Can HAPS Defer Terrestrial Capacity Expansion?

Mohamed Amine Mbarek
Politecnico di Torino
Turin, Italy
mohamed.mbarek@polito.it

Michela Meo
Politecnico di Torino
Turin, Italy
michela.meo@polito.it

Daniela Renga
Politecnico di Torino
Turin, Italy
daniela.renga@polito.it

***Abstract*—Mobile traffic growth pushes operators toward terrestrial densification, either by deploying additional sites or by expanding existing ones through additional spectrum, carriers, or radio equipment. This work focuses on the latter case, where capacity upgrades are fixed infrastructure commitments whose justification depends not only on traffic magnitude, but also on whether excess demand is persistent across time and concentrated across space. We ask whether a High Altitude Platform Station (HAPS) can defer terrestrial capacity expansion by reducing the ground-side support capacity required under the same demand. We compare two matched benchmarks: a terrestrial-only case that identifies the required sector-level support activations, and a HAPS-assisted case that identifies the activations that remain after aerial capacity is available. Their difference quantifies the terrestrial support requirement avoided by HAPS and serves as a proxy for capacity-expansion pressure. Using real traffic traces on a sectorized macro network, we show that a modest HAPS beam budget removes most sector-slot support activations and reduces the recurrence and spatial footprint of the remaining support needs. The results indicate that HAPS can shift terrestrial upgrade pressure from recurrent capacity expansion toward more occasional support requirements.**

***Index Terms*—High Altitude Platform Stations (HAPS), terrestrial densification, capacity expansion, radio access networks, network planning.**

I. Introduction

The continued growth of mobile traffic forces operators to provision additional capacity in terrestrial radio access networks. This capacity increase is not only an aggregate-throughput problem; it is also a deployment problem, because additional spectrum, radio equipment, antennas, and site resources become fixed network assets once introduced. In 5G densification, capacity scaling is commonly associated with additional spectrum, cell splitting, and spatial multiplexing, and is further supported by spectral aggregation, backhaul densification, and interference management [1]. In practice, operators may expand existing sites through new bands, carrier aggregation, Massive Multiple Input Multiple Output (MIMO), sectorization, or small-cell layers, while new sites are often avoided when existing sites can still be upgraded because site deployment remains costly and difficult [2].

This creates a planning challenge under time-varying demand. A sector that exceeds its installed capacity in a few isolated peak slots does not imply the same reinforcement need as a sector that repeatedly requires additional capacity. Similarly, a small set of recurring constrained sectors and a large number of simultaneously constrained sectors lead to different expansion decisions. Aggregate traffic volume alone cannot capture these differences. The relevant questions are where additional terrestrial capacity is needed, how often it is needed, and whether the need is recurrent enough to justify fixed ground-side reinforcement.

High Altitude Platform Stations (HAPS) offer a different way to support such traffic peaks. Operating from the stratosphere, HAPS can provide wide-area aerial capacity without being permanently tied to a specific terrestrial site. They have been studied as nodes in vertical heterogeneous networks, supporting high-capacity access and data services across urban, suburban, and remote areas [3]. The HAPS-Super Macro Base Station (SMBS) concept further positions HAPS as a metropolitan support layer for dense areas where demand may be difficult to accommodate through fixed densification alone [4], while other works have considered HAPS for sustainable radio access network (RAN) operation [5]. These properties motivate a planning question: beyond the traffic carried by HAPS, can its availability reduce the terrestrial capacity additions that would otherwise be required?

This work studies this question by comparing two capacity-addition outcomes under the same demand conditions. We first determine the support requirement of the terrestrial network when it operates alone. We then introduce HAPS support and recompute the ground-side support requirement that remains after aerial capacity is available. Their difference quantifies the terrestrial support requirement avoided by HAPS, which we interpret as a reduction in capacity-expansion pressure. In this sense, HAPS is evaluated not only as an additional source of carried traffic, but as a flexible aerial layer that can reduce, delay, or reshape terrestrial densification needs. The main contributions of this paper are:

- We use support-carrier activation as a sector-level proxy for terrestrial capacity-expansion pressure, allowing HAPS support to be evaluated through avoided ground-side reinforcement.
- We formulate matched terrestrial-only and HAPS-assisted optimization benchmarks that compute, under the same traffic demand, the required and remaining support activations.
- We analyze the avoided support requirement in terms of total activation volume, simultaneity, and persistence, distinguishing temporary traffic peaks from recurrent capacity pressure.

II. Related Work

One line of work addresses the radio design of HAPS-assisted cellular systems, including coverage formation, cell

979-8-3195-0167-7/26/$31.00 ©2026 IEEE

configuration, and coexistence with terrestrial networks. Analytical models have been developed to characterize HAPS beam footprints, per-user capacity, spectral efficiency, and area spectral efficiency for extended cellular coverage [6]. Dynamic HAPS cell optimization has also been considered for non-uniform user distributions, with antenna-parameter optimization used to adapt the HAPS cell layout to spatial demand variations [7]. More recent work incorporates coexistence with terrestrial mobile networks by optimizing HAPS antenna parameters under both coverage and interference-related objectives [8]. Spectrum-sharing studies further analyze the coverage probability and transmission capacity of HAPS and terrestrial networks under shared-spectrum operation, showing the importance of HAPS height, density, and directional antennas for limiting interference [9]. These studies provide the radio-level basis for HAPS-assisted service, defining how aerial coverage can be formed, optimized, and coordinated with terrestrial systems.

A second line of work investigates HAPS integration for sustainable RAN operation. HAPS-mounted aerial BSs have been considered in urban mobile networks to offload terrestrial clusters, reduce grid-energy demand, lower operational cost, support renewable-energy sizing, and enable BS deactivation under resource-on-demand operation [10]. Similar HAPS-integrated RAN models use real traffic patterns to quantify energy-consumption reductions obtained by offloading terrestrial traffic and switching selected ground BSs to sleep mode [11]. HAPS has also been introduced in cell-switching strategies, where the aerial platform acts as a serving entity for users displaced from switched-off terrestrial cells [12]. These studies show that HAPS can modify terrestrial network operation, especially from an energy-efficiency perspective, with outputs such as energy consumption, cost, renewable supply, and BS activity.

