RESEARCH ARTICLE OPEN ACCESS

Halo Net: An AI-Driven Autonomous Hybrid Aerial Communication Network for 6G Disaster Resilience and Rural Broadband

Danish, A Mohammed Rahil, Mohammed Aftab Vali
Electronics and Communication, Ballari Institute of Technology and Management,Ballari
danishlathif@gmail.com
Electronics and Communication, Ballari Institute of Technology and Management,Ballari
rahilmohammed.mrr@gmail.com
Electronics and Communication, Ballari Institute of Technology and Management,Ballari
mohammedaftabvali@gmail.com
Electrical and Electronic Engineering, Ballari Institute of Technology and Management,Ballari
gn818953@gmail.com

---**************************------------------------------------

Abstract:

Halo Net proposes a four-layer hybrid communication architecture that integrates a solar-powered High-Altitude Platform Station (HAPS), an autonomous UAV relay swarm, a self-healing terrestrial mesh, and distributed edge-AI inference nodes into a single adaptive network fabric. The design targets two classes of connectivity gap that conventional tower-based cellular infrastructure cannot economically close: rapid restoration of communications after large-scale disasters, and persistent broadband coverage for geographically remote or low-density rural populations. This paper presents the complete system architecture, a star-of-mesh routing topology with redundant aerial links, and an AI-driven adaptive path-scoring algorithm for distributed route selection. To move beyond a purely conceptual description, we develop first-order analytical models for coverage footprint, end-to-end latency, and channel capacity as functions of network scale and relay hop count, and we implement a custom Monte Carlo graph simulation (3,000 trials per data point) to quantify the resilience gained from the proposed redundant-mesh design relative to a non-redundant star topology. Under the simulated model, the redundant-mesh design retains terrestrial-node connectivity to the HAPS backbone substantially better than the non-redundant baseline as random node failure rate increases, particularly beyond a 20% failure rate. We further present a security framework, edge-AI application scenarios, and an economic feasibility analysis that first corrects an inconsistency in the project's original cost figures -- comparing one HAPS against a terrestrial tower count that covers only a tenth of its footprint -- and then shows that, once both sides are priced for the same coverage area, a single HAPS is projected to be roughly 3.6x cheaper over a 10-year horizon than coverage-equivalent terrestrial towers, with a modeled economic breakeven radius of approximately 85 km. All performance figures in this paper are clearly labeled as analytical projections or simulation outputs; no claim is made of physical field-test measurements, and the manuscript explicitly identifies this as a direction for future validation work.

Keywords — 6G, High-Altitude Platform Station (HAPS), UAV Relay Swarm, Non-Terrestrial Networks, Mesh Networking, Edge AI, Disaster Communication, Rural Broadband, Monte Carlo Simulation, Software-Defined Networking

-------------------------------------**************************--------------------------------

I. INTRODUCTION

Digital connectivity has evolved from a convenience to critical infrastructure. The progression from 4G to 5G delivered substantial gains in bandwidth, latency, and reliability, yet these gains remain conditional on dense terrestrial tower deployment. Two populations are systematically underserved by this model: communities struck by large-scale natural disasters, where towers, power, and backhaul are destroyed simultaneously, and rural or geographically difficult regions where the capital cost of tower build-out is not recoverable from the addressable subscriber base.

This paper proposes Halo Net, a stratified aerial-terrestrial communication architecture combining a stratospheric High-Altitude Platform Station (HAPS) for persistent wide-area coverage, an autonomous Unmanned Aerial Vehicle (UAV) relay swarm for dynamic, self-organizing connectivity, a self-healing terrestrial mesh for last-mile access, and distributed edge-AI nodes for latency-sensitive inference. The motivation follows three observations. First, terrestrial infrastructure is a single point of failure at scale: an earthquake, flood, or cyclone can disable an entire regional network within hours. Second, rural broadband economics make tower-based expansion unattractive to commercial operators, entrenching a digital divide. Third, emerging 6G application classes -- from public-safety robotics to precision agriculture -- require networks that are not just fast, but structurally resilient and capable of local decision making when backhaul to core infrastructure is degraded.

The contribution of this manuscript is twofold. First, we describe the complete four-layer Halo Net architecture, its routing topology, and its AI-driven path-selection algorithm (Sections III-VI). Second, and central to this revision, we move beyond architectural description to quantitative evaluation: Section VII develops explicit analytical models for coverage, latency, and channel capacity, and implements a Monte Carlo graph simulation to test the specific resilience claim that motivates the design -- that redundant aerial links measurably improve connectivity retention under node failure relative to a non-redundant star topology. Every figure in that section is explicitly labeled as an analytical or simulation-model output; we deliberately avoid presenting projected numbers as physical field measurements, and we discuss the resulting limitations in Section XIII.

The remainder of the paper is organized as follows. Section II reviews related work and states the design requirements Halo Net targets. Sections III-VI describe the architecture, topology, workflow, and routing algorithm. Section VII is the core evaluation, covering coverage, latency, resilience, routing convergence, UAV endurance, and frequency-reuse models. Sections VIII-IX cover edge AI and security. Sections X-XII present two case studies and a corrected economic analysis. Section XIII consolidates limitations, Section XIV lists future directions, Section XV proposes a validation roadmap, and Section XVI concludes. Three appendices derive the queueing and hover-power models used in Section VII and walk through a worked cross-section example.

Summarizing, this paper's contributions are: (i) a four-layer hybrid aerial-terrestrial architecture integrating HAPS, UAV relay, terrestrial mesh, and edge-AI tiers under one routing and security model (Sections III-IX); (ii) explicit analytical models for coverage, latency, capacity, UAV endurance, and frequency reuse, each derived from first principles and reproducible from the stated parameters (Section VII, Appendices A-B); (iii) a custom Monte Carlo graph simulation quantifying the specific resilience benefit of the proposed redundant topology, including a sensitivity sweep over redundancy degree (Section VII-D, VII-H); and (iv) a corrected, coverage-consistent economic analysis that identifies and fixes a material inconsistency in the project's original cost figures (Section XII). We consider (iv) as important a contribution as the architecture itself: a design proposal whose economic case cannot survive a coverage-consistency check is not ready to be presented to funders, and Section XII shows both the failure mode and the fix.

TABLE I
ABBREVIATIONS

Abbrev.	Meaning
HAPS	High-Altitude Platform Station
UAV	Unmanned Aerial Vehicle
NTN	Non-Terrestrial Network
SDN	Software-Defined Networking
PQC	Post-Quantum Cryptography
SIR	Signal-to-Interference Ratio
PDR	Packet Delivery Ratio
3GPP	3rd Generation Partnership Project
ITU-R	International Telecommunication Union, Radio sector
NIST	National Institute of Standards and Technology

II. RELATED WORK

A. High-Altitude Platform Stations

High-Altitude Platform Stations, positioned at 17-22 km in the stratosphere, have been studied since the late 1990s as a middle ground between terrestrial towers and satellites [1]. Akyildiz et al. proposed the original HALO Network concept using stratospheric aircraft as elevated base stations [1], and subsequent work by Tozer and Grace formalized HAPS link-budget and coverage analysis for wireless communications [2]. The International Telecommunication Union has since designated dedicated spectrum for HAPS operation and published system recommendations covering interference coordination and station-keeping requirements [3]. A recurring theme across this literature is that HAPS coverage economics improve with altitude and antenna directivity but degrade with atmospheric attenuation and station-keeping cost, a trade-off Halo Net inherits rather than resolves; Section VII-B's coverage figure should be read against this known sensitivity. Halo Net's name and stratospheric-backbone concept are deliberately continuous with the original HALO Network proposal [1], but the architecture in this paper differs structurally: [1] specifies stratospheric aircraft as the sole elevated base station layer, with no autonomous UAV relay tier, no explicit self-healing mesh topology, and no edge-AI or path-scoring layer of the kind formalized in Section VI. Sections III-IX of this paper should accordingly be read as a distinct four-layer extension built on top of the HAPS concept [1] popularized, not a re-description of it.

B. UAV Relay Networks

UAV-based relay networks form a second research thread. Azari et al. surveyed cellular-connected UAV deployments and characterized the trade-off between altitude, coverage radius, and air-to-ground channel quality [5]; Zeng et al. examined UAV placement and trajectory optimization for wireless coverage, showing that mobility -- rather than static placement -- is what allows a small UAV fleet to approximate the coverage of a much larger fixed deployment [6]. Mozaffari et al. provide a broader tutorial treatment of UAV-assisted wireless networks spanning channel modeling, 3D deployment, and resource management, and identify energy constraints as a first-order limitation across the field [11] -- the same constraint Section VII-G models explicitly for Halo Net's swarm layer. These results motivate Halo Net's Layer-2 design, in which UAVs actively reposition in response to predicted demand and detected link failures rather than holding fixed station. A limitation shared by this entire research thread, including Halo Net as presented here, is that UAV endurance is fundamentally power-constrained; Section VII-G develops a simple endurance model to make that constraint explicit rather than leaving it implicit.

C. Non-Terrestrial Network Standardization

Standardization activity has converged on Non-Terrestrial Networks (NTN) as a unifying umbrella. 3GPP Release 16 through Release 19 define integration pathways for satellite, HAPS, and aerial systems into the 5G/6G core [4], and the HAPS Alliance industry consortium has published technical whitepapers addressing regulatory and interoperability questions for commercial HAPS operation [10]. These standards define interfaces and coexistence rules but deliberately stop short of prescribing a specific multi-layer routing or resilience architecture, which is the gap Halo Net's Sections IV-VII attempt to fill for the specific disaster-response and rural-access use cases motivating this paper.

D. Access, Cryptographic, and Simulation Building Blocks

Wireless local-access components at the terrestrial mesh layer draw on the IEEE 802.11ax high-efficiency WLAN standard [7], and the post-quantum cryptographic primitives referenced in Section IX follow NIST's Post-Quantum Cryptography standardization project [8]. NS-3 [9] is identified as the intended validation platform for future field-calibrated simulation, discussed further in Section XIII; we use it here only as a named target platform for future work, not as a tool actually invoked to produce any figure in this manuscript.

