

LOW-NOISE VERTIPORT PROCEDURE PLANNING USING DEEP REINFORCEMENT LEARNING

Yikang Wang[1], Jiang Guo[1], Yingjie He[1], Fangyi Wu[1] & Zhenyu Gao[1]

[1]The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong

Abstract

Urban vertiport operations require eVTOL arrival procedures that are dynamically feasible while managing aircraft noise for nearby communities. Deep reinforcement learning (DRL) can generate noise-aware eVTOL trajectories, but the results do not directly translate into standard, reviewable procedures. In this work, we develop a policy-to-procedure translation framework that converts low-noise DRL agent behavior into reviewable vertiport operations-planning objects, including low-noise trajectory clusters, eco-acoustic arrival gate candidates, and extracted standard procedures. We demonstrate the framework using eVTOL noise models and two real-world Hong Kong urban models under a shared 360-degree simulation. On the two urban models, the low-noise trajectories reduce mean maximum sound exposure level (SEL) by 3.95 dB and 2.94 dB, respectively, and Gini coefficient by 0.097 and 0.059, respectively, with mean flight-time increases of 29.85 and 17.71 seconds. The resulting standardized procedures contain 15 gates per scenario, along with 8 and 13 planning cells, respectively, and replay-consistent fixes for selected cases. These results demonstrate that learned low-noise behavior can be converted into bounded, reviewable procedure evidence, providing an early-stage tool for vertiport arrival design.

Keywords: Urban Air Mobility, Vertiport Procedures, Deep Reinforcement Learning, Aircraft Noise

1. Introduction

Urban Air Mobility (UAM) and electric vertical takeoff and landing (eVTOL) aircraft are expected to support passenger transport, cargo delivery, and emergency-response services in low-altitude urban airspace [1]. Their integration into dense cities, however, is constrained by aircraft noise emissions, infrastructure siting, and the need to assess received aircraft noise near people and buildings [2, 3, 4]. These issues are especially acute near vertiports, where arrival paths pass close to buildings and residents [5]. Low-noise vertiport procedure design therefore plays a key role in UAM noise mitigation. Low-noise arrival procedure design in a vertiport environment is technically difficult because several coupled objects must be handled together: three-dimensional urban geometry, guidance-level motion limits, aircraft noise source directivity, SEL-consistent exposure accumulation, and receptor-level exposure concentration. Deep reinforcement learning (DRL) is an attractive noise-aware trajectory planning approach for this setting, as a learned policy can adapt long-horizon route choices to buildings, approach direction, and cumulative noise exposure in ways that fixed geometric rules cannot easily prescribe. The central challenge is not only to find a quieter trajectory but also to render learned low-noise behavior as standardized candidate procedure geometry.

A DRL-based trajectory planning approach can generate feasible solutions, but procedure design requires standardized artifacts that can be named, compared, replayed, and revised. To address this missing translation layer, in this work we develop a policy-to-procedure framework for low-noise vertiport arrival procedure design. The framework first trains and evaluates acoustic-aware DRL policies

in an urban landing environment. It then converts successful rollouts into reviewable procedure-candidate artifacts, such as low-noise trajectory clusters, eco-acoustic arrival gate candidates, and extracted standard procedures.

The framework is evaluated in two Hong Kong case-study scenarios: Tseung Kwan O (TKO) and Hong Kong Central. The scenes feature different building distributions and receptor counts but share the same target, fixed-start protocol, and metric definitions. In each scenario, we generate low-noise trajectories, compare them against noise-agnostic baselines, and extract low-noise procedures around the vertiport.

The paper makes three contributions. First, it formulates low-noise vertiport arrival design as a policy-to-procedure translation workflow in which learned behavior becomes standardized procedural artifacts. Second, it defines a traceable artifact chain: fixed-start rollouts are grouped into route families, summarized by arrival gates and hotspot cells, converted into waypoints, and checked through replays. Third, it applies the workflow to two Hong Kong urban cases, Tseung Kwan O (TKO) and Hong Kong Central, yielding comparisons between a reachability-oriented reference and a retained noise-aware candidate.

2. Related Work and Positioning

UAM noise assessment links aircraft source characteristics, flight trajectory, event duration, receiver location, and the built environment. Prior UAM studies have examined feasibility, infrastructure constraints, airspace concepts, and public acceptance [6, 7]. Noise-focused studies use receptor- and grid-based maps to quantify received sound in urban settings [8, 4, 9], and vertiport-environment models improve the efficiency and spatial detail of that assessment [10]. These studies motivate the receptor-level acoustic layer used here and establish the need to report where acoustic exposure is concentrated around a vertiport.

Trajectory optimization and learning methods generate feasible motion under aircraft and environmental constraints. Convex and optimal-control formulations support constrained landing and terminal-area trajectory design [11]. Arrival-management and learning-based studies address adaptive terminal-area behavior [12]. Recent low-noise UAM studies also combine DRL with acoustic prediction models for urban landing trajectories [13]. This literature supplies path-generation methods; procedure review still needs a representation of repeated behavior as route structures, gates, programmed fixes, and replay checks.

The exposure-distribution component monitors the terminal receptor-energy vector for spatial concentration. Alpha-fairness provides a standard form for efficiency-distribution trade-offs [14], and related UAM routing work has used distribution-aware routing under risk-aware constraints [15]. In this paper, exposure refers to SEL-consistent cumulative acoustic energy at receptors. Receptor-energy Gini reports the concentration of terminal acoustic energy across receptors. Alpha-fair shaping increases the marginal penalty assigned to receptors with higher accumulated energy during policy generation. These quantities describe hotspot aversion within a single landing event; demographic equity, land-use justice, and certified operational fairness require social, regulatory, and repeated-operation data outside the present procedure-evidence package.