The closest line of work considers HAPS as a response to dense and time-varying terrestrial traffic demand. HAPS-SMBS architectures have been compared with conventional RAN densification in dense urban scenarios, showing that aerial capacity can complement terrestrial deployments when demand is dynamic and difficult to accommodate through fixed infrastructure alone [13]. HAPS-assisted offloading has also been evaluated under realistic and upscaled traffic traces, using packet delay, user denial, energy consumption, and aggregate infrastructure expansion as performance metrics [14]. In contrast, the present work makes sector-level support-carrier activation the quantity being minimized and analyzes the simultaneity, recurrence, and spatial footprint of the remaining terrestrial support requirement.

Overall, existing HAPS studies establish how aerial coverage can be formed, coordinated, and used for offloading or energy-efficient operation. However, terrestrial densification is usually treated as a motivation or as an aggregate consequence of traffic growth, rather than being evaluated through the sector-level terrestrial support requirements that remain under the same demand. This paper therefore evaluates HAPS assistance through the support-carrier activations that remain necessary in the terrestrial network.

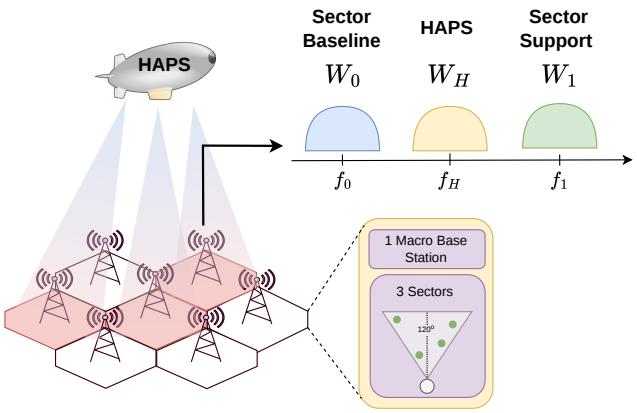


Fig. 1: System model overview.

III. System Model

We consider a macro-cellular downlink assisted by a HAPS. Terrestrial capacity reinforcement is evaluated at the sector level: each deployed sector has an installed carrier that is always available, while an additional support carrier may be activated when the installed capacity is insufficient. The aerial layer provides a limited number of directional downlink beams that can reduce the need for these terrestrial support activations.

A. Network Layout and Radio Resources

Fig. 1 sketches the considered system model. Let $\mathcal{B}$ denote the set of macro base stations (MBSs), with $|\mathcal{B}| = B$. Each MBS $b \in \mathcal{B}$ is located at $\mathbf{p}_b = (x_b, y_b, h_{\mathrm{BS}})$, where h_{BS} is the MBS antenna height. The sites are placed on a regular hexagonal lattice with inter-site distance D_{ISD}. Each MBS has three directional sectors, indexed by $s \in \mathcal{S} = \{1, 2, 3\}$, with azimuth boresights separated by $120°$. We denote a sector by $i = (b, s)$, and the set of all sectors by $\mathcal{I} = \mathcal{B} \times \mathcal{S}$. The service region of sector $i = (b, s)$ is denoted by $\mathcal{A}_i$. In the numerical evaluation, users are placed within these directional sector regions according to the user-level realization procedure described in Section V-C. Each sector operates an installed terrestrial carrier (f_0, W_0), where f_0 is the carrier frequency and W_0 is the bandwidth. This carrier represents the capacity already available in the deployed network. A terrestrial support carrier (f_1, W_1) can additionally be activated on sector i. Let $z_i(t) \in \{0, 1\}$ indicate whether this support carrier is active at time t. Thus, $z_i(t) = 1$ represents a sector-level terrestrial capacity reinforcement; activating $z_i(t)$ makes the support carrier available only in sector i, not across the whole network.

The HAPS is deployed at altitude H above the service region and operates on carrier (f_H, W_H). It can form at most $K_{\max}$ simultaneous downlink beams. We denote by $\mathcal{A}_b^{\mathrm{H}}$ the ground footprint of the candidate beam centered on MBS b. The candidate beam directions are fixed to the MBS

locations; hence, the aerial-layer decision is whether to support a given macro site, not where to place the beam. This keeps the benchmark focused on terrestrial reinforcement avoidance rather than beam-placement optimization. Let $a_b(t) \in \{0,1\}$ indicate whether the candidate beam associated with MBS b is active at time t. The beam budget is

$$\sum_{b\in\mathcal{B}} a_b(t) \le K_{\max}. \tag{1}$$

When $a_b(t) = 1$, users anchored to any of the three sectors of MBS b may receive additional downlink capacity from the corresponding beam.

B. Traffic Demand and User Association

Time is divided into slots $t \in \mathcal{T}$. At each slot, the active users are denoted by $\mathcal{U}(t)$. Each user $u \in \mathcal{U}(t)$ has a traffic demand $\lambda_u(t)$ in bit/s and a ground position $\mathbf{q}_u(t)$. Users are first associated with the terrestrial layer. Let $\sigma(u,t) \in \mathcal{I}$ denote the serving sector of user u at time t. The association is based on the strongest received terrestrial reference signal from the installed carrier. The aerial layer does not redefine the cellular association; it provides additional capacity to users already anchored to the terrestrial network. The set of users associated with sector i at time t is

$$\mathcal{U}_i(t) = \{u \in \mathcal{U}(t) : \sigma(u,t) = i\}. \tag{2}$$

The aggregate demand attached to sector i is

$$\Lambda_i(t) = \sum_{u\in\mathcal{U}_i(t)} \lambda_u(t). \tag{3}$$

The system model assumes that $\lambda_u(t)$, $\mathbf{q}_u(t)$, and $\sigma(u,t)$ are known at each slot. The numerical evaluation later describes how these quantities are generated from sector-level traffic traces. This separates the planning formulation from the traffic-generation procedure used in the experiments.

