A Two-Stage Hierarchical Optimization framework for Efficient Generative Urban Design

Xiao Hu
Department of Architecture, School of Design
Shanghai Jiao Tong University
Shanghai, 200240, China
huxiao1009@sjtu.edu.cn

Dayi Lai *
Department of Architecture, School of Design
Shanghai Jiao Tong University
Shanghai, 200240, China
dayi_lai@sjtu.edu.cn

Abstract

Performance-driven design (PDD) is frequently constrained by the high computational cost associated with optimizing a large number of design variables against multiple objectives with varying computational costs. This paper presents a novel two-stage hybrid optimization framework to address this challenge. The framework first utilizes a Genetic Algorithm (GA) to efficiently conduct a global search, identifying a candidate pool of designs that adhere to fundamental constraints such as zoning and economic viability. Subsequently, Bayesian Optimization (BO) is deployed to perform a targeted local search, refining these candidates by optimizing computationally intensive metrics, such as outdoor thermal comfort. A case study involving a residential block design demonstrates the effectiveness of the method. Our framework improved Outdoor Thermal Comfort Autonomy (OTCA) to 72. 5% from a baseline of 63. 8%, while simultaneously reducing the required computational time by 83.8% relative to a benchmark-accelerated GA. This work provides a scalable and efficient paradigm for complex architecture optimization, enabling a more effective balance between economic, regulatory, and environmental performance in the pursuit of sustainable urbanism.

1 Introduction

Conventional architectural design often relies on experiences rather than systematic performance evaluation, with assessments typically deferred to later stages, limiting adaptability and sustainability[11]. Research shows that early-stage performance optimization yields greater environmental and economic benefits at a lower cost [3], underscoring its critical role. Consequently, generative design (GD) and performance-driven design (PDD)[12] have emerged, integrating parametric modeling, performance evaluation, and heuristic optimization to improve early-stage outcomes[9].

However, PDD faces two major challenges. First, as design problems scale, the search space expands dramatically, making heuristic optimization inefficientespecially when performance simulations are computationally expensive[12, 5]. Surrogate models can reduce evaluation time[6] but require large datasets and do not address intrinsic search complexity. Second, current studies often ignore practical objective hierarchies. In real projects, mandatory requirements (e.g., regulations, solar access) take precedence over optional yet costly objectives (e.g., outdoor thermal comfort)[10]. Treating all objectives equally increases computational burden and reduces feasibility for practice.

This study proposes a two-stage optimization framework based on objective importance and computational complexity. Stage 1 addresses mandatory but simple objectives to broadly explore and filter

*Corresponding author. Email: dayi_lai@sjtu.edu.cn

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: UrbanAI: Harnessing Artificial Intelligence for Smart Cities.

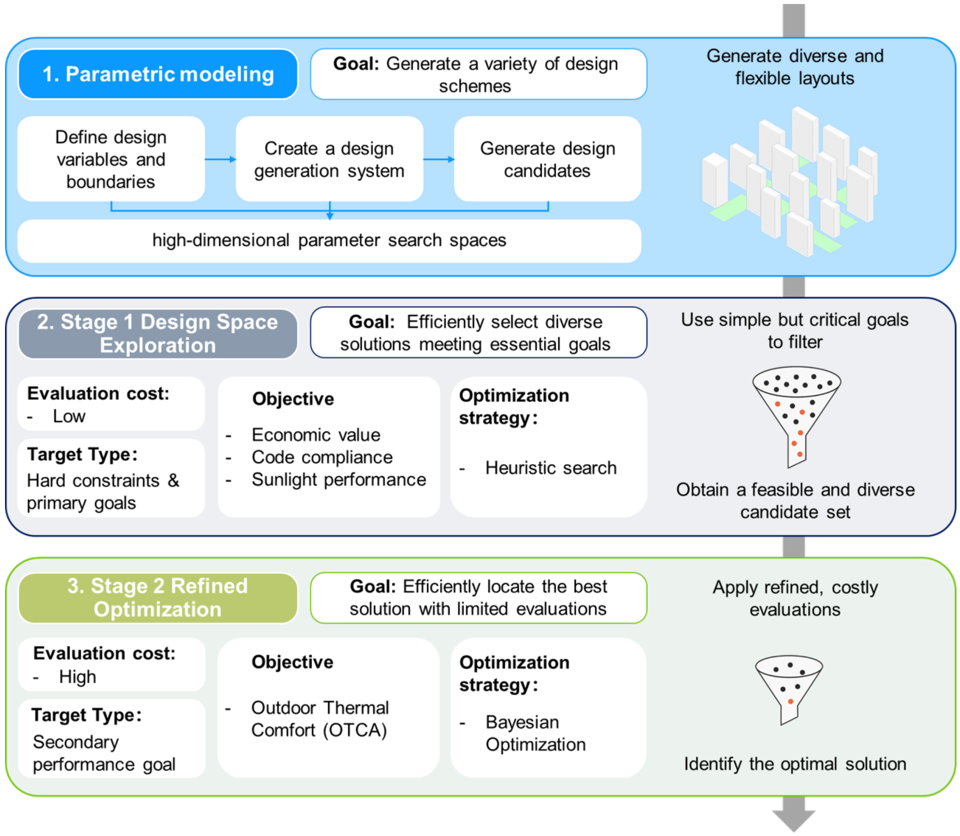

Figure 1: Overview of the two-stage optimization workflow.

feasible solutions; Stage 2 refines these using costly performance criteria. The framework combines hierarchical optimization with surrogate modeling and Bayesian optimization to reduce redundant evaluations, control computational cost, and enhance practical applicability in urban block design.