This paper connects low-noise trajectory generation with procedure review. Acoustic and trajectory studies provide the tools for evaluating and optimizing motion. Distribution-aware formulations provide a way to penalize concentration in receptor-level cumulative exposure. The workflow developed here organizes the next layer of evidence: recurrent route families, arrival-gate locations, hotspot-derived inspection cells, route-fix geometry, and follower replay metrics.

3. Framework

Figure 1 gives the reading order for the workflow. The training layer produces fixed-start rollout records with aircraft states, feasibility outcomes, cumulative exposure histories, and terminal acoustic metrics. The translation layer converts those records into procedure-candidate objects: route families, Acoustic Arrival Gate candidates, hotspot-derived candidate planning cells, and programmed route fixes. The validation layer replays selected fixes and reports feasibility, acoustic, exposure-distribution, and geometric quantities under the same metric definitions. The results section follows

this chain from fixed-start performance to trajectory morphology, gate extraction, hotspot-cell diagnosis, route-fix specification, and follower replay.

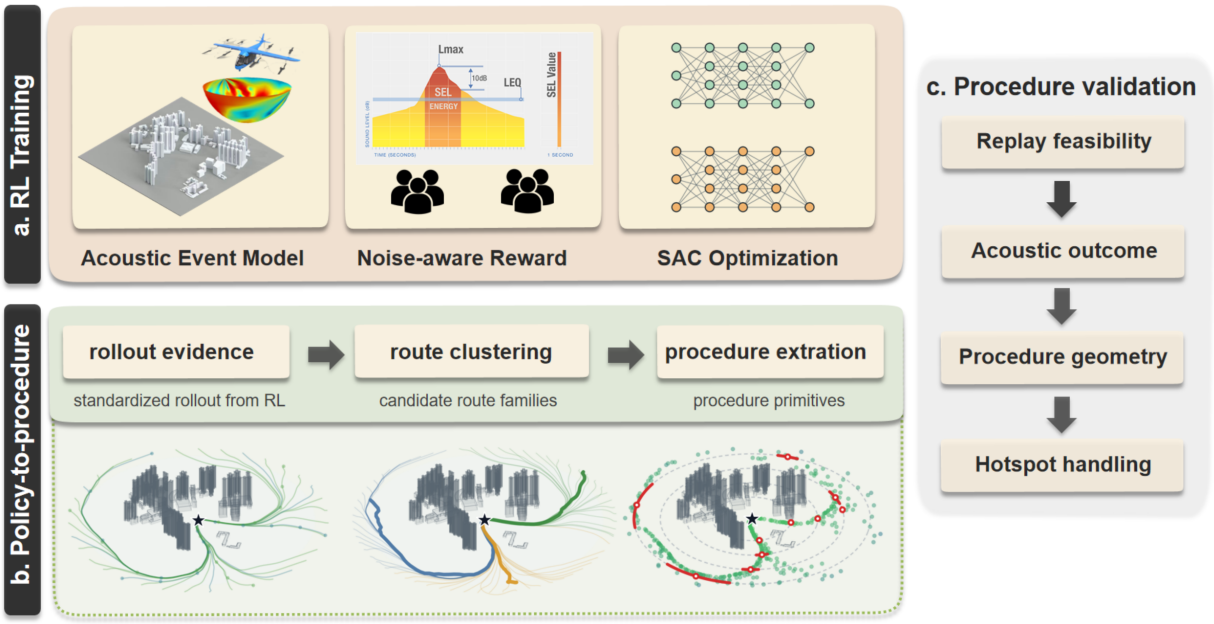


Figure 1 – Framework for policy-to-procedure translation in low-noise vertiport arrivals. The training layer generates fixed-start rollout records with urban geometry, an event-level acoustic interface, noise-aware reward shaping, and Soft Actor-Critic (SAC) policy optimization. The acoustic block uses directional source and free-field propagation terms for event-level route comparison. The translation layer converts rollout records into route families, Acoustic Arrival Gate candidates, hotspot-derived candidate planning cells, and programmed route fixes. The validation layer evaluates the replay consistency, acoustic outcome, hotspot response, and geometric traceability of the simplified procedure geometry.

3.1 Landing Episode and Receptor Acoustic Metrics

The operating domain is a bounded three-dimensional urban scene with building obstacles, a vertiport target, and receptor locations on building surfaces. An episode begins from a fixed-start aircraft state and ends when the aircraft reaches the target region, leaves the domain, collides, or reaches the time limit. The target region is a procedure-design capture region near the vertiport. Final touchdown modeling is left outside this abstraction. At decision step t, the aircraft position is $\mathbf{p}_t$ and the commanded velocity is $\mathbf{v}_t$. The guidance-level motion update is

$$\mathbf{p}_{t+1} = \mathbf{p}_t + \mathbf{v}_t \Delta t. \tag{1}$$

The action-mapping layer enforces fixed approach speed, non-climbing motion, descent-angle limits, and heading-rate limits. The implementation uses $\Delta t = 0.2$ s, fixed approach speed 20.6 m/s, a 5°–15° descent-angle envelope, a 10°/s heading-rate limit, a 50 m target-capture radius, and a 1000-step timeout. The 360-start audit applies the same deterministic initial-position set and initial-velocity rule to every compared policy within a scenario. This motion envelope makes the rollout geometry suitable for gate extraction, hotspot-cell analysis, and replay validation. Aircraft-specific inner-loop control closure remains a later design layer.

