaicharoenkul et al. [20] extend the orienteering problem with time windows (OPTW) to the multi-objective orienteering problem with time windows, restaurant selection, and compulsory POIs (MOPTW-RSCP), generating Pareto-optimal itinerary sets via branch-and-cut and greedy algorithms and ranking solutions through an angle dominance method. Pilato et al. [21] model social dimension trade-offs in planning, balancing congestion risk, pandemic safety, and social engagement, through Doc2Vec and FastText interest representations, demonstrating modular deployability on the Yelp dataset. While technically powerful, these approaches generally depend on large-scale historical trajectory corpora or specialized hardware, limiting practical deployment in data-scarce settings and constraining real-time user customization. Our system complements these methods with a computationally lightweight, interpretable GA framework that requires no historical trajectory data and remains directly configurable by users at runtime.

A complementary strand of research examines how real-world contextual factors, particularly weather and sustainability, and emerging LLM-based planning tools reshape the itinerary planning problem. Zhang et al. [9] incorporate weather as a travel context feature in an encoder-decoder route generation model, recognizing that adverse conditions affect outdoor activity suitability; however, weather is treated as a broad categorical input rather than a fine-grained, intensity-calibrated modifier, and outdoor POIs risk wholesale exclusion even during mild precipitation with no mechanism for user-declared weather sensitivity. Barua and Kaiser [22] propose an AI-driven route optimization system built on a microservices architecture that integrates machine learning, GA heuristics, and explicit sustainability objectives, demonstrating balanced optimization of cost, travel time, user preferences, and carbon footprint. Regarding the growing influence of LLM-based tools in travel planning, Volchek and Ivanov [23] evaluate ChatGPT as a travel planner and document significant factual inaccuracies, temporal constraint violations, and specificity deficits in generated itineraries, underscoring the limitations of language model-based planning for precision-critical applications. Seyfi et al. [24] investigate tourist adoption barriers for generative artificial intelligence (GAI)-based travel planning through innovation resistance theory (IRT) and personality theory, identifying perceived complexity, accuracy skepticism, and trust deficits as persistent obstacles. These findings collectively highlight that fine-grained weather integration and user trust remain unresolved challenges in intelligent itinerary planning. Our approach directly targets the coarse weather-handling limitation of Zhang et al. [9] by implementing a precipitation sensitivity model that distinguishes nine weather intensity levels, including drizzle, light shower rain, light rain, moderate rain, and extreme weather, among others, while maintaining a transparent, user-configurable GA architecture that avoids the accuracy and trust concerns associated with fully AI-generated itineraries.

3. Overview Framework

This section presents the complete methodology of the proposed itinerary planning system, illustrated in Figure 1. This framework composes of 5 parts: data collection, user preference collection, weather data retrieval, POIs score calculation and genetic algorithm for itinerary generation.

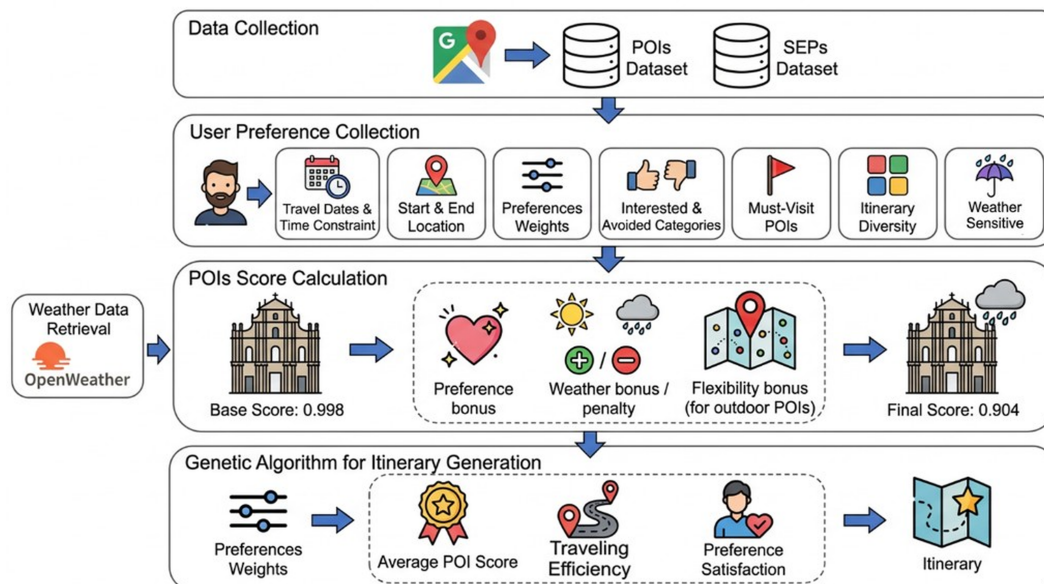


Figure 1. Overview of our proposed framework.

3.1. Data Collection

POI Dataset. The POI dataset is a pre-compiled database covering 144 attractions across the Macao Special Administrative Region. Attraction name, category, geographic coordinates (latitude and longitude), sub-region, weekly closure dates, daily operating hours, indoor/outdoor attribute, and total review count were collected from Google Maps. Recommended visit durations are simulated estimates

used for experimental purposes, as standardized empirical duration data are not publicly available at the per-attraction level for this region. Pre-compiling the dataset eliminates repeated API queries during itinerary generation, ensuring consistent data quality and deterministic reproducibility across experimental runs. Each POI is pre-assigned a `location_code` based on its geographic sub-region to enable efficient proximity-based filtering during POI scoring and real-time adaptation. Attractions are additionally stratified into ten quality tiers according to their review volume relative to the full dataset distribution. The collected POI attributes collectively serve as the primary data source for both the POI scoring pipeline and the genetic algorithm's candidate pool construction.

SEP Dataset. The start/end point (SEP) dataset contains 336 pre-configured origin and destination locations spanning airports, ferry terminals, border crossings, metro stations, hotels, and major transit hubs. Each record includes location name, type, sub-region, and geographic coordinates. Pre-populating this dataset eliminates the need for free-form address entry by users, reduces geocoding latency during itinerary generation, and ensures that all selected departure and arrival points are routable within Macao's transportation network.

3.2. User Preference Collection

As shown in Figure 1, the user interface collects seven categories of structured user input that collectively parameterize the behavior of all downstream modules.

Travel Dates & Time Constraint define the travel period and the available daily time window. These parameters trigger the system to retrieve per-day weather forecasts from the OpenWeatherMap API and establish the total time budget within which the combined visit durations and inter-POI travel times of all selected attractions must fit.

Start & End Location specify the geographic coordinates of the daily departure and arrival points, selected from the SEP dataset. These coordinates serve as fixed anchors for route construction, bounding the POI sequence from both ends within each travel day.

Preference Weights are three user-adjustable integer sliders for POI quality, traveling efficiency, and preference satisfaction, each taking an integer value in $[1, 10]$ with the three values constrained to sum to 10. Each slider value is divided by 10 to yield the corresponding normalized weight coefficient used directly in the GA fitness function: $\alpha = s_\alpha/10$ (POI quality), $\beta = s_\beta/10$ (traveling efficiency), and $\gamma = s_\gamma/10$ (preference satisfaction), where $s_\alpha, s_\beta, s_\gamma$ denote the integer slider inputs. Since $s_\alpha + s_\beta + s_\gamma = 10$, the derived coefficients satisfy $\alpha + \beta + \gamma = 1$ with each lying in $[0.1, 1.0]$. This zero-sum normalization ensures that increasing emphasis on one objective necessarily reduces the weight on others, making trade-offs explicit and mathematically consistent. The weight coefficients are determined once per planning session and govern the GA fitness function directly, enabling users with diverse preference profiles to obtain correspondingly distinct optimized itineraries from the same candidate POI set without requiring any reconfiguration of the underlying algorithm.

Interested & Avoided Categories capture user preferences over POI types. Users designate attraction categories as *interested* via the Interested panel and mark specific categories as *avoided* via the Avoided panel in the interface. These designations are forwarded to the POI scoring pipeline, which elevates the priority of interested-category attractions and applies hard exclusion constraints for avoided categories, and to the GA fitness function, which uses them to guide evolutionary selection towards preference-aligned itinerary compositions. A post-optimization coverage check further consumes the interested-category list to verify and, if necessary, enforce the presence of each designated category in the final itinerary.

Must-Visit POIs specify individual attractions that the user mandates to appear in the final itinerary. These selections are collected as hard constraints and passed to the GA initialization module, where they receive the highest scheduling priority. If the aggregate required visit duration of all must-visit POIs exceeds the available daily time budget, the system raises a scheduling infeasibility notification before optimization begins.

Itinerary Diversity is a binary toggle that signals the post-evolution solution selection module to enforce categorical diversity across each travel day. This setting is consumed during the final itinerary

assembly stage to ensure that the selected solution maintains a balanced distribution of attraction types.

Weather Sensitive is a binary toggle that specifies whether retrieved weather forecasts should influence POI scoring. This setting is forwarded to the POI scoring pipeline together with the per-day weather data, where it jointly determines which weather-adaptive scoring strategy is applied to each travel day.