2 Methodology

2.1 Framework of this study

The proposed methodology adopts a two-stage optimization framework for residential block design (Figure 1). Built on a Rhino-Grasshopper parametric platform, the process begins with a Genetic Algorithm (GA) to broadly explore the search space and identify feasible solutions based on primary, low-cost objectives such as regulatory compliance, solar access, and value potential. These solutions form the basis for Stage II, where Bayesian Optimization (BO) targets a computationally intensive objectiveannual outdoor thermal comfort. This hierarchical approach balances global exploration with high-fidelity refinement, ensuring efficiency and practical applicability.

2.2 Stage I: design space exploration

The first stage aims to efficiently narrow a broad parameter space by prioritizing mandatory objectives, thereby reducing exploration of invalid designs. A genetic algorithm (GA) is used to optimize three strongly constrained objectives: regulatory compliance, solar access, and value potential. These objectives can be evaluated using closed-form or fast algorithms, allowing rapid assessment of large solution sets. All objectives are converted into a minimization form as:

$$f = 1 - S \tag{1}$$

where S denotes the normalized satisfaction score.

Stage 1 employs the Wallacei plugin on Grasshopper to perform multi-objective optimization using the Non-Dominated Sorting Genetic Algorithm II (NSGA-II)[2]. NSGA-II ranks solutions into Pareto fronts via fast non-dominated sorting, preserves diversity through crowding distance, and maintains quality with elitism. The optimization targets three mandatory objectives(Section A.2)—regulatory compliance (f_1), solar access (f_2), and value potential (f_3)—all formulated as minimization functions:

$$\min F(X) = \big[f_1, f_2, f_3\big] \tag{2}$$

The algorithm was configured with a population size of 50 and 400 generations (20,000 individuals), using a crossover probability of 0.8, mutation probability of 0.1, and distribution indices of 10 to enhance solution diversity.

Rather than directly adopting all Pareto solutions, candidate selection prioritized designs with balanced performance across all three objectives, avoiding solutions that excel in some metrics but underperform in others. Specifically, only layouts ranking within the top 15% for each objective were retained. This strategy ensures compliance with mandatory design requirements, reduces the computational burden of subsequent thermal performance optimization, and improves overall efficiency.

2.3 Stage II: refined optimization

Due to the computational cost of repeated CFD and radiation simulations, we employ *Bayesian Optimization (BO)* with a *Gaussian Process (GP)* surrogate to reduce evaluation demands. The GP predicts OTCA performance(Section A.3), while the Upper Confidence Bound (UCB) acquisition function guides sampling to balance exploration and exploitation.

BO proceeds iteratively under a 20-simulation budget: the GP is updated after each iteration, UCB identifies the next candidate, and CFD-based OTCA is computed. This active learning process converges toward the global optimum with significantly fewer simulations compared to exhaustive search, enabling efficient refinement of design solutions.

3 Result

3.1 Stage I: broad search by GA

The first-stage optimization was performed on a parametric residential block generation platform developed in Rhino–Grasshopper, using NSGA-II implemented via Wallacei. Three objectives were considered: f_1 (regulatory compliance), f_2 (solar access), and f_3 (value potential). The evolutionary process ran for 400 generations with 50 individuals per generation, yielding 20,000 solutions(Section A.4).

After filtering, the top 15% of high-performing solutions were selected(Section A.6), resulting in a total of 328 candidates. These solutions exhibit strong performance across all three objective functions. In the Stage II, optimization will proceed based on these 328 solutions to identify the best one in terms of outdoor thermal performance.

3.2 Stage II: finer optimization by BO

Based on the 328 high-performing solutions obtained from the optimization stage I, a second-stage *Bayesian Optimization (BO)* process was implemented to maximize outdoor thermal comfort (OTCA). Using the initial dataset(Section A.5), a BO process with an evaluation budget of 20 iterations was executed, employing the Upper Confidence Bound acquisition function to balance exploration and exploitation(Figure. 2). The initial samples exhibited OTCA values between 0.60 and 0.66 (mean $= 0.638$), which served as prior knowledge for the GPR model. During optimization, the algorithm progressively improved performance, achieving a substantial improvement at the 11$^{\text{th}}$ iteration with an OTCA of 0.725(Section A.7), as shown in Fig. 10. Visualization of the predictive landscape at the final iteration (Fig. 11) revealed strong clustering of evaluated points around high-value regions, confirming the effectiveness of the adaptive sampling strategy.

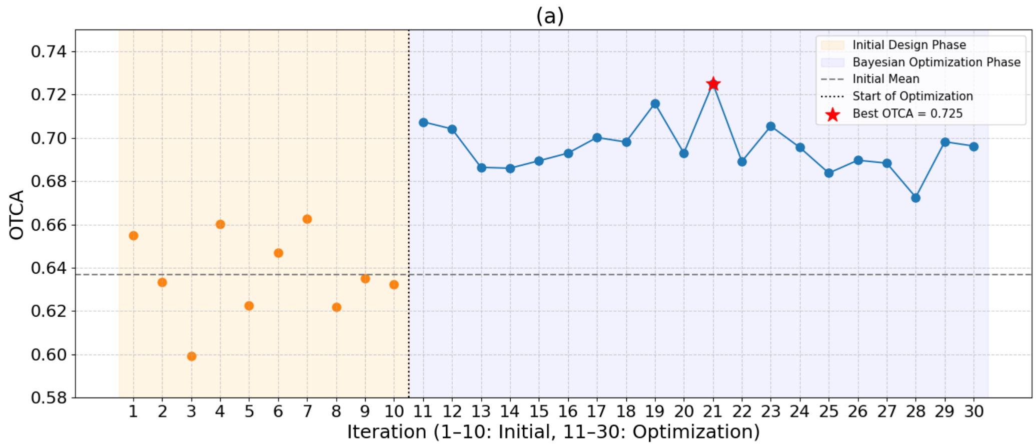

Figure 2: OTCA evolution with design phase highlighting.