The acoustic module provides event-level route comparison. It queries a fixed-operating-state directional source spectrum for a lift-plus-cruise UAM vehicle on a 500 ft reference surface, interpolates the heading-aligned aircraft-to-receptor direction, applies spherical spreading to the actual receptor distance, and sums frequency bands with A-weighting to obtain $L_{A,i,t}$ [16]. This interface captures source directivity and distance effects for each aircraft–receptor pair at each time step.

Each landing episode is treated as one acoustic event. For receptor i, the simulator stores linear acoustic energy and reports sound exposure level (SEL) after logarithmic conversion. With time step

Δt and reference duration $T_0 = 1$ s, the discrete receptor-energy update is

$$e_{i,t} = \frac{\Delta t}{T_0} 10^{L_{A,i,t}/10}, \qquad E_{i,t+1} = E_{i,t} + e_{i,t}. \tag{2}$$

At terminal step N, receptor-level event SEL is reported as

$$\mathrm{SEL}_i = 10 \log_{10}(E_{i,N}). \tag{3}$$

The main acoustic metric is the mean, over evaluated starts, of the maximum receptor SEL within an episode. This value represents the event-level hotspot SEL averaged over the fixed-start set. The distributional metric is the Gini coefficient of the terminal receptor-energy vector,

$$G(\mathbf{E}) = \frac{\sum_{i=1}^{R} \sum_{j=1}^{R} |E_i - E_j|}{2R \sum_{i=1}^{R} E_i}, \tag{4}$$

where R is the number of receptors. Energy Gini is computed for each start from the terminal receptor-energy vector and then averaged over the 360 starts.

3.2 Policy Generator and Noise-Aware Reward

The policy generator must produce continuous guidance commands under state-dependent feasibility constraints and long-horizon acoustic cost. Soft Actor-Critic (SAC) is selected for this role because it supports continuous actions and stochastic exploration during training [17]. The policy outputs a continuous directional action, and the action-mapping layer converts it to a feasible velocity command. The baseline observation contains position, previous velocity, and target-relative displacement. The exposure-history observation adds cumulative exposure scale, exposure concentration, and hotspot-direction features. These added features allow the policy to respond to the receptor-energy pattern created by earlier route choices.

Each scenario uses two policy roles. The conventional reference is a reachability-oriented SAC policy with noise shaping disabled. The retained candidate is a noise-aware SAC policy selected after the fixed-start audit for artifact extraction. Both policies use the same motion envelope, termination rules, receptors, acoustic model, and 360 initial approach directions within a scenario. This paired design makes each reported acoustic, geometric, hotspot, and replay change traceable to the arrival behavior used by the retained candidate.

The training objective combines progress, time, terminal success or failure, and a noise-aware acoustic shaping term. Let $\mathbf{E}_t$ denote the cumulative receptor-exposure vector and let Γ be an exposure normalization scale. The alpha-fairness-inspired impact potential is

$$\Phi_\alpha(\mathbf{E}_t) = \sum_{i=1}^{R} \frac{1}{1+\alpha} \left(\frac{E_{i,t}}{\Gamma} \right)^{1+\alpha}, \qquad \alpha \geq 0. \tag{5}$$

With $\alpha = 0$, the potential is proportional to total normalized cumulative exposure. Larger α increases the marginal cost assigned to receptors with higher accumulated exposure. The dense acoustic shaping term uses the potential difference

$$r_t^{\mathrm{noise}} = -\lambda_\alpha \left[\gamma \Phi_\alpha(\mathbf{E}_{t+1}) - \Phi_\alpha(\mathbf{E}_t) \right], \tag{6}$$

following the standard reward-shaping form [18]. Here λ_α is the acoustic shaping weight, γ is the discount factor used in policy optimization, and Γ normalizes cumulative receptor exposure before the alpha-fair power is applied. SAC and receptor-energy-aware reward shaping supply the behavior source for the translation stage.

3.3 Translation from Rollouts to Procedure Artifacts

The translation stage starts from a fixed-start rollout dataset. Each record contains the aircraft trajectory, feasibility outcome, terminal receptor energies, and event-level metrics for one initial approach direction. Failed records remain in the feasibility statistics. Successful records provide the geometry

used for artifact extraction. Successful trajectories are resampled into a common target-centered horizontal frame and grouped by approach morphology to form inspectable route families.
Acoustic Arrival Gates are extracted from repeated crossings of circular rings around the vertiport. Successful trajectory points are sampled every fifth simulator step and projected into a target-centered horizontal frame. Four ring radii are used for both scenarios: 900, 700, 500, and 300 m. At each ring, points with horizontal range within ± 35 m of the ring radius are retained and binned by azimuth using 5° bins. For each ring, the active-bin support threshold is

$$T_r = \max\{3,\ Q_{0.65}(c_b : c_b > 0)\}, \tag{7}$$

where c_b is the sampled-point count in azimuth bin b and $Q_{0.65}$ is the 65th percentile over non-empty bins. A gate candidate is a circularly contiguous component of bins with $c_b \geq T_r$. It records angular sector, lateral width, altitude band, point and episode support, and mean step SEL. The same ring set, radial tolerance, bin width, sampling stride, and support rule are used in TKO and Central. In Fig. 2, sampled points show rollout support and red arcs show retained gate candidates.
Hotspot-derived candidate planning cells are built from receptor SEL under the conventional reference. Terminal receptor energy is averaged over successful fixed-start episodes and converted to mean receptor SEL. The hotspot set is defined by the 90th percentile of the conventional-reference receptor SEL distribution within each scenario. Hotspot receptors are assigned to a common 100 m horizontal grid using the integer cell indices $\lfloor x/100 \rfloor$ and $\lfloor y/100 \rfloor$. A cell is retained when it contains at least one top-decile hotspot receptor. Coverage is computed by applying the same grid assignment to all receptors. These cells act as diagnostic objects for local inspection. Representative route families are then compressed into programmed route fixes and corridor cross-sections. The resulting specification contains route candidates, fix sequences, cell associations, and metric links that trace each procedure artifact back to the rollout evidence.