3.3. Weather Data Retrieval

Weather data are retrieved from the OpenWeatherMap API at two distinct operational stages. During the *itinerary generation phase*, the API is queried once per planned travel day to obtain the per-day weather forecast. Two data elements are extracted from each response: (1) the official weather condition code (`weather_id`), which classifies the prevailing atmospheric state according to the standardized OpenWeatherMap weather code taxonomy; and (2) the `alerts` field, which carries any active official meteorological warnings issued by national weather services. Both elements are forwarded to the POI scoring pipeline together with the user's Weather Sensitivity setting. Currently, this study does not simulate weather conditions that do not occur in Macao; therefore, the weather conditions described below do not represent all the weather data available on OpenWeatherMap. During the *real-time execution phase*, the API is queried each time the user confirms completion of a POI visit and is about to proceed to the next destination. The data retrieved at this stage capture any deterioration in atmospheric conditions relative to the forecast used during itinerary generation, specifically transitions from non-rainy to rainy conditions or increases in precipitation intensity. These real-time observations are passed to the weather monitoring module, which evaluates whether an outdoor-to-indoor substitution recommendation should be issued for the next scheduled POI.

3.4. POIs Score Calculation

The POI scoring pipeline integrates data from the three preceding collection stages: the pre-compiled POI attributes, including quality tiers, `location_code`, and indoor/outdoor classification, from Data Collection; the interested/avoided category designations, must-visit specifications, and Weather Sensitivity setting from User Preference Collection; and the per-day `weather_id` and `alerts` from Weather Data Retrieval. These inputs are jointly processed to assign each candidate attraction a scenario-specific composite score. The complete mathematical formulation is presented in Section 4.1; the following paragraphs provide a procedural overview to contextualize the subsequent formal treatment.

Filter Closed & Avoided POIs. Before scoring begins, the system applies availability filtering. Attractions closed on the day of visit are identified by matching each POI's weekly closure field against the corresponding travel weekday and removed from the candidate pool. Attractions whose type or name matches the user's avoided list are similarly excluded, and any remaining POIs whose category does not appear in the user's interested or neutral preferences are also removed.

Base Score Calculation. Each POI receives a base score combining its ten-level popularity rating (80% weight) and its log-normalized review count (20% weight), normalized to $[0, 1]$. The level-based rating provides a coarse but robust primary ranking signal; the log-normalized review count differentiates POIs within the same level. The log-normalized review count is derived by applying a decimal logarithmic transformation to the number of reviews. Review counts in the dataset exhibit a highly skewed distribution, ranging from a minimum of 10 to a maximum of 26,897. Raw review counts are transformed using a base-10 logarithm normalized against the dataset maximum, compressing the range into a perceptually uniform scale consistent with the Weber–Fechner law. This prevents massively reviewed landmarks from dominating the scoring and obscuring quality differences among the majority of attractions.

Determine Weather Scoring Strategy. Based on the retrieved `weather_id`, the presence of any `alerts`, and the user's weather sensitive setting, each travel day is assigned to one of nine weather conditions and a corresponding scoring strategy drawn from three qualitatively distinct strategies.

The strategy applied under each weather condition is summarized in Table 1; the full mathematical formulation is presented in Section 4.1.

Extreme weather activation is based on the presence of any entry in the OpenWeatherMap `alerts` field, rather than on custom numerical precipitation thresholds, delegating hazard-level assessment to the issuing meteorological authority. This three-strategy branching structure enables differentiated scoring under qualitatively distinct atmospheric conditions.

Flexibility Bonus. For outdoor POIs under rainfall or extreme weather, a flexibility bonus partially offsets the rain penalty based on the density of nearby indoor alternatives within the same sub-region. The complete formulation is presented in Section 4.1.

Preference Score. User category preferences are applied at two complementary stages: as a score multiplier within the POI scoring pipeline, and as a preference coefficient within the GA fitness function. Avoided-category POIs are fully excluded from the candidate pool. Specific coefficients and scenario-dependent adjustments are detailed in Section 4.1.4.

Table 1. Weather condition scoring strategies

Weather Condition	Scoring Strategy			
	Indoor POI		Outdoor POI	
	Pref. bonus	Pref. bonus	Flex. bonus	Rain penalty
No weather-sensitive in non-extreme condition	✓	✓	–	–
Clear	✓	✓ [†]	–	–
Drizzle	✓	✓	✓	15%
Light shower rain	✓	✓	✓	25%
Heavy drizzle	✓	✓	✓	30%
Light rain	✓	✓	✓	30%
Shower rain	✓	✓	✓	40%
Moderate rain	✓	✓	✓	50%
Extreme rain	✓*	✓*	✓*	60%
Extreme weather	✓*	✓*	✓*	60%

[†] *Clear*: Outdoor POIs additionally receive an outdoor log-bonus.

* *Extreme rain and official alert*: Indoor POIs receive a 10% score boost; preference bonus applies at 50% effectiveness.

Outdoor POIs: Preference bonus at 30% effectiveness; applies flexibility bonus in extreme weather condition.

3.5. Genetic Algorithm for Itinerary Generation

The genetic algorithm receives as its primary inputs the scenario-specific ranked candidate pools produced by the POI scoring pipeline, the preference weight sliders (α , β , γ), must-visit POI list, and daily time constraints from User Preference Collection, and the Itinerary Diversity flag also from User Preference Collection. Operating on these inputs, it evolves a population of 50 candidate multi-day itineraries over up to 100 generations, jointly optimizing POI quality, route compactness, and preference satisfaction as defined by the user-specified weights. Each individual is encoded as an ordered list of daily routes:

$$\begin{aligned} \text{Individual} &= [\text{POI}_1, \text{POI}_2, \dots], && \leftarrow \text{Day 1} \\ &[\text{POI}_3, \text{POI}_4, \dots], && \leftarrow \text{Day 2} \\ &\vdots && \leftarrow \text{Day } N \end{aligned}$$

Candidate Pool Preparation. A per-day candidate pool is constructed for each travel day by (i) removing POIs closed on the corresponding weekday via closure-field matching; (ii) excluding POIs in the user's avoided categories or explicitly avoided by name; (iii) excluding any POI whose category does not appear in the user's interested or neutral preferences; and (iv) removing POIs already allocated to earlier days to prevent duplicate visits across the multi-day itinerary. Each surviving POI is scored using the scenario-specific scoring function described in Section 4.1, producing a ranked candidate list for that day.

Must-Visit Allocation. Must-visit POIs are allocated before any other candidates. If the total must-visit count $|M| \leq D$ (number of travel days), at most one must-visit POI is placed per day. If $|M| > D$, must-visit POIs are distributed as evenly as possible: the first days each receive $\lceil |M|/D \rceil$ must-visit POIs and the last day absorbs the remainder. Must-visit POIs are inserted in descending score order, and each insertion is accepted only if the required visit duration fits within the remaining time budget less a reserve buffer of $\min(30, 0.1 \times T_{\text{available}})$ minutes. If the last day cannot accommodate all remaining must-visit POIs, initialization raises an infeasibility error.

Time-Aware POI Filling. After must-visit allocation, the remaining time budget is filled in two priority-ordered passes. In the first pass, at least one POI from each interested category is inserted, selected by descending score from the per-day candidate pool, to guarantee category coverage. In the second pass, additional POIs are drawn from the top-scored candidates until the remaining available time falls below 25 minutes. The 25-minute floor preserves a minimum buffer for transit to the end point, and the filling target is set to utilize 85% of available daily time, leaving a margin for transit variability. Visit duration is assigned according to the user's interest level: interested-category POIs and must-visit POIs use d_{max} (the Google Maps maximum recommended visit duration), reflecting the expectation of in-depth engagement; all other POIs use d_{min} (the minimum recommended duration), enabling higher throughput of diverse attractions within the same time window. Both d_{max} and d_{min} are sourced from the pre-compiled POI dataset.

Transportation mode and inter-POI travel time are estimated during route construction as follows: for inter-POI distances of 999 metres or less, walking is assumed with a 5-minute buffer to account for GPS coordinate imprecision and pedestrian detour paths; for distances exceeding 999 metres, public bus transport is assumed with a 10-minute buffer for waiting time. Consistent with Macao's tourism context, in which visitors predominantly use walking and public buses, these two modes are exclusively supported. Inter-POI distances are computed using the Haversine great-circle distance formula, and the resulting estimates are incorporated into the temporal feasibility check during candidate filling.

Fitness Evaluation. Each individual is evaluated by a weighted fitness function combining three normalized sub-objectives: POI quality (S_{quality}), traveling efficiency (S_{travel}), and preference satisfaction (S_{pref}). Weather compatibility is pre-embedded in per-POI scores before fitness evaluation; schedule conflicts are pre-empted at the candidate pool level by removing closed POIs prior to scoring. The complete fitness formulation is detailed in Section 4.2.

Tournament Selection. Parent selection uses k -tournament selection with $k = 5$. Five individuals are sampled without replacement from the current population, and the highest-fitness individual is selected as a parent. This process is repeated independently to obtain two parents, maintaining population diversity since lower-fitness individuals retain a non-zero selection probability. The choice of $k = 5$ is deliberate: with a population size of 50, a tournament pool of 5 corresponds to 10% of the population. Binary tournament ($k = 2$) provides insufficient selection pressure, making it difficult to consistently favour genuinely superior individuals and slowing convergence; at the other extreme, large tournament sizes (e.g., $k \geq 20$) concentrate selection pressure excessively on the current best individuals, reducing population diversity and increasing the risk of premature convergence to a local optimum. A tournament size of $k = 5$ strikes a balance between these extremes, providing strong but not overwhelming selection pressure appropriate for a population of this size.

Crossover & Mutation. A day-level uniform crossover is applied with crossover rate 0.8. For each travel day $d \in \{1, \dots, D\}$, the daily sub-route is independently inherited from either Parent 1 or Parent 2 with equal probability (0.5 each). Offspring 2 is constructed by the complementary assignment: wherever Offspring 1 takes its daily route from Parent 1, Offspring 2 takes it from Parent 2, and vice versa. Operating at the day level preserves the internal coherence of each daily sub-route, including its time budget, POI ordering, and transit structure, while enabling recombination of daily plans across parents.