4 Discussion

Despite achieving significant reductions in computational cost and enabling multi-objective optimization, the proposed two-stage framework has several limitations. First, its linear and unidirectional structure de-pends on the global search capability of the first-stage genetic algorithm; insufficient coverage of the optimal region at this stage may constrain the effectiveness of the subsequent Bayesian optimization. Second, the study isolates residential blocks from their surrounding urban context, neglecting potential interactions with neighboring morphology and environmental factors, which limits scalability in real-world applications. Third, the considered design variables and performance metrics are restricted, excluding factors such as building orientation and energy consumption. Future work should incorporate feedback mechanisms be-tween stages, broader contextual information, and additional performance indicators to support more comprehensive and context-sensitive optimization strategies.

5 Conclusion

This study proposes a two-stage hybrid optimization framework that integrates *Genetic Algorithms (GA)* and *Bayesian Optimization (BO)* to address the challenge of high-dimensional, multi-objective problems with costly evaluations. The framework decomposes the optimization into sequential tasks, combining GAs global exploration with BOs efficient sampling to reduce redundant evaluations while maintaining solution quality.

A residential block case study demonstrates its effectiveness: Stage 1 explored 20,000 design schemes and selected 328 high-performing candidates, while Stage 2 identified an optimal solution within 20 iterations, achieving a spatial OTCA of 72.5%, markedly higher than the 63.8% cluster-based average. Compared to accelerated single-stage GA, the proposed method achieved a substantial reduction in computational time while improving robustness.

The main contributions of this study are as follows: it proposes a generalized framework capable of handling computationally expensive multi-objective optimization problems, introduces a hierarchical strategy that reflects the logic of practical design iterations, and demonstrates the applicability of this approach in early-stage design by efficiently identifying high-performance regions within large and complex parameter spaces.

References

[1] Peter Bröde, Dusan Fiala, Krzysztof Błażejczyk, Ingvar Holmér, Gerd Jendritzky, Bernhard Kampmann, Birger Tinz, and George Havenith. Deriving the operational procedure for the universal thermal climate index (utci). *International journal of biometeorology*, 56(3):481–494, 2012.

[2] Kalyanmoy Deb, Amrit Pratap, Sameer Agarwal, and TAMT Meyarivan. A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE transactions on evolutionary computation*, 6(2):182–197, 2002.

[3] Yaxing Du, Cheuk Ming Mak, and Yantong Li. A multi-stage optimization of pedestrian level wind environment and thermal comfort with lift-up design in ideal urban canyons. *Sustainable Cities and Society*, 46:101424, 2019.

[4] Patrick Kastner and Timur Dogan. Eddy3d: A toolkit for decoupled outdoor thermal comfort simulations in urban areas. *Building and Environment*, 212:108639, 2022.

[5] Hangxin Li and Shengwei Wang. Coordinated optimal design of zero/low energy buildings and their energy systems based on multi-stage design optimization. *Energy*, 189:116202, 2019.

[6] Xueqing Li, Weisheng Lu, Ziyu Peng, Yi Zhang, and Jianxiang Huang. Generative design of walkable urban cool spots using a novel heuristic gan$\times$ gan approach. *Building and Environment*, 266:112027, 2024.

[7] Ministry of Housing and Urban-Rural Development of the People's Republic of China. Standard for Urban Residential Area Planning and Design, 2021. Issue: GB 50180-2018 Place: Beijing, China Published: Translated official edition.

[8] Negin Nazarian, Juan A Acero, and Leslie Norford. Outdoor thermal comfort autonomy: Performance metrics for climate-conscious urban design. *Building and environment*, 155:145–160, 2019.

[9] ShanShan Wang, Yun Kyu Yi, and NianXiong Liu. Multi-objective optimization (moo) for high-rise residential buildings layout centered on daylight, visual, and outdoor thermal metrics in china. *Building and Environment*, 205:108263, 2021.

[10] Zhaoji Wu, Zhe Wang, Jack CP Cheng, and Helen HL Kwok. A knowledge-informed optimization framework for performance-based generative design of sustainable buildings. *Applied Energy*, 367:123318, 2024.

[11] Ran Zhang, Xiaodong Xu, Peifan Zhai, Ke Liu, Lingyu Kong, and Wei Wang. Agile and integrated workflow proposal for optimising energy use, solar and wind energy potential, and structural stability of high-rise buildings in early design decisions. *Energy and Buildings*, 300:113692, 2023.

[12] Xinkai Zhang, Xiaoyu Wang, Sihong Du, Shuai Tian, Ariel Jia, Yu Ye, Naiping Gao, Xiaoming Kuang, and Xing Shi. A systematic review of urban form generation and optimization for performance-driven urban design. *Building and Environment*, 253:111269, 2024.

A Technical Appendices and Supplementary Material

A.1 Generation of residential block

As shown in Figure 3, the residential block is located within a 200 m Œ 200 m site, bounded by arterial roads to the north and east. A predefined number of buildings (i) and green spaces (j) are placed within building setbacks, each represented by a central coordinate $P(x, y)$. The proximity to arterial roads can reduce the value of the building market due to traffic noise.