3.4 Procedure-Follower Replay

A procedure-follower controller replays the programmed route fixes for representative candidates. The replay controller receives the fix sequence as its reference geometry and uses the same motion envelope as the rollout simulator. The replay check has two tasks. It tests the replay consistency of the simplified geometry after compression from a continuous policy path to route fixes. It also measures the maximum SEL and receptor-energy Gini of the replayed procedure against the values obtained from the learned trajectory family. Replay provides local feasibility and consistency evidence for selected representative procedure cases.

4. Case-Study Setup

The two case-study scenarios use the same target and evaluation protocol with different Hong Kong urban morphologies. Table 1 summarizes the scene assets. TKO is a dense residential benchmark. Hong Kong Central provides a second morphology with more processed buildings, more receptors, and a taller maximum building. Both scenes are shifted into a common $2000\,\mathrm{m} \times 2000\,\mathrm{m} \times 500\,\mathrm{m}$ operating domain, with target position $[1000, 1000, 100]$ m.

Table 1 – Scene and evaluation setup. Receptors are the points over which cumulative acoustic energy and terminal SEL are evaluated.

Scenario	Raw buildings	Processed buildings	Receptors	Maximum height (m)	Evaluation starts	Target position (m)
TKO	157	110	330	186.0	360	$[1000, 1000, 100]$
Central	227	159	477	415.8	360	$[1000, 1000, 100]$

For each scenario, the evaluation compares a conventional reference with a retained noise-aware candidate over the same 360 initial approach directions. The comparison is paired within each scenario. Start states, target position, receptor set, acoustic model, termination rules, and metric definitions are fixed before evaluation. Reported reductions are computed against the within-scenario

reference, making the acoustic, receptor-level exposure distribution, gate, hotspot, and replay quantities traceable to changes in arrival behavior.
Table 2 defines the compared policy roles and the retained-candidate selection rule. The conventional reference is the reachability-oriented SAC reference policy for that scenario, with noise shaping disabled. The retained candidate is a noise-aware SAC policy retained after the fixed-start audit. Selection first required completion of all 360 starts. Among feasible candidates, lower mean maximum receptor SEL was prioritized, and lower Energy Gini was used as the distributional check. This rule defines the case-study artifact candidate. Broader training-performance ranking requires multi-seed model sweeps.

Table 2 – Compared policy roles and retained-candidate definition. All rows use the same scene, starts, motion envelope, acoustic model, termination rules, and metric definitions within each scenario.

Role	Configuration	Use in evidence chain
Conventional reference	Reachability-oriented SAC; $\alpha = 0$, $\lambda_\alpha = 0$	Within-scenario comparison point for fixed-start acoustic, receptor-level exposure distribution, gate, hotspot, and replay metrics.
Retained candidate	Noise-aware SAC; $\alpha = 1$, $\lambda_\alpha = 15$	Case-study candidate used to extract gates, hotspot-derived candidate planning cells, route fixes, and selected representative replays.

5. Results

The results follow the procedure-evidence chain in Fig. 1. Table 3 reports fixed-start reachability, acoustic exposure, receptor-energy concentration, and cost metrics. Figure 2 and Table 4 extract recurrent route geometry and gate artifacts. Figure 3 localizes the baseline hotspot cells and their change under the retained candidate. Table 5, Fig. 4, and Table 6 test the route-fix representation through selected follower replays.

5.1 Fixed-Start Acoustic, Receptor-Level Exposure Distribution, and Cost Trade-Offs

Table 3 reports the fixed-start results. The success-rate column verifies that the reference and candidate remain feasible over all 360 initial approach directions. Mean maximum SEL is computed by taking the highest receptor SEL within each event and averaging that value over the start set. Energy Gini measures the concentration of terminal acoustic energy across receptors. The retained candidates reduce mean maximum receptor SEL and receptor-energy concentration in both scenarios, with longer flight time and path length.

Table 3 – Fixed-start evaluation summary. Reductions are relative to the within-scenario conventional reference; positive time and path changes are costs.

Scenario and policy	Success rate	Mean max. SEL (dB)	SEL reduction (dB)	Energy Gini	Gini reduction	Flight time (s)	Time increase (s)	Path increase (m)
TKO reference	1.000	83.429	0.000	0.527	0.000	49.553	0.000	0.000
TKO candidate	1.000	79.480	3.949	0.430	0.097	79.401	29.848	614.867
Central reference	1.000	81.669	0.000	0.432	0.000	51.980	0.000	0.000
Central candidate	1.000	78.730	2.940	0.372	0.059	69.692	17.712	364.863

5.2 Trajectory Morphology and Acoustic Arrival Gates

Figure 2 shows the geometric counterpart of the fixed-start trade-off in Table 3. Panels (a) and (c) show reference trajectories that concentrate many starts into direct approach bundles. Panels (b) and (d) show retained-candidate trajectories that route more traffic around the urban mass before entering the target region. The bottom row converts these bundles into gate evidence by plotting sampled rollout locations, recurrent crossing arcs, and the vertiport. The figure turns trajectory morphology into geometry that can be reviewed as a procedure candidate.

