Journal of Architectural and Planning Research
37:4 (Winter, 2023) 28

TYPOLOGY-CONDITIONED PERFORMANCE, ROBUSTNESS, AND SELECTION LOGIC OF GENERATIVE HOTEL LAYOUT ALGORITHMS: A DETAILED COMPARATIVE STUDY ACROSS 50 ADAPTIVE-REUSE FLOOR PLANS

L. Faggion
R. Furlan

Comparative studies of generative space-planning methods are most useful when they move beyond an overall ranking and explain how performance changes across recurrent spatial conditions. This article presents a secondary, typology-conditioned reanalysis of a 50-case benchmark for adaptive hotel planning in existing buildings. The benchmark evaluates three algorithmic strategies for room allocation within fixed floor boundaries: a general-rules tessellation method, a self-organization method, and a mixed corridor-attraction method. The underlying benchmark used a common brief with a target room area of $25\,\mathrm{m}^2$ *and corridor width of* $1.6\,\mathrm{m}$*, and it assessed outcomes across I-, L-, C-, T-, and H-type plan families. The present study reanalyzes the published case-level reviewed room counts for all 50 benchmark cases and treats reviewed capacity as the primary outcome because it reflects post-validation usable yield under a common proof-of-concept brief. The article augments whole-sample comparisons with typology-specific descriptive statistics, repeated-measures nonparametric tests, pairwise dominance analysis, winner-share analysis, a leave-one-case-out sensitivity check, and a detailed appendix of case-level rankings. Across the full benchmark, the general-rules method produced the highest reviewed total (1588 rooms), followed by the mixed method (1468) and self-organization (1307). The overall difference was statistically significant (Friedman* $\chi^2(2) = 50.51$*,* $p < 0.001$*, Kendall's* $W = 0.505$*), the same hierarchy persisted in each typological subset, and the pooled ordering was preserved in all 50 leave-one-case-out re-estimations. However, the margin of superiority varied substantially. The mixed method remained locally competitive in several I-, C-, L-, and T-type cases, whereas the general-rules method was especially dominant in H-type plans. These findings show that typology-conditioned benchmarking yields a more operational basis for algorithm selection than pooled comparison alone, while also indicating that the evidence should be interpreted within the reviewed-capacity criterion used by the source benchmark. For early-stage adaptive hotel planning under that criterion, tessellation should be treated as the default baseline, the mixed method as the principal flexible alternative, and self-organization as an exploratory supplement rather than a primary capacity-maximizing routine.*

https://doi.org/10.66033/japr2022-403

INTRODUCTION

Adaptive reuse projects increasingly depend on computational workflows that can generate, test, and compare large numbers of spatial alternatives in a short time. In hotel conversion, this need is especially acute. A single floor plate may support materially different room yields depending on corridor structure, subdivision logic, and how the layout algorithm interprets corners, branches, and perimeter access. Because room count is tightly connected to revenue potential, generative design methods are especially relevant at the feasibility stage, where project teams must compare options before detailed design begins (Caetano et al., 2020; Rahbar et al., 2019, 2022).

Recent work has demonstrated the relevance of parametric, rule-based, agent-based, and hybrid methods in architectural layout generation, but direct comparative evidence remains limited (Chen, 2012; Mostafavi et al., 2022; Stieler et al., 2022). The hotel-layout benchmark presented by Cudzik and Kruk (2022) is therefore particularly valuable. It compares three distinct computational strategies within one controlled design problem: a general-rules tessellation method, a self-organization method, and a mixed corridor-attraction method. The benchmark uses 50 floor boundaries drawn from five recurrent typological families and evaluates both quantitative outputs and qualitative preferences. Within that benchmark, the general-rules method achieved the strongest overall performance, while the mixed method occupied an intermediate position and self-organization trailed in aggregate terms (Cudzik & Kruk, 2022).

That whole-sample finding is important, but it does not fully resolve the practical decision problem. Architects and analysts do not select algorithms for average cases; they select algorithms for particular floor boundaries with identifiable geometric characteristics. A pooled ranking may therefore understate two kinds of variation. First, the strength of an algorithm may depend on the typological structure of the plan. Second, methods that appear weaker in aggregate may nonetheless become locally attractive in specific plan families. From the standpoint of deployment, these issues matter as much as the overall ranking.

This article develops a more detailed interpretation of the benchmark by focusing on typology-conditioned performance and selection logic. Instead of asking only which algorithm performs best overall, it asks a broader set of decision-oriented questions. Does the superiority of the general-rules method remain stable within each recurrent floor-plan family? How close does the mixed method come to the best-performing solution under different typological conditions? How often do local reversals occur, and what do they imply for practical workflow design? Which method is most robust when one considers both mean yield and case-level dominance?