Building geometry is defined by an apartment unit configuration T, with four typologies considered: t_1 (1 core with 2 units, 1C2U), t_2 (2C3U), t_3 (2C4U_H), and t_4 (2C4U_V). Building height h_i is determined by the number of floors l_i and floor height h_f. Typologies are assigned distinct market values. Green spaces are characterized by their area a_j and aspect ratio r_j, following design guidelines. Their configuration and distribution also influence the market value of adjacent buildings (Tab. 1).

Table 1: Design variables and their ranges

Category	**Design variables**	**Symbol**	**Description and Value range**
Buildings	Number of buildings	i	Total number of buildings in the site, $i \in [8, 15]$.
	Building position	$P_i^B = (x_i^B, y_i^B)$	Coordinates of the centroid of building i, $x_i^B, y_i^B \in [-100, 100]$.
	Building height	$h_i = l_i \cdot h_f$	Total height of building i, where l_i is the number of floors and h_f is the floor height, $l_i \in [12, 26],\ h_f \in \{2.7, 3.0, 3.3\}$.
	Unit configuration	$t_i^B \in T$	Residential typology configuration of building i, $T = \{t_1, t_2, t_3, t_4\}$.
Green spaces	Number of green spaces	j	Total number of green spaces in the site, $j \in [2, 8]$.
	Green space position	$P_j^G = (x_j^G, y_j^G)$	Coordinates of the centroid of green space j, $x_j^G, y_j^G \in [-100, 100]$.
	Green space form	$F_j^G = (a_j, r_j)$	Shape parameters: a_j is the area; r_j is the aspect ratio, $a_j \in [400, 2500],\ 1.0 \le r_j \le 3.0$.

A.2 Design objective of stage I

Regulatory Compliance

Compliance is evaluated based on Floor Area Ratio (FAR), Building Density (BD), and Green Space Ratio (GR), as required by *Standard for Urban Residential Area Planning and Design*[7] . Acceptable ranges are set as $[3.0, 3.5]$ for FAR, $[0.17, 0.20]$ for BD, and $[0.35, 0.40]$ for GR. Each indicator is scored using a piecewise linear satisfaction function, and the compliance score is the average of the three:

$$S(X) = \begin{cases} 1, & X \in [X_{\text{low}}, X_{\text{high}}] \\ \max\left(0, 1 - \dfrac{X_{\text{low}} - X}{X_{\text{high}} - X_{\text{low}}}\right), & X < X_{\text{low}} \\ \max\left(0, 1 - \dfrac{X - X_{\text{high}}}{X_{\text{high}} - X_{\text{low}}}\right), & X > X_{\text{high}} \end{cases} \tag{3}$$

$$S_{\text{regulatory compliance}} = \frac{S_{\text{FAR}} + S_{\text{BD}} + S_{\text{GR}}}{3}, \qquad f_1 = 1 - S_{\text{regulatory compliance}} \tag{4}$$

Solar Access

Solar access is evaluated through two metrics: P_{win} represents the proportion of south-facing window area that receives at least two hours of direct sunlight on the winter solstice, while P_{site} denotes the proportion of ground area within the site that, between 8:00 and 16:00 on the winter solstice, receives at least two hours of direct sunlight. Based on these definitions, the solar access satisfaction index is formulated as:

$$S_{\text{solar access}} = \frac{P_{\text{win}} + P_{\text{site}}}{2}, \qquad f_2 = 1 - S_{\text{solar access}} \tag{5}$$

Value Potential

Economic value is estimated by aggregating adjusted building prices. Adjustments are based on unit type, greenery conditions, and noise proximity to arterial roads (Table 2).

Table 2: Unit price adjustment factors for residential value potential

Factor	Category	Adjustment coefficient
Unit type	one core with two units (1C2U)	1.2
	two cores with three units (2C3U)	1.1
	two cores with four units in a horizontal arrangement (2C4U_H)	1.0
	two cores with four units in a vertical arrangement (2C4U_V)	0.9
Landscape	Greenery on both north and south sides	1.2
	Greenery on only one of the north/south sides	1.1
	No greenery on north or south sides	1.0
Noise proximity	Distance to arterial road < 30 m	0.9
	Distance 30–60 m	0.95
	Distance > 60 m	1.0

The satisfaction score is defined as:

$$S_{\text{value potential}} = \begin{cases} \frac{V_{\text{act}}}{V_{\text{ref}}}, & \text{if } V_{\text{act}} \leq V_{\text{ref}} \\ 1, & \text{if } V_{\text{act}} > V_{\text{ref}} \end{cases} \tag{6}$$

$$f_3 = 1 - S_{\text{value potential}} \tag{7}$$

A.3 Performance simulation: annual outdoor thermal comfort

The second stage focuses on outdoor thermal comfort, which varies spatially and temporally with solar radiation, air temperature, and wind. To capture these dynamics, we adopt *Outdoor Thermal Comfort Autonomy (OTCA)* [8], a metric that quantifies the proportion of comfortable hours annually. OTCA is based on the *Universal Thermal Climate Index (UTCI)* [1], computed from air temperature, mean radiant temperature (T_{mrt}), wind speed, and relative humidity. While air temperature and humidity are derived from EPW weather data, T_{mrt} and wind fields are geometry-dependent and obtained via Eddy3D [4], which integrates OpenFOAM for CFD and Radiance for radiation analysis.