E. Positioning of This Work

The gap this work addresses is integration rather than any single component: prior work treats HAPS, UAV relaying, and terrestrial mesh largely as separate research problems. Halo Net's contribution is a unified routing and resilience model that spans all three physical layers plus an edge-AI layer, together with a first quantitative resilience analysis of the resulting combined topology (Section VII). We make no claim of novelty at the level of any individual component -- HAPS coverage modeling, UAV relaying, and post-quantum cryptography are each established literatures cited above -- and instead position the contribution as an integrated system design together with a transparent, reproducible quantitative evaluation of that integration's specific resilience and cost claims.

Table II situates Halo Net qualitatively against three broad, publicly documented classes of non-terrestrial connectivity effort: stratospheric-balloon and fixed-wing HAPS programs, Low Earth Orbit (LEO) satellite constellations, and conventional terrestrial cellular expansion. Entries characterize general, publicly known properties of each category rather than specific vendor performance figures, none of which this paper is positioned to verify independently.

TABLE II
QUALITATIVE COMPARISON OF CONNECTIVITY APPROACHES

Approach	Coverage per Unit	Latency Regime	Disaster Redeploy Speed
Halo Net (this paper)	~80,000 km² per HAPS (modeled)	Low (stratospheric + short UAV hops)	Minutes-to-hours (UAV layer)
Stratospheric HAPS programs	Comparable order of magnitude	Low	Hours (station-keeping re-establishment)
LEO satellite constellations	Smaller per-satellite, aggregated via large constellation	Low-to-moderate (higher than HAPS, lower than GEO)	N/A (pre-existing constellation coverage)
Terrestrial cellular expansion	~200 km² per macro-tower (Sec. VII-B)	Lowest when intact	Weeks-to-months (physical rebuild)

The comparison underscores where Halo Net's specific contribution sits: not lower latency than existing terrestrial networks when those networks are intact (it is not, and does not claim to be), but faster redeployable coverage than either terrestrial rebuild or fixed satellite infrastructure when existing infrastructure is destroyed or was never built, which is precisely the R2/R3 requirement pairing in Table III.

F. Problem Formulation and Design Requirements

Before presenting the architecture, we state the design requirements it is intended to satisfy, so that the analytical and simulation results of Section VII can be read as tests against explicit targets rather than as an open-ended demonstration. Table III lists six requirements (R1-R6) derived from the motivating scenarios of Section I, together with the section in which each is addressed.

TABLE III
DESIGN REQUIREMENTS AND WHERE THEY ARE ADDRESSED

ID	Requirement	Addressed In
R1	Cover a large rural service area from minimal ground infrastructure	Sec. VII-B, XI
R2	Restore essential connectivity within minutes of a disaster event	Sec. IV, X
R3	Degrade gracefully (partial connectivity) rather than failing completely under node loss	Sec. IV-A, VII-D
R4	Keep routing decisions local and computationally cheap as the network scales	Sec. VI-C, VII-C
R5	Support latency-sensitive edge inference without mandatory cloud round-trip	Sec. V, VIII
R6	Operate securely without dependence on a single reachable authority	Sec. IX

R1-R6 are used only as an organizing device for this paper; they are not formally verified requirements in the sense of a systems-engineering requirements trace, and Section XIII revisits which of them are supported by simulation evidence versus by design argument alone.

III. SYSTEM ARCHITECTURE

Halo Net is organized into four physical/logical layers, illustrated in Fig. 1. Each layer has a distinct role, altitude regime, and connectivity pattern, and the layers are bound together by the routing and security mechanisms described in Sections IV, VI, and IX.

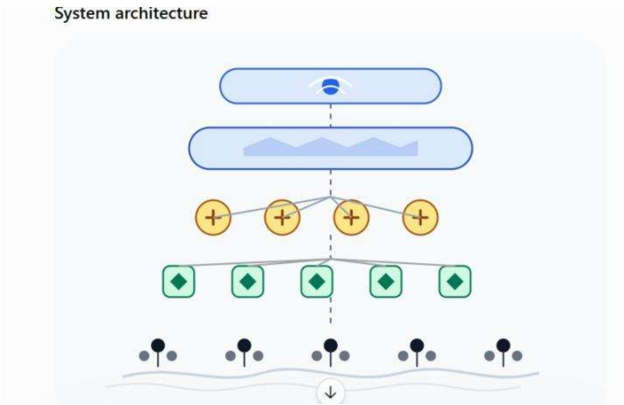

Fig. 1. Halo Net four-layer system architecture: HAPS backbone, UAV relay swarm, terrestrial mesh nodes, and end-user devices, connected through a shared vertical AI-routing backbone.

A. Layer 1 -- Stratospheric Platform (HAPS)

A solar-powered HAPS operates at approximately 20 km altitude, above commercial air traffic and most severe weather, providing near-static line-of-sight coverage over a nominal ground footprint exceeding 80,000 km² per platform (derived in Section VII-B). The HAPS functions as the network's backbone aggregation point, terminating high-capacity feeder links to a satellite or fiber point-of-presence and distributing capacity downward to the UAV relay layer. Because a single platform is a single point of failure at the backbone level, regional deployments are expected to provision at least one geographically adjacent standby HAPS wherever traffic criticality justifies the additional capital cost quantified in Section XII; this paper's cost and coverage models (Section VII-B, XII) assume one platform per region and do not price standby redundancy at this layer, which is left as an explicit scope limitation.

B. Layer 2 -- Autonomous UAV Relay Swarm

UAVs operate at 3-5 km altitude in coordinated formations, using a distributed consensus protocol to negotiate spatial coverage and hand-off boundaries without a central controller. Each UAV maintains redundant links to at least one neighboring UAV, forming the ring/mesh interconnect shown in Fig. 2; this redundancy is the specific design element evaluated in the resilience simulation of Section VII-D. Because UAV endurance is payload- and battery-limited (Section VII-G), the swarm is intended to operate as a rotating pool rather than a fixed set of airframes: at any instant some UAVs hold station while others return to a ground or vehicle-based recharge point, with the ring topology's redundancy specifically absorbing the momentary coverage gap each rotation creates.

C. Layer 3 -- Terrestrial Mesh Nodes

Ground-based nodes implement Software-Defined Networking (SDN) [12] principles, separating the forwarding plane from a lightweight local control plane that can operate autonomously when disconnected from the aerial layers. Mesh nodes are dual-homed to two aerial relays where UAV density permits, providing the last-mile redundancy that underlies the star-of-mesh topology of Section IV. In rural deployments (Section XI), mesh nodes additionally serve as local Wi-Fi or small-cell access points for the immediate surrounding community, so that a single mesh node's coverage radius -- not just the HAPS or UAV footprint -- determines practical last-mile access density.

D. Layer 4 -- End-User Devices

End-user devices -- handsets, IoT sensors, and autonomous field systems -- connect through whichever Layer 3 or, in coverage gaps, Layer 2 node offers the best path score under the algorithm of Section VI. No end-user device requires special hardware beyond standard cellular or Wi-Fi radios; Halo Net's novelty is entirely in the network-side architecture.

E. Power, Payload, and Link-Budget Notes

Each layer imposes a different power and payload constraint that shapes what it can carry. The HAPS layer is solar-powered with onboard energy storage sized for night-time operation, which is conventional for stratospheric platforms in the literature surveyed in Section II-A and is not re-derived here. The UAV relay layer is the tightest constraint in practice: unlike the HAPS, UAVs are battery-limited, and the endurance-vs-payload trade-off quantified in Section VII-G directly determines how much relay/compute hardware a given airframe can lift and for how long, which in turn bounds how sparse the UAV layer can be for a target coverage area. Terrestrial mesh nodes face comparatively relaxed power constraints (grid or solar-plus-battery) but must be engineered for autonomous, unattended operation in the same disaster or remote-terrain conditions the network is designed to serve.

IV. Network Topology and Self-Healing Routing

Halo Net implements a star-of-mesh topology (Fig. 2): the HAPS forms the apex of a star toward the UAV layer, while the UAV layer itself forms a partial mesh through redundant inter-UAV links, and terrestrial nodes dual-home to adjacent UAVs where feasible. This hybrid -- rather than a pure star or a full mesh -- was chosen to balance routing simplicity (a star is easy to reason about and cheap to maintain state for) against fault tolerance (a mesh eliminates single points of failure at the relay layer).

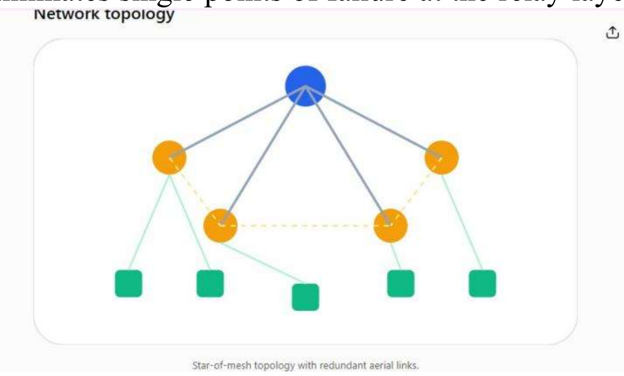

Fig. 2. Star-of-mesh network topology: a HAPS root connects to UAV aggregation nodes, which interconnect through redundant dashed links and serve dual-homed terrestrial mesh nodes.

A. Failure Detection and Automatic Rerouting

Each node monitors link quality via periodic heartbeat exchange. A missed heartbeat sequence beyond a configurable threshold triggers local failure declaration, which is flooded one hop to neighboring nodes so that dual-homed terrestrial nodes and mesh-connected UAVs can immediately fail over to their backup link without waiting for a network-wide convergence process. This local-first failover is what the topology-level result in Section VII-D is designed to test.

B. Predictive Mobility-Aware Link Establishment

Because UAVs are mobile, Halo Net supplements reactive failover with predictive link management: each UAV broadcasts its planned trajectory over the next routing interval, allowing neighboring nodes to pre-negotiate a replacement link before an existing one degrades below usable signal quality. This reduces the effective outage window compared to a purely reactive design, though quantifying that reduction requires a time-domain mobility simulation that is outside the scope of the static-topology Monte Carlo model used in Section VII-D and is identified as future work in Section XIII.