The contribution of the article is deliberately bounded. It does not introduce a new generator, a new benchmark, or a new participant-level preference dataset. Instead, it extracts more decision-relevant evidence from an existing controlled comparison. First, it reframes comparative evaluation from a pooled benchmark exercise into a morphology-aware selection problem. Second, it expands the analytical depth of the existing benchmark by adding typology-specific descriptive and inferential statistics, case-level winner shares, pairwise dominance measures, and a leave-one-case-out sensitivity check. Third, it provides a more transparent basis for design-method selection by distinguishing default, flexible, and exploratory roles for the three algorithmic families while keeping those roles explicitly tied to the reviewed-capacity criterion. Fourth, it offers a reproducible secondary-analysis structure for future comparative studies in architectural computing, where the practical value of an algorithm depends not only on global averages but also on how performance behaves across recurring classes of design situations.

LITERATURE AND CONCEPTUAL FRAMING

Generative design in architectural space planning

Generative design in architecture encompasses a wide range of computational paradigms, including parametric modeling, shape grammars, rule-based systems, agent-based simulations, optimization-driven workflows, and data-intensive learning systems (Caetano et al., 2020). These approaches differ not only in implementation but also in epistemic orientation. Rule-based systems favor explicitness, control, and interpretability. Agent-based and self-organizing systems emphasize local interaction, emergence, and exploratory variation. Hybrid systems aim to combine the advantages of both by pairing global guidance with local adaptability (Rahbar et al., 2022; Stieler et al., 2022).

This diversity has become especially visible in floor-plan generation. Ligler (2021) showed how visual computation can support hotel-related plan generation through strongly articulated spatial logics. Rahbar et al. (2019) and Rahbar et al. (2022) explored data-driven and hybrid approaches to synthetic space allocation. Mostafavi et al. (2022) emphasized that floor-plan generation should not be judged only by raw numerical performance, because representation, topology, and designer interpretation also influence quality. Meanwhile, Mukkavaara and Sandberg (2020) and Reitberger (2022) argued that generative tools are most useful when they support early-phase decision-making rather than simply automate one isolated step.

Within this field, hotel adaptation is a particularly informative test domain. It combines repetitive units, well-defined minimum requirements, and strong economic pressure for efficient spatial use. At the same time, adaptive reuse imposes irregular and pre-existing boundaries that are often less forgiving than new-build hotel planning. This makes hotel conversion a strong empirical setting for testing how different algorithmic logics respond to the same constraint structure.

Why typology-conditioned interpretation matters

Comparative algorithm studies often stop at aggregate performance totals. That is understandable, but it can be misleading when the benchmark includes heterogeneous cases. If plan typology influences the effectiveness of an algorithm, then a single global ranking is analytically incomplete. In practice, a method should be judged not only by how much it produces on average, but also by where it is reliable, where it is merely competitive, and where it should be used only for exploratory purposes.

The underlying benchmark was explicitly designed around five recurrent plan families—I, L, C, T, and H—and it reported case-level reviewed room counts for each algorithm across all 50 cases (Cudzik & Kruk, 2022). This structure makes it possible to ask questions that are more actionable than the global comparison alone. Typology-conditioned evaluation turns the benchmark into a design-support instrument. It allows one to infer whether an algorithm can be deployed as a baseline, whether it should be reserved for certain geometric conditions, and whether its main value lies in exploratory rather than production-oriented use.

MATERIALS AND METHODS

Benchmark basis and analytical scope

The benchmark analyzed in this article consists of 50 floor-boundary cases used for adaptive hotel planning. The source study implemented three algorithmic strategies in Rhinoceros 7, Grasshopper, and C# components, and it evaluated them under a common brief defined by a corridor width of 160 cm and target room area of 25 m^2. The benchmark included I-, L-, C-, T-, and H-type plans and also incorporated a supplementary survey involving 86 architecture students who rated selected outputs on a 10-point scale (Cudzik & Kruk, 2022). The present study is therefore a secondary analysis of published case-level benchmark data rather than a new experimental campaign. The source study deliberately simplified the planning problem to room distribution and horizontal circulation, leaving vertical circulation, structural constraints, windows, and other contextual components outside the proof-of-concept evaluation frame (Cudzik & Kruk, 2022).

The present article uses the case-level *reviewed* room counts reported for all 50 cases as the primary analytical dataset. This choice follows the logic of the source study itself. In the benchmark, the reviewed count represents the most credible post-validation capacity measure because it incorporates expert revision of proof-of-concept outputs and excludes clearly dysfunctional behavior from substantive interpretation (Cudzik & Kruk, 2022). Because only benchmark-wide survey means rather than the full participant-level survey matrix are available in the present dataset, the qualitative findings are discussed contextually and used only to triangulate the quantitative interpretation. Claims in the present article are therefore limited to comparative performance within this benchmark and under the reviewed-capacity criterion used by the source study.

Case structure and outcome variables

Each case i belongs to one typology $t \in \{I, L, C, T, H\}$ and was evaluated by all three algorithms $a \in \{G, M, S\}$, where G denotes general rules, M denotes the mixed corridor-attraction method, and S denotes self-organization. Let $y_{i,a}$ denote the reviewed room count for case i under algorithm a.