Wind simulations use steady-state RANS with the $k\omega$ SST model for eight wind directions within a 600 m domain. Annual wind conditions are reconstructed by interpolating these results with hourly EPW data. T_{mrt} is calculated using a Radiance-based two-phase daylight simulation, considering both longwave and shortwave components. These workflows generate high-resolution UTCI maps, from which OTCA is computed:

$$OTCA = \frac{\sum_{k=1}^{N} \sum_{h_r=h_i}^{h_f} TC_{k,h_r}}{N \times n} \tag{8}$$

where

$$TC_{k,h_r} = \begin{cases} 1, & \text{if } UTCI \in [9^\circ\text{C}, 26^\circ\text{C}] \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

The spatial OTCA is defined as the proportion of outdoor areas remaining within the thermal comfort range during occupied hours. In this studies a 45% threshold is adopted.

A.4 Optimization process of genetic optimization

As shown in Figure. 4(a), the optimization followed a typical evolutionary pattern. During the initial ~ 50 generations, all objectives improved rapidly, indicating efficient global exploration. Subsequently, improvements slowed as the algorithm shifted to local exploitation. Minor fluctuations occurred due to stochastic operators such as mutation, which help avoid local optima. Overall, f_1 decreased from approximately 0.62 to 0.35, f_2 from 0.35 to 0.25, and f_3 (expressed as $1 - V_{\text{act}}/V_{\text{ref}}$) from 0.00 to -0.35. Figure. 4(b) shows that f_1 and f_3 formed narrow unimodal peaks, while f_2 remained more dispersed, reflecting stronger trade-offs with other objectives.

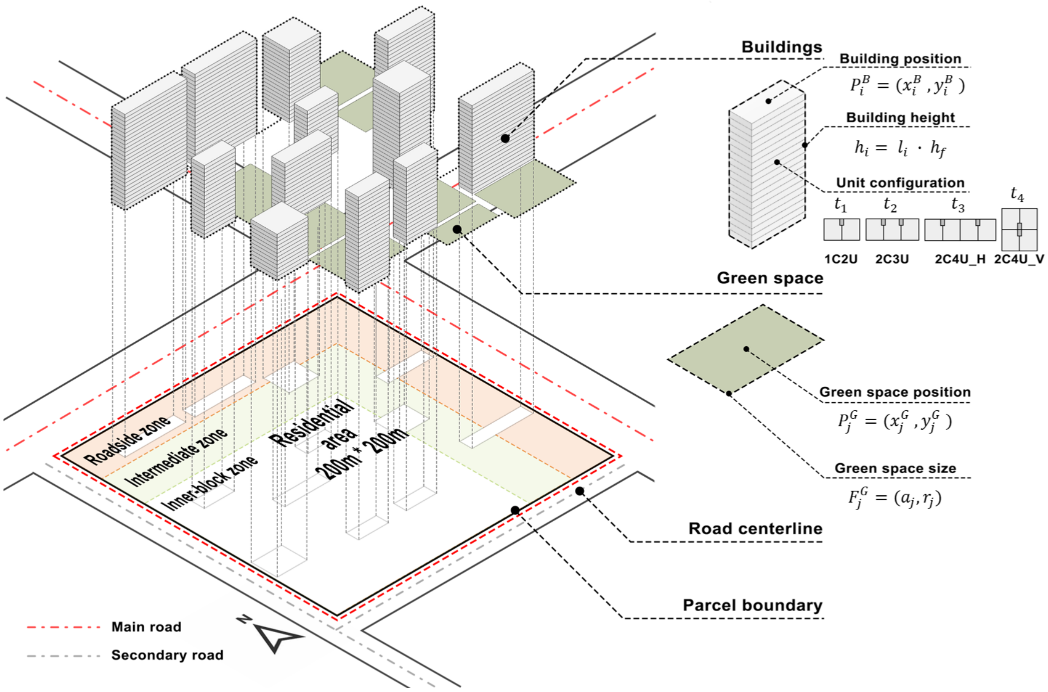

Figure 3: Generation system and design variables.

A.5 Clustering based on key features

To initialize BO, feasible solutions from Stage 1 are clustered via K-Means using key geometric parameters (Tab. 3). From each cluster, centroid and edge points are selected for initial simulations, ensuring diversity. Feature importance analysis using Random Forest reduces dimensionality, retaining the top five influential variables for GP modeling.

Table 3: Key Feature Parameters for Clustering

Feature	Name	Description	Calculation
BC	Building count	Total number of discrete building entities within the site	n
SFA	South-facing façade area	Total façade area of south	$\sum_{i=1}^{n} h_i w_i$
CBH	Cumulative building height	Sum of heights of all buildings in the site	$\sum_{i=1}^{n} h_i$
MAR	Mean aspect ratio	Average ratio of plan length to plan width across buildings	$\frac{1}{n}\sum_{i=1}^{n} \frac{w_i}{d_i}$
BSF	Building shape factor	Ratio of total external building surface area to building volume	$\frac{\sum_{i=1}^{n} A_i^{\text{surf}}}{\sum_{i=1}^{n} w_i d_i h_i}$
MPAR	Mean perimeterarea ratio	Average ratio of building footprint perimeter to footprint area	$\frac{1}{n}\sum_{i=1}^{n} \frac{2(w_i+d_i)}{w_i d_i}$
VHV	Vertical height variation	Standard deviation of building heights	$\sigma(h_1, h_2, \ldots, h_n)$
OSR	Open space ratio	Proportion of open ground area (non-built) to total floor area	$\frac{A_{\text{site}} - \sum_{i=1}^{n} w_i d_i}{\sum_{i=1}^{n} w_i d_i l_i}$
PP	Plan porosity	Ratio of unbuilt volume to total building envelope volume	$\frac{h_{\max} A_{\text{site}} - \sum_{i=1}^{n} w_i d_i h_i}{h_{\max} A_{\text{site}}}$