Mutation is applied independently to each day of an offspring at rate 0.1. Rather than applying all three mutation operators simultaneously or selecting between them based on a fixed schedule, one operator is chosen uniformly at random from the following three, and that single operator is applied to the day:

- **Replace:** A randomly selected POI in the day's route is substituted with a new candidate drawn from the scored candidate pool. The replacement POI must not already appear elsewhere in the day's route, preventing duplicate visits.
- **Add:** A new POI is appended to the day's route, drawn from the top-10 highest-scored candidates not already in the route. This operator is only eligible when the current route contains fewer than 5 POIs, ensuring that route length remains manageable.
- **Remove:** A randomly selected POI is deleted from the day's route. Must-visit POIs are excluded from removal eligibility, guaranteeing that user-mandated attractions are preserved across all mutations.

The random uniform selection among the three operators ensures that each form of structural perturbation (substitution, expansion, and reduction) has equal opportunity to contribute to population diversity, preventing the search from becoming biased towards any single type of modification. Must-visit POIs are protected in all operators.

Elite Preservation. An elitism strategy retains the top 10 individuals (20% of the population of 50) unchanged at the start of each generation. The remaining 40 positions are filled with offspring produced by crossover and mutation, guaranteeing that the best-observed fitness is monotonically non-decreasing across generations. The 20% retention rate is a deliberate design choice: retaining too few elites (e.g., 1–2 individuals, or around 2–4%) risks losing high-quality solutions when the population undergoes large generational shifts; conversely, retaining too many (e.g., 50% or more) restricts the proportion of the population subject to crossover and mutation, reducing the exploratory capacity of the evolutionary search and slowing convergence. With a population of 50, preserving 10 individuals leaves 40 slots for offspring, maintaining a sufficient evolutionary pressure while guaranteeing that the best solutions observed so far are never lost.

Convergence Check. The evolutionary cycle terminates under either of the following conditions: (i) the maximum generation limit of 100 is reached; or (ii) the improvement in best fitness over the preceding 20 generations falls below a stagnation threshold. Specifically, starting from generation 21, if the absolute difference between the best fitness recorded at the current generation g and the best fitness recorded 20 generations earlier falls below $\varepsilon = 0.001$, i.e., $|F_g^* - F_{g-20}^*| < \varepsilon$, and the algorithm terminates early. This sliding-window criterion avoids premature termination caused by short-term fitness plateaus while efficiently halting the search once no meaningful improvement has been observed over a sustained period.

Post-Evolution Schedule Optimization and Solution Guarantee. After evolution terminates, the fitness-ranked population is passed to the Schedule Optimizer module, which iterates through candidates in descending fitness order. Each candidate individual is converted to a formatted itinerary via the route construction pipeline, which applies a greedy nearest-neighbor heuristic (NNH) to finalize the POI visit sequence within each day: starting from the day's departure point, NNH iteratively selects the geographically closest unvisited POI until all POIs are visited, after which the designated end point is appended. Inter-POI distances for NNH sequencing are computed using the Haversine

great-circle distance formula, which also supplies the d_{actual} value used in the traveling efficiency sub-objective S_{travel} .

Following route sequencing, the visit window of every POI, defined as the interval [arrival_time, departure_time], is validated against the attraction's operating hours [open_time, close_time]. The first candidate whose complete itinerary satisfies all schedule constraints is returned as the final solution. If no fully schedule-valid individual is found after exhausting the population, the highest-fitness individual is returned with a set of human-readable violation warnings listing the conflicting POI-time pairs.

The selected itinerary is then inspected for any interested category that is entirely absent. For each missing category, the highest-scoring available POI within that category is inserted into the travel day with the greatest remaining time capacity, subject to: (i) the receiving day having a weather compatibility score of at least 0.3; and (ii) the insertion not violating the minimum time buffer. This guarantee ensures that the final output honours all user-specified category preferences regardless of the evolutionary trajectory.

4. Formal Algorithmic Specification

4.1. POIs Score Calculation

The POI scoring formulation assigns each candidate attraction a composite score that integrates four information streams: (1) intrinsic attraction quality, encoded as a base score combining a ten-tier popularity rating (primary signal) and a log-normalized review count (secondary differentiator); (2) prevailing weather conditions, expressed through penalty and bonus multipliers calibrated to nine distinct weather conditions; (3) geographic flexibility, captured by a flexibility bonus that reflects the density of nearby indoor alternatives; and (4) user category preferences, applied as a multiplier that elevates or suppresses a POI's effective score relative to the user's declared interested and avoided designations.

The nine weather conditions recognised by the system (clear, drizzle, heavy drizzle, light shower rain, light rain, shower rain, moderate rain, extreme rain, and extreme weather) are mapped to one of three qualitatively distinct scoring strategies. Strategy 1 covers clear-day or weather-insensitive conditions, in which outdoor POIs receive a logarithmic competitiveness bonus relative to the weather-agnostic baseline. Strategy 2 covers active rainfall under weather-sensitive mode (drizzle through moderate rain), in which outdoor POIs incur intensity-proportional score penalties partially offset by a flexibility bonus. Strategy 3 covers unconditional extreme weather triggered by official meteorological warnings (extreme rain accompanied by an active alert, or extreme weather), in which outdoor POI scores are suppressed by a fixed 60% penalty and indoor POI scores receive a 10% boost regardless of user weather sensitivity settings. Anchoring strategy activation to the standardized OpenWeatherMap `weather_id` taxonomy and the official `alerts` field, rather than to custom numerical thresholds, delegates severity classification to authoritative meteorological sources, ensuring objectivity and experimental reproducibility. The multiplicative structure of the formulation preserves the relative ordering of POIs within the same quality tier while allowing weather conditions and preference alignment to modulate inter-tier rankings dynamically. All final scores are clipped to $[0, 1]$ and sorted in descending order to produce the ranked candidate list consumed by the GA initialization stage. The complete formulae for all three scoring strategies are presented below.

$$\text{Strategy 1 (Clear/Weather-sensitive, outdoor): } \text{score}_1 = \text{base_score} \times (1 + b_{\text{pref}} + b_{\text{outdoor}}) \quad (1)$$

$$b_{\text{outdoor}} = 0.08 \times \log_{10}(\text{base_score} \times 10 + 1) \quad (2)$$

$$\text{Strategy 2 (Rainfall, outdoor): } \text{score}_2 = \text{base_score} \times r_{\text{rain}} \times (1 + b_{\text{pref}} + b_{\text{flex}}) \quad (3)$$

Strategy 3 (Extreme weather, outdoor): $\text{score}_{3,\text{out}} = \text{base_score} \times 0.4 \times (1 + b_{\text{pref}} \times 0.3 + b_{\text{flex,ex}})$ (4)

Strategy 3 (Extreme weather, indoor): $\text{score}_{3,\text{in}} = \text{base_score} \times 1.1 \times (1 + b_{\text{pref}} \times 0.5)$ (5)

4.1.1. Base Score Calculation & Review Count Log Normalization

Review counts in the dataset exhibit a highly skewed distribution, ranging from a minimum of 10 to a maximum of 26,897 (The Venetian Macao). If applied directly as a quality signal, this skew would cause a small number of massively reviewed landmarks to dominate the scoring, obscuring genuine quality differences among the majority of attractions. Google Maps star ratings were not incorporated as a primary signal because, for attractions with few reviews, ratings are statistically unreliable, while for high-review attractions, quality is already adequately reflected by review volume. To compress this range into a perceptually meaningful scale, review counts are transformed using a base-10 logarithm normalized against the dataset maximum, consistent with the Weber–Fechner law, which states that human perception of stimulus magnitude is proportional to the logarithm of the stimulus intensity, such that the perceptual difference between 100 and 1,000 reviews is treated as equivalent to that between 1,000 and 10,000:

$$\text{reviews_count} = \frac{\log_{10}(\text{reviews})}{\log_{10}(26,897)} \quad (6)$$

After transformation, $\text{reviews_count} \in [0, 1]$. Each POI in the dataset is assigned a popularity_rating on a ten-tier scale derived from its review count relative to all attractions in the dataset; this tiered rating serves as the primary quality signal and provides a coarse but robust ranking that is not distorted by individual outlier review counts. The log-normalized review count (log_factor) is retained as a secondary differentiating signal within the same tier. The base score is computed as a weighted sum of these two components and normalized to $[0, 1]$:

$$\text{base_score} = 0.8 \times \frac{\text{popularity_rating}}{10} + 0.2 \times \text{reviews_count} \quad (7)$$

This base score serves as the foundational input for all subsequent scenario scoring calculations.

4.1.2. Rain Penalty

Rain penalties apply exclusively to outdoor POIs under Scoring Strategy 2 (active rainfall with weather sensitivity enabled). Rainfall intensity is classified using the official OpenWeatherMap weather_id code, rather than custom numerical precipitation thresholds, ensuring that classification aligns with standardized meteorological definitions. The penalty rates corresponding to each of the six rainfall conditions (drizzle through moderate rain) are specified in Table 1. Extreme weather detection operates independently and takes precedence over user weather sensitivity settings: even if the user has not enabled weather sensitivity, a non-empty alerts field unconditionally activates Scoring Strategy 3 for all POIs on the affected travel day. This design delegates hazard-level assessment to the issuing meteorological authority, ensuring objectivity and standardization. The three-strategy branching structure enables the system to apply differentiated scoring under qualitatively distinct atmospheric conditions, overcoming the limitation of a single scoring function that cannot simultaneously accommodate preferences for sunny-day outdoor exploration and the safety imperatives of severe weather.