The analysis uses six complementary performance views:

1. *Reviewed room count* $y_{i,a}$ as the primary case-level performance metric;
2. *Typology-conditioned central tendency*, reported through mean, median, and standard deviation within each plan family;
3. *Winner share*, defined as the proportion of cases within a typology for which a method is best or tied-best;
4. *Pairwise dominance*, defined as the share of cases in which one algorithm strictly exceeds another;
5. *Relative typology gain*, reported as

$$G_{a,b,t} = \frac{\bar{y}_{t,a} - \bar{y}_{t,b}}{\bar{y}_{t,b}} \times 100, \tag{1}$$

   where $\bar{y}_{t,a}$ is the mean reviewed room count of algorithm a in typology t;
6. *Leave-one-case-out stability*, used to test whether the pooled ranking depends on any single influential floor boundary.

Because each floor boundary was processed by all three algorithms, the data have a repeated-measures structure. That structure is central to the statistical design.

Statistical analysis

Since room counts vary substantially across cases and include several extreme values, rank-based repeated-measures tests were preferred to Gaussian modeling. The following procedure was used:

1. A global Friedman test was applied across all 50 cases.
2. Typology-specific Friedman tests were then computed separately for I, L, C, T, and H cases.
3. Pairwise Wilcoxon signed-rank tests were used for the full benchmark comparison.
4. Bonferroni-adjusted p-values were reported for the three pairwise tests.
5. Kendall's W was used as an omnibus effect-size index for the Friedman tests.
6. A leave-one-case-out sensitivity analysis was run across all 50 cases to verify that the pooled ordering was not produced by a single influential boundary.

For three repeated measures and N cases, Kendall's W was computed as

$$W = \frac{\chi_F^2}{N(k-1)}, \tag{2}$$

where χ_F^2 is the Friedman statistic and $k = 3$ is the number of algorithms.

In addition to significance testing, the article emphasizes descriptive decision metrics. This is intentional. In architectural computing, a statistically significant result is not automatically the most practically useful result. Designers also need to know how often a method wins, how often it nearly wins, and in which plan families a second-ranked approach remains close enough to justify exploratory use.

Reproducibility note

The analytical dataset used here consists of the case-level reviewed room counts reported in the benchmark table of Cudzik and Kruk (2022). These values are reproduced in Appendix . All descriptive statistics, dominance rates, inferential tests, and figures in the present manuscript were recomputed directly from that matrix, and all reported values were cross-checked against the case totals before the manuscript was finalized. No imputation or case-wise transformation was applied. The article therefore offers a transparent and reproducible reanalysis of the benchmark at the level of reported case outcomes rather than an opaque narrative summary.

RESULTS

Whole-benchmark performance

Across the full benchmark, the general-rules method produced 1588 reviewed rooms, the mixed method produced 1468, and self-organization produced 1307. On a per-case basis, the average reviewed room count was 31.76 for general rules, 29.36 for the mixed method, and 26.14 for self-organization. Median reviewed room counts showed the same ordering: 25.5, 24.0, and 19.0, respectively. The boxplot in Figure 1 visualizes this distributional hierarchy and confirms that the ordering is not created by one narrow part of the range alone.

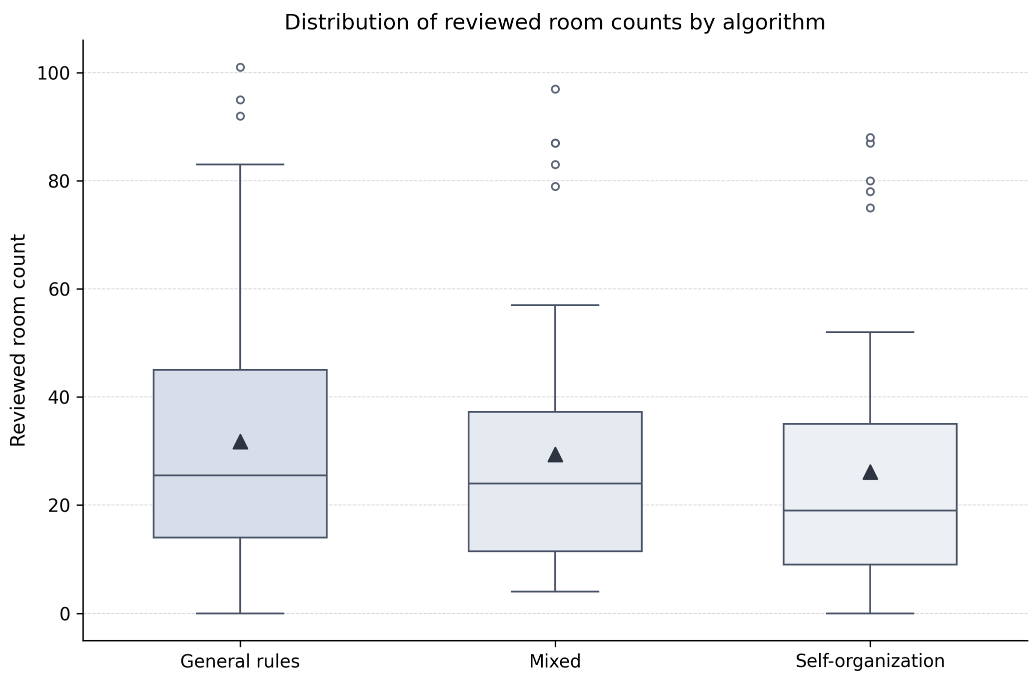

Figure 1: Distribution of reviewed room counts by algorithm across all 50 cases. Triangles indicate means.