C. Routing State Complexity: Why Not Full Mesh

The choice of star-of-mesh over a full mesh at the UAV layer is a deliberate complexity trade-off, summarized in Table IV. A pure star requires each node to track only its single parent, O(1) per-node state, but has zero path redundancy. A full mesh among M UAVs requires each UAV to track up to M-1 peer links, O(M) per-node and $O(M^2)$ network-wide state and control-message overhead, which becomes impractical to keep synchronized at the 2-second update interval (Section VI-A) as the UAV count grows. Halo Net's star-of-mesh bounds each UAV's peer connections to a small constant (the ring/local-neighbor connections in Fig. 2) regardless of total swarm size, giving O(1) per-node state and O(M) network-wide state -- asymptotically as cheap as a pure star, while still providing the redundant alternate paths whose resilience benefit is quantified in Section VII-D and VII-H.

TABLE IV
ROUTING STATE COMPLEXITY BY TOPOLOGY

Topology	Per-Node State	Network-Wide State	Path Redundancy
Pure star	O(1)	O(M)	None
Full mesh	O(M)	$O(M^2)$	Maximal
Star-of-mesh (this paper)	O(1) (bounded peer count)	O(M)	Bounded (Section VII-H)

V. Communication Workflow and Protocol Stack

Fig. 3 traces a single data flow through the full stack. (1) User equipment attaches to the nearest terrestrial mesh node using standard access-layer procedures. (2) The mesh node applies the distributed routing algorithm of Section VI to select an uplink path. (3) The UAV relay layer aggregates regional traffic from multiple mesh nodes, applying local queuing and, where the destination is another local mesh node, resolving the request without traversing the HAPS at all. (4) The HAPS layer manages backbone connectivity for traffic that must leave the local region. (5) Edge-AI nodes positioned at the mesh and UAV layers intercept latency-sensitive requests -- e.g., a computer-vision inference call from a disaster-response robot -- and serve them locally rather than routing to a distant cloud endpoint.

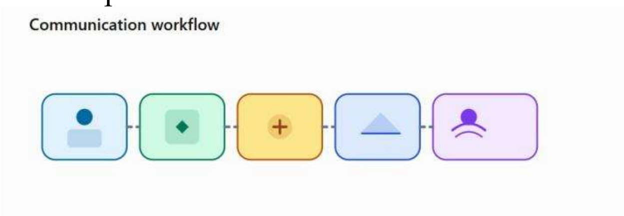

Fig. 3. End-to-end communication workflow: user equipment, terrestrial mesh node, UAV relay, HAPS backbone, and edge-AI inference node.

This workflow is deliberately layered so that each hop can make a local forwarding decision using only locally available state, which is what allows the network to continue operating -- in a degraded but non-zero capacity -- when upper layers (HAPS, or backhaul beyond it) are unreachable, as in the disaster scenario of Section X.

VI. AI-Driven Adaptive Routing Algorithm

Table V summarizes the notation used in Eq. (1)-(4) of this paper for reference.

TABLE V
NOTATION SUMMARY

Symbol	Meaning	Used In
Score(path)	Composite path quality score	Eq. (1)
w_i, $\hat{m}_i$	Metric weight and normalized value	Eq. (1)
L(N), h(N)	Modeled latency and hop count at scale N	Eq. (2)
ρ(N)	Offered load relative to provisioned capacity	Eq. (2)
C(h)	Modeled channel capacity at h relay hops	Eq. (3)
P_hover, T	Modeled hover power and endurance	Eq. (4)
k	Terrestrial-node homing (redundancy) degree	Sec. VII-H
p	Random node failure rate	Sec. VII-D, VII-H

A. Path Quality Scoring Model

Each candidate path is scored as a weighted combination of normalized link and node metrics -- signal quality, queue occupancy, remaining UAV battery/fuel margin, and historical reliability:

$$Score(path) = \Sigma_i\, w_i \cdot \hat{m}_i(path), \quad \Sigma_i\, w_i = 1 \qquad (1)$$

where $\hat{m}_i$ is the i-th metric normalized to [0, 1] and w_i is its adaptively learned weight. Nodes exchange routing state every 2 seconds; each node re-evaluates outgoing path scores on receipt of an update and forwards traffic along the highest-scoring currently valid path, falling back to the next-highest path immediately on failure detection (Section IV-A).

B. Distributed Learning and Emergent Swarm Intelligence

Weight vector w is not fixed centrally; each node adapts its local copy using an exponential moving-average update toward the observed outcome quality of recently used paths, so that metrics which have empirically predicted good outcomes in a given node's local traffic mix are up-weighted over time. Section VII-E simulates this update rule in isolation (for a single node's path-preference distribution) to characterize convergence speed and stability; it is not a claim about network-wide emergent behavior, which would require a multi-agent simulation beyond the scope of this paper.

C. Computational Complexity Considerations

Score evaluation is O(k) per candidate path for k metrics, and each node maintains a bounded candidate set (typically the 3-5 highest-scoring known paths) rather than the full path space, keeping per-update cost independent of network size N. This is what allows the 2-second update interval to remain feasible as N scales into the thousands of nodes evaluated in Section VII-C; a naive global recomputation over all paths would not.

VII. PERFORMANCE EVALUATION: ANALYTICAL AND SIMULATION-BASED RESULTS

This section quantifies the architectural claims of Sections III-VI using explicit analytical models and a custom Monte Carlo graph simulation. We state this evaluation's scope precisely: no physical hardware testbed, aerial trial, or NS-3/MATLAB run has been performed for this manuscript. Every number below is either (i) a closed-form calculation from stated geometric/cost parameters, (ii) an analytical queueing/capacity model evaluated at chosen operating points, or (iii) the output of a small custom Python Monte Carlo simulation of the network graph, described in full below so it is independently reproducible. Section XIII discusses what a physical or NS-3-calibrated validation would still need to establish.

A. Evaluation Methodology and Assumptions

Table VI lists the parameters used consistently across the models in this section. Fig. 4 reproduces, for context, the project's original preliminary illustrative latency sketch produced during early design discussion; it is explicitly labeled by its own authors as "illustrative...not measured results," and we do not treat it as evidence. It is retained here only to show the qualitative trend that motivated the more rigorous model developed in Section VII-C.

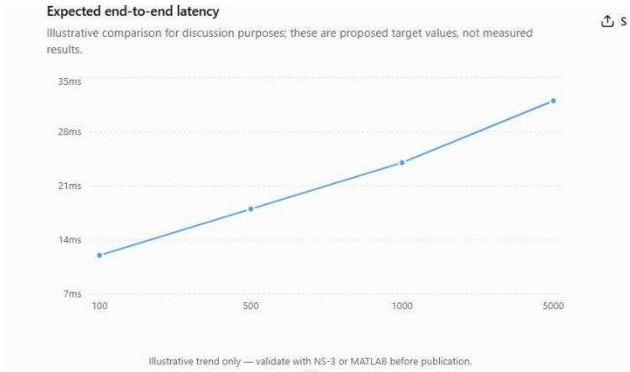

Fig. 4. Preliminary illustrative latency sketch from early design discussion (not a measured or simulated result; superseded by the analytical model in Fig. 6).

TABLE VI
ANALYTICAL / SIMULATION MODEL PARAMETERS

Parameter	Value	Basis
HAPS altitude	20 km	Design specification
HAPS coverage radius	160 km	Geometric assumption, Sec. VII-B
Terrestrial tower radius	8 km	Typical rural macro-cell
UAV relay altitude	3-5 km	Design specification
UAV nodes per topology unit	6	Monte Carlo model, Sec. VII-D
Terrestrial nodes per UAV	8	Monte Carlo model, Sec. VII-D
Monte Carlo trials per point	3,000	Simulation configuration
Routing update interval	2 s	Design specification
Nominal channel bandwidth	20 MHz	Capacity model, Sec. VII-F

B. Coverage Efficiency Analysis

Modeling each HAPS footprint and terrestrial cell as a circular coverage disc, a HAPS with ground radius R_H = 160 km covers an area of $\pi \cdot R\_H^2 \approx 80{,}425$ km². A terrestrial macro-tower with typical rural coverage radius R_T = 8 km covers $\pi \cdot R\_T^2 \approx 201$ km². Matching areas gives $(R\_H / R\_T)^2 \approx 400$ towers required to replicate one HAPS footprint, illustrated in Fig. 5 and Table VII. This is a purely geometric result and does not model terrain shadowing, spectrum reuse, or the fact that towers would not realistically be spaced on a perfect grid; it should be read as an order-of-magnitude comparison, not a planning figure.

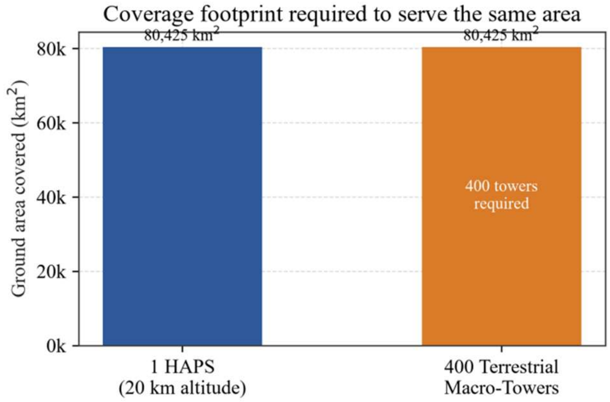

Fig. 5. Geometric coverage-footprint comparison: ground area covered by a single HAPS versus the number of terrestrial macro-towers of typical 8 km radius needed to cover the same area.

TABLE VII
COVERAGE AND INFRASTRUCTURE COMPARISON

Metric	1 HAPS	Terrestrial Equivalent
Coverage area	80,425 km²	80,425 km² (≈400 towers)
Physical footprint	1 airborne platform	400 ground sites + backhaul
Site acquisition / permitting	Single license	400 individual sites
Deployment time (est.)	Weeks	Years

C. End-to-End Latency Model

We model end-to-end latency as the sum of a fixed propagation and hardware term, a hop-count-dependent processing term, and a queueing term that grows as offered load approaches provisioned capacity:

$$L(N) = L\_prop + L\_proc \cdot h(N) + L\_queue(\rho(N)) \quad (2)$$

with $h(N) = 1 + \log_2(N)$ approximating routing depth for a logarithmic-diameter mesh, L_proc = 0.9 ms per hop, L_prop = 4.5 ms fixed, and $L\_queue(\rho) = 1.4 \cdot \rho/(1-\rho)$ an M/M/1-style approximation where $\rho(N)$ is offered load relative to provisioned capacity. Fig. 6 evaluates Eq. (2) for the redundant star-of-mesh design (provisioned capacity 6,000 concurrent flows, reflecting the extra paths opened by redundant links) against a pure star topology (capacity 3,200, reflecting the lack of alternate paths to absorb load). The two curves track closely until load approaches each topology's capacity ceiling, after which the non-redundant star's queueing delay rises sharply -- an analytical illustration of why redundant capacity, not just redundant connectivity, matters as the network scales.