C. Terrestrial and Aerial Link Rates

For each user, the optimization uses peak rates computed from the corresponding terrestrial and aerial link budgets. Let N_0 denote the thermal noise power spectral density, so that the thermal noise power over bandwidth W is $N_0 W$. Let $\ell \in \{0,1\}$ denote the installed and support terrestrial carriers, with carrier frequency f_ℓ, bandwidth W_ℓ, and sector transmit power $P_{\mathrm{tx}}^{\mathrm{T},\ell}$. For a user u associated with sector i, the received terrestrial power is

$$P_{i,u}^{\mathrm{T},\ell}(t) = P_{\mathrm{tx}}^{\mathrm{T},\ell} G_{i,u}^{\mathrm{T}}(t) 10^{-\mathrm{PL}_{i,u}^{\mathrm{T},\ell}(t)/10}, \tag{4}$$

where $G_{i,u}^{\mathrm{T}}(t)$ is the sector antenna gain and $\mathrm{PL}_{i,u}^{\mathrm{T},\ell}(t)$ is the terrestrial path loss. The path loss and line-of-sight (LOS) probability follow the 3GPP urban macro (UMa) model[1], while the sector gain is computed from the 3GPP three-dimensional antenna pattern, including horizontal and vertical off-boresight attenuation and array gain [15]. The terrestrial SINR is

$$\gamma_{i,u}^{\mathrm{T},\ell}(t) = \frac{P_{i,u}^{\mathrm{T},\ell}(t)}{I_{i,u}^{\mathrm{T},\ell}(t) + N_0 W_\ell}, \tag{5}$$

where $I_{i,u}^{\mathrm{T},\ell}(t)$ is the co-channel terrestrial interference. In the evaluation, interference is computed considering only the dominant first-tier neighbouring sites and assuming a frequency reuse factor of 1, thereby capturing the dominant macro-cellular interference contributions while neglecting marginal interference from distant cells. The corresponding peak terrestrial rate is

$$c_{i,u}^{\mathrm{T},\ell}(t) = W_\ell \log_2\left(1 + \gamma_{i,u}^{\mathrm{T},\ell}(t)\right). \tag{6}$$

For the aerial layer, the candidate beam associated with MBS b operates on carrier (f_H, W_H) with transmit power $P_{\mathrm{tx}}^{\mathrm{H}}$. Its directional gain toward user u, denoted by $G_{b,u}^{\mathrm{H}}(t)$, is computed from the beam footprint and the angular position of the user using an aperture-based antenna model [16]. The received aerial-link power is

$$P_{b,u}^{\mathrm{H}}(t) = P_{\mathrm{tx}}^{\mathrm{H}} G_{b,u}^{\mathrm{H}}(t) 10^{-\mathrm{PL}_{b,u}^{\mathrm{H}}(t)/10}, \tag{7}$$

where $\mathrm{PL}_{b,u}^{\mathrm{H}}(t)$ is the HAPS-to-user path loss. The large-scale aerial channel follows the 3GPP Non-Terrestrial Network (NTN)/HAPS model, including free-space loss, elevation-dependent LOS probability[2], shadow fading, and clutter loss [17]. In the main benchmark, each active HAPS beam provides one normalized downlink resource pool of bandwidth W_H to users anchored to its target MBS region, while inter-beam interference is not explicitly modeled. This system-level abstraction corresponds to a noise-limited HAPS access link, which may be approached through resource partitioning, beam isolation, or coordinated beam scheduling. The aerial-link SNR is

$$\gamma_{b,u}^{\mathrm{H}}(t) = \frac{P_{b,u}^{\mathrm{H}}(t)}{N_0 W_H}. \tag{8}$$

The corresponding peak aerial rate is

$$c_{b,u}^{\mathrm{H}}(t) = W_H \log_2\left(1 + \gamma_{b,u}^{\mathrm{H}}(t)\right). \tag{9}$$

The quantities $c_{i,u}^{\mathrm{T},\ell}(t)$ and $c_{b,u}^{\mathrm{H}}(t)$ are peak rates obtained when the corresponding carrier or beam resource is fully allocated to user u. The resource allocation variables introduced in Section IV determine what fraction of these rates is actually delivered.

IV. Problem Formulation

Using the support-carrier model introduced above, this section defines the two capacity-addition benchmarks used for comparison. The terrestrial-only benchmark computes the support activations required by the ground network alone, whereas the HAPS-assisted benchmark computes the support activations that remain after aerial capacity is available. Their difference gives the avoided terrestrial support requirement.

[1]The UMa path-loss expressions are taken from Table 7.4.1-1, while the UMa LOS-probability model is taken from Table 7.4.2-1 of [15].

[2]The LOS probability is a function of the elevation angle and the deployment environment, as reported in Table 6.6.1-1 of [17].

A. Terrestrial-Only Capacity Addition Benchmark

The terrestrial-only benchmark identifies the minimum set of sectors requiring support activation under fixed terrestrial association. All decisions are made independently at each time slot. To simplify notation, the time index t is omitted throughout this section. Let $\mathcal{I}$ denote the set of terrestrial sectors, and let $\mathcal{U}_i$ denote the set of users associated with sector i, with $\mathcal{U} = \bigcup_{i\in\mathcal{I}} \mathcal{U}_i$. Each user u has demand λ_u. The peak rates achievable by user u from the installed and support carriers of sector i are denoted by $c_{i,u}^{\mathrm{T},0}$ and $c_{i,u}^{\mathrm{T},1}$, respectively. For each sector $i \in \mathcal{I}$, the binary variable $z_i^D \in \{0,1\}$ indicates whether the support carrier is activated in the terrestrial-only benchmark. Users receive fractions α_u^0 and α_u^1 of the installed and support carriers, respectively. The terrestrial-only capacity addition problem is formulated as

$$\min_{\boldsymbol{z}^D,\boldsymbol{\alpha}^0,\boldsymbol{\alpha}^1} \quad \sum_{i\in\mathcal{I}} z_i^D \tag{10a}$$

$$\text{s.t.} \quad c_{i,u}^{\mathrm{T},0}\alpha_u^0 + c_{i,u}^{\mathrm{T},1}\alpha_u^1 \geq \lambda_u, \quad \forall i \in \mathcal{I},\ \forall u \in \mathcal{U}_i, \tag{10b}$$