The overall repeated-measures comparison was statistically significant (Friedman $\chi^2(2) = 50.51$, $p < 0.001$), with Kendall's $W = 0.505$, indicating a moderate-to-strong omnibus effect. Pairwise Wilcoxon signed-rank tests showed that each algorithm differed significantly from the others after Bonferroni correction. General rules exceeded self-organization most strongly, while the mixed method also outperformed self-organization with a smaller but still statistically robust margin. Table 1 summarizes these results.

Table 1: Overall repeated-measures comparison across the 50-case benchmark.

Test	Statistic	Raw p	Adj. p	Interpretation
Friedman test	$\chi^2(2) = 50.51$	< 0.001	—	Strong global difference
Gen. rules vs. self-org.	$W = 52.5$	< 0.001	< 0.001	General rules superior
General rules vs. mixed	$W = 101.0$	< 0.001	< 0.001	General rules superior
Mixed vs. self-org.	$W = 221.0$	< 0.001	0.002	Mixed superior

The omnibus effect size was Kendall's $W = 0.505$. Bonferroni correction was applied to the three pairwise Wilcoxon tests.

A leave-one-case-out sensitivity check further strengthened this interpretation. In all 50 re-estimations, the pooled ordering remained General rules > Mixed > Self-organization, and the Friedman statistic remained strongly significant, ranging from $\chi^2 = 48.66$ to $\chi^2 = 54.67$ with all $p < 10^{-10}$. The benchmark-wide hierarchy is therefore not driven by any single influential floor boundary or outlying high-yield case.

These whole-benchmark findings matter for two reasons. First, they confirm that the pooled superiority of general rules is not marginal. Second, they establish a robust baseline against which typology-conditioned variation can be interpreted.

Typology-conditioned central tendency

Table 2 reports typology-conditioned descriptive statistics. General rules produced the highest mean reviewed count in every plan family. However, the *size* of the advantage changed across typologies. The general-rules mean was only modestly higher than the mixed-method mean in I-, L-, C-, and T-type plans, but the gap widened markedly in H-type plans. This indicates that typological structure affects not the ordering of the three algorithms, but the intensity of their separation.

Table 2: Typology-conditioned descriptive statistics for reviewed room count.

Typology	Cases	Algorithm	Mean	Median	SD	Best-or-tied-best share (%)
I	16	General rules	26.50	20.00	20.32	87.50
		Mixed	25.00	18.50	20.45	37.50
		Self-organization	19.19	11.00	19.63	6.25
L	10	General rules	18.20	16.50	13.29	80.00
		Mixed	16.70	14.50	10.08	10.00
		Self-organization	13.40	11.50	12.19	20.00
C	12	General rules	52.67	48.00	32.96	75.00
		Mixed	49.50	46.00	32.18	25.00
		Self-organization	47.00	42.00	30.05	8.33
T	7	General rules	25.14	26.00	7.56	85.71
		Mixed	23.43	23.00	6.63	14.29
		Self-organization	22.57	23.00	7.28	14.29
H	5	General rules	34.80	35.00	17.28	100.00
		Mixed	28.60	30.00	15.24	0.00
		Self-organization	28.80	28.00	14.15	0.00

Figure 2 presents the typology-conditioned means graphically. The chart confirms that a pooled ranking would conceal important variation in relative distance. The general-rules method is clearly dominant in H-type plans, while the mixed method remains much closer in I- and C-type plans.