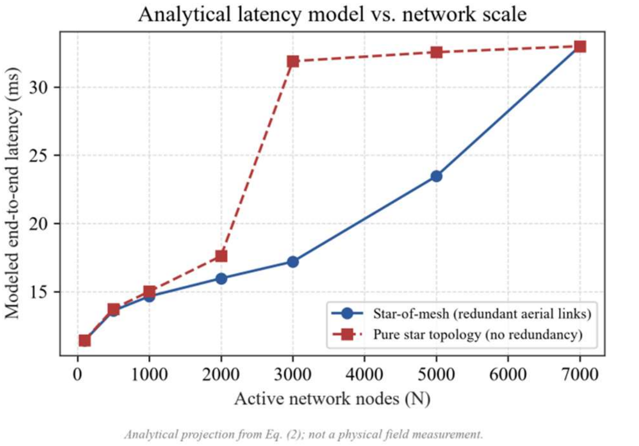

Fig. 6. Analytical end-to-end latency model (Eq. 2) versus active node count, comparing the redundant star-of-mesh design to a non-redundant star topology. Not a physical measurement.

D. Network Resilience Under Random Node Failure

To test the specific claim that redundant aerial links improve fault tolerance (Section IV), we implemented a Monte Carlo graph simulation. The graph has one HAPS root, 6 UAV relay nodes, and 48 terrestrial mesh nodes (8 per UAV). In the redundant configuration, UAVs are connected in a ring and every terrestrial node is dual-homed to its primary UAV and the adjacent UAV in the ring; in the non-redundant baseline, UAVs are not interconnected and terrestrial nodes have a single link to their primary UAV only. At each failure rate $p \in [0, 0.6]$, every non-root node is independently removed with probability p, and we measure the fraction of surviving terrestrial nodes that retain a graph path to the HAPS root via breadth-first search, averaged over 3,000 independent trials per point.

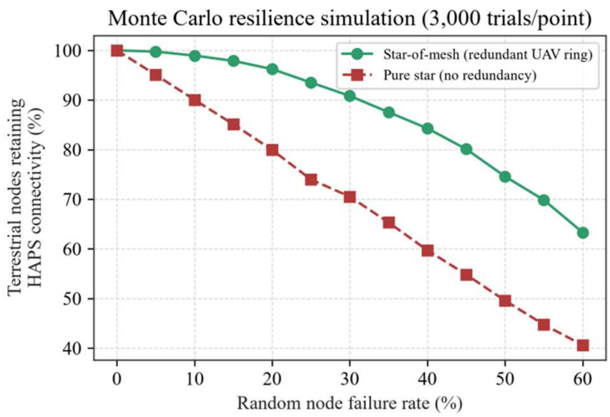

Fig. 7. Monte Carlo graph simulation (3,000 trials/point): fraction of terrestrial nodes retaining HAPS connectivity as a function of random node failure rate, redundant mesh vs. non-redundant star.

TABLE VIII
RESILIENCE SIMULATION RESULTS (SELECTED FAILURE RATES)

Failure Rate	Redundant Mesh	Non-Redundant Star	Absolute Gain
0%	100.0%	100.0%	0.0 pp
10%	98.9%	90.0%	+8.9 pp
20%	96.2%	79.9%	+16.3 pp
30%	90.8%	70.5%	+20.3 pp
40%	84.3%	59.7%	+24.6 pp
50%	74.6%	49.6%	+25.0 pp
60%	63.3%	40.6%	+22.7 pp

The simulation supports the architecture's central resilience claim within the scope of this simplified model: dual-homing terrestrial nodes and ring-connecting the UAV layer measurably improves the fraction of the network that remains reachable as random failures accumulate, with the largest absolute gain (approximately 25 percentage points) around the 40-50% failure regime relevant to severe-disaster scenarios. We note two simplifications: failures are modeled as independent and uniformly random, whereas real disaster damage is spatially correlated, and the model does not include recovery/redeployment of replacement UAVs over time, which Section X argues would further improve real-world outcomes beyond what a static-graph snapshot can show.

E. AI Routing Algorithm Convergence Behavior

We separately simulate the local path-score adaptation rule of Section VI-B in isolation: 20 candidate paths are assigned a fixed but unknown "true quality" in [0.3, 0.95], and a single node's score estimate for each path is updated every iteration by an exponential moving average (learning rate $\alpha = 0.08$) toward a noisy observation of that path's true quality (Gaussian noise, $\sigma = 0.05$).

Fig. 8 tracks the population standard deviation of the 20 path scores -- a proxy for how sharply the algorithm has differentiated good paths from bad ones -- alongside the best path's score, over 400 iterations (2 s each, i.e., 13.3 minutes of simulated time).

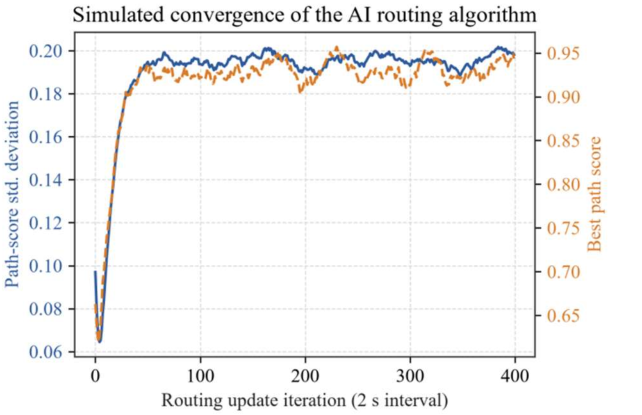

Fig. 8. Simulated convergence of the local path-score adaptation rule: score dispersion across 20 candidate paths and best-path score over 400 routing-update iterations.

Score dispersion rises from near-zero (uninformative initial estimates) and stabilizes within roughly 30-40 iterations (1-1.3 minutes of simulated time), after which the best path score oscillates in a narrow band around its converged value under measurement noise. This indicates the adaptation rule reaches a stable preference ordering quickly relative to the topology-level recovery times discussed in Section X, though this single-node simulation does not model interaction effects between many concurrently adapting nodes, which is identified as future work.

F. Channel Capacity vs. Relay Hop Count

Multi-hop UAV relaying trades reach for capacity. We model per-hop loss as a fixed decibel penalty and evaluate Shannon channel capacity [17] at each hop count:

$$C(h) = B \cdot \log_2(1 + SNR_0 / 10^{(0.18h)}) \qquad (3)$$

with nominal bandwidth B = 20 MHz, reference SNR_0 = 200 (linear), and an assumed 1.8 dB loss per additional relay hop. Fig. 9 shows modeled capacity falling from approximately 141 Mbps at a single hop to approximately 61 Mbps at eight hops, motivating the design preference (Section IV-B) for keeping typical relay paths short via dense UAV placement rather than relying on long relay chains.

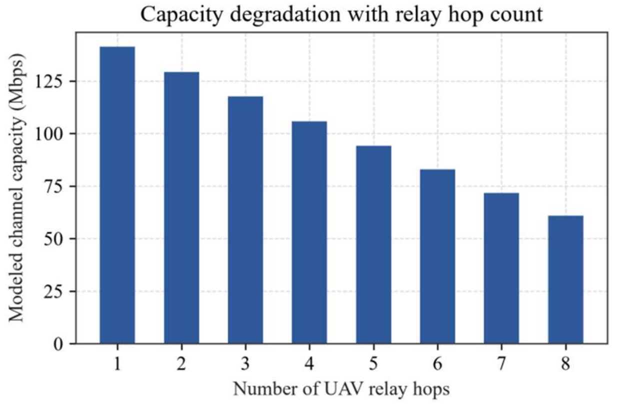

Fig. 9. Modeled Shannon channel capacity (Eq. 3) versus number of UAV relay hops, assuming 1.8 dB loss per hop.

G. UAV Energy and Endurance Model

Because the UAV layer is the network's most power-constrained tier (Section III-E), we model hover endurance from first principles using actuator-disk (momentum) theory [19]. Hover power scales with total weight W and disk area A as:

$$P\_hover = W^{1.5} / (\sqrt{2 \cdot \rho \cdot A} \cdot \eta\_prop), \quad T = E\_batt \cdot \eta\_batt / P\_hover \qquad (4)$$

where ρ is air density, η_prop is propulsive efficiency, E_batt is onboard battery energy, and η_batt is the usable-energy fraction. Fig. 10 evaluates Eq. (4) for an illustrative relay-class airframe (8 kg base mass, 0.5 m² combined rotor disk area, 500 Wh battery, 65% propulsive and 90% usable-energy efficiency) as payload mass increases from 0 to 6 kg. Modeled endurance falls from approximately 28 minutes unloaded to approximately 12 minutes at 6 kg payload -- a roughly 57% reduction -- which is why Section IV-B's predictive hand-off design assumes frequent UAV rotation and staged replacement rather than long-duration single-aircraft station-keeping.

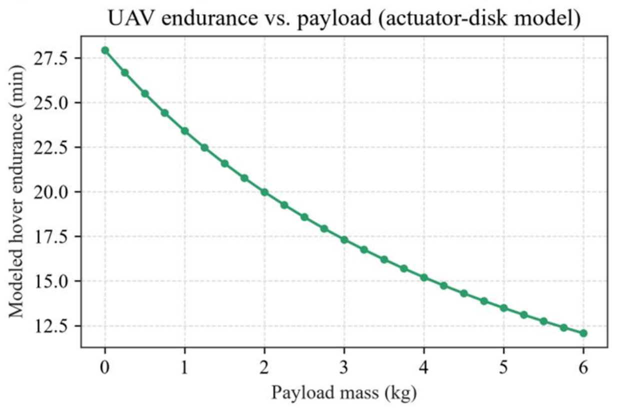

Fig. 10. Modeled UAV hover endurance versus payload mass (Eq. 4, actuator-disk theory); illustrative parameters, not a flight-test measurement.