4.1.3. Flexibility Bonus

The flexibility bonus partially offsets rain penalties for outdoor POIs located in regions with abundant nearby indoor alternatives. The rationale is that an outdoor attraction surrounded by accessible indoor options presents a lower effective weather risk to visitors, since itinerary adjustments can be made on-site with minimal disruption. The flexibility factor f_{flex} quantifies this substitutability, and the bonus b_{flex} is further scaled by the logarithm of the base score to ensure that only higher-quality outdoor attractions benefit meaningfully from geographic advantage; low-quality outdoor POIs receive only marginal compensation. The calculation is as follows:

$$f_{\text{flex}} = \min\left(\frac{\text{indoor_alternatives}}{10}, 1.0\right) \quad (8)$$

$$\text{Rainfall: } b_{\text{flex}} = 0.12 \times f_{\text{flex}} \times \log_{10}(\text{base_score} \times 10 + 1) \quad (9)$$

$$\text{Extreme weather: } b_{\text{flex,ex}} = 0.18 \times f_{\text{flex}} \times \log_{10}(\text{base_score} \times 10 + 1) \quad (10)$$

The higher coefficient under extreme weather (0.18 vs. 0.12) ensures that high-quality outdoor POIs in geographically flexible regions retain a non-trivial selection probability even under severe conditions, reflecting the principle that meteorological warnings signal elevated risk rather than a blanket prohibition on outdoor activity. Tying the bonus to the logarithm of the base score ensures that only attractions of substantive quality benefit meaningfully from geographic advantage, preventing low-quality POIs from inflating their scores through proximity to indoor alternatives.

4.1.4. Preference Score

User category preferences are encoded as a three-tier multiplier applied at two complementary stages. At the POI scoring stage, attractions belonging to the user's interested categories receive a 15% linear bonus ($b_{\text{pref}} = 0.15$) on their base score, elevating their position in the ranked candidate pool used to initialise the GA population. The effective weight of this bonus varies by weather scenario to reflect the competing priority of safety. Under Scoring Strategies 1 and 2, the 15% preference bonus applies at full effectiveness to all POIs. Under Scoring Strategy 3 (extreme weather), the bonus is attenuated: for outdoor POIs it is reduced to 30% of its nominal value, since the mandatory 60% weather penalty already substantially suppresses outdoor scores and retaining a high preference weight would risk allowing strongly preferred outdoor POIs to out-compete indoor alternatives, undermining the safety-oriented design intent; for indoor POIs it is retained at 50%, balancing the need to differentiate among the enlarged indoor candidate pool while preventing over-prioritisation of niche categories at the expense of overall quality. The asymmetric attenuation (30% outdoor, 50% indoor) encodes the system's explicit priority hierarchy: weather safety first, broad quality second, personal preference third.

At the GA fitness evaluation stage, the same preference signal is applied through the preference satisfaction objective S_{pref} , where each POI i is assigned a preference coefficient:

$$p_i = \begin{cases} 1.5 & \text{if } t_i \in \mathcal{C}_{\text{interested}} \\ 1.0 & \text{if } t_i \in \mathcal{C}_{\text{neutral}} \\ 0.3 & \text{if } t_i \in \mathcal{C}_{\text{excluded}} \end{cases} \quad (11)$$

This drives the evolutionary selection process to favor itinerary compositions with higher interested-category coverage. The two preference mechanisms operate independently and complement each other: the candidate-pool bonus ensures preferred categories are represented early in population initialization, while the fitness multiplier ensures itineraries containing preferred POIs are selectively retained through evolution. In all scenarios, POIs belonging to excluded categories are

assigned a zero score and are entirely removed from the candidate pool; this hard exclusion rule takes precedence over all other scoring mechanisms.

4.2. Genetic Algorithm Fitness Function

The fitness function F governs the evolutionary selection pressure that drives the GA towards high-quality itinerary solutions. It is formulated as a weighted linear combination of three normalized sub-objectives: POI quality (S_{quality}), traveling efficiency (S_{travel}), and preference satisfaction (S_{pref}). Each sub-objective is weighted by the corresponding normalized coefficient (α , β , or γ) derived from the user's integer slider input by dividing by 10 (as described in Section 3.2). This design enables users with diverse preference profiles to derive correspondingly distinct optimized itineraries from the same candidate pool without requiring algorithm reconfiguration. Normalizing each sub-objective to $[0, 1]$ before weighting ensures commensurability across objectives that differ in natural units (quality scores, route distances, and category coverage ratios), preventing any single objective from dominating solely by virtue of scale. Since the three integer slider values sum to 10, the derived weight coefficients satisfy $\alpha + \beta + \gamma = 1$, making trade-offs explicit: increasing one weight necessarily reduces the others, ensuring all user configurations produce a valid, interpretable fitness landscape:

$$F = \alpha \cdot S_{\text{quality}} + \beta \cdot S_{\text{travel}} + \gamma \cdot S_{\text{pref}} \quad (12)$$

The weight coefficient α encodes the user's priority on POI quality; β encodes priority on traveling efficiency (i.e., route compactness); and γ encodes priority on personal preference alignment. The correspondence between each weight coefficient and its sub-objective, together with its derivation from the integer slider input, is summarized in Table 2.

Table 2. Mapping of user-defined slider inputs to normalized weight coefficients and fitness sub-objectives. Each integer slider value $s \in [1, 10]$; the weight coefficient equals $s/10 \in [0.1, 1.0]$.

Slider Label	Integer Input	Weight Coefficient	Sub-Objective
Score	$s_{\alpha} \in [1, 10]$	$\alpha = s_{\alpha}/10$	POI Quality (S_{quality})
Distance	$s_{\beta} \in [1, 10]$	$\beta = s_{\beta}/10$	Traveling Efficiency (S_{travel})
Preference	$s_{\gamma} \in [1, 10]$	$\gamma = s_{\gamma}/10$	Preference Satisfaction (S_{pref})

Two additional quality dimensions, namely weather–POI compatibility and schedule adherence, are handled outside the GA fitness function. Weather compatibility is embedded directly into per-POI scores during the POI Score Calculation stage (Section 4.1): the final score of each POI already encodes the scenario-specific rain penalty, flexibility bonus, and extreme-weather suppression before any individual is evaluated. Schedule conflicts are pre-empted at the candidate pool level: POIs closed on the travel day are removed before scoring (Section 4.1.1), ensuring that individuals in the initial population and all evolved offspring are largely free of temporal infeasibility; residual conflicts are resolved post-evolution by the Schedule Optimizer. Each sub-objective is described below.

4.2.1. POI Quality (S_{quality})

This sub-objective measures the average intrinsic attraction quality of the POIs selected in the itinerary. It is defined as the mean base score over all included POIs:

$$S_{\text{quality}} = \frac{1}{|P|} \sum_{p \in P} S_{\text{base}}(p) \quad (13)$$

Here P is the set of all POIs in the itinerary, $|P|$ is its cardinality, and $S_{\text{base}}(p)$ is the base score of POI p as defined in Section 4.1.1. Maximizing S_{quality} drives the GA to favor candidate sets composed of high-popularity, well-reviewed attractions.

4.2.2. Traveling Efficiency (S_{travel})

This sub-objective penalizes unnecessary detours relative to the minimum achievable point-to-point traversal distance. A detour ratio ρ is computed as the ratio of the NNH-sequenced route distance to the theoretical minimum distance (consecutive straight-line links in the unoptimized fixed order), and S_{travel} is derived from ρ via a diminishing-return mapping:

$$\rho = \frac{d_{\text{actual}}}{d_{\text{min}}} \quad S_{\text{travel}} = \frac{1}{1 + 0.5(\rho - 1)} \quad S_{\text{travel}} \in [0.1, 1.0] \quad (14)$$

Since $\rho \geq 1$ by construction, $S_{\text{travel}} = 1$ when the actual route equals the minimum distance and decreases monotonically as detours increase, with the lower bound of 0.1 preventing complete collapse. d_{actual} is the total NNH-sequenced route distance; d_{min} is the sum of consecutive straight-line distances between the start point, POIs in their original (unoptimized) order, and the end point. Both distances are computed using the Haversine great-circle distance formula, which calculates the shortest over-the-surface distance between two geographic coordinates by accounting for the curvature of the Earth given the latitude and longitude differences of each pair of points. A mean Earth radius of $R = 6,371$ km is used throughout.

4.2.3. Preference Satisfaction (S_{pref})

This sub-objective quantifies how well the itinerary reflects the user's declared category preferences and must-visit requirements. It combines two components: the mean preference coefficient over all itinerary POIs, and the fraction of must-visit POIs covered:

$$S_{\text{pref}} = 0.7 \cdot \bar{p} + 0.3 \cdot \frac{|M_{\text{covered}}|}{|M|} \quad (15)$$

Here \bar{p} is the mean preference coefficient over all POIs in the itinerary, where each POI i carries coefficient p_i drawn from the three-tier scheme defined in Section 4.1.4: $p_i = 1.5$ for interested, 1.0 for neutral, and 0.3 for excluded categories. The second term, $|M_{\text{covered}}|/|M|$, is the proportion of user-specified must-visit POIs that appear in the itinerary. Must-visit coverage is assigned the smaller weight (0.3) because the population initialization strategy already allocates must-visit POIs with the highest priority; this component serves as a soft verification check to reward and maintain full must-visit coverage across generations, rather than being the primary optimization driver.