$$\sum_{u\in\mathcal{U}_i} \alpha_u^0 \leq 1, \qquad \forall i \in \mathcal{I}, \tag{10c}$$

$$\sum_{u\in\mathcal{U}_i} \alpha_u^1 \leq z_i^D, \qquad \forall i \in \mathcal{I}, \tag{10d}$$

$$\alpha_u^0, \alpha_u^1 \geq 0, \qquad \forall u \in \mathcal{U}, \tag{10e}$$

$$z_i^D \in \{0,1\}, \qquad \forall i \in \mathcal{I}. \tag{10f}$$

The objective in (10a) minimizes the number of sectors requiring support-carrier activation. Constraint (10b) enforces demand satisfaction for every user. Constraint (10c) limits the installed carrier of each sector to one normalized resource unit. Constraint (10d) makes the support-carrier resource available only when the corresponding sector is activated. Hence, if $z_i^D = 0$, no support-carrier resource can be allocated in sector i; if $z_i^D = 1$, users in that sector may share one normalized support carrier.

The benchmark in (10) is separable across sectors because user association is fixed and terrestrial carrier resources are local to each sector. Thus, each sector can be evaluated independently. If the installed carrier alone is feasible, then $z_i^D = 0$. Otherwise, the support carrier is activated and feasibility is checked with $z_i^D = 1$. If demand cannot be satisfied even after activating the support carrier, the sector is classified as exceeding the modeled terrestrial capacity-addition capability.

For the HAPS-assisted case, aerial support is restricted to users whose demand is not fully served by the installed terrestrial carrier. Let δ_u^0 denote the residual demand of user u under the installed-carrier-only allocation. We define

$$e_u = \mathbf{1}\{\delta_u^0 > 0\}. \tag{11}$$

Thus, only baseline-overflow users are eligible for HAPS assistance, while users fully served by the installed carrier remain served by the terrestrial layer. This overflow-only eligibility is a modeling policy used to isolate HAPS as an auxiliary capacity source, rather than an unrestricted joint terrestrial–aerial association rule. The residual demand δ_u^0 is obtained by solving, independently for each sector, an installed-carrier-only allocation that maximizes the total served demand subject to the normalized installed-carrier resource constraint.

B. HAPS-Assisted Capacity Addition Problem

We next recompute the terrestrial capacity-addition requirement when HAPS support is available. The terrestrial association remains fixed: a user anchored to sector $i = (b, s)$ may still receive terrestrial resources from that sector, while the HAPS can provide additional downlink capacity only through the candidate beam associated with MBS b.

Let $z_i^H \in \{0,1\}$ indicate whether sector i still requires the terrestrial support carrier after HAPS assistance. Let $a_b \in \{0,1\}$ indicate whether the HAPS beam centered on MBS b is active, and let $\phi_u \geq 0$ denote the fraction of the corresponding beam resource assigned to user u. We denote by $b(u)$ the MBS anchoring user u, and use e_u from (11) to indicate whether user u is eligible for aerial support. The HAPS-assisted capacity-addition problem is formulated as

$$\min_{\substack{\boldsymbol{z}^H,\boldsymbol{a},\boldsymbol{\alpha}^0,\\ \boldsymbol{\alpha}^1,\boldsymbol{\phi}}} \quad \sum_{i\in\mathcal{I}} z_i^H \tag{12a}$$

$$\text{s.t.} \quad c_{i,u}^{\mathrm{T},0}\alpha_u^0 + c_{i,u}^{\mathrm{T},1}\alpha_u^1 + c_{b(u),u}^{\mathrm{H}}\phi_u \geq \lambda_u, \qquad \forall i \in \mathcal{I},\ u \in \mathcal{U}_i, \tag{12b}$$

$$\sum_{u\in\mathcal{U}_i} \alpha_u^0 \leq 1, \qquad \forall i \in \mathcal{I}, \tag{12c}$$

$$\sum_{u\in\mathcal{U}_i} \alpha_u^1 \leq z_i^H, \qquad \forall i \in \mathcal{I}, \tag{12d}$$

$$\sum_{b\in\mathcal{B}} a_b \leq K_{\max}, \tag{12e}$$

$$\sum_{u:b(u)=b} \phi_u \leq a_b, \qquad \forall b \in \mathcal{B}, \tag{12f}$$

$$0 \leq \phi_u \leq e_u, \qquad \forall u \in \mathcal{U}, \tag{12g}$$

$$\alpha_u^0, \alpha_u^1 \geq 0, \qquad \forall u \in \mathcal{U}, \tag{12h}$$

$$z_i^H \in \{0,1\}, \quad a_b \in \{0,1\}, \quad \forall i \in \mathcal{I},\ b \in \mathcal{B}. \tag{12i}$$

The objective in (12a) minimizes the number of sectors that still require terrestrial support-carrier activation after HAPS assistance is made available. Constraint (12b) enforces demand satisfaction by combining the installed terrestrial carrier, the terrestrial support carrier, and the HAPS beam resource. Constraints (12c) and (12d) preserve the sector-local terrestrial resource structure of the terrestrial-only benchmark: each sector has one normalized installed carrier, and one normalized support carrier can be used only if the residual activation variable z_i^H is equal to one.

The HAPS layer introduces the only coupling across terrestrial sectors. Constraint (12e) limits the number of MBS regions that can be simultaneously supported by the HAPS, while (12f) assigns one normalized HAPS beam resource to each active MBS-level beam. This beam resource is shared by

the users anchored to the corresponding MBS. Constraint (12g) restricts aerial allocation to the baseline-overflow users defined in (11), so that HAPS resources are used to serve residual demand left by the installed terrestrial carrier. The variables a_b therefore select where HAPS support is available, whereas the objective continues to measure the remaining ground-side capacity additions through $\{z_i^H\}$. Thus, the HAPS layer acts as a targeted auxiliary capacity source while preserving the fixed terrestrial association.

If a slot remains infeasible even with all modeled terrestrial and HAPS resources, the corresponding demand realization is classified as exceeding the modeled capacity-addition capability.