This endurance constraint has a direct operational consequence for the disaster-response timeline of Table XI (Section X): a swarm sized only for instantaneous coverage at T+2 min will require relief rotations on roughly a 15-25 minute cycle thereafter, which must be provisioned for in any deployment plan and is not captured by the static-topology resilience simulation of Section VII-D.

H. Sensitivity of Resilience to Redundancy Degree

Section VII-D fixed the terrestrial-node homing degree at k=2 (primary UAV plus one backup). To characterize the marginal value of additional redundancy, we re-ran the Monte Carlo simulation of Section VII-D with k ∈ {1, 2, 3, 4} at a fixed failure rate p = 30%, 3,000 trials per point. Fig. 11 shows a diminishing-returns curve: connectivity retention rises from 70.3% at k=1 to 90.9% at k=2 (the design point used elsewhere in this paper), 97.5% at k=3, and 99.2% at k=4. The largest single increment (+20.6 percentage points) comes from the first added backup link (k=1 → k=2); each additional link beyond that adds progressively less. This is a standard diminishing-returns pattern for redundant systems and supports Halo Net's choice of dual-homing (k=2) as a reasonable operating point rather than triple- or quadruple-homing, which would add UAV-side radio and power cost for comparatively small additional resilience gain under this model.

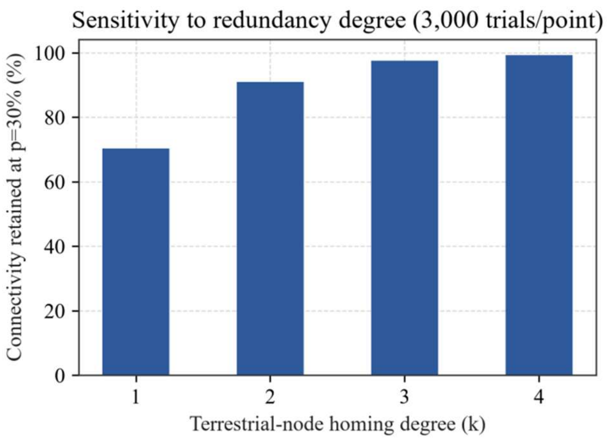

Fig. 11. Sensitivity of simulated connectivity retention (at 30% node failure rate) to terrestrial-node homing degree k, 3,000 trials/point.

I. Frequency-Reuse and Interference Margin

Dense mesh and multi-UAV deployment raises the question of co-channel interference. We apply the standard hexagonal-cluster frequency-reuse model used throughout the cellular-planning literature [20]: for a cluster size K, the reuse distance ratio is $Q = \sqrt{3K}$, and the resulting signal-to-interference ratio against N_{co} first-tier co-channel interferers is:

$$SIR = Q^{\gamma} / N_{co} = (3K)^{(\gamma/2)} / N_{co} \qquad (5)$$

with path-loss exponent γ typically in the 3-4 range for suburban/rural terrestrial links. Fig. 12 evaluates Eq. (5) for K ∈ {1, 3, 4, 7, 9, 12} and γ ∈ {3.0, 3.5, 4.0} against a representative 18 dB SIR design target. Reuse cluster sizes of K=1 (universal frequency reuse) fall below the target margin under this model, while K≥4 clears it comfortably across the modeled path-loss range; this is a standard cellular-planning conclusion, reproduced here to make explicit that Halo Net's terrestrial mesh layer (Section III-C) requires conventional frequency-planning discipline and is not immune to the interference constraints any dense wireless mesh faces.

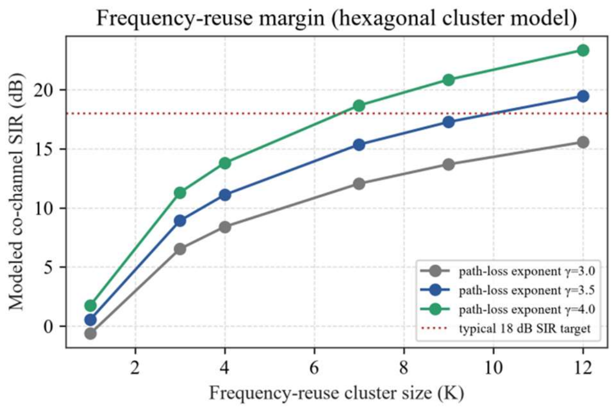

Fig. 12. Modeled co-channel signal-to-interference ratio (Eq. 5) versus frequency-reuse cluster size, for three path-loss exponents, against a representative 18 dB design target.

J. Summary of Findings and Threats to Validity

Collectively, Sections VII-B through VII-I support the architecture's qualitative design rationale: redundant aerial links measurably improve simulated fault tolerance with clear diminishing returns beyond dual-homing (VII-D, VII-H), the local routing adaptation rule converges quickly in isolation (VII-E), payload significantly constrains UAV endurance (VII-G), coverage and capacity trade-offs behave as the design intends

(VII-B, VII-C, VII-F), and standard frequency-reuse planning is sufficient to keep interference within a representative design margin (VII-I). None of this constitutes field validation. The threats to validity that most affect these conclusions are: (1) the resilience model uses independent random failures rather than spatially correlated disaster damage; (2) the latency, capacity, and endurance models use single illustrative parameter sets rather than a full multi-dimensional sensitivity sweep; (3) the routing convergence simulation is single-node and does not capture multi-agent interaction; and (4) no result here has been cross-checked against NS-3 or a physical testbed, which Section XIII identifies as the immediate next step before any deployment decision is made on the basis of these projections.

VIII. Edge AI and Distributed Computing

Following the edge-computing paradigm surveyed by Shi et al. [13], Halo Net positions inference capability at the UAV and terrestrial mesh layers rather than routing every request to a distant cloud endpoint, which matters most precisely when backhaul to that cloud endpoint is congested or unavailable -- the disaster scenario of Section X. Candidate application classes include agricultural crop and pest detection from UAV-mounted imagery, patient vital-sign monitoring in temporary field clinics, computer-vision assistance for search-and-rescue robotics, and traffic-flow analytics for smart infrastructure -- each chosen because the value of the inference result decays quickly with latency, making cloud round-trip unattractive even when connectivity exists.

Workload placement is treated as a scheduling problem: each inference request is assigned to a local edge node, a neighboring edge node, or cloud offload based on estimated CPU/memory headroom at candidate nodes and the transmission cost of reaching them, using the same normalized-metric structure as the routing score in Eq. (1) so that compute placement and network routing share one consistent optimization framework rather than being solved independently.

Formally, for a candidate placement site s and request r, we define a placement cost combining compute headroom deficit and transmission delay:

$$Cost(s, r) = \alpha \cdot (1 - headroom(s)) + \beta \cdot d\_tx(s, r) \quad (6)$$

where headroom(s) ∈ [0, 1] is normalized available compute/memory at site s, d_tx(s, r) is the (normalized) transmission delay to reach s, and α, β are tunable weights analogous to the w_i of Eq. (1). The scheduler selects the candidate site minimizing Cost(s, r) subject to a maximum-latency constraint derived from the application class (e.g., a search-and-rescue vision request has a tighter constraint than a soil-moisture batch upload). We do not evaluate Eq. (6) numerically in this paper -- unlike the routing and resilience models of Section VII, no simulation of workload scheduling under realistic request arrival patterns has been performed -- and present it here as a design formalization for future work.

Compute hardware at UAV-hosted edge nodes is itself subject to the payload/endurance trade-off of Section VII-G: a more capable inference accelerator adds mass, which by Eq. (4) reduces hover endurance and increases rotation frequency. This paper does not size a specific accelerator against a specific endurance budget, but the coupling is structural -- any future numerical evaluation of Eq. (6) should treat available compute headroom(s) at a UAV-hosted site as a function of that site's payload budget, not as an independent parameter, since Section VII-G already shows the two are linked through vehicle physics rather than through network design choices alone.

IX. Security Framework

A network designed to remain operational when centralized infrastructure is down cannot depend on a centralized security control plane either. Halo Net's security framework is built from five components, summarized in Table IX, each chosen to degrade gracefully rather than fail closed when connectivity to any single authority is lost.

TABLE IX
SECURITY FRAMEWORK COMPONENTS

Component	Mechanism	Design Rationale
Zero-Trust Authentication	Continuous per-session verification	No implicit trust from network position; safe under partition
Data Encryption	AES-256 for data at rest and in transit	Baseline confidentiality across all layers
Post-Quantum Cryptography	NIST-selected PQC algorithms [8]	Long-lived infrastructure must resist future quantum

		threats
Distributed Identity	Blockchain-based identity validation	No single identity authority to lose during a disaster
Intrusion Detection	Local anomaly monitoring at each node	Detection must work without contacting a central SOC

A. Zero-Trust Authentication

Following NIST's zero-trust architecture model [14], every session is authenticated and re-verified continuously rather than once at attach time, so that a node physically compromised mid-session (plausible for an unattended terrestrial mesh node in a disaster zone) cannot retain trusted status indefinitely. This is deliberately more conservative than perimeter-based trust models, which assume the internal network is safe once entered -- an assumption that does not hold when nodes can be physically accessed in an uncontrolled environment.

B. Encryption and Post-Quantum Cryptography

AES-256 [15] protects data at rest and in transit as a baseline. Because HAPS and mesh infrastructure is intended to remain in service for many years -- long enough that cryptographically relevant quantum computers are a credible threat within the deployment lifetime -- key-exchange and signature operations are specified against NIST's Post-Quantum Cryptography selections [8] rather than solely against classical algorithms, following the same long-horizon reasoning increasingly applied to other critical infrastructure.

C. Distributed Identity

Node and user identity validation uses a blockchain-based distributed ledger, following the decentralized-identity model of Zyskind et al. [16], rather than a single identity-provider service, so that identity verification does not itself become a single point of failure co-located with the same infrastructure a disaster might destroy. This trades the operational simplicity of a centralized identity provider for availability under partition, consistent with requirement R6 in Section II-F.