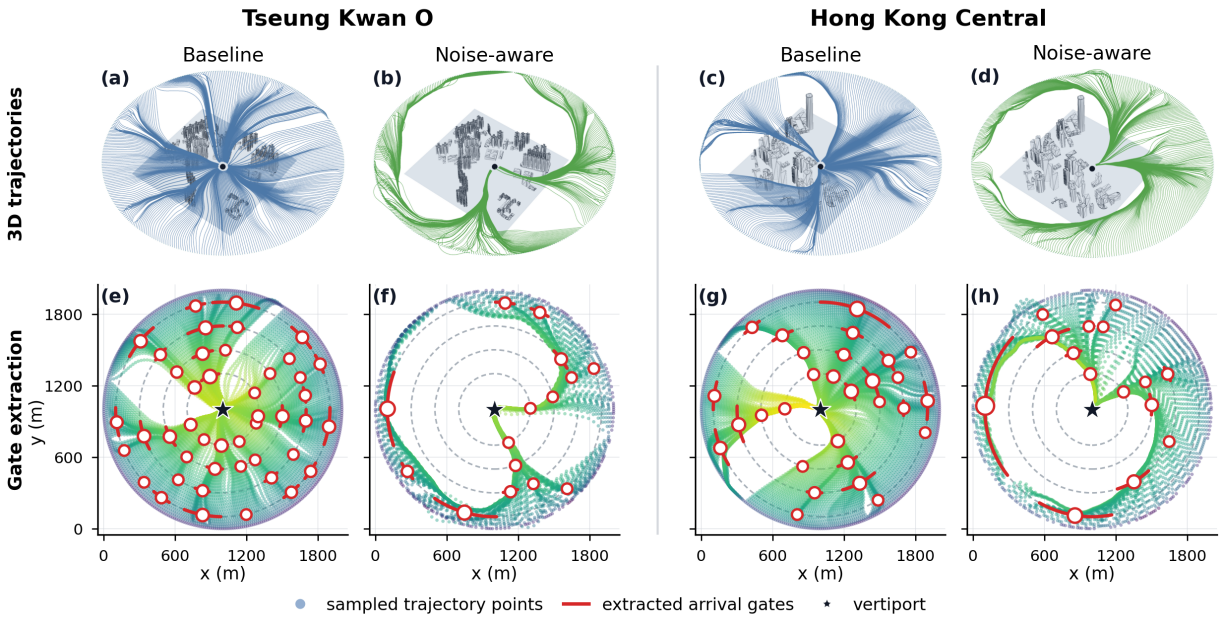


Figure 2 – Dual-scenario trajectory morphology and Acoustic Arrival Gate extraction. Panels (a–d) show the full fixed-start trajectory bundles for reference and candidate policies in TKO and Hong Kong Central. Panels (e–h) project sampled trajectory points onto the horizontal plane and overlay extracted Acoustic Arrival Gate candidates as red arcs. The gates summarize recurrent crossing regions and provide procedure-candidate geometry for later inspection.

Table 4 summarizes the extracted gate and hotspot-cell artifacts. The arrival-gate count gives the number of recurrent crossing sectors retained for review. Total gate angle approximates angular coverage around the vertiport. Mean lateral width gives the horizontal spread of each retained crossing, and mean altitude band gives the vertical tolerance observed in the rollout evidence. The planning-cell columns link the gate geometry to receptor-level acoustic diagnostics by reporting the number of monitored hotspot cells and the SEL reduction inside and outside those cells.

Table 4 – Extracted gate and hotspot-derived candidate planning-cell artifacts for retained candidates. Cell receptors are the receptors inside the hotspot-derived candidate planning cells.

Scenario	Arrival gates	Total gate angle (deg)	Mean lateral width (m)	Mean altitude band (m)	Planning cells	Cell receptors / all receptors	SEL reduction inside / outside (dB)
TKO	15	175.0	161.2	11.5	8	60 / 330	3.48 / 1.10
Central	15	270.0	229.2	11.9	13	99 / 477	3.54 / 2.49

5.3 Hotspot-Based Candidate Planning Cells

Figure 3 links gate-level geometry to the baseline hotspot receptors used for candidate-cell review. The left column shows the hotspot cells extracted from the conventional reference, with printed values giving cell-level baseline SEL summaries. The right column shows cell-level SEL reduction after applying the retained candidate. Larger positive values indicate stronger relief at a baseline hotspot. Negative values mark local cells where the candidate increases SEL. The cells cover all top-decile baseline hotspot receptors and contain a limited share of the receptor set: 60 of 330 receptors in TKO and 99 of 477 receptors in Central. Mean reductions inside the cells exceed the outside-cell reductions in both scenarios, making the cells localized diagnostics with explicit receptor support.