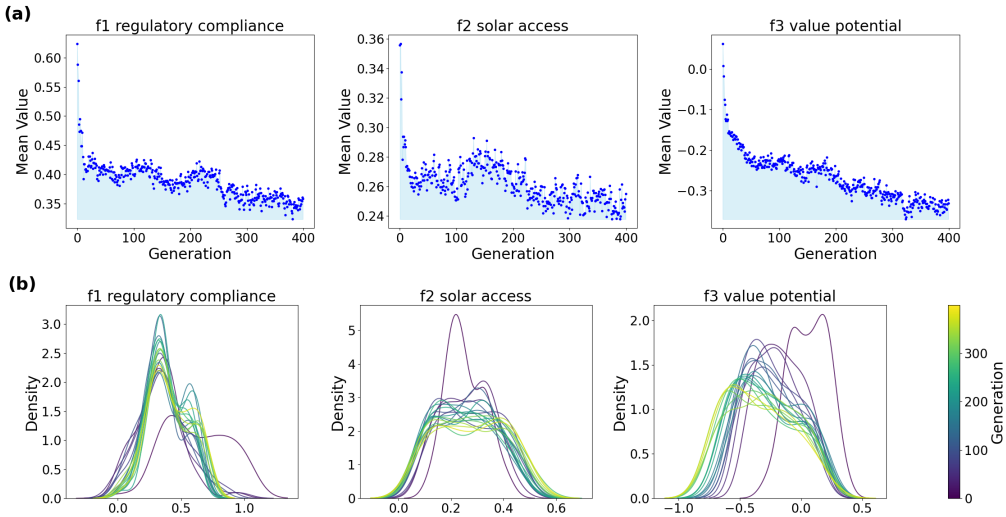

Figure 4: Visualization of optimization process of GA.

To efficiently represent the solution space, K-means clustering was applied to nine morphological features, with the optimal cluster number ($k = 5$) determined using the Elbow Method. Principal Component Analysis (PCA) was then employed to reduce dimensionality and visualize clusters, from which two representative points per cluster (one centroid and one edge point) were selected, yielding ten candidates for detailed annual OTCA simulations. These simulations provided initial data for constructing a *Gaussian Process Regression (GPR)* surrogate model and enabled feature importance analysis via SHAP values(Figure. 6), which indicated *Mean Aspect Ratio*, *Vertical Height Variation*, *Plan Porosity*, and *South-Facing Façade Area* as the most influential features.

A.6 Filtering strategy

To reduce complexity for the second stage, a filtering strategy was applied. Solutions ranking in the top 15% for all three objectives were retained, resulting in 328 high-quality candidates ($\approx 1.64\%$ of the total). These points, highlighted in Figure. 5, cluster in the most desirable region of the Pareto front, corresponding to the knee area that offers balanced trade-offs among the objectives. Correlation analysis reveals that f_1 and f_2 exhibit a moderate positive correlation ($\rho = 0.655$), f_1 and f_3 a moderate negative correlation ($\rho = -0.711$), and f_2 and f_3 a very strong negative correlation ($\rho = -0.859$), indicating a pronounced trade-off between solar access and value potential.

This filtering strategy ensures that the second-stage optimization operates within a high-quality solution space characterized by balanced performance across all objectives, thereby accelerating convergence and establishing a robust foundation for further refinement.

A.7 Optimal solution

Ultimately, the BO approach identified an optimal residential block layout (Figure. 7) delivering a spatial OTCA of 72.5%, demonstrating its capability to efficiently explore a complex design space and significantly enhance outdoor thermal comfort within a limited computational budget.

A.8 Validation of the Two-Stage Optimization Framework

To verify whether the proposed two-stage optimization method can identify the optimal solution within a limited number of iterations, we conducted a validation experiment. Considering the high computational cost of full-year OTCA evaluations, it was infeasible to compute OTCA for all candidate solutions. Instead, we employed 328 candidate solutions obtained from the first stage and replaced OTCA with a computationally less expensive metricthe site-averaged UTCI between 12:00 and 14:00 on July 21. The optimization objective was set to minimize this UTCI. To rigorously assess the performance of Bayesian Optimization (BO), we precomputed