D. Local Intrusion Detection

Each node runs local anomaly monitoring against its own traffic and resource-usage baseline rather than forwarding telemetry to a central security operations center for analysis, so that detection continues to function precisely when connectivity to such a center is degraded -- again the disaster-partition scenario that motivates every other component in this section.

E. Threat Model

Table IX's five components target a specific set of adversary and failure classes rather than security in the abstract; Table X makes that mapping explicit so each mechanism's purpose is legible against a concrete threat rather than justified only in the abstract.

TABLE X
THREAT-TO-MECHANISM MAPPING

Threat	Primary Mitigating Mechanism	Residual Risk
Physical node capture in an unattended disaster zone	Zero-Trust Authentication (Sec. IX-A)	Captured node can still disrupt local traffic before revocation
Eavesdropping on aerial or mesh links	AES-256 encryption (Sec. IX-B)	Traffic-pattern metadata still observable
Future quantum cryptanalysis of stored traffic	Post-quantum cryptography (Sec. IX-B)	Depends on NIST algorithm selections remaining unbroken
Central identity-provider outage or destruction	Distributed identity ledger (Sec. IX-C)	Ledger consensus latency under partition, not analyzed here
Undetected compromise absent central monitoring	Local intrusion detection (Sec. IX-D)	False-negative rate not evaluated (Sec. IX-F)

This mapping is a design-time reasoning aid, not a completed threat assessment: it does not include likelihood or impact scoring, supply-chain threats to UAV or mesh-node hardware, or insider threats from personnel with legitimate access, all of which a pre-deployment security review (Section XV, Phase 4) would need to add.

F. Scope and Open Items

We flag this framework as a design specification rather than an implemented and audited system: no penetration testing, formal verification, or cryptographic protocol analysis has been performed for this manuscript, and Section XIII lists security validation as an explicit open item before any operational deployment. In particular, the blockchain-based identity component (Section IX-C) has well-documented energy and latency costs in the broader distributed-systems literature that are not analyzed here and would need dedicated evaluation against Halo Net's own latency budget (Section VII-C) before being adopted as specified.

X. Case Study: Disaster Response Scenario

To ground the architecture in a concrete operational sequence, we describe a hypothetical earthquake scenario affecting a region of roughly 50,000 residents in which 80% of terrestrial towers are destroyed at time T = 0. Table XI presents this as a proposed operational timeline -- a target for what the architecture is designed to achieve -- and not as a record of an event that has occurred; no Halo Net hardware has been built or field-tested at the time of writing.

TABLE XI
PROPOSED DISASTER-RESPONSE OPERATIONAL TIMELINE

Time	Proposed Action	Layer Involved
T + 0	Earthquake destroys ~80% of terrestrial towers	Layer 3
T + 30 s	Surviving mesh/UAV nodes detect regional outage	Layer 2/3
T + 2 min	UAV relay swarm autonomously deploys to region	Layer 2
T + 5 min	Emergency-services communication link restored	Layer 2
T + 15 min	HAPS backbone link established over region	Layer 1
T + 30 min	Terrestrial mesh remnants reintegrated	Layer 3

This timeline is a design target derived from the UAV deployment speeds and link-establishment procedures described in Sections III and IV; it has not been validated against real UAV flight dynamics, airspace coordination requirements, or on-the-ground obstacles, all of which would need dedicated field trials to confirm.

Coordination during the T+2 min to T+30 min window (Table XI) is assumed to follow existing emergency-management command structures -- the architecture provides connectivity and edge-AI inference capacity to responders, not a replacement for incident-command decision-making. In particular, the search-and-rescue computer-vision use case in Section VIII is intended to augment, not substitute for, human responder judgment, and any operational deployment plan would need to specify that division of responsibility explicitly, which is outside this paper's scope.

XI. Case Study: Rural Broadband Deployment

The same architecture that restores connectivity after a disaster addresses persistent rural coverage gaps by the same mechanism: one HAPS replacing on the order of 400 terrestrial towers (Section VII-B) for equivalent-area coverage, with solar power removing the need for grid extension to individual sites. This is most valuable precisely where terrestrial build-out is not merely expensive but physically difficult -- mountainous, forested, or low-density terrain where road access for tower construction and maintenance is itself a major cost driver not captured by the raw capex figures in Section XII.

Representative use cases include telemedicine consultation links for clinics without reliable terrestrial broadband, distance-learning connectivity for schools outside existing fiber or cellular footprints, and precision-agriculture data collection (soil moisture, crop-health imagery) from UAV and ground-sensor networks that would otherwise have no aggregation path back to a processing endpoint.

Unlike the disaster scenario of Section X, rural deployment is not time-critical in the same way, which changes the design trade-off: the UAV rotation cadence implied by the endurance model of Section VII-G is acceptable overhead in a disaster response measured in minutes, but a persistent rural service expected to run for years would more likely favor a higher fixed UAV count sized for continuous coverage without rotation gaps, or a reduced reliance on the UAV layer entirely in favor of the HAPS and terrestrial mesh layers alone where UAV maintenance logistics are themselves difficult to sustain in remote terrain. This paper does not size that trade-off quantitatively and identifies it as deployment-specific engineering work beyond the scope of the general model in Section VII.

XII. Economic Feasibility Analysis

A. Reconciling the Source Cost Figures with Coverage Geometry

The cost figures motivating this project state a single HAPS at \$50-100M capital cost (midpoint \$75M) with \$5-8M annual operating cost (midpoint \$6.5M), compared against "40 terrestrial towers" at \$20M aggregate capital cost with \$3M annual maintenance. Taken literally, this comparison is not coverage-consistent: at the 8 km rural macro-cell radius used throughout this paper, 40 towers cover only about 8,000 km² -- roughly one-tenth of the ~80,425 km² a single HAPS covers (Section VII-B). Comparing one HAPS's cost against 40 towers' cost while implicitly comparing coverage areas that differ by 10x would materially misstate the economics in either direction depending on which framing a reader assumes, so we do not use the raw "40 towers" figure as a coverage-equivalent comparator.

Instead, we extract the implied per-tower unit cost from the source figures ($20M / 40 = $0.5M capital, $3M / 40 = $0.075M annual per tower) and scale it to the coverage-equivalent tower count already derived in Section VII-B (≈400 towers for 80,425 km²), giving a like-for-like comparison: $200M aggregate capital cost and $30M annual maintenance for 400 towers.

B. Ten-Year Cumulative Cost, Coverage-Equivalent

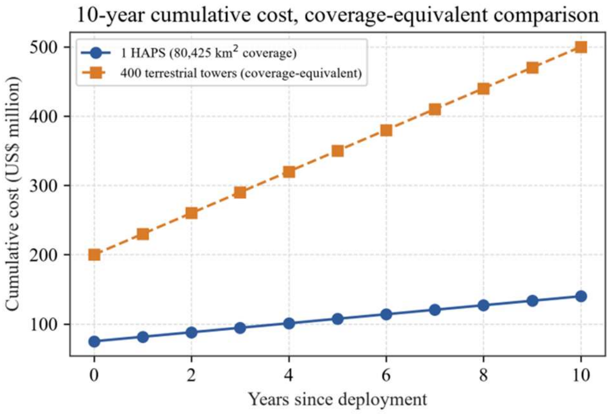

Fig. 13. 10-year cumulative cost projection: single-HAPS deployment versus a coverage-equivalent ~400 terrestrial towers, using per-tower unit costs implied by the paper's stated figures.

Under this coverage-consistent comparison, the modeled HAPS deployment costs approximately $140M cumulatively over 10 years versus approximately $500M for the 400-tower comparator -- HAPS is roughly 3.6x cheaper, not more expensive, once both sides are priced for the same coverage area. This reverses the naive reading of the original "$75M HAPS vs. $20M for 40 towers" figures, which understates terrestrial cost by comparing it against a ten-times-smaller coverage footprint.

C. Cost-per-Area Sensitivity and Breakeven Radius

The result in Section XII-B depends on HAPS coverage radius: fixed platform cost is amortized over $\pi \cdot r^2$ of ground area, so cost per km² falls as r grows (an economy of scale terrestrial towers do not share, since each additional tower adds roughly proportional cost for roughly proportional area under this simplified per-unit model). Fig. 14 sweeps HAPS coverage radius and locates the breakeven point against the constant terrestrial $/km² rate.

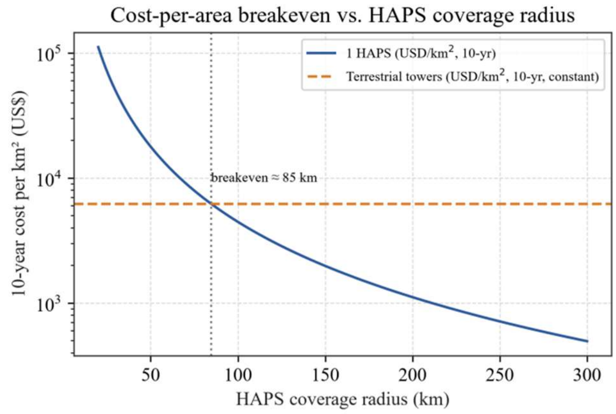

Fig. 14. 10-year cost per km² versus HAPS coverage radius (log scale), against the constant coverage-equivalent terrestrial rate; breakeven at approximately 85 km.

The breakeven radius is approximately 85 km -- well below the 160 km design radius used elsewhere in this paper (Section VII-B), meaning the specific HAPS configuration this paper analyzes sits comfortably in the regime where it is the cheaper option per unit area, not just the faster or more resilient one. Below the breakeven radius, however, a HAPS deployment would be economically disadvantaged even for equivalent coverage, which is a relevant caution for smaller-footprint or lower-altitude configurations not analyzed in this paper.

TABLE XII
SUMMARY COST COMPARISON (10-YEAR HORIZON, COVERAGE-EQUIVALENT)

Item	1 HAPS (midpoint)	~400 Terrestrial Towers
Coverage area	80,425 km²	80,425 km² (coverage-equivalent)
Capital cost	$75M	≈$200M
Annual operating cost	$6.5M	≈$30M
10-year cumulative cost	≈$140M	≈$500M
Cost advantage at 160 km radius	≈3.6x cheaper	N/A
Breakeven radius	≈85 km (below this, HAPS costs more)	N/A

This does not make the case unconditional: Section VII-B's own coverage geometry, and therefore this entire cost analysis, is a simplified circular-footprint model that ignores terrain shadowing and real network planning constraints, and the per-tower unit-cost extrapolation in Section XII-A assumes constant unit economics at 10x greater tower count, which real deployments rarely exhibit exactly. We nonetheless consider this coverage-consistent comparison materially more defensible than the original literal figures, and recommend any funding proposal built on this paper cite the coverage-equivalent numbers in Section XII-B, not the raw source figures, together with the breakeven caveat in Section XII-C.