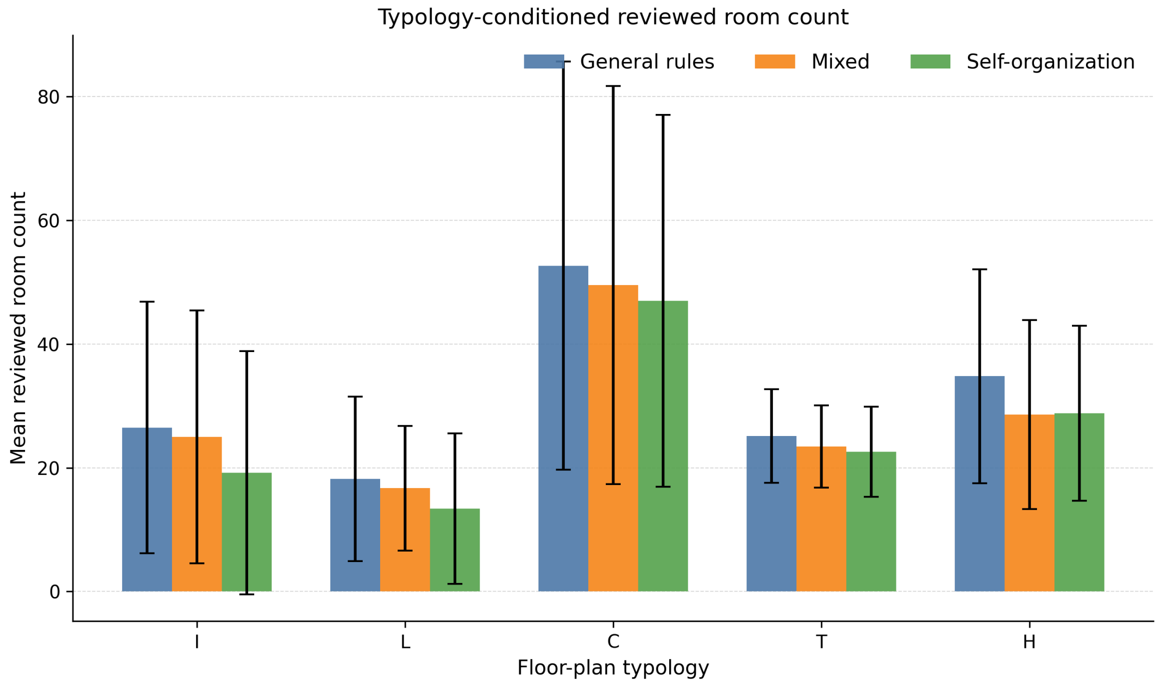

Figure 2: Mean reviewed room count by typology and algorithm. Error bars indicate one standard deviation.

To make these typology-level differences more interpretable, Table 3 reports mean percentage gains. The general-rules method outperformed the mixed method by 6.0% in I-type plans, 9.0% in L-type plans, 6.4% in C-type plans, 7.3% in T-type plans, and 21.7% in H-type plans. In contrast, the mixed method outperformed self-organization by a much narrower margin in most plan families, and in H-type plans the two iterative methods were essentially tied at the mean level.

Table 3: Relative mean gains by typology. Positive values indicate that the numerator algorithm has the higher mean reviewed count.

Typology	G vs. M (%)	G vs. S (%)	M vs. S (%)
I	6.00	38.09	30.28
L	8.98	35.82	24.63
C	6.40	12.06	5.32
T	7.32	11.39	3.81
H	21.68	20.83	−0.69

Typology-specific significance tests

The Friedman tests remained statistically significant in each typological subset, indicating that algorithm choice matters not only in the pooled benchmark but also after stratification by plan family. As shown in Table 4, effect sizes ranged from moderate in L-type plans to very strong in H-type plans. Although the H subset contains only five cases and should therefore be interpreted cautiously, it still shows the clearest internal differentiation because the general-rules method dominates every instance in that subset.

Table 4: Typology-specific Friedman tests on reviewed room count.

Typology	Cases	Friedman χ^2	p-value	Kendall's W
I	16	18.32	< 0.001	0.573
L	10	7.54	0.023	0.377
C	12	12.13	0.002	0.506
T	7	7.46	0.024	0.533
H	5	7.89	0.019	0.789

This result is analytically important. It means that the global ordering is not an artifact of combining heterogeneous cases into one sample. The hierarchy remains present when the benchmark is partitioned into the very typological groups that structure the original dataset.

Winner shares and pairwise dominance

Mean values are useful, but they do not fully capture how a method behaves in practice. For decision support, the share of cases in which an algorithm is best or tied-best is often more informative. Figure 3 shows these winner shares by typology. General rules was best or tied-best in all H-type cases, in 87.5% of I-type cases, in 80.0% of L-type cases, in 75.0% of C-type cases, and in 85.7% of T-type cases.