C. Computational Complexity

Both formulations are solved independently at each time slot. Although the terrestrial-only benchmark in (10) is written with binary support-activation variables, it decomposes across sectors because user association is fixed and terrestrial carrier resources are local to each sector. Thus, the value of z_i^D can be determined independently for each sector by checking whether the installed carrier alone is feasible and, if not, whether the installed-plus-support carrier allocation is feasible. These checks involve only continuous allocation variables and linear constraints, so the terrestrial-only benchmark is evaluated as a collection of sector-level linear programs.

The HAPS-assisted formulation in (12) is a mixed-integer linear program. The continuous resource-allocation variables are α_u^0, α_u^1, and ϕ_u, while z_i^H and a_b are binary variables representing residual support-carrier activation and HAPS beam activation, respectively. The main network-level coupling is introduced by the HAPS beam budget in (12e), which limits the number of MBS regions that can be simultaneously supported. For a slot with $U = |\mathcal{U}|$ active users, the formulation contains $|\mathcal{I}| + |\mathcal{B}|$ binary variables, $3U$ continuous allocation variables, and $O(U + |\mathcal{I}| + |\mathcal{B}|)$ linear constraints. We solve the Mixed-Integer Linear Programming (MILP) offline for each traffic snapshot using Gurobi and report solutions certified optimal within the prescribed optimality tolerance.

V. Simulation Setup

We evaluate the capacity-addition formulations using the sectorized macro-cellular model introduced in Section III. The simulation setup specifies: (i) the terrestrial and HAPS radio configuration, (ii) the sector-level traffic trace used to instantiate the aggregate demands $\Lambda_i(t)$, and (iii) the user-level realization used to obtain $\mathcal{U}_i(t)$, $\lambda_u(t)$, and $\mathbf{q}_u(t)$ for the link-budget and allocation models.

A. Deployment and Radio Configuration

The terrestrial deployment follows the sectorized layout of Section III-A. We consider $B = 32$ macro base stations placed on a regular hexagonal lattice. Each macro site is divided into three 120° sectors, so that the simulated network contains $|\mathcal{I}| = 96$ sectors. The inter-site distance fixes the sector footprint used for user placement, terrestrial association, and link-budget computation. The same macro-site locations define the candidate HAPS beam centers, since each HAPS beam is associated with one MBS region. The main deployment and radio parameters are listed in Table I. The transmit-power values in Table I are transmitter powers before antenna gain; the terrestrial and HAPS link budgets therefore differ through their respective directional gains and propagation models.

TABLE I: Deployment and radio parameters.

Parameter	Value
Number of MBSs, B	32
Sectors per MBS	3
Total sectors, $\lvert\mathcal{I}\rvert$	96
Simulation slots, $\lvert\mathcal{T}\rvert$	1461
Inter-site distance, D_{ISD}	1096.51 m
Sector radius	633.07 m
MBS height, h_{BS}	25 m
UE height, h_{UT}	1.5 m
Installed carrier, (f_0, W_0)	2 GHz, 40 MHz
Installed TX power, $P_{\mathrm{tx}}^{\mathrm{T},0}$	40 dBm
Support carrier, (f_1, W_1)	2.6 GHz, 20 MHz
Support TX power, $P_{\mathrm{tx}}^{\mathrm{T},1}$	40 dBm
HAPS altitude, H	20 km
HAPS carrier, (f_H, W_H)	2.1 GHz, 20 MHz
HAPS beam TX power, $P_{\mathrm{tx}}^{\mathrm{H}}$	40 dBm
Thermal noise PSD, N_0	-174 dBm/Hz

TABLE II: Traffic demand statistics.

Quantity	Mean	95th pct.	Max
Sector-slot demand	68.31 Mbps	170.57 Mbps	292.24 Mbps
MBS-slot demand	204.94 Mbps	424.15 Mbps	642.87 Mbps
Network-slot demand	6.56 Gbps	9.26 Gbps	10.91 Gbps

B. Traffic Trace and Scaling

The sector demands $\Lambda_i(t)$ are obtained from a 96×1461 sector-time traffic matrix, consistent with the $|\mathcal{I}| = 96$ sectors and $|\mathcal{T}| = 1461$ hourly slots reported in Table I. Each matrix entry gives the aggregate traffic demand of one sector during one hourly slot.

The trace is derived from real-world 4G traffic data collected by an Italian Mobile Network Operator in Milan in 2015. Following [14], the trace is scaled to contemporary demand levels using peak and 5th-percentile statistics from a 2020 Chinese MNO dataset of 1460 BSs, while preserving the spatial and temporal structure of the original Milan data. Table II summarizes the resulting demand levels.

C. User-Level Realization

The optimization variables allocate resources to individual users, whereas the traffic trace provides aggregate sector demands $\Lambda_i(t)$. Therefore, each sector-slot demand is converted into a user-level realization before the link rates are computed.

For each sector $i = (b, s)$ and slot t, the number of users is drawn as

$$N_i(t) \sim \mathrm{Poisson}\left(\frac{\Lambda_i(t)}{\bar{\lambda}}\right), \tag{13}$$

where $\bar{\lambda} = 10$ Mbps is the reference mean demand per user. The corresponding user set $\mathcal{U}_i(t)$ is generated inside the sector footprint $\mathcal{A}_i$. User positions are drawn from a Gaussian spatial

distribution oriented along the sector boresight and truncated to the 120° sector region. To obtain heterogeneous per-user demands while preserving the aggregate sector traffic, we draw positive lognormal weights $x_u(t)$ for all $u \in \mathcal{U}_i(t)$ and assign

$$\lambda_u(t) = \Lambda_i(t) \frac{x_u(t)}{\sum_{v \in \mathcal{U}_i(t)} x_v(t)}, \qquad u \in \mathcal{U}_i(t). \tag{14}$$

Thus, the original sector-level trace is preserved exactly:

$$\sum_{u \in \mathcal{U}_i(t)} \lambda_u(t) = \Lambda_i(t), \qquad \forall i \in \mathcal{I},\ t \in \mathcal{T}. \tag{15}$$

The generated positions and demands are then used to compute the terrestrial and HAPS rates, and users are associated with the installed terrestrial carrier according to the strongest received reference signal as defined in Section III-B.