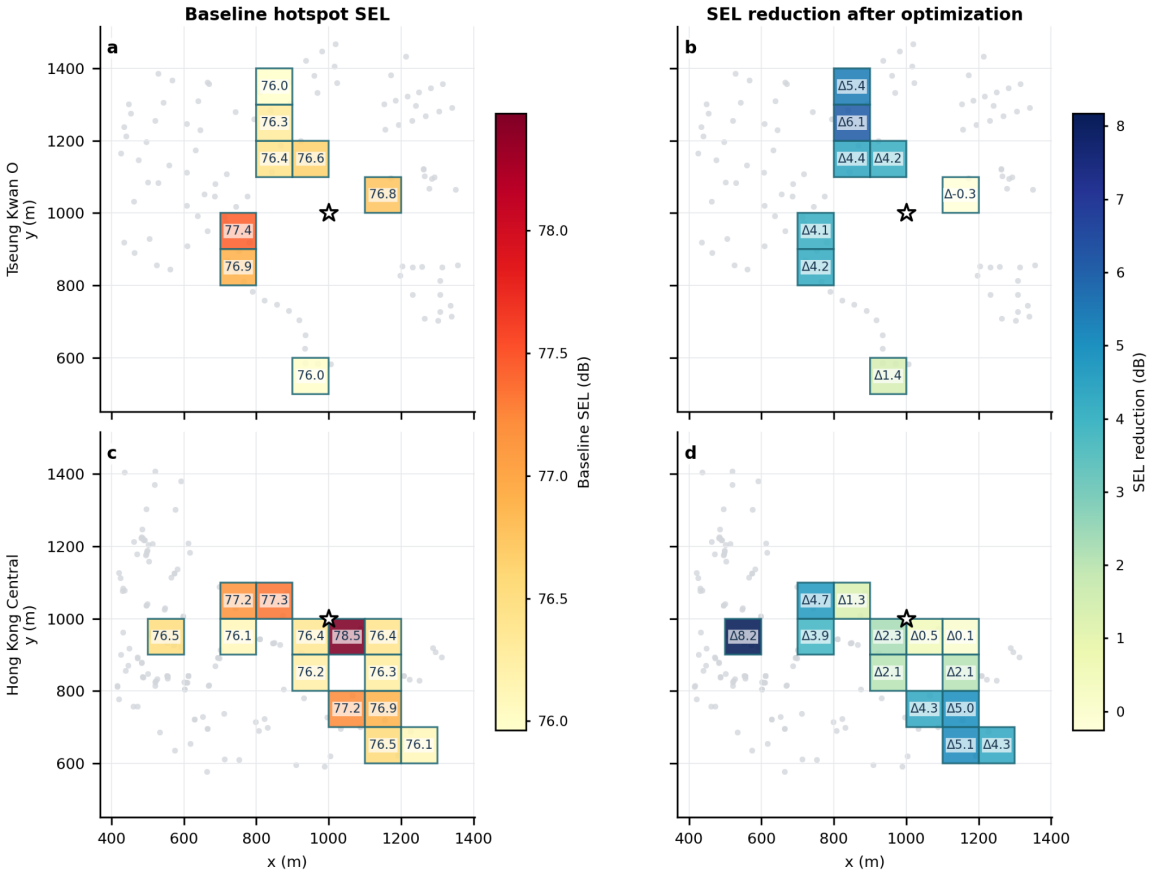


Figure 3 – Baseline hotspot cells and retained-candidate SEL change. Panels (a,c) show hotspot-derived candidate planning cells derived from the conventional reference in TKO and Hong Kong Central. Panels (b,d) show the cell-level SEL change after applying the retained candidate. Positive values indicate SEL reduction at the corresponding cell; negative values indicate local SEL increase. The gray dots are receptors and the star marks the vertiport.

The hotspot-cell result converts receptor-level acoustic concentration into spatial objects that can be inspected with arrival gates and route fixes. In TKO, the retained candidate gives strong reductions in the northern hotspot group and leaves one eastern cell nearly unchanged. In Central, reductions appear across a larger set of cells, with the strongest relief on the western and southern sides of the scene. These cell-level differences identify where later design iterations should examine path placement, monitoring assumptions, and local constraints.

5.4 Procedure Extraction and Replay Validation

The final evidence step converts selected route families into programmed route fixes and tests them with a follower. Table 5 summarizes the extracted specifications. A route candidate is a representative family from the successful rollout set. A route-fix representation is the simplified waypoint sequence used by the follower. Corridor cross-sections provide local geometry checks along that sequence. Monitoring cells link each route candidate back to the hotspot-cell analysis. Each scenario has five representative route candidates and 15 corridor cross-sections.

Table 5 – Procedure-specification summary for retained candidates. Route candidates are representative route families selected from successful rollouts. Route-fix representations are simplified waypoint sequences used for follower replay. Monitoring cells are the hotspot-derived cells associated with the evidence package.

Scenario	Route candidates	Corridor cross sections	Route-fix representations	Monitoring cells
TKO	5	15	5	8
Central	5	15	5	13

Figure 4 visualizes the compression step. The blue curves are original policy outputs, the orange dashed lines are programmed route fixes, and the green curves are procedure-follower replays. Close overlap between the orange and green curves indicates that the follower can track the simplified fix sequence. Differences between the blue and orange curves show the simplification introduced when continuous learned behavior is converted into procedure-scale geometry. The panels contain representative routes from selected families. The full 360-start population is reported through fixed-start metrics.