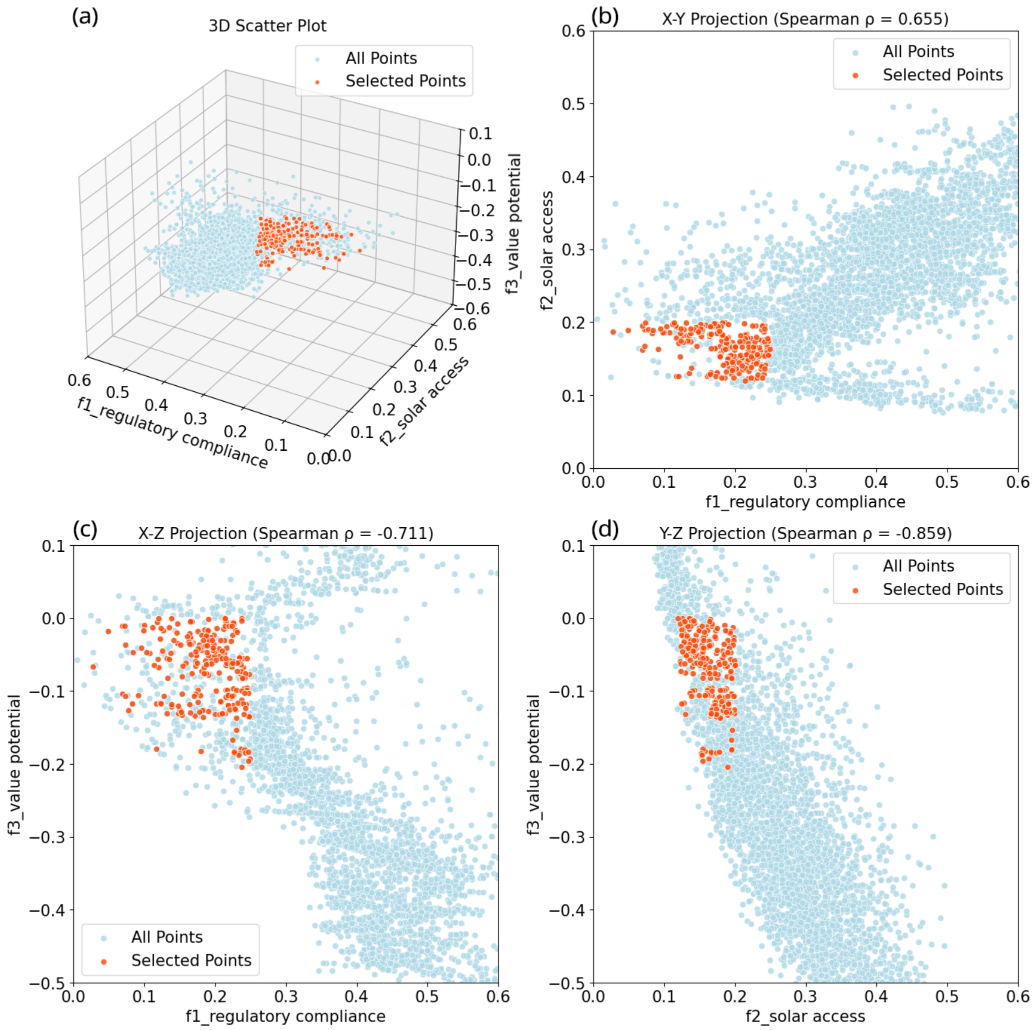

Figure 5: Space distribution of solutions generated on the optimization process.

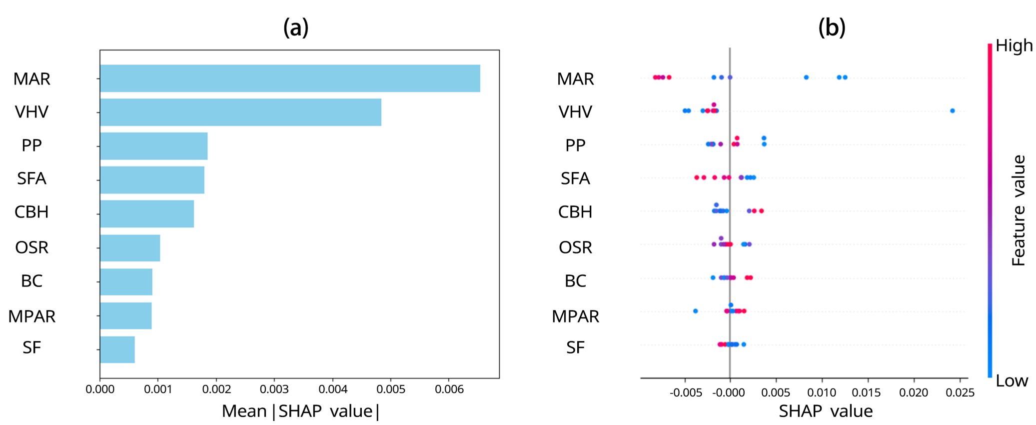

Figure 6: Feature Importance on OTCA via SHAP Value Analysis.

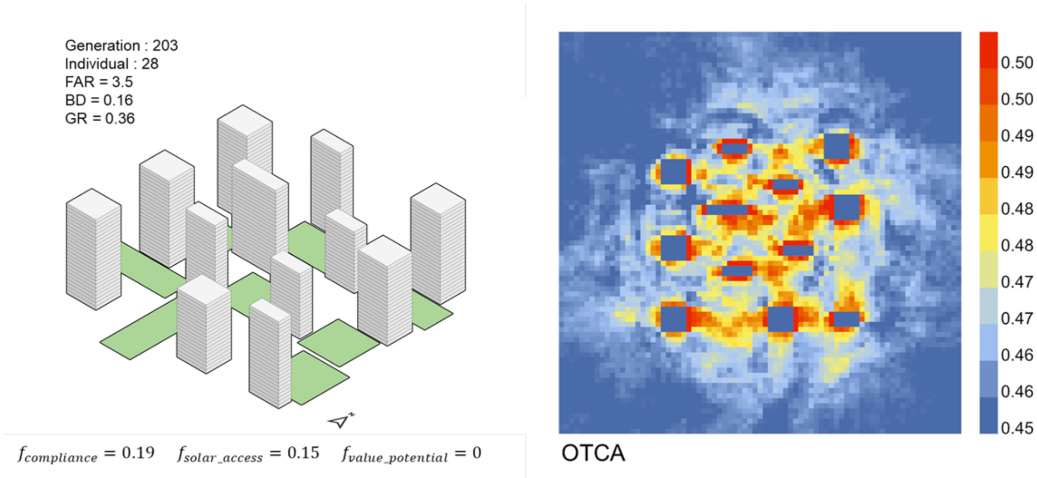

Figure 7: Optimal residential block design and OTCA results based on Bayesian optimization.