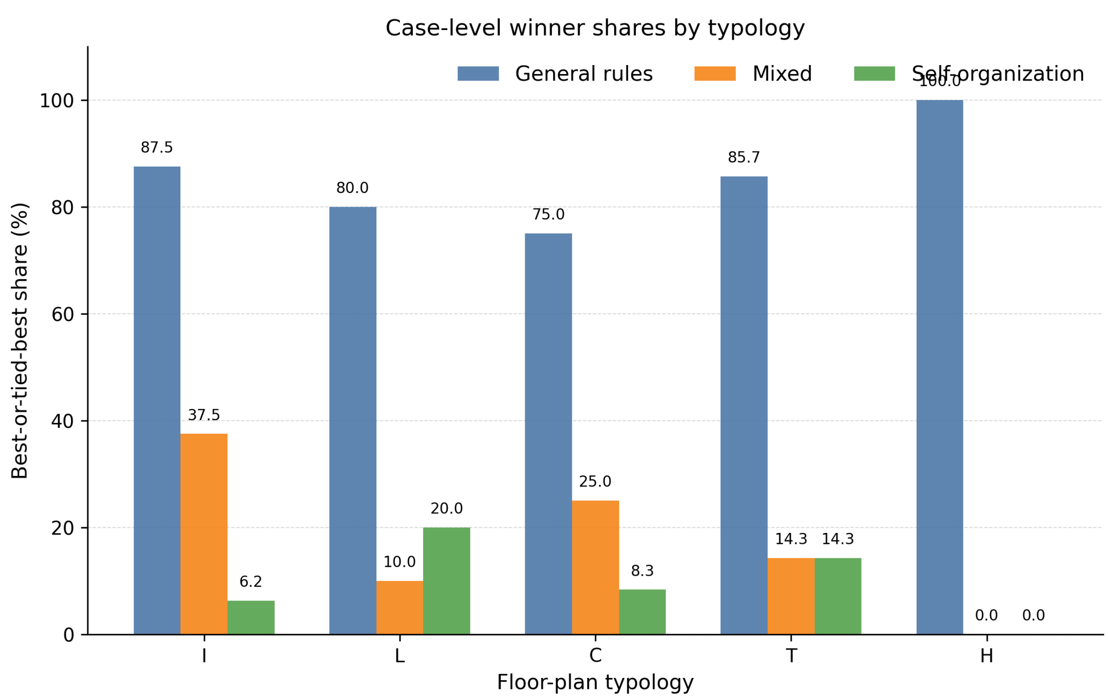

Figure 3: Share of cases in which each algorithm was best or tied-best within each typology.

The mixed method was most competitive in I-type plans, where it was best or tied-best in 37.5% of cases, and in C-type plans, where it reached 25.0%. Self-organization rarely achieved best-or-tied-best status and never did so in H-type plans.

To complement winner shares, Table 5 reports pairwise dominance rates and paired mean differences. General rules exceeded self-organization in 88% of all cases and exceeded the mixed method in 74% of all cases. The mixed method exceeded self-organization in 68% of cases. These frequencies are helpful because they translate the benchmark into a simple operational question: if a planner chooses one method before seeing the result, how often is that choice likely to outperform an alternative on the reviewed capacity criterion?

Table 5: Pairwise dominance and paired differences across all 50 benchmark cases.

Comparison	A > B (%)	A = B (%)	A < B (%)	Mean diff.
Gen. rules vs. self-org.	88	6	6	5.62
General rules vs. mixed	74	14	12	2.40
Mixed vs. self-org.	68	10	22	3.22

Local exceptions and their design meaning

Despite the consistency of the global hierarchy, several local reversals are noteworthy. The mixed method uniquely outperformed general rules in selected C-, I-, L-, and T-type cases. For example, it surpassed general rules in Case 23 (C), Case 16 (I), Case 17 (L), and Case 5 (T), and it tied general rules in additional cases. These reversals are not numerous enough to displace the general-rules method as the benchmark leader, but they are important because they identify the space in which a flexible alternative remains substantively credible.

Self-organization performed best in only a very small number of cases, including Case 45 and Case 41, and it tied for first in several others. This pattern supports a more differentiated interpretation than the simple label of "weakest algorithm." Self-organization is weak as a default capacity-maximization method, but not irrelevant. Its localized successes suggest that its main contribution may lie in exploratory search and the generation of atypical alternatives that a stricter corridor-led method might not discover.

Case-level ranking transparency

Appendix reports the reviewed counts and case-level winners for all 50 cases. This appendix serves two functions. First, it makes the typology-conditioned interpretation fully transparent. Second, it helps distinguish between two kinds of algorithmic strength: consistent baseline strength and opportunistic local strength. General rules exemplifies the first pattern. The mixed method shows a combination of baseline adequacy and occasional local advantage. Self-organization exhibits the second pattern only rarely, but those rare cases are still informative for exploratory workflows.

DISCUSSION

Main empirical interpretation

The most important result of this detailed analysis is not merely that general rules achieves the highest total reviewed capacity. More importantly, within the limits of the benchmark, it does so *consistently across all five typological families* and retains the same pooled ordering under leave-one-case-out sensitivity. This makes the benchmark result operationally stronger than a pooled ranking alone. A method that remains first in every typological subset is not just high-performing; it is robust in a way that matters for deployment.

The general-rules approach appears to benefit from the explicit corridor logic and strong geometric order built into tessellation-based subdivision. This order reduces ambiguity in how rooms relate to circulation and perimeter access, which in turn yields higher and more stable reviewed capacity. The source benchmark also observed that general-rules outputs aligned more closely with architectural expectations in the qualitative assessment, with the highest overall satisfaction score among the three methods (Cudzik & Kruk, 2022). The present reanalysis is consistent with that observation: the algorithm that dominates quantitatively is also the one whose outputs appear most legible and dependable. At the same time, the present evidence supports comparative robustness on reviewed capacity rather than comprehensive architectural superiority across all possible performance criteria.

Why the mixed method remains important

The mixed corridor-attraction method is not simply "second best." In several typologies it remains competitively close to the best-performing method, and in a limited number of cases it becomes the unique winner. This pattern is especially clear in I- and C-type plans, where the mixed method reaches relatively high winner shares and keeps the mean gap to general rules small.

This has a practical implication. If a project team seeks a method that retains corridor guidance but allows more local adaptability than pure tessellation, the mixed method becomes the most defensible secondary choice. In workflow terms, it is the strongest candidate for a parallel-run strategy: use general rules as the primary baseline, then evaluate the mixed method in cases where geometric irregularity suggests that a more flexible response might unlock additional yield or spatial variety.