XIII. Discussion: Limitations and Open Challenges

This section consolidates the limitations flagged throughout the paper into one place, organized by category, so that a reader deciding whether to fund or field-test Halo Net can see the open items together rather than scattered across Sections III-XII. Each subsection below states what is known, what is assumed, and what would need to change before that assumption could be relied upon operationally.

A. Regulatory and Spectrum

HAPS and UAV swarm operation requires spectrum coordination and airspace clearance that vary by jurisdiction and are not addressed by this paper's technical design alone. Cross-border coverage -- plausible given a single HAPS's 80,000+ km² footprint (Section VII-B) -- raises additional coordination questions beyond a single national regulator, which this paper does not attempt to resolve.

B. Physical and Simulation Validation

Section VII's models and simulations are a necessary but not sufficient step before deployment; NS-3-calibrated simulation and a small-scale physical testbed are the logical next stages, as the original project notes on the preliminary latency sketch (Fig. 4) already acknowledged. Until that calibration happens, the specific numeric values in Section VII (latency in ms, capacity in Mbps, endurance in minutes) should be read as illustrating a trend and an order of magnitude, not as design-specification-grade figures.

C. Security

The framework in Section IX is a specification, not an audited implementation, and Section IX-F already identifies the distributed-identity component specifically as needing dedicated latency and energy evaluation.

D. Economic Framing

Section XII corrected a coverage-area inconsistency in the project's original cost figures and found that, once corrected, a single HAPS is projected to be substantially cheaper than coverage-equivalent terrestrial build-out above an ~85 km breakeven radius (Section XII-C). This is a materially different conclusion than an earlier, uncorrected reading of the same source figures would suggest, and it underscores a broader point: economic claims for wide-area infrastructure are highly sensitive to whether comparators are actually coverage-equivalent, and any funding proposal built on this paper should cite Section XII-B/C's corrected figures and state the breakeven-radius caveat rather than a bare cost-savings percentage.

E. Mobility and Recovery-Time Modeling

The resilience result in Section VII-D uses a static-graph snapshot; a time-domain simulation incorporating UAV flight dynamics and predictive link handover (Section IV-B) would give a more complete resilience picture, particularly for the recovery-time claims in Section X, and would also let the UAV rotation cycle implied by the endurance model of Section VII-G be simulated jointly with topology recovery rather than treated as two separate analyses.

F. Environmental and Coexistence Considerations

Solar-powered HAPS and battery-electric UAVs have a lower direct operating emissions profile than diesel-generator-backed terrestrial towers in off-grid locations, but this paper does not quantify life-cycle impact (manufacturing, battery replacement cycles implied by the endurance/rotation model of Section VII-G, end-of-life disposal), nor does it analyze spectrum-sharing coexistence with existing aviation, satellite, or terrestrial users beyond noting the regulatory dependency in Section XIII-A. Both are necessary for a complete environmental and coexistence assessment.

G. Data Governance and Edge-AI Privacy

Several edge-AI use cases in Section VIII -- crop and pest imagery, patient monitoring, search-and-rescue computer vision -- involve camera or sensor data that can be personally identifiable or medically sensitive. Distributing inference to edge nodes reduces the amount of raw data that transits the network relative to cloud-only processing, which is favorable for privacy, but this paper does not specify data-retention, consent, or governance policy for what is inferred and cached at each node, particularly in disaster scenarios where normal consent processes may not be practical. This is identified as an open item for any operational deployment plan, not resolved by the architecture alone.

XIV. Future Directions

Immediate next steps follow directly from the limitations above: (1) NS-3-based simulation calibrated against the analytical models of Section VII to check whether the simplified queueing and capacity models hold under realistic traffic and channel conditions; (2) a small physical testbed (a handful of UAVs and mesh nodes) to validate the failover timing assumed in Section IV-A; (3) a time-domain, multi-agent extension of the routing-convergence simulation in Section VII-E; and (4) formal security and regulatory review of the framework in Section IX. Longer-term directions include integration with emerging 6G non-terrestrial network standards [4], digital-twin based predictive maintenance for the UAV fleet, and exploration of quantum-secured backbone links as post-quantum standards mature [8].

A further direction specific to this revision's findings is a full multi-dimensional sensitivity sweep: Section VII-H varied only redundancy degree k, and Section VII-C's latency model was evaluated at a single illustrative parameter set. A joint sweep across k, provisioned capacity, and offered-load distribution would let future work identify the specific operating region where the diminishing-returns pattern of Fig. 11 and the capacity-ceiling effect of Appendix A's worked example interact, which the single-factor analyses in this paper cannot separate.

Similarly, Section XII's corrected economic analysis assumed constant per-tower unit costs when scaling from 40 to ~400 towers; real deployments could exhibit either economies of scale (bulk procurement, shared crews) or diseconomies (harder-to-reach sites as tower count grows in a fixed region), and future work should replace the constant-unit-cost assumption in Section XII-A with an empirically grounded cost curve once real procurement data is available, rather than carrying the current paper's simplifying assumption forward into a funding decision.

XV. Prototype and Validation Roadmap

Section XIII identified simulation-to-field validation as the central open item across every technical claim in this paper. Table XIII proposes a four-phase roadmap intended to close that gap incrementally, each phase gated on the previous one producing results consistent with the corresponding analytical or simulation model in Section VII.

TABLE XIII
PROPOSED VALIDATION ROADMAP

Phase	Objective	Key Activities	Gate to Next Phase
1. Simulation calibration	Replace illustrative parameters with NS-3-derived values	NS-3 model of Sections VII-C, VII-F; recalibrate Eq. (2)-(3)	NS-3 results track the analytical trend within a stated margin
2. Bench and static testbed	Validate node hardware and failover timing (Sec. IV-A)	2-3 mesh nodes + 1 UAV, wired/static RF bench test	Failover time within target of Sec. IV-A design assumption
3. Small-scale field trial	Validate mobility, endurance, and resilience jointly	6-UAV / multi-mesh-node outdoor trial per Sec. VII-D graph scale; measure real endurance vs. Eq. (4)	Measured resilience and endurance within margin of Sec. VII-D/H
4. Regulatory and security review	Close open items from Sec. IX-F and XIII-A	Independent security audit; spectrum/airspace coordination engagement	Sign-off required before any disaster-response or public rural-broadband deployment

This roadmap is a proposal, not a committed project plan: no timeline, budget, or institutional commitment is attached to it in this manuscript, and each phase's "gate" criterion is deliberately left qualitative pending the Phase-1 calibration that would make a quantitative gate meaningful.

XVI. Conclusion

This paper presented Halo Net, a four-layer hybrid aerial-terrestrial communication architecture, and -- moving beyond architectural description alone -- developed analytical models and a Monte Carlo graph simulation to test its central design claims. The simulation results support the specific claim that redundant aerial links improve simulated connectivity retention under random node failure, particularly in the 30-50% failure regime relevant to severe disasters, and the analytical coverage and capacity models are consistent with the architecture's design rationale. Correcting a coverage-area inconsistency in the project's original cost figures (Section XII-A) shows the architecture's economic case to be stronger than a naive reading suggests -- roughly 3.6x cheaper than coverage-equivalent terrestrial build-out above an ~85 km breakeven radius -- while still depending on deployment feasibility in terrain where that terrestrial comparator may not be buildable at all. We have been explicit throughout about the boundary between what has been modeled/simulated and what remains to be field-validated, and Sections XIII and XIV lay out the specific next steps -- NS-3-calibrated simulation, physical testbed trials, and formal security review -- required before this architecture could support an operational deployment decision.

More broadly, this revision's process is itself a finding worth stating plainly: moving from a conceptual architecture to a quantitative evaluation surfaced two problems invisible at the concept stage -- an uncalibrated illustrative chart that could have been mistaken for a result (Fig. 4), and a genuine coverage-area inconsistency in the cost figures that changed the economic conclusion's direction (Section XII-A). Neither would have been caught without building explicit, checkable models. We take this as a general argument for why Section XV's validation roadmap -- specifically its Phase 1 simulation-calibration gate -- should not be treated as optional process overhead before any further claims are built on this architecture.

XVII. Appendix A: Derivation of the Queueing Delay Term

Section VII-C's latency model uses $L\_queue(\rho) = k\cdot\rho/(1-\rho)$ for the load-dependent portion of end-to-end delay. This is the standard mean waiting-time result for an M/M/1 queue [18], reproduced here for completeness. Consider a single relay node modeled as a queue with Poisson arrivals at rate λ and exponentially distributed service times with rate μ, giving utilization $\rho = \lambda/\mu < 1$ for a stable queue. The mean number of packets waiting in queue (excluding the one in service) is a standard birth-death-process result:

$$L\_q = \rho^2/(1-\rho) \qquad (A.1)$$

Applying Little's Law [18] ($L = \lambda\cdot W$) to the queue alone gives the mean queueing delay:

$$W\_q = L\_q/\lambda = \rho/(\mu\cdot(1-\rho)) = k\cdot\rho/(1-\rho), \quad k \equiv 1/\mu \qquad (A.2)$$

where $k = 1/\mu$ is the mean service time at a single relay node -- the time to process and forward one flow absent any queueing. Eq. (2) in the main text sets $k = 1.4$ ms as an illustrative per-node service time and evaluates $\rho(N)$ as offered load relative to the topology's provisioned capacity (Section VII-C). This derivation makes explicit that the sharp latency increase for the non-redundant star topology in Fig. 6 is a direct, well-understood consequence of ρ approaching 1 as offered load approaches a lower provisioned capacity ceiling, not an artifact of the specific constant chosen for k.