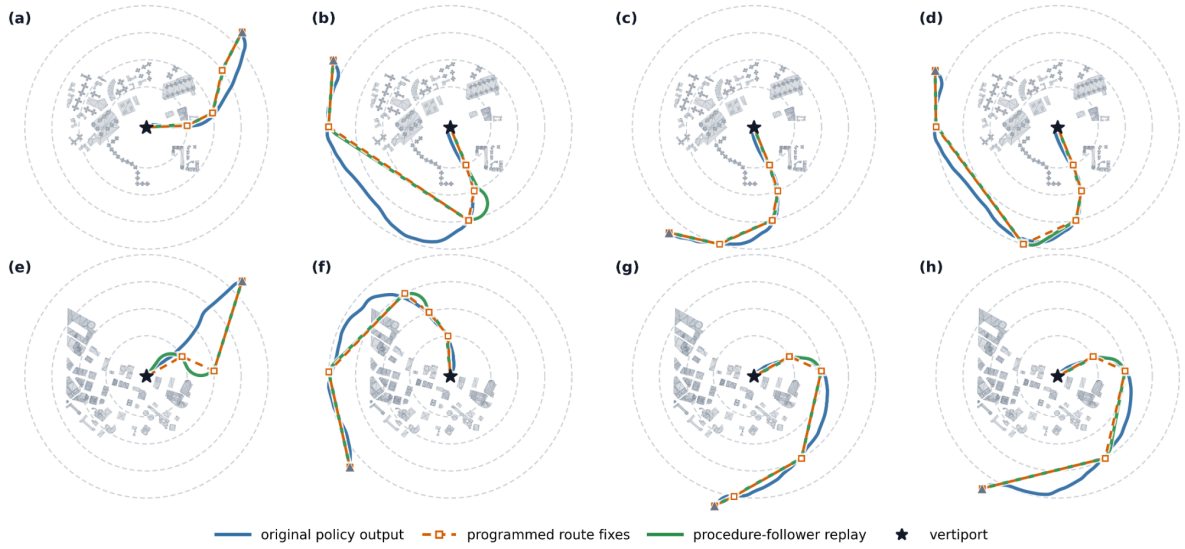


Figure 4 – Policy-to-procedure path comparison for representative route candidates. Blue curves show original policy outputs, orange dashed lines show programmed route fixes, and green curves show procedure-follower replays. The orange-green comparison checks replay consistency under a controller separate from the DRL policy; the blue-orange comparison shows the geometric compression introduced by procedure extraction.

Table 6 reports the acoustic effect of the compression in Fig. 4. The baseline rows give reference values for the same selected representative cases. The original-policy rows show the learned behavior before route-fix simplification. The procedure-replay rows show the independently followed fix sequence. Replay remains feasible in these selected cases and keeps substantial SEL reduction relative to the selected-case baseline. The replayed Central cases show larger SEL and Gini shifts than TKO, indicating a need for further route-fix tuning before operational-tolerance assessment.

Table 6 – Policy-to-procedure replay metrics for selected representative cases. Values are mean ± standard deviation across the representative cases shown in Fig. 4. This table is selected-case replay evidence and is separate from the 360-start summary used in the abstract and Table 3.

Scenario	Method	Flight time (s)	Mean max. SEL (dB)	Energy Gini	SEL reduction vs selected-case baseline (dB)
TKO	Baseline	50.8 ± 3.5	85.3 ± 1.6	0.530 ± 0.045	–
TKO	Original policy	99.8 ± 38.6	79.3 ± 1.4	0.391 ± 0.073	6.06
TKO	Procedure replay	94.1 ± 36.4	79.7 ± 1.1	0.415 ± 0.059	5.60
Central	Baseline	54.9 ± 5.6	83.1 ± 3.2	0.460 ± 0.095	–
Central	Original policy	92.6 ± 30.0	79.3 ± 1.0	0.369 ± 0.104	3.83
Central	Procedure replay	92.8 ± 20.4	80.7 ± 0.7	0.413 ± 0.097	2.39

The replay result completes the traceability chain from learned policy behavior to procedure evidence. Selected representative route families remain replay-consistent under the tested cases. The replay metrics also define the next design task: route-fix placement, corridor width, altitude windows, and follower logic need to be tuned together as the procedure candidate moves from evidence generation to operational tolerance assessment.

6. Discussion

The main result is the traceable conversion from learned low-noise behavior to procedure-review evidence. The framework gives procedure designers concrete objects to inspect: trajectory families, arrival gate candidates, hotspot-derived cells, programmed route fixes, and replay metrics. Each artifact has a clear role. A route family shows repeated behavior, a gate shows a recurrent crossing region, a planning cell identifies localized receptor-energy concentration, and a replay row shows the consistency of simplified geometry under a separate follower.

The paired case studies make the operational trade-off explicit. In both TKO and Central, retained candidates reduce mean maximum receptor SEL and receptor-energy Gini relative to within-scenario conventional references. These reductions come with longer flight time and greater path length. Reporting acoustic, receptor-level exposure distribution, geometry, hotspot, and replay quantities in the same evidence package keeps each low-noise candidate tied to its operational cost.

The two scenarios test the coherence of the artifact pipeline across different urban morphologies. TKO and Central share the same target, fixed-start protocol, and metric definitions. Their building layouts and receptor counts differ. The extracted gates, planning cells, and replay rows show that the same workflow can populate comparable review artifacts in both scenes. The numerical values are within-scenario evidence. Cross-city difficulty, geographic transfer, and training-recipe ranking require broader experiments.

The planning cells serve as screening artifacts. They localize baseline hotspot receptors, show where the retained candidate changes receptor SEL, and create spatial units for later monitoring or route-placement discussion. Regulatory restrictions, separation standards, navigation performance, weather margins, vehicle heterogeneity, and schedule capacity require additional analyses beyond the present evidence package.

7. Limitations and Future Work

The current evidence is a case-study framework demonstration. It uses deterministic fixed-start evaluation and one retained candidate per scenario. Receptor-level exposure distribution refers to cumulative acoustic exposure concentration within single landing events. Demographic equity, repeated-operation equity, and certified operational fairness require social, operational, and regulatory data beyond this single-event setup. Future work should add multi-seed training, threshold sensitivity for gate and planning-cell extraction, and direct comparisons with hand-designed, graph-search, and optimal-control procedure candidates.

The acoustic model captures event-level receptor SEL accumulation and directional source effects with simplified free-field propagation. A certification-oriented version of the workflow should integrate higher-fidelity reflection, diffraction, shielding, ground impedance, weather effects, navigation uncertainty, separation assurance, and operational scheduling.

The procedure-follower replay is a first validation layer. The current follower is intentionally simpler than the DRL policy. The next procedure-extraction stage should optimize waypoint placement, corridor width, altitude windows, and follower logic jointly, then validate under wind and navigation-error profiles that match operational assumptions.