5. Results and Discussion

This section presents the experimental validation of the proposed system, structured around two complementary objectives: verifying that the weather-adaptive scoring mechanism produces itinerary compositions consistent with prevailing atmospheric conditions, and confirming that the user-defined preference weight sliders (α , β , γ) each produce the intended optimization behavior relative to the default configuration. All experiments were conducted on the Macao POI dataset comprising 142 attractions (69 outdoor, 73 indoor), of which 41 are classified as high-quality (popularity rating ≥ 8).

5.1. Weather-Adaptive Scoring Validation

As described in Table 1, POI scoring applies one of three scoring strategies, determined jointly by the retrieved forecast across nine weather conditions and the user's Weather Sensitivity setting. A dedicated experiment was designed for each scenario, comparing itineraries generated with weather-adaptive scoring enabled against a weather-agnostic baseline that applies only base scores and preference bonuses. All experiments in this subsection use 20 diverse preference configurations with a fixed daily capacity of 6 POIs per itinerary.

1) Scoring Strategy 1: Clear-Day Outdoor Bonus Validation. Under Scoring Strategy 1 (clear or weather-insensitive conditions), outdoor POIs receive a logarithmic score bonus that amplifies their competitiveness relative to indoor alternatives. To validate this mechanism, itineraries generated with the outdoor bonus enabled were compared against a weather-agnostic baseline across 20 preference

configurations with a fixed capacity of 6 POIs. The experimental results are shown in Figure 2. With the outdoor bonus active, the average number of outdoor POIs selected per itinerary increased from 2.80 to 2.88, an improvement of 0.08 attractions, while the average high-quality POI count remained stable at 2.75 (versus 2.76 in the baseline) and the total POI count held at 6.00 under both conditions, confirming that the outdoor bonus elevates outdoor POI competitiveness on clear days without displacing high-quality attractions or reducing overall itinerary completeness.

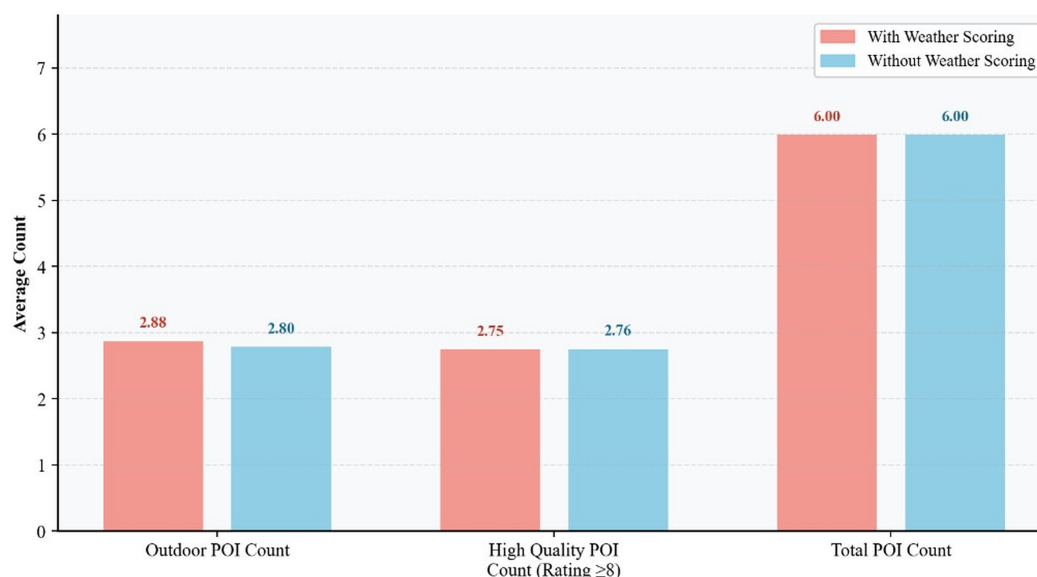
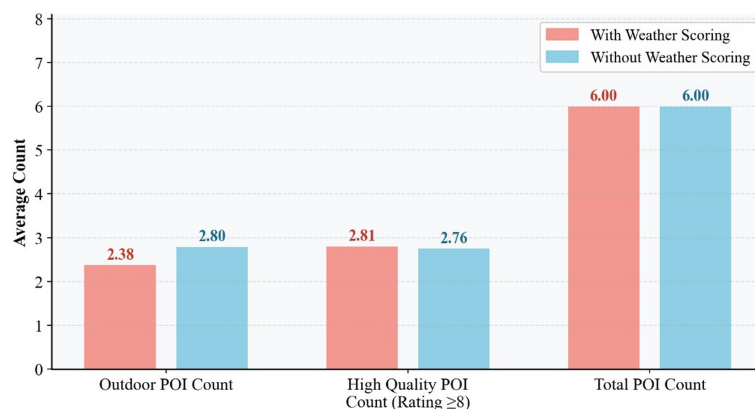


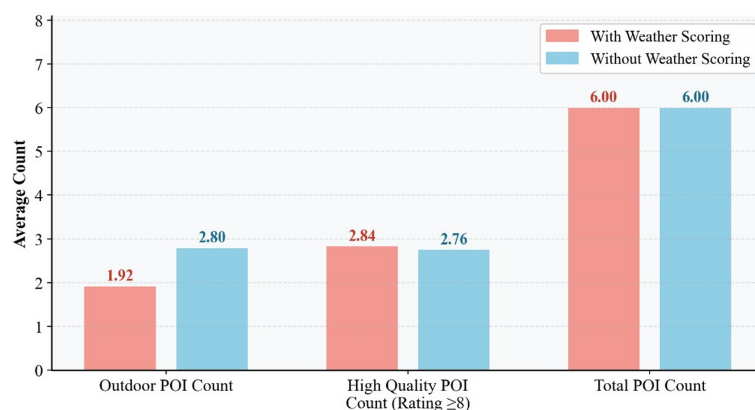
Figure 2. Sunny day outdoor POI bonus validation.

2) Scoring Strategy 2: Rainfall Penalty and Flexibility Bonus Validation. Under Scoring Strategy 2 (active rainfall with Weather Sensitivity enabled), outdoor POIs incur intensity-proportional score penalties partially offset by a flexibility bonus reflecting the density of nearby indoor alternatives. To validate this mechanism, comparative experiments were conducted across 20 preference configurations under three representative rainfall levels drawn from the nine weather conditions (light rain: 30% penalty; moderate rain: 50% penalty; heavy rain: 70% penalty). The experimental results are shown in Figure 3.

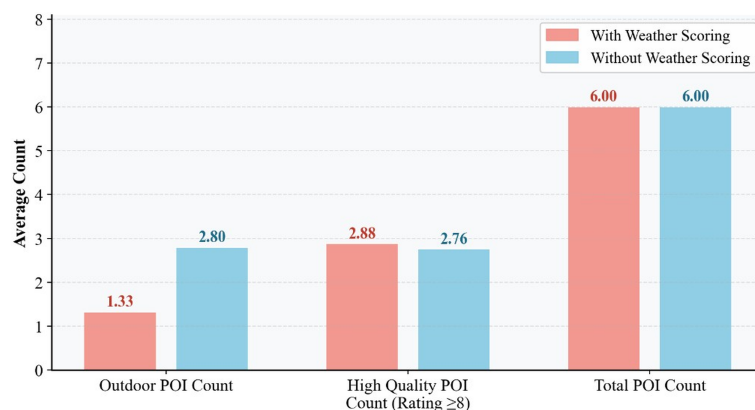
The results demonstrate a progressive, intensity-proportional reduction in outdoor POI count: the weather-scoring group averaged 2.38, 1.92, and 1.33 outdoor POIs under light, moderate, and heavy rain respectively, compared with a stable baseline of 2.80 across all three conditions. Critically, the average high-quality POI count was maintained or marginally improved under rainfall scoring (2.81, 2.84, and 2.88 vs. 2.76 for the baseline), confirming that the flexibility bonus effectively compensates for rain penalties on high-quality outdoor attractions situated in geographically dense areas. The total POI count remained at 6.00 for both profiles under all rainfall levels, indicating that weather-adaptive scoring reallocates selection pressure from outdoor to indoor POIs without reducing itinerary completeness. These results collectively validate the graduated penalty design: as rainfall intensifies, outdoor POI representation decreases proportionally, yet the overall itinerary quality is preserved through the selective retention of high-quality outdoor attractions with abundant indoor substitution options nearby.



(a) Comparison in light rain



(b) Comparison in moderate rain



(c) Comparison in heavy rain

Figure 3. Comparison in three levels of rainfall

3) Scoring Strategy 3: Extreme Weather Suppression Validation. Under Scoring Strategy 3, triggered unconditionally by an active official meteorological alert (corresponding to the extreme rain or extreme weather conditions in Table 1), the system applies a 60% score suppression to all outdoor POIs (score multiplied by $\times 0.4$) and a 10% boost to all indoor POIs ($\times 1.1$), regardless of user weather sensitivity settings. To validate this mechanism, itineraries generated under a simulated official alert were compared against the weather-agnostic baseline across 20 preference configurations. Given the documented frequency of typhoon events in Macao, the primary objective was to confirm that outdoor POIs would be significantly suppressed under such conditions. The experimental results are shown in Figure 4.

Under the extreme weather profile, the average indoor POI count increased from 3.20 to 4.42 (+1.22), implying a corresponding reduction in outdoor POI representation from 2.80 to 1.58. The

average high-quality POI count rose marginally from 2.76 to 2.84, and the total POI count remained at 6.00 for both profiles. These results confirm that the unconditional nature of Scoring Strategy 3 activation, anchored to the presence of official alerts rather than user settings, reliably suppresses outdoor POI selection under officially declared hazardous conditions, substantially shifting itinerary composition towards indoor alternatives while preserving overall itinerary quality.