VI. Evaluation Metrics

The optimization problems return, for each time slot t, the terrestrial-only support activation vector $\boldsymbol{z}^{\mathrm{D}}(t)$ and the support activation vector after HAPS assistance $\boldsymbol{z}^{\mathrm{H}}(t)$. We first measure the slot-level terrestrial support requirement as

$$K_{\mathrm{D}}(t) = \sum_{i \in \mathcal{I}} z_i^{\mathrm{D}}(t), \qquad K_{\mathrm{H}}(t) = \sum_{i \in \mathcal{I}} z_i^{\mathrm{H}}(t). \tag{16}$$

Here, $K_{\mathrm{D}}(t)$ is the number of sectors requiring support-carrier activation in the terrestrial-only benchmark, while $K_{\mathrm{H}}(t)$ is the number of sectors that still require support-carrier activation after HAPS assistance. To distinguish persistent capacity needs from occasional peaks, we also compute the support-activation duty cycle of each sector:

$$d_i^{\mathrm{D}} = \frac{1}{|\mathcal{T}|} \sum_{t \in \mathcal{T}} z_i^{\mathrm{D}}(t), \qquad d_i^{\mathrm{H}} = \frac{1}{|\mathcal{T}|} \sum_{t \in \mathcal{T}} z_i^{\mathrm{H}}(t). \tag{17}$$

The quantities d_i^{D} and d_i^{H} measure how often sector i requires support activation before and after HAPS assistance, respectively. For a recurrence threshold η, we count the number of sectors whose support-activation duty exceeds that threshold:

$$N_{\mathrm{D}}(\eta) = \#\{i : d_i^{\mathrm{D}} \geq \eta\}, \qquad N_{\mathrm{H}}(\eta) = \#\{i : d_i^{\mathrm{H}} \geq \eta\}. \tag{18}$$

The reduction from $N_{\mathrm{D}}(\eta)$ to $N_{\mathrm{H}}(\eta)$ quantifies the effectiveness of HAPS in reducing recurring support needs.

VII. Results

We first report an aggregate view of the support-activation volume over the full traffic trace. Fig. 2 sums the number of required support-carrier activations over all time slots and sectors. In the terrestrial-only benchmark, the network requires 6621 sector-slot support activations. With HAPS assistance, this count decreases rapidly as the beam budget increases. Already with $K_{\max} = 4$, HAPS avoids 69.6% of the terrestrial-only activations, and with $K_{\max} = 8$ the avoided fraction reaches 95.5%. Beyond this point, the marginal gain becomes smaller: $K_{\max} = 10$ and $K_{\max} = 12$ avoid 98.9% and 99.8% of the activations, respectively. This indicates that most of the aggregate support demand is absorbed with a moderate number of HAPS beams, while additional beams mainly remove the small remaining activation volume. This count is useful to quantify how much the HAPS reduces the total use of terrestrial support capacity, but it should not be interpreted as the number of distinct sectors selected for permanent expansion.

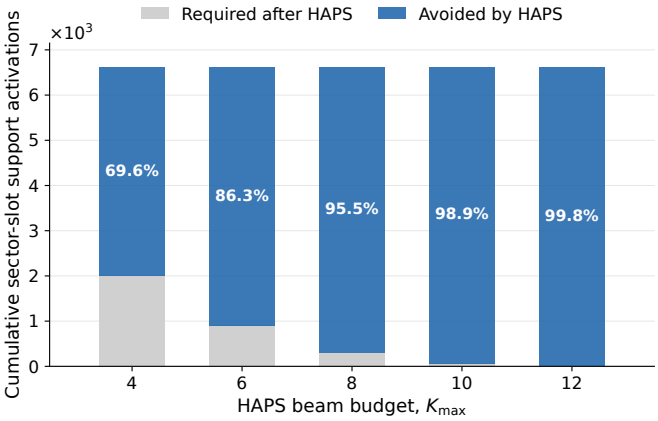


Fig. 2: Cumulative sector-slot support activations over the traffic trace.

Fig. 3 provides a conditional view of the HAPS contribution using the support requirements defined in (16). For each value of the terrestrial-only support requirement $K_{\mathrm{D}}(t)$, the figure reports the mean remaining requirement $K_{\mathrm{H}}(t)$ after HAPS assistance. The dashed diagonal is the no-reduction reference, for which the HAPS layer would leave $K_{\mathrm{H}}(t) = K_{\mathrm{D}}(t)$. Therefore, the vertical gap between each curve and this reference indicates the support-carrier activations avoided by HAPS under the same terrestrial-only support condition.

The curves show that the HAPS layer almost eliminates the remaining support requirement when the terrestrial-only support requirement is low or moderate. As $K_{\max}$ increases, this near-zero region extends to larger values of $K_{\mathrm{D}}(t)$, meaning that additional beams allow the HAPS to absorb more simultaneous support needs without requiring ground-side support-carrier activation. For larger values of $K_{\mathrm{D}}(t)$, the remaining support requirement increases again. This indicates the limiting regime of the modeled HAPS configuration: when support needs are more geographically distributed or when several MBS regions require assistance in the same slot, the finite beam budget and the fixed one-beam-per-MBS structure prevent all support activations from being removed.

The high-$K_{\mathrm{D}}(t)$ region should be interpreted as a limiting trend rather than as a precise ranking between beam budgets, because very large terrestrial-only support requirements occur less frequently in the trace. Nevertheless, this region is informative: even large $K_{\max}$ values can leave a nonzero $K_{\mathrm{H}}(t)$ when the support need is highly simultaneous. This suggests that further reductions would require either more active beams or a different beam design.