8. Conclusion

This paper presented a policy-to-procedure framework for low-noise vertiport arrival design. Evaluated acoustic-aware DRL policies supplied candidate low-noise behavior. The translation workflow converted that behavior into Acoustic Arrival Gate candidates, hotspot-derived candidate planning cells, programmed route fixes, selected-case replay evidence, and auditable metrics. In TKO and Hong Kong Central case studies, retained candidates completed all 360 fixed-start arrivals and reduced mean maximum receptor SEL and receptor-energy Gini relative to within-scenario conventional references, with explicit flight-time costs. The framework provides a reproducible evidence package for reviewing learned low-noise procedure candidates and a basis for future robustness, propagation, and certification-oriented extensions.

9. Contact Author Email Address

mailto: ywangxm@connect.ust.hk

10. Copyright Statement

The authors confirm that they, and/or their company or organization, hold copyright on all of the original material included in this paper. The authors also confirm that they have obtained permission, from the copyright holder of any third party material included in this paper, to publish it as part of their paper. The authors confirm that they give permission, or have obtained permission from the copyright holder of this paper, for the publication and distribution of this paper as part of the ICAS proceedings or as individual off-prints from the proceedings.

References

[1] Zhenyu Gao, Yue Yu, Qinshuang Wei, Ufuk Topcu, and John-Paul Clarke. Noise-aware and equitable urban air traffic management: An optimization approach. *Transportation Research Part C: Emerging Technologies*, 165:104740, 2024.

[2] Laurie A. Garrow, Brian J. German, and Caroline E. Leonard. Urban Air Mobility: A comprehensive review and comparative analysis with autonomous and electric ground transportation for informing future research. *Transportation Research Part C: Emerging Technologies*, 132:103377, 2021.

[3] Adam P. Cohen, Susan A. Shaheen, and Emily M. Farrar. Urban Air Mobility: History, ecosystem, market potential, and challenges. *IEEE Transactions on Intelligent Transportation Systems*, 22(9):6074–6087, 2021.

[4] Stephen A. Rizzi, Dennis L. Huff, Donald D. Boyd, Paul Bent, Brenda S. Henderson, Kyle A. Pascioni, D. C. Sargent, D. L. Josephson, M. Marsan, H. He, and R. Snider. Urban Air Mobility noise: Current practice, gaps, and recommendations. Technical Report NASA/TP-2020-5007433, NASA, 2020.

[5] Zhenyu Gao, Alex Porcayo, and John-Paul Clarke. Developing virtual acoustic terrain for urban air mobility trajectory planning. *Transportation Research Part D: Transport and Environment*, 120:103794, 2023.

[6] Parker D. Vascik and R. John Hansman. Scaling constraints for Urban Air Mobility operations: Air traffic control, ground infrastructure, and noise. In *Proc. 2018 Aviation Technology, Integration, and Operations Conference*, AIAA Paper 2018-3849, 2018.

[7] Aleksandar Bauranov and Jasenka Rakas. Designing airspace for Urban Air Mobility: A review of concepts and approaches. *Progress in Aerospace Sciences*, 125:100726, 2021.

[8] Mathias Basner, Charlotte Clark, Anna Hansell, James I. Hileman, Sabine Janssen, Kevin Shepherd, and Victor Sparrow. Aviation noise impacts: State of the science. *Noise and Health*, 19(87):41–50, 2017.

[9] Jin Li, Hok K. Ng, Yifeng Zheng, and Sergio Gutierrez-Nolasco. Noise exposure maps for Urban Air Mobility. In *Proc. AIAA AVIATION 2021 Forum*, AIAA Paper 2021-3203, 2021.

[10] Farhan Yunus, Damiano Casalino, Francesco Avallone, and Daniele Ragni. Efficient prediction of Urban Air Mobility noise in a vertiport environment. *Aerospace Science and Technology*, 139:108410, 2023.

[11] Y. Wu, S. Deniz, Z. Wang, and D. Huang. A convex approach to high-fidelity landing trajectory optimization for Advanced Air Mobility. In *Proc. AIAA SCITECH 2024 Forum*, AIAA Paper 2024-2484, 2024.

[12] Martin Waltz, Ostap Okhrin, and Michael Schultz. Self-organized free-flight arrival for Urban Air Mobility. *Transportation Research Part C: Emerging Technologies*, 167:104806, 2024.

[13] Younghoon Kim, Jeongwoo Ko, Kukhwan Yu, and Soogab Lee. Low-noise trajectory optimization of Urban Air Mobility in the urban environment using deep reinforcement learning. *Journal of the Acoustical Society of America*, 158(1):476–494, 2025.

[14] Jeonghoon Mo and Jean Walrand. Fair end-to-end window-based congestion control. *IEEE/ACM Transactions on Networking*, 8(5):556–567, 2000.

[15] Yue Yu, Zhenyu Gao, Sarah H. Q. Li, Qinshuang Wei, John-Paul Clarke, and Ufuk Topcu. Alpha-fair routing in Urban Air Mobility with risk-aware constraints. In *Proc. 2024 American Control Conference*, pages 2248–2253, 2024.

[16] Stephen A. Rizzi, S. J. Letica, Donald D. Boyd, and Leonard V. Lopes. Prediction of noise-power-distance data for Urban Air Mobility vehicles. *Journal of Aircraft*, 61(1):166–182, 2024.

[17] Tuomas Haarnoja, Aurick Zhou, Pieter Abbeel, and Sergey Levine. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In *Proc. 35th International Conference on Machine Learning*, pages 1861–1870, 2018.

[18] Andrew Y. Ng, Daishi Harada, and Stuart Russell. Policy invariance under reward transformations: Theory and application to reward shaping. In *Proc. 16th International Conference on Machine Learning*, pages 278–287, 1999.