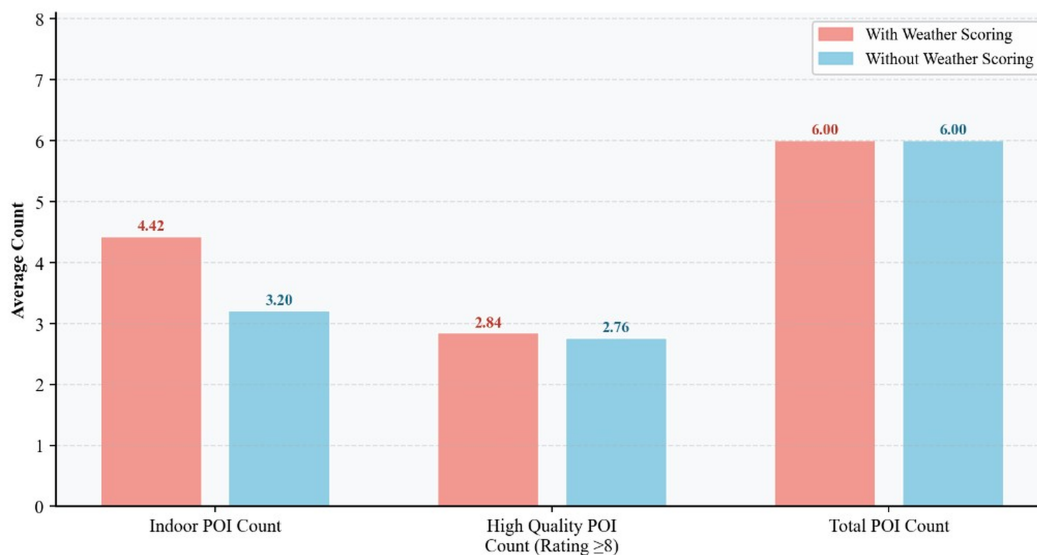


Figure 4. Extreme weather scoring validation.

5.2. User Preference Weight Sensitivity Analysis

As specified in Table 2, the normalized weight coefficients α , β , and γ govern the fitness sub-objectives S_{quality} , S_{travel} , and S_{pref} respectively, and are derived from the corresponding integer slider inputs by dividing by 10. Throughout this subsection, slider configurations are expressed as integer tuples $(s_{\alpha}:s_{\beta}:s_{\gamma})$; for example, the default configuration (4:3:3) corresponds to weight coefficients $\alpha = 0.4$, $\beta = 0.3$, $\gamma = 0.3$.

To evaluate whether each weight coefficient produces the intended optimization behavior under diverse user scenarios, a dedicated experiment was designed for each, comparing a profile that maximizes the target slider (set to 8, with the remaining two each set to 1) against the system default (4:3:3) across 20 preference configurations. Each preference configuration corresponds to a distinct set of interested POI category types that a real user would designate through the Interested panel in the web application's frontend interface (described in Section 3.2). For example, a user interested primarily in museums would select only "Museum", while a user interested in a broader range of religious sites might select "Catholic Church", "Buddhist Temple", and "Taoist Temple" simultaneously. The 20 configurations used in this study span single-category, multi-category, and unconstrained (no preference) cases, covering the full range of use patterns observed in the Macao context; the complete list is shown in the first column of Table 3.

5.2.1. Quality Weight (α) Sensitivity

As specified in Table 2, α is the weight coefficient governing the POI quality sub-objective S_{quality} , which is computed as the mean base score over all itinerary POIs. To validate its effect, a quality-focused profile (slider 8:1:1, i.e. $\alpha = 0.8$, $\beta = 0.1$, $\gamma = 0.1$) was compared against the default profile (slider 4:3:3, i.e. $\alpha = 0.4$, $\beta = 0.3$, $\gamma = 0.3$) across 20 preference configurations with a fixed capacity of 6 POIs, yielding 20 experimental groups. The results are presented in Table 3. Under the high- α profile, the average POI score across all 20 configurations increased from 0.798 to 0.902, an improvement of 0.104 (+13.0%), directly confirming that amplifying the quality weight causes the GA to strongly favour higher-scored attractions. The average total POI count remained at 6.00 for both profiles, indicating that the quality-driven selection pressure reallocates capacity towards higher-rated attractions without

reducing itinerary completeness. This result is consistent with the probability-weighted selection model: under high- α , the fitness landscape disproportionately rewards individuals containing top-tier POIs, concentrating evolutionary selection pressure on the high-quality end of the candidate pool.

Table 3. Average POI score under the Default and Quality-Priority profiles across 20 preference configurations.

Preference Setting	Avg. POI Score (Default slider 4:3:3)	Avg. POI Score (Quality slider 8:1:1)
No preference	0.890	0.890
Museum	0.856	0.940
Park	0.759	0.931
Tourist attraction	0.802	0.860
Historical landmark	0.723	0.889
Garden	0.883	0.956
Shopping (no museum)	0.875	0.887
Museum+Landmark	0.679	0.806
Park+Garden	0.769	0.898
Attraction+Church	0.840	0.840
Catholic Church	0.758	0.950
Buddhist Temple	0.747	0.865
Taoist Temple	0.764	0.899
Beach	0.843	0.933
Observation Deck	0.867	0.931
Religious (3 faiths)	0.668	0.929
Park+Garden+Beach	0.704	0.938
Museum+Landmark+Church	0.839	0.875
Attraction+Museum (No temples)	0.869	0.869
Attraction+Park+Museum	0.831	0.955
Average	0.798	0.902

5.2.2. Traveling Efficiency Weight (β) Sensitivity

As specified in Table 2, β is the weight coefficient governing the traveling efficiency sub-objective S_{travel} , which penalises route detours relative to the minimum point-to-point traversal distance via the detour ratio ρ (see Section 4.2). A lower ρ indicates a more geographically compact, efficient itinerary: $\rho = 1$ corresponds to a route with no excess distance above the theoretical minimum traversal, while larger values reflect progressively more circuitous routing. Minimizing ρ is therefore the intended optimization outcome when β is amplified.

To validate its effect, a travel-efficiency-focused profile (slider 1:8:1, i.e. $\alpha = 0.1, \beta = 0.8, \gamma = 0.1$) was compared against the default profile (slider 4:3:3, i.e. $\alpha = 0.4, \beta = 0.3, \gamma = 0.3$) across 20 preference configurations. Table 4 presents the detour ratio ρ for each configuration under both profiles. Under the high- β profile, ρ is reduced in 18 out of 20 configurations, with an average decrease from 1.264 to 1.127 ($\Delta = -0.137$). The two configurations with the largest absolute reductions are Taoist Temple ($\rho: 1.379 \rightarrow 1.057, \Delta = -0.322$) and Observation Deck ($\rho: 1.448 \rightarrow 1.146, \Delta = -0.302$), which are categories with geographically dispersed candidates where route compactness optimisation has the greatest leverage. The two configurations exhibiting no change (No Preference and Historical Landmark)

correspond to candidate pools that are already spatially compact under both profiles, leaving no margin for further distance reduction. These results confirm that amplifying β consistently drives the GA towards more compact itinerary routes, with the magnitude of improvement scaling with the geographic dispersion of the candidate POI set. Figure 5 provides a visual comparison of representative routes under the two configurations; orange routes correspond to the Travel-Efficiency profile (slider 1:8:1, $\alpha = 0.1, \beta = 0.8, \gamma = 0.1$) and blue routes to the Default profile (slider 4:3:3, $\alpha = 0.4, \beta = 0.3, \gamma = 0.3$).

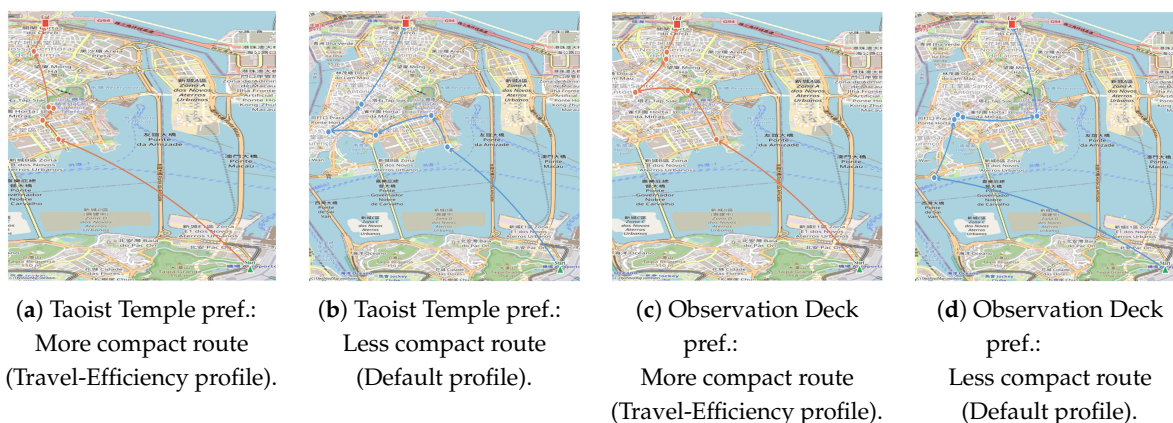


Figure 5. Representative route maps for the two preference configurations with the largest detour ratio reduction.