What to do with self-organization

Self-organization remains the weakest performer on reviewed capacity, mean yield, median yield, and winner share. Yet it should not be dismissed altogether. The source study already noted that the self-organizing process can produce irregular but potentially inspiring corridor formations and may reveal configurations that deterministic rules do not readily anticipate (Cudzik & Kruk, 2022). The present analysis supports that view. The method rarely dominates, but its isolated successes suggest a distinct role within the design process.

That role is exploratory rather than primary. Self-organization is most useful when the purpose of computation is not to maximize capacity under a fixed brief, but to widen the search space, challenge established assumptions, or identify unusual possibilities for later refinement by more controlled methods. In other words, the value of self-organization may lie less in direct feasibility optimization and more in divergence generation.

Implications for human–machine collaboration

A detailed comparative analysis also clarifies the meaning of human–machine collaboration in architectural computing. Generative systems do not merely replace design labor; they distribute different kinds of intelligence across different phases of decision-making (Mukkavaara & Sandberg, 2020; Reitberger, 2022). In the present benchmark, the three algorithms are not interchangeable tools competing for one identical role. They map more naturally to three distinct workflow functions:

1. *General rules* as the default baseline for reliable capacity generation under a fixed hotel brief;
2. *Mixed corridor attraction* as the flexible alternative when a team wants to test a more adaptive yet still corridor-aware strategy;
3. *Self-organization* as an exploratory supplement for widening the search space and provoking nonstandard spatial options.

This differentiated interpretation is preferable to a winner-takes-all narrative. It preserves the empirical reality that one method is strongest overall while still recognizing that the weaker methods may hold workflow value when their role is defined appropriately.

Methodological contribution

The article also contributes methodologically. It demonstrates that comparative benchmark studies in architecture can gain substantial interpretive value by introducing typology-conditioned analysis, repeated-measures statistics, winner-share logic, case-level transparency, and basic sensitivity testing. These additions do not create new primary evidence, but they do provide a more decision-oriented way of reading the data already available.

That point extends beyond hotel planning. Any benchmark involving heterogeneous plan classes—whether in housing, offices, healthcare, logistics, or urban block design—may benefit from the same shift. Instead of asking only which method has the best global score, researchers should also ask where a method is most stable, where a runner-up remains close enough to matter, how sensitive the ranking is to influential cases, and what role each algorithm should play in an applied workflow.

Limitations

Three limitations should be acknowledged. First, the evidence comes from a single published benchmark, and the typology labels used here are a coarse proxy for geometry. They distinguish major plan families, but they do not capture continuous boundary descriptors such as convexity, perimeter-to-area ratio, branch depth, or local concavity. A future study using these descriptors across multiple datasets could turn the present typology-conditioned interpretation into a more predictive geometry-aware selection model with stronger external validity.

Second, the present article focuses on reviewed room count as the principal outcome. This is justified because room yield is central to early hotel feasibility and because the source benchmark treated reviewed counts as the most credible comparison metric. Nevertheless, reviewed capacity does not exhaust the design-quality question. Corridor legibility, room regularity, evacuation logic, daylight quality, user experience, and computational cost remain important evaluation dimensions that were outside the present reanalysis.

Third, the available qualitative dataset is aggregated rather than participant-level. The benchmark reports strong overall preference for the general-rules method, but the present study cannot reconstruct rater-level models of preference formation or test whether qualitative preference changes by typology. Future work with the full survey matrix could connect quantitative efficiency and human-centered evaluation more explicitly.

Future research directions

The next research step should combine the present typology-conditioned logic with continuous geometric measures. That would permit the construction of a predictive meta-selection model in which the boundary itself suggests the most promising generative strategy before any full layout generation occurs. A second direction would examine robustness under changing briefs, for example different room-size targets or corridor-width requirements. A third would integrate additional validation criteria such as vertical circulation, structural constraints, code compliance, and façade access.

Together, these directions would move the field from comparative benchmarking toward decision-support systems that recommend not only layouts, but also the most appropriate generative strategy for a given design condition.

CONCLUSION

This detailed reanalysis shows that the comparative strength of generative hotel-layout algorithms is both hierarchical and condition-sensitive. The general-rules tessellation method remains the strongest approach in the full benchmark and in every typological subset, confirming its position as the most reliable baseline for early-stage adaptive hotel planning under the reviewed-capacity criterion. The mixed corridor-attraction method consistently occupies second place overall but becomes locally competitive in several plan families, especially I- and C-type cases, making it the most credible flexible alternative. Self-organization rarely achieves top performance on reviewed capacity, yet it retains value as an exploratory mechanism for widening the search space.