Fig. 4 moves the analysis from slot-level support requirements to sector-level recurrence. Using the support-activation

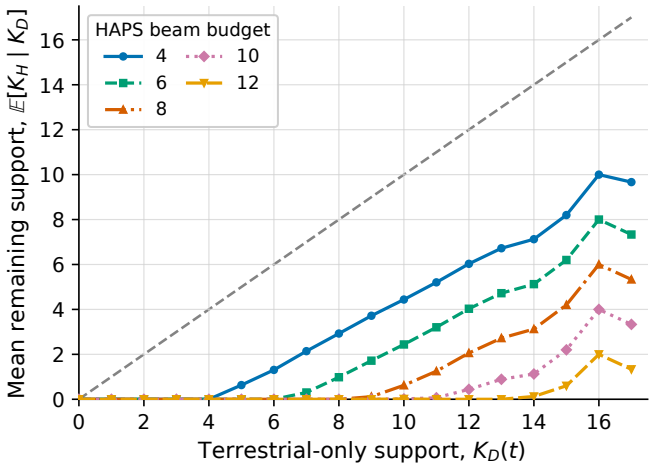


Fig. 3: Remaining support requirement after HAPS assistance as a function of the terrestrial-only support requirement.

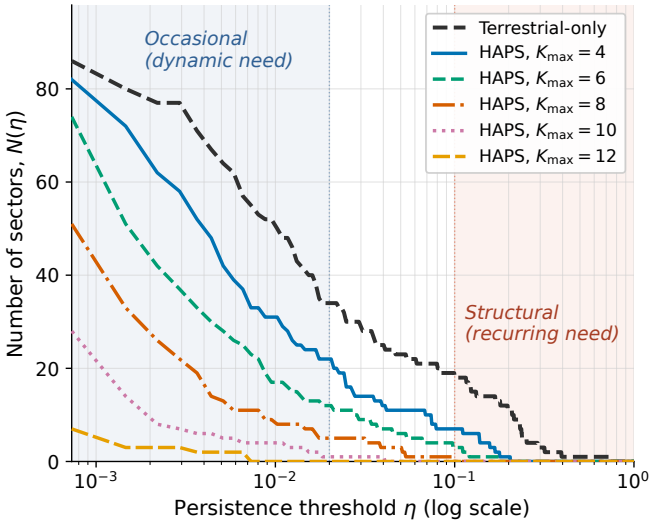


Fig. 4: Number of sectors with support-activation duty above threshold η.

duties defined in (17), the figure reports $N_{\mathrm{D}}(\eta)$ and $N_{\mathrm{H}}(\eta)$, i.e., the number of sectors whose support-activation duty is at least η. Low values of η capture sectors that require support only occasionally, while larger values of η identify sectors whose support need appears repeatedly over the trace.

The terrestrial-only curve shows that support requirements are not limited to a small number of isolated events. A large number of sectors require support at low recurrence thresholds, and several sectors remain above higher values of η, indicating repeated inability of the installed carrier to satisfy the traffic demand. HAPS assistance shifts the curves downward for all beam budgets. At low η, this means that fewer sectors experience even occasional support needs. At higher η, the reduction is more significant from a capacity-planning perspective, because it shows that HAPS reduces the number of sectors that repeatedly require terrestrial support.

The effect becomes stronger as $K_{\max}$ increases. With small beam budgets, some recurrent support needs remain, because the HAPS can assist only a limited number of MBS regions in each slot. As the beam budget increases, the HAPS-assisted curves approach zero over a wider range of η, showing that many sectors that would otherwise appear as recurring support candidates are converted into occasional or negligible support cases. The threshold η is used as a recurrence-analysis parameter, not as a universal deployment rule. The purpose of Fig. 4 is to show how HAPS changes the recurrence profile of support needs, not to prescribe a fixed expansion threshold.

Fig. 5 complements the recurrence analysis by showing where the support-activation duty is concentrated. In the terrestrial-only benchmark, recurrent support needs are spatially localized: only a subset of sectors exhibits high activation duty, but those sectors repeatedly require additional terrestrial capacity over the trace. After HAPS assistance, this spatial footprint is strongly reduced. With $K_{\max} = 6$, only weak residual activation remains in a small number of sectors, while with $K_{\max} = 8$, the recurrent support footprint is almost eliminated. This confirms that HAPS does not only reduce the aggregate number of support activations; it also reduces the set of locations that repeatedly appear as terrestrial capacity-addition candidates.

VIII. Discussion

By reducing the remaining support-carrier activations, the aerial layer gives operators more time to distinguish between temporary traffic peaks and capacity needs that persist over time. This distinction is important for network planning. A sector that requires support only occasionally may be handled through operational tuning or short-term radio optimization. A sector that repeatedly exceeds its installed capacity is a stronger candidate for terrestrial reinforcement. The recurrence analysis therefore gives a more planning-relevant view than aggregate traffic relief alone: it indicates whether HAPS changes the set of sectors that continue to appear as capacity-addition candidates.

More broadly, the proposed evaluation suggests a design objective for future HAPS-assisted RAN studies: minimizing the terrestrial capacity additions that remain necessary after aerial support. This objective can guide more advanced architectures and control strategies, including beam placement, beam-footprint selection, beam scheduling, power and bandwidth allocation, terrestrial–aerial user association, and interference-aware coordination.

Limitations and extensions: The benchmark uses fixed HAPS beam centers, a single HAPS platform, a fixed beam budget, and a noise-limited HAPS access abstraction; per-beam transmit power is fixed without an aggregate platform-wide power constraint. The per-slot problems are solved offline with known demand and link conditions and omit beam-reconfiguration and handover overheads. The results therefore quantify capacity-relief potential under the modeled resources rather than an economic comparison of HAPS and terrestrial