5.2.3. Preference Satisfaction Weight (γ) Sensitivity

As specified in Table 2, γ is the weight coefficient governing the preference satisfaction sub-objective S_{pref} , which rewards itineraries that include POIs from the user's interested categories and cover must-visit requirements. To validate its effect, a preference-focused profile (slider 1: 1:8, i.e. $\alpha = 0.1, \beta = 0.1, \gamma = 0.8$) was compared against the default profile (slider 4: 3:3, i.e. $\alpha = 0.4, \beta = 0.3, \gamma = 0.3$) across 20 preference configurations with a fixed capacity of 5 POIs per itinerary. The primary evaluation metric is the count of preferred POIs, that is, attractions belonging to the user's declared interested categories, included in the final itinerary. Table 5 reports the results across all 20 configurations. Under the high- γ profile, the average number of preferred POIs increased from 2.70 to 3.60 (+0.90, +33.3%), demonstrating that amplifying γ effectively drives the GA to select more POIs aligned with the user's declared category interests. Among the 19 configurations with non-empty interested categories, the average preferred POI count rises from 2.84 to 3.79, confirming a consistent and meaningful improvement. The total POI count remained at 5.00 for both profiles, indicating that the enhanced preference alignment is achieved through category-selective substitution within the same capacity budget rather than through changes to itinerary size. Notably, configurations with geographically or categorically underrepresented interested types (e.g., Beach, Observation Deck) show the largest absolute gains, reflecting the stronger selection pressure exerted by high- γ fitness weighting on categories that are otherwise easily displaced by high-scoring generic attractions under the default configuration.

Table 4. Detour ratio ρ under the Default and Travel-Efficiency profiles across 20 preference configurations.

Preference Setting	ρ (Default slider 4:3:3)	ρ (Travel-Efficiency slider 1:8:1)
No preference	1.207	1.207
Museum	1.288	1.071
Park	1.313	1.108
Tourist attraction	1.369	1.112
Historical landmark	1.057	1.057
Garden	1.301	1.072
Shopping (no museum)	1.197	1.106
Museum+Landmark	1.261	1.076
Park+Garden	1.218	1.186
Attraction+Church	1.327	1.159
Catholic Church	1.275	1.157
Buddhist Temple	1.240	1.149
Taoist Temple	1.379	1.057
Beach	1.286	1.114
Observation Deck	1.448	1.146
Religious (3 faiths)	1.204	1.145
Park+Garden+Beach	1.243	1.136
Museum+Landmark+Church	1.244	1.174
Attraction+Museum (no temples)	1.220	1.175
Attraction+Park+Museum	1.195	1.141
Average	1.264	1.127

5.3. Weather–Preference Scoring Mechanism Compatibility

The experiments in the preceding subsections validate each component of the scoring framework in isolation. This subsection examines whether the weather-adaptive scoring mechanism and the user preference scoring mechanism interact in a theoretically coherent manner, specifically whether the relative competitiveness of preferred outdoor POIs versus neutral indoor POIs degrades in a graduated, expectation-consistent fashion as precipitation intensifies. The experiment uses representative analytical POIs with a base score of 0.7 and a flexibility factor of 0.8. A base score of 0.7 was selected because, within the Macao dataset distribution, this value corresponds to popularity ratings of approximately 8–9, which represent the primary competitive tier actually recommended by the system rather than tail-end extremes. A flexibility factor of 0.8 corresponds to approximately 8 nearby indoor alternatives, consistent with high-density sub-regions in the dataset. Figure 6 plots the final scores of four representative POI categories (preferred outdoor, neutral outdoor, preferred indoor, and neutral indoor) across all weather intensity levels defined in Table 1.

Table 5. Preferred POI count under the Default and Preference-Satisfaction profiles across 20 preference configurations.

Preference Setting	N_{pref} (Default slider 4:3:3)	N_{pref} (Preference slider 1:1:8)
No preference	0	0
Museum	3	3
Park	4	4
Tourist attraction	3	4
Historical landmark	1	3
Garden	0	3
Shopping (no museum)	4	4
Museum+Landmark	5	5
Park+Garden	3	5
Attraction+Church	5	5
Catholic Church	3	4
Buddhist Temple	1	3
Taoist Temple	1	2
Beach	0	1
Observation Deck	0	2
Religious (3 faiths)	4	5
Park+Garden+Beach	4	4
Museum+Landmark+Church	4	5
Attraction+Museum (no temples)	4	5
Attraction+Park+Museum	5	5
Average (19 with preferences)	2.84	3.79

Several structural intersections in Figure 6 confirm the expected priority hierarchy. First, the preferred outdoor curve remains above the preferred indoor curve for clear-day and drizzle conditions, crossing to below only as precipitation enters the light rain range, indicating that outdoor attractions retain competitive advantage under mild adverse conditions rather than being immediately suppressed. Second, the preferred outdoor curve crosses the neutral indoor curve slightly before the light rain threshold, signalling the point at which weather severity begins to override personal preference, consistent with the intended design priority of weather safety above category alignment. Third, the neutral outdoor and neutral indoor curves intersect near the drizzle-to-light-rain transition, confirming that the penalty gradient is calibrated such that weather suppression begins to matter substantively only once precipitation becomes objectively meaningful. Collectively, these intersection patterns confirm that the two scoring mechanisms are mutually compatible and that their combined effect produces a coherent, graduated trade-off between user preference and weather safety across the full precipitation spectrum.

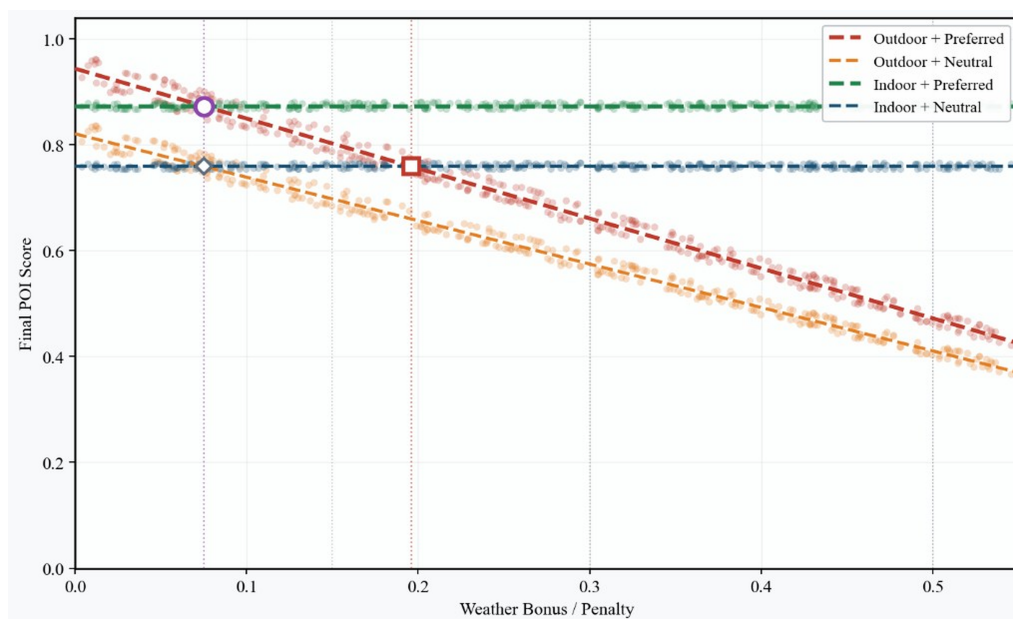


Figure 6. Trend lines for the four categories of POIs.

5.4. Discussion

The experimental results collectively demonstrate that the proposed multi-objective GA framework achieves its two principal design objectives. With respect to weather-adaptive scoring, the three-strategy mechanism, which maps nine distinct weather conditions to three qualitatively differentiated scoring strategies, produces itinerary compositions that respond to precipitation in a proportional, safety-consistent manner: outdoor POI representation increases modestly on clear days (+0.08 on average, from 2.80 to 2.88, relative to the weather-agnostic baseline), decreases progressively as rainfall intensifies (from 2.38 under light rain to 1.33 under heavy rain), and is substantially suppressed under official extreme weather alerts (reduced from 2.80 to 1.58, with indoor POI count rising from 3.20 to 4.42). Critically, these weather-driven reallocation effects are achieved without sacrificing overall itinerary quality or completeness: high-quality POI counts are maintained or marginally improved across all rainfall conditions and under extreme weather (2.76 to 2.84), validating the flexibility bonus as an effective mechanism for selectively preserving high-quality outdoor attractions in geographically dense areas. With respect to user-participatory weight adjustment, amplifying each weight coefficient produces measurable, monotone improvements in its respective sub-objective: raising α to 0.8 (slider 8:1:1) increases the average POI score from 0.798 to 0.902 (+13.0%); raising β to 0.8 (slider 1:8:1) reduces mean detour ratio from 1.264 to 1.127 ($\Delta = -0.137$) with reductions observed in 18 out of 20 configurations; and raising γ to 0.8 (slider 1:1:8) increases preferred POI count from 2.70 to 3.60 (+33.3%). These improvements are realized without coupling effects that would degrade non-target objectives, confirming the functional independence of the three sub-objectives, namely POI quality (S_{quality}), traveling efficiency (S_{travel}), and preference satisfaction (S_{pref}), as implemented. The compatibility analysis further confirms that the preference and weather scoring mechanisms interact in a theoretically consistent manner, with the weather-to-preference priority crossover occurring at the expected precipitation intensity level.