UTCI values for all 328 solutions, enabling direct comparison between the BO-identified optimum and the true global minimum.

A.9 Progression of Bayesian Optimization and Sampling Behavior

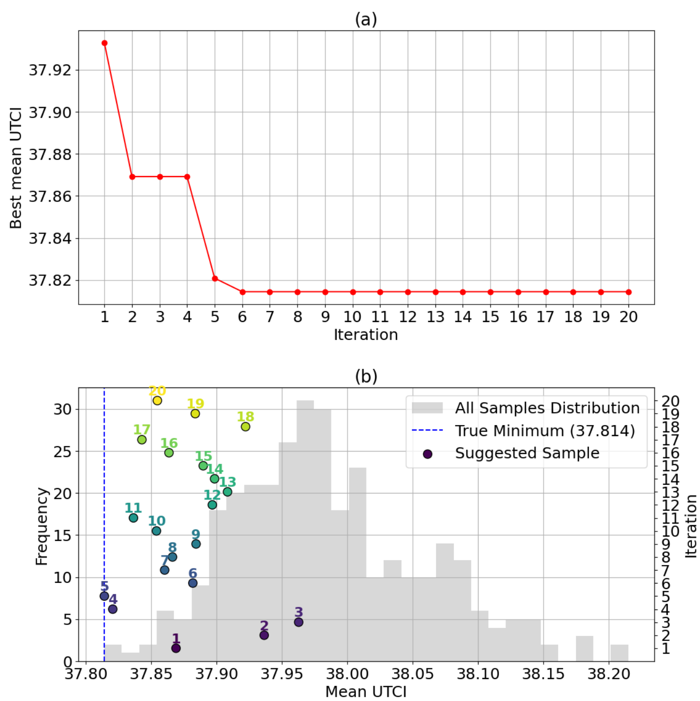

Figure 8: Progression of Bayesian optimization for UTCI minimization

Figure. 8(a) presents the evolution of the best-found value during the BO process. A sharp decrease occurred between iterations 1 and 5, after which the best value stabilized, reaching 37.81 °C at iteration 5. Figure. 8(b) shows the sampling distribution across 20 iterations alongside the full distribution of UTCI values for all 328 solutions. Notably, the global minimum was identified at the fifth iteration, demonstrating that BO successfully located the optimal solution with minimal evaluations.

All experiments were conducted on a workstation equipped with an Intel Core i5-13600K processor (14 cores: 6 performance-cores and 8 efficiency-cores, 24 MB cache, up to 5.10 GHz), 32 GB DDR5 memory, running Windows 11 Professional.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introduction accurately summarize the contributions, including the proposed method, experimental setup, and main findings.

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: We explicitly discuss several limitations, including (1) dependency on the first-stage genetic algorithms global search capability, (2) neglect of surrounding urban context, and (3) restricted set of design variables and performance metrics. We also outline potential future directions to address these issues (see Section 4).

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [NA]

   Justification: The paper does not include formal theoretical results; it focuses on algorithmic development and empirical evaluation.

   Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: All computational and parameter details are described in the corresponding sections of the text.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [Yes]

   Justification: The author will provide all relevant materials and data necessary to fully replicate the experimental procedures.

   Guidelines:

   - The answer NA means that paper does not include experiments requiring code.
   - Please see the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.

- While we encourage the release of code and data, we understand that this might not be possible, so No is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

   Answer: [NA]

   Justification: Although this paper does not utilize training and testing sets, all experimental settings discussed herein have been reported.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [NA]

   Justification: The core of this paper is to propose a two-stage optimization framework, which does not involve this part of the content.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).
   - It should be clear whether the error bar is the standard deviation or the standard error of the mean.
   - It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
   - For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
   - If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

Justification: The configuration of the computer used is reported in the text.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
- The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
- The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: The research presented in this paper was conducted in full accordance with the NeurIPS Code of Ethics.

   Guidelines:

   - The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [No]

    Justification: This work is primarily theoretical and methodological, and does not have immediate societal applications.

    Guidelines:

    - The answer NA means that there is no societal impact of the work performed.
    - If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
    - The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
    - The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
    - If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

    Answer: [NA]

Justification: The research presented in this paper does not involve the release of pretrained language models, generative models, or large-scale scraped datasets that could pose a high risk of misuse.

Guidelines:

- The answer NA means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [NA]

Justification: This work does not make use of any existing datasets, pretrained models, or external code.

Guidelines:

- The answer NA means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

Answer: [NA]

Justification: This work does not introduce new datasets, code, or models.

Guidelines:

- The answer NA means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

Answer:[NA]

Justification: This work does not involve crowdsourcing or research with human subjects.

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.

- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [NA]

    Justification: This work does not involve crowdsourcing or research with human subjects.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigorousness, or originality of the research, declaration is not required.

    Answer:[NA]

    Justification: The core method development in this research does not involve large language models (LLMs). Any usage of LLMs was limited to minor writing or editing assistance.

    Guidelines:

    - The answer NA means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy (`https://neurips.cc/Conferences/2025/LLM`) for what should or should not be described.