The broader contribution of the article lies in how the benchmark is interpreted. Pooled totals alone are not sufficient for practical algorithm selection. Typology-conditioned analysis reveals whether an approach is globally strong, locally competitive, or mainly exploratory, and the leave-one-case-out check shows that the benchmark-wide ordering is not an artifact of a single influential case. For architectural computing, that distinction matters because methods are deployed in workflows, not abstract rankings. For adaptive hotel planning, the decision logic that follows from the present study is therefore conditional rather than universal: use general rules as the baseline, test the mixed method when flexibility is desirable, and reserve self-organization for exploratory diversification rather than primary capacity optimization, while recognizing that broader design-quality criteria still require additional validation.

CASE-LEVEL REVIEWED ROOM COUNTS AND WINNERS

Table 6 reports the reviewed room counts used in the present reanalysis. These values are transcribed from the benchmark table in Cudzik and Kruk (2022) and are provided here to make all secondary calculations transparent.

Table 6: Case-level reviewed room counts and case winners.

Case	Type	General rules	Mixed	Self-organization	Winner
01	I	10	10	8	General rules / Mixed

Continued on next page

Table 6 – continued from previous page

Case	Type	General rules	Mixed	Self-organization	Winner
02	L	11	10	7	General rules
03	L	19	16	14	General rules
04	C	29	25	25	General rules
05	T	11	13	10	Mixed
06	C	25	26	23	Mixed
07	T	34	33	31	General rules
08	I	22	21	18	General rules
09	H	47	39	41	General rules
10	I	10	10	8	General rules / Mixed
11	L	0	8	9	Self-organization
12	I	18	16	13	General rules
13	I	18	14	0	General rules
14	C	92	87	80	General rules
15	C	95	87	87	General rules
16	I	75	79	75	Mixed
17	L	17	18	15	Mixed
18	L	7	6	5	General rules
19	L	19	17	15	General rules
20	L	12	11	9	General rules
21	H	22	17	17	General rules
22	C	38	35	32	General rules
23	C	49	50	44	Mixed
24	I	45	34	36	General rules
25	L	45	38	45	General rules / Self-organization
26	T	26	23	23	General rules
27	I	28	26	24	General rules
28	I	48	44	44	General rules
29	C	101	97	88	General rules
30	I	32	31	28	General rules
31	L	36	30	0	General rules
32	I	57	57	0	General rules / Mixed
33	C	54	49	52	General rules
34	H	35	30	28	General rules
35	T	22	19	20	General rules
36	I	28	26	25	General rules
37	H	56	47	45	General rules
38	I	12	11	9	General rules
39	I	6	6	4	General rules / Mixed
40	L	16	13	15	General rules
41	I	5	5	7	Self-organization
42	I	10	10	8	General rules / Mixed
43	T	23	21	18	General rules
44	C	5	4	0	General rules
45	C	14	8	15	Self-organization
46	C	47	43	40	General rules
47	H	14	10	13	General rules
48	T	29	26	29	General rules / Self-organization
49	T	31	29	27	General rules
50	C	83	83	78	General rules / Mixed

REFERENCES

Caetano, I., Santos, L., & Leitão, A. (2020). Computational design in architecture: Defining parametric, generative, and algorithmic design. *Frontiers of Architectural Research*, *9*(2), 287–300.

Chen, L. (2012). Agent-based modeling in urban and architectural research: A brief literature review. *Frontiers of Architectural Research*, *1*(2), 166–177.

Cudzik, J., & Kruk, J. (2022). Environmental impact of construction. methods of conscious shaping architecture in terms of ecological solutions. *Przestrzeń i Forma*, 121–136.

Ligler, H. (2021). Reconfiguring atrium hotels: Generating hybrid designs with visual computations in Shape Machine. *Automation in Construction*, *132*, 103923.

Mostafavi, P., Merkin, V. G., Provornikova, E., Sorathia, K., Arge, C., & Garretson, J. (2022). High-resolution simulations of the inner heliosphere in search of the kelvin–helmholtz waves. *The Astrophysical Journal*, *925*(2), 181.

Mukkavaara, J., & Sandberg, M. (2020). Architectural design exploration using generative design: Framework development and case study of a residential block. *Buildings*, *10*(11), 201.

Rahbar, M., Mahdavinejad, M., Bemanian, M., Davaie Markazi, A. H., & Hovestadt, L. (2019). Generating synthetic space allocation probability layouts based on trained conditional-GANs. *Applied Artificial Intelligence*, *33*(8), 689–705.

Rahbar, M., Mahdavinejad, M., Markazi, A. H., & Bemanian, M. (2022). Architectural layout design through deep learning and agent-based modeling: A hybrid approach. *Journal of Building Engineering*, *47*, 103822.

Reitberger, M. (2022). *Regulation of abcb1 expression is a potential therapeutic target in drug resistant pancreatic cancer* [Doctoral dissertation].

Stieler, D., Schwinn, T., Leder, S., Maierhofer, M., Kannenberg, F., & Menges, A. (2022). Agent-based modeling and simulation in architecture. *Automation in Construction*, *141*, 104426.

AUTOBIOGRAPHICAL SKETCHES

L. Faggion, Charles Darwin University; laurafaggion35@gmail.com.

R. Furlan, Lusail University.

Manuscript revisions completed 17 November 2023.