Several limitations warrant acknowledgement. The current POI dataset is limited to 144 attractions in Macao, and the system's generalization to larger-scale, multi-city deployment has not been empirically validated. Recommended visit durations are simulated estimates, and deviations between estimated and actual visit times may affect schedule feasibility in practice. The system does not currently model structured meal-time breaks within daily itineraries; lunch scheduling is treated implicitly as a flexible interval within the available time budget rather than as an explicit planning objective. Additionally, the system evaluates itinerary quality through objective algorithmic metrics rather than subjective user satisfaction measures; A/B testing and user-study validation remain directions for

future work. Future improvements could extend the scoring framework to incorporate real-time crowd density data, integrate multi-modal transport options beyond walking and bus, and support collaborative planning for group travellers with heterogeneous preferences.

6. Conclusions

This paper proposes a multi-objective genetic algorithm-based itinerary planning system, deployed as a Flask web application and evaluated on a real-world dataset of 144 attractions in Macao. The system addresses two structural limitations common to existing GA-based approaches: the absence of user participation in itinerary scoring logic, and the reliance on coarse, non-graduated weather treatment that risks wholesale exclusion of high-quality outdoor attractions under mild precipitation.

The system operates through an integrated five-stage pipeline. First, a pre-compiled POI dataset and user inputs, including travel dates, start/end locations, must-visit POIs, preference category designations, and three integer sliders whose values are normalized to produce weight coefficients α (POI quality), β (traveling efficiency), and γ (preference satisfaction), are collected. Second, per-day weather forecasts are retrieved via the OpenWeatherMap API. Third, a weather-adaptive scoring pipeline assigns each candidate attraction a composite score by combining its intrinsic quality (base score), a weather-dependent penalty or bonus calibrated to one of nine recognized weather conditions, a geographic flexibility bonus reflecting the density of nearby indoor alternatives, and a user preference multiplier. The nine weather conditions are mapped to three scoring strategies: Strategy 1 for clear or weather-insensitive days, Strategy 2 for active rainfall (drizzle through moderate rain) with a graduated intensity-proportional penalty, and Strategy 3 for unconditional extreme weather triggered by official meteorological alerts. Fourth, the GA evolves a population of 50 candidate multi-day itineraries over up to 100 generations, optimizing a fitness function $F = \alpha \cdot S_{\text{quality}} + \beta \cdot S_{\text{travel}} + \gamma \cdot S_{\text{pref}}$ whose weights are directly governed by the user sliders. The GA employs k -tournament selection ($k = 5$), day-level uniform crossover (rate 0.8), randomized mutation (rate 0.1), elite preservation of the top 10 individuals, and a sliding-window early-termination criterion. Fifth, the best schedule-valid solution is returned, with post-evolution insertion guaranteeing coverage of all interested categories.

The principal innovations of this work are (i) a nine-condition, three-strategy weather-adaptive scoring framework that implements a graduated rain penalty with a flexibility bonus, overcoming binary weather exclusion while preserving high-quality outdoor attractions under mild adverse conditions; and (ii) a user-participatory evaluation architecture that maps user-defined weight sliders directly to GA fitness sub-objectives, enabling diverse traveler profiles to produce correspondingly distinct itineraries without algorithm reconfiguration. Experimental results confirm that weather-adaptive scoring shifts itinerary composition proportionally with precipitation intensity while maintaining overall quality, and that each preference slider produces its intended, direction-correct effect on the corresponding sub-objective without degrading the others.

The current system does not explicitly model structured meal-time or lunch breaks within daily itineraries; meal time is treated implicitly as a flexible interval within the available time budget rather than as a discrete planning objective. Incorporating structured lunch scheduling, including proximity-aware restaurant selection, optimal timing relative to attraction visit sequences, and dietary preference filtering, represents a natural extension of the present framework and is reserved as a direction for future work. Additional future directions include large-scale multi-city validation, incorporation of real-time crowd density and queuing time data, multi-modal transport support, and controlled user studies combining A/B testing and preference elicitation.

Author Contributions: Conceptualization, S.L.; investigation, S.L.; data curation, S.L.; writing original draft, S.L.; writing review and editing, S.L. and R.C.; supervision, R.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the .

Data Availability Statement: The dataset presented in this article can be collected on your own, since all the information in it was gathered using Google Maps.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Cárdenas-García, P. J.; Sánchez-Rivero, M.; Pulido-Fernández, J. I.. Does tourism growth influence economic development?. *Journal of travel Research* **2015**, *54*(2), 206–221.
2. Sharma, G. D.; Thomas, A.; Paul, J.. Reviving tourism industry post-COVID-19: A resilience-based framework. *Tourism management perspectives* **2015**, *37*, 100786.
3. Zhu, C.; Hu, J. Q.; Wang, F.; Xu, Y.; Cao, R.. On the tour planning problem. *Annals of Operations Research* **2012**, *192*(1), 67–86.
4. Yochum, P.; Chang, L.; Gu, T.; Zhu, M.. An adaptive genetic algorithm for personalized itinerary planning. *IEEE Access* **2020**, *8*, 88147–88157.
5. Şehab, M.; Turan, M.. An enhanced genetic algorithm solution for itinerary recommendation considering various constraints. *PeerJ Computer Science* **2024**, *10*, e2340.
6. Anshari, M. R.; Baizal, Z. K. A.. N-days tourist route recommender system in Yogyakarta using genetic algorithm method. *JUPI* **2023**, *8*(3), 736–743.
7. Cena, F.; Console, L.; Micheli, M.; Vernerero, F.. Including the Temporal Dimension in the Generation of Personalized Itinerary Recommendations. *IEEE Access* **2024**, *12*, 112794–112809.
8. Huang, T.; Gong, Y. J.; Zhang, Y. H.; Zhan, Z. H.; Zhang, J.. Automatic planning of multiple itineraries: A niching genetic evolution approach. *IEEE Transactions on Intelligent Transportation Systems* **2019**, *21*(10), 4225–4240.
9. Halder, S.; Lim, K. H.; Chan, J.; Zhang, X.. Deep learning of dynamic POI generation and optimisation for itinerary recommendation. *ACM Transactions on Recommender Systems* **2025**, *3*(4), 1–29.
10. Zhang, J.; Ma, M.; Gao, X.; Chen, G.. Encoder-decoder based route generation model for flexible travel recommendation. *IEEE Transactions on Services Computing* **2024**, *17*(3), 905–920.
11. Wong, C. U. I.; Qi, S.. Tracking the evolution of a destination's image by text-mining online reviews-the case of Macau. *Tourism management perspectives* **2017**, *23*, 19–29.
12. Yixuan, W.; Jiayu, W.; Tian, C.. Multi-Scenario analysis of rooftop greening regulation on runoff effects based on adaptive Evaluation: A case study of Macau, China. *Ecological Indicators* **2024**, *163*, 111856.
13. Kim, S.; Park, J. H.; Lee, D. K.; Son, Y. H.; Yoon, H.; Kim, S.; Yun, H. J.. The impacts of weather on tourist satisfaction and revisit intention: A study of South Korean domestic tourism. *Asia Pacific Journal of Tourism Research* **2017**, *22*(9), 895–908.
14. Steiger, R.; Abegg, B.; Jänicke, L.. Rain, rain, go away, come again another day. Weather preferences of summer tourists in mountain environments. *Atmosphere* **2016**, *7*(5), 63.
15. Tenemaza, M.; Luján-Mora, S.; De Antonio, A.; Ramirez, J.. Improving itinerary recommendations for tourists through metaheuristic algorithms: an optimization proposal. *IEEE Access* **2020**, *8*, 79003–79023.
16. Choi, K. C.; Li, S.; Lam, C. T.; Wong, A.; Lei, P.; Ng, B., ; Siu, K. M.. Genetic algorithm for tourism route planning considering time constrains. *International Journal of Engineering Trends and Technology* **2022**, *70*(1), 170–178.
17. Chen, B. H.; Han, J.; Chen, S.; Yin, J. L.; Chen, Z.. Automatic itinerary planning using triple-agent deep reinforcement learning. *IEEE Transactions on Intelligent Transportation Systems* **2022**, *23*(10), 18864–18875.
18. Noguchi, T.; Fukada, K.; Bao, S.; Togawa, N.. Hybrid subQUBO Annealing With a Correction Process for Multi-Day Intermodal Trip Planning. *IEEE Access* **2025**, *13*, 19716–19727.
19. Lim, K. H.; Chan, J.; Karunasekera, S.; Leckie, C.. Personalized itinerary recommendation with queuing time awareness. In *Proceedings of the 40th international ACM SIGIR conference on research and development in information retrieval* **2017**, pp, 325–334.
20. Choachaicharoenkul, S.; Coit, D.; Wattanapongsakorn, N.. Multi-objective trip planning with solution ranking based on user preference and restaurant selection. *IEEE Access* **2022**, *10*, 10688–10705.
21. Pilato, G.; Persia, F.; Ge, M.; D'Auria, D.. Social sensing for personalized orienteering mediating the need for sociality and the risk of covid-19. *IEEE Transactions on Technology and Society* **2022**, *3*(4), 323–332.
22. Barua, B.; Kaiser, M. S.. Optimizing Travel Itineraries with AI Algorithms in a Microservices Architecture: Balancing Cost, Time, Preferences, and Sustainability. *arXiv preprint* **2024**, *2410*, 17943.

23. Volchek, K.; Ivanov, S.. ChatGPT as a travel itinerary planner. In ENTER e-Tourism Conference. *Springer Nature Switzerland* **2024**, pp, 365–370.
24. Seyfi, S.; Kim, M. J.; Nazifi, A.; Murdy, S.; Vo-Thanh, T.. Understanding tourist barriers and personality influences in embracing generative AI for travel planning and decision-making. *International Journal of Hospitality Management* **2025**, *126*, 104105.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.