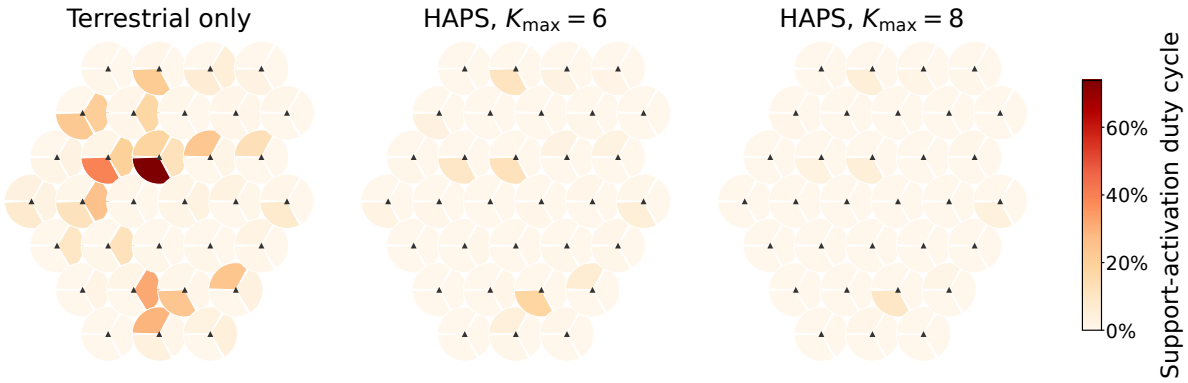


Fig. 5: Spatial distribution of the sector support-activation duty cycle over the full traffic trace. The left panel reports d_i^{D} for the terrestrial-only benchmark, while the middle and right panels report d_i^{H} after HAPS assistance with $K_{\max} = 6$ and $K_{\max} = 8$, respectively.

CAPEX/OPEX. Adaptive beam design, multi-HAPS coordination, interference-aware resource allocation, and explicit long-term investment modeling remain extensions.

IX. Conclusion

This paper evaluated HAPS assistance as a mechanism for reducing terrestrial capacity-expansion pressure under time-varying traffic demand. We formulated matched terrestrial-only and HAPS-assisted benchmarks in which support-carrier activation represents a sector-level capacity addition. The results show that a limited HAPS beam budget can substantially reduce both the aggregate support-activation volume and the number of sectors whose support needs recur over time. Thus, the value of HAPS is not only the traffic it carries, but also the reduction in persistent ground-side support commitments.

Acknowledgment

This work was supported by the European Union under the Italian National Recovery and Resilience Plan (NRRP) of NextGenerationEU, partnership on "Telecommunications of the Future" (PE00000001 - program "RESTART", focused project R4R and project 6GWINET).

References

[1] Naga Bhushan et al. "Network densification: the dominant theme for wireless evolution into 5G". In: *IEEE Communications Magazine* 52.2 (2014), pp. 82–89. DOI: 10.1109/MCOM.2014.6736747.

[2] European Telecommunications Standards Institute (ETSI). *Environmental Engineering (EE); Best Practice to Assess Energy Performance of Future RAN Deployment*. ETSI Technical Report ETSI TR 103 541 V1.1.1. Sophia Antipolis, France: European Telecommunications Standards Institute (ETSI), May 2018.

[3] Gunes Karabulut Kurt et al. "A Vision and Framework for the High Altitude Platform Station (HAPS) Networks of the Future". In: *IEEE Communications Surveys Tutorials* 23.2 (2021), pp. 729–779. DOI: 10.1109/COMST.2021.3066905.

[4] Md Sahabul Alam et al. "High Altitude Platform Station Based Super Macro Base Station Constellations". In: *IEEE Communications Magazine* 59.1 (2021), pp. 103–109. DOI: 10.1109/MCOM.001.2000542.

[5] Daniela Renga and Michela Meo. "Can High Altitude Platform Stations Make 6G Sustainable?" In: *IEEE Communications Magazine* 60.9 (2022), pp. 75–80. DOI: 10.1109/MCOM.002.2101048.

[6] Steve Chukwuebuka Arum, David Grace, and Paul Daniel Mitchell. "Delivering Extended Cellular Coverage and Capacity Using High-Altitude Platforms". In: *Electronics* 11.9 (2022), p. 1508. DOI: 10.3390/electronics11091508.

[7] Yohei Shibata et al. "Two-Step Dynamic Cell Optimization Algorithm for HAPS Mobile Communications". In: *IEEE Access* 10 (2022), pp. 68085–68098. DOI: 10.1109/ACCESS.2022.3186003.

[8] Yohei Shibata et al. "HAPS Cell Design Method for Coverage Extension Considering Coexistence on Terrestrial Mobile Networks". In: *IEEE Access* 12 (2024), pp. 55506–55520. DOI: 10.1109/ACCESS.2024.3390116.

[9] Zhiqing Wei et al. "Spectrum Sharing Between High Altitude Platform Network and Terrestrial Network: Modeling and Performance Analysis". In: *IEEE Transactions on Communications* 71.6 (2023), pp. 3736–3751. DOI: 10.1109/TCOMM.2023.3262305.

[10] Michela Meo, Daniela Renga, and Felice Scarpa. "Integrating Aerial Base Stations for Sustainable Urban Mobile Networks". In: *Proc. IEEE Global Communications Conference (GLOBECOM)*. 2022, pp. 1727–1733. DOI: 10.1109/GLOBECOM48099.2022.10001387.

[11] Tailai Song et al. "High Altitude Platform Stations: The New Network Energy Efficiency Enabler in the 6G Era". In: *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*. 2024, pp. 1–6. DOI: 10.1109/WCNC57260.2024.10571153.

[12] Görkem Berkay Koç et al. "HAPS-Enabled Sustainability Provision in Cellular Networks Through Cell-Switching". In: *Proc. IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)*. 2023, pp. 294–299. DOI: 10.1109/BlackSeaCom58138.2023.10299797.

[13] Cihan Emre Kement et al. "Sustaining Dynamic Traffic in Dense Urban Areas with High Altitude Platform Stations (HAPS)". In: *IEEE Communications Magazine* 61.7 (2023), pp. 150–156. DOI: 10.1109/MCOM.001.2200584.

[14] Mohamed Amine Mbarek et al. "HAPS for Mitigating RAN Traffic Peaks amid Rising Demand and B5G Evolution". In: *Proc. International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM)*. 2025, pp. 763–766. DOI: 10.1109/MSWiM67937.2025.11308964.

[15] 3GPP. *Study on channel model for frequencies from 0.5 to 100 GHz*. Technical Report TR 38.901. Version 17.0.0, Release 17. 3rd Generation Partnership Project (3GPP), 2022.

[16] C. A. Balanis. *Antenna Theory: Analysis and Design*. 3rd. Wiley, 2005.

[17] 3GPP. *Study on New Radio (NR) to support non-terrestrial networks*. Technical Report. Release 15. 3rd Generation Partnership Project (3GPP), 2020.