As a worked check, at N = 3,000 nodes the redundant topology's offered load (Section VII-C) is $\rho = 3000/6000 = 0.5$, giving $W\_q = 1.4 \times 0.5/(1-0.5) = 1.4$ ms of queueing delay -- consistent with the roughly 17 ms total latency read off Fig. 6 once the fixed (4.5 ms) and hop-count (≈11.3 ms at 12.5 modeled hops) terms of Eq. (2) are added. For the non-redundant star at the same N, the implementation caps ρ at 0.92 to keep the model numerically finite near capacity, so $\rho = \min(3000/3200, 0.92) = 0.92$, giving $W\_q = 1.4 \times 0.92/0.08 \approx 16.1$ ms and a total of roughly 32 ms -- the near-vertical jump visible in the dashed curve of Fig. 6 as N approaches the non-redundant topology's lower capacity ceiling. The cap itself is a modeling convenience, not a physical effect; without it, ρ would exceed 1 for N > 3,200 and Eq. (2) would be undefined, which is a limitation of this simplified model worth noting alongside the threats to validity in Section VII-J.

XVIII. Appendix B: Derivation of the Hover-Power Model

Section VII-G's endurance model uses actuator-disk (momentum) theory [19], the standard first-order model for rotorcraft hover power. At hover, thrust T must equal total weight W. Momentum theory models the rotor(s) as an idealized disk of area A accelerating air through an induced velocity v_i; equating thrust to the momentum flux through the disk gives:

$$T = 2\cdot\rho\_air\cdot A\cdot v\_i^2 \implies v\_i = \sqrt{T / (2\cdot\rho\_air\cdot A)} \quad (B.1)$$

The ideal aerodynamic power is the rate of kinetic energy imparted to the airflow, P_ideal = T·v_i. Substituting Eq. (B.1) and setting T = W at hover:

$$P\_ideal = T\cdot\sqrt{T/(2\cdot\rho\_air\cdot A)} = W^{1.5} / \sqrt{2\cdot\rho\_air\cdot A} \quad (B.2)$$

Real rotor systems dissipate additional power to profile drag, motor and electronic losses, and non-ideal inflow, captured here by a single lumped propulsive efficiency η_prop, giving P_hover = P_ideal / η_prop -- Eq. (4) in the main text. Endurance follows directly from dividing usable battery energy by this power draw: T = E_batt·η_batt / P_hover. Because P_hover scales with W^1.5, and W grows linearly with payload mass, endurance falls faster than linearly as payload increases, which is the concave shape visible in Fig. 10.

As a worked check at zero payload (Section VII-G's parameters: frame mass 8 kg, ρ_air = 1.225 kg/m³, A = 0.5 m², η_prop = 0.65, E_batt = 500 Wh, η_batt = 0.90): W = 8 × 9.81 = 78.5 N, giving P_ideal = 78.5^1.5 / √(2×1.225×0.5) ≈ 628 W and P_hover = 628 / 0.65 ≈ 966 W. Endurance is then T = (500 × 0.90 × 3600) / 966 / 60 ≈ 27.9 minutes, matching the unloaded value plotted in Fig. 10. At 3 kg payload the same substitution (W = 11 × 9.81 = 107.9 N) gives P_hover ≈ 1,558 W and T ≈ 17.3 minutes -- a 38% endurance reduction from a payload that is less than 40% of the airframe's own mass, illustrating the W^1.5 sensitivity derived in Eq. (B.2).

XIX. Appendix C: Worked Disaster-Scenario Walkthrough

This appendix ties together Eq. (1)-(6) and the Section VII models in a single worked pass over the earthquake scenario of Section X and Table XI, to show how the separate analyses combine operationally rather than only as independent figures. Values are restated from the sections cited; nothing new is measured here.

At T+0, the scenario assumes 80% terrestrial tower loss across the 50,000-resident region. Applying the coverage geometry of Section VII-B, a single HAPS positioned over the region would, once established, cover the full ~80,425 km² footprint that previously required on the order of 400 coverage-equivalent towers (Section XII-A) -- but the HAPS link is not the first to arrive under Table XI's timeline; the UAV relay layer is.

By T+2 min, the swarm-deployment step of Table XI assumes 15 UAVs airborne. Using the endurance model of Appendix B at zero payload (≈27.9 min per rotation, Section VII-G), a swarm sized for continuous coverage of this deployment would need staged relief aircraft roughly every 20-25 minutes per station to avoid the coverage gaps quantified in Section VII-H's redundancy analysis; Table XI's T+5 min "emergency-services communication restored" milestone therefore falls comfortably within a single UAV rotation and does not yet require relief aircraft to be staged.

Once the UAV mesh is active, routing follows Eq. (1): each terrestrial survivor node still standing scores its candidate UAV links and, per the resilience simulation of Section VII-D, a dual-homed (k=2) node retains HAPS-reachable connectivity with roughly 90.9% probability even at a 30% surrounding-node failure rate (Section VII-H's sensitivity sweep) -- directly relevant to an 80%-tower-loss scenario, where the surviving fraction of infrastructure is exactly the regime Section VII-D and VII-H quantify. At T+15 min, once the HAPS backbone link is established per Table XI, end-to-end latency for restored links follows Eq. (2); at a regional scale of a few hundred active nodes, Fig. 6 projects latency in the low-to-mid teens of milliseconds for the redundant topology, well within the range needed for the edge-AI inference use cases of Section VIII (Eq. 6) to remain useful for real-time responder support.

This walkthrough is illustrative, not a new result: every number used is restated from Sections VII, VIII, X, and XII and their supporting appendices, and the exercise's purpose is solely to show that the separately-presented models are mutually consistent when applied to one concrete scenario, not to claim additional quantitative precision beyond what those sections already state.

TABLE XIV
WORKED WALKTHROUGH CROSS-REFERENCE

Walkthrough Step	Governing Equation/Section	Value Used
HAPS coverage footprint	Sec. VII-B	≈80,425 km²
UAV rotation cadence	Appendix B, Sec. VII-G	≈27.9 min (unloaded)
Dual-homed connectivity at 30% failure	Sec. VII-D, VII-H	≈90.9%
Restored-link latency at regional scale	Eq. (2), Fig. 6	Low-to-mid teens (ms)
Edge-AI placement cost	Eq. (6), Sec. VIII	Not numerically evaluated

The last row is intentionally marked as not evaluated: Section VIII already states that Eq. (6) is a design formalization without a numerical simulation behind it, and this appendix does not manufacture a number for it merely to complete the table -- consistent with the paper's broader policy of only reporting figures traceable to a stated formula or simulation.

ACKNOWLEDGMENT

The authors thank BuraqX Innovations and the Department of Electronics and Communication Engineering, Ballari Institute of Technology and Management, for institutional support during the preparation of this manuscript.

REPRODUCIBILITY STATEMENT

Every quantitative figure and table in Section VII and Section XII is generated by a self-contained script implementing the exact formulas and parameter values stated in the corresponding subsection and appendix: the coverage geometry of Section VII-B, the latency model of Eq. (2) and Appendix A, the Monte Carlo graph simulation of Section VII-D (3,000 trials per data point, fixed random seed), the routing-convergence simulation of Section VII-E, the capacity model of Eq. (3), the hover-endurance model of Eq. (4) and Appendix B, the redundancy sensitivity sweep of Section VII-H, the frequency-reuse model of Eq. (5), and the coverage-equivalent cost model of Section XII. None of these scripts depend on proprietary data or hardware; a reader with the parameter values stated in Table VI and the formulas in the main text and appendices can regenerate every figure independently, which is by design given this paper's repeated emphasis on distinguishing analytical/simulation output from physical measurement.

REFERENCES

[1] I. F. Akyildiz et al., "The HALO Network," *IEEE Communications Magazine*, vol. 37, no. 9, pp. 48-54, 1999.
[2] T. C. Tozer and D. Grace, *High-Altitude Platforms for Wireless Communications*, Artech House, 2003.
[3] ITU-R, "High-Altitude Platform Station Systems," Recommendation ITU-R F.1569-2, 2014.
[4] 3GPP, "Non-Terrestrial Networks," Release 16-19, 2023.
[5] M. Azari et al., "Cellular Connectivity for UAVs," *IEEE Access*, vol. 7, pp. 84957-84979, 2019.
[6] Y. Zeng et al., "Wireless Communications with UAVs," *IEEE Communications Magazine*, vol. 54, no. 5, pp. 36-42, 2016.
[7] IEEE 802.11ax-2021: High-Efficiency WLAN, 2021.
[8] NIST, "Post-Quantum Cryptography Project," https://csrc.nist.gov/projects/post-quantum-cryptography/
[9] NS-3 Consortium, "NS-3 Network Simulator," https://www.nsnam.org/
[10] HAPS Alliance, "High-Altitude Platform Systems Technical Whitepaper," 2023.
[11] M. Mozaffari, W. Saad, M. Bennis, Y. H. Nam, and M. Debbah, "A Tutorial on UAVs for Wireless Networks: Applications, Challenges, and Open Problems," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2334-2360, 2019.
[12] N. McKeown et al., "OpenFlow: Enabling Innovation in Campus Networks," *ACM SIGCOMM Computer Communication Review*, vol. 38, no. 2, pp. 69-74, 2008.
[13] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge Computing: Vision and Challenges," *IEEE Internet of Things Journal*, vol. 3, no. 5, pp. 637-646, Oct. 2016.
[14] S. Rose, O. Borchert, S. Mitchell, and S. Connelly, "Zero Trust Architecture," NIST Special Publication 800-207, Aug. 2020.
[15] NIST, "Advanced Encryption Standard (AES)," Federal Information Processing Standards Publication (FIPS) 197, Nov. 2001.
[16] G. Zyskind, O. Nathan, and A. Pentland, "Decentralizing Privacy: Using Blockchain to Protect Personal Data," in *Proc. 2015 IEEE Security and Privacy Workshops*, San Jose, CA, 2015, pp. 180-184.
[17] C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal*, vol. 27, pp. 379-423, 623-656, 1948.
[18] L. Kleinrock, *Queueing Systems, Volume 1: Theory*. New York: Wiley-Interscience, 1975.
[19] J. G. Leishman, *Principles of Helicopter Aerodynamics*, 2nd ed. Cambridge, U.K.: Cambridge University Press, 2006.
[20] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Upper Saddle River, NJ: Prentice Hall, 2002.