Integrating Terrestrial and Non-Terrestrial Networks in 6G: A Review of Architectures, AI-Driven Techniques, and Sustainability Strategies

Thavavel Vaiyapuri [1,*] and Karthiyayini Murugesan [2]
[1] College of Computer Engineering and Sciences, Prince Sattam bin Abdulaziz University, Al Kharj, Saudi Arabia
[2] Department of Computer Science and Engineering, Sethu Institute of Technology, Kariapatti, India
Email: t.thangam@psau.edu.sa (T.V.), karthiyayinisaravanan5@gmail.com (K.M.)

Manuscript received September 25, 2025; revised November 15, 2025; accepted November 27, 2025
*Corresponding author

***Abstract*—The integration of Terrestrial Networks (TNs) and Non-Terrestrial Networks (NTNs) is a foundational requirement for Sixth Generation (6G), enabling ubiquitous, resilient, and globally inclusive connectivity. However, existing surveys typically analyze this integration by concentrating on individual dimensions—such as architectural design, control and virtualization mechanisms, or Artificial Intelligence (AI)—while giving limited attention to sustainability considerations. This paper addresses this gap by introducing a unified architecture–AI–sustainability triadic framework, which forms the core contribution of the review. First, the paper provides a structured architectural synthesis that clarifies how different integration models influence the design and operational behavior of TN–NTN systems. Second, it consolidates the role of AI in enabling intelligent, adaptive, and context-aware network operation across integrated space–air–ground environments. Third, it advances sustainability as a primary design principle by synthesizing emerging strategies aimed at improving energy and carbon efficiency in future 6G infrastructures. By examining these three dimensions collectively, the review offers a coherent and comprehensive perspective on TN–NTN convergence, identifies persistent challenges including interoperability limitations and standardization gaps, and outlines future research directions needed to develop resilient, intelligent, and environmentally responsible 6G ecosystems aligned with United Nation Sustainable Development Goals (UN SDGs).**

***Index Terms*—6G networks, energy-aware design, green communications, intelligent orchestration, sustainable connectivity, TN–NTN convergence**

I. Introduction

The advent of Sixth Generation (6G) communication systems is poised to transform global connectivity by extending networks beyond traditional terrestrial infrastructure into the domain of Non-Terrestrial Networks (NTNs), including satellites, High-Altitude Platform Stations (HAPS), and Unmanned Aerial Vehicles (UAVs) [1]. The convergence of Terrestrial Networks (TNs) and NTNs offers unprecedented opportunities for achieving global coverage, enhancing capacity, and supporting a wide spectrum of mission-critical and emerging applications. This paradigm shift is not only a technological evolution but also a societal imperative, as it directly impacts digital inclusion, disaster management, autonomous mobility, remote healthcare, and sustainable development [2].

Over the last few years, a growing body of surveys and reviews has reflected the rapid progress in TN–NTN integration [3]. Some contributions have centered on architectural frameworks and control-plane convergence, while others have explored Artificial Intelligence (AI)- and Machine Learning (ML)-enabled orchestration, Internet of Things (IoT) integration, virtualization, or Third Generation Partnership Project (3GPP) - driven standardization efforts [4]. These studies collectively highlight the feasibility and promise of TN–NTN convergence. However, they often remain domain-specific or unbalanced, emphasizing one aspect disproportionately while overlooking others. More importantly, the dimension of sustainability—encompassing energy awareness, carbon reduction, and long-term resilience—remains either underrepresented or treated as a peripheral issue, despite its growing relevance in the 6G era [5].

This review seeks to address these gaps by advancing a triadic framework that positions architecture, artificial intelligence, and sustainability strategies as equally vital and mutually reinforcing dimensions of TN–NTN integration. Unlike previous works that present linear or siloed narratives, this study emphasizes the interdependence of the three pillars: architectural coupling defines the degree of flexibility available to AI-driven optimization, while both must operate in alignment with sustainable energy and carbon-aware practices. Through this lens, the review not only maps the state of the art but also highlights the pressing need for integrated solutions that combine technical scalability, adaptive intelligence, and environmental responsibility.

The importance of this perspective lies in its comprehensiveness and forward-looking orientation. By

doi: 10.18178/ijeetc.15.1.1-18

systematically categorizing research contributions under the three dimensions and avoiding overlaps, the review enables a clearer identification of synergies, as well as gaps (e.g., limited attention to sustainability-aware orchestration) [6]. At the same time, it underscores the practical significance of TN–NTN integration: enabling equitable global access, ensuring resilient services in extreme environments such as oceans and disaster zones, and aligning 6G with global sustainability frameworks such as the UN SDGs [7]. Through this unique vantage point, the paper unifies fragmented research into a coherent roadmap for building resilient, intelligent, and sustainable 6G connectivity. It extends beyond technical integration to embrace the operational and societal dimensions that will ultimately define the success of TN–NTN systems in practice. The main contributions are as follows:

1) Proposes a structured taxonomy of 6G-oriented TN–NTN architectures, clarifying how coupling depth, multi-layer composition, and cooperative mechanisms shape integration outcomes.
2) Synthesizes AI techniques across key 6G TN–NTN functions—including resource management, mobility, slicing, routing, and computation placement—highlighting their roles in intelligent orchestration.
3) Positions sustainability as a central 6G design requirement and examines energy- and carbon-aware strategies that connect architectural choices with AI-driven operation.
4) Introduces a unified triadic perspective that links architecture, AI-enabled optimization, and sustainability, offering an integrated foundation for developing resilient and environmentally responsible 6G TN–NTN system

II. Background

A. Evolution from 5G to 6G: The Role of TN and NTN

Fifth-Generation (5G) networks have marked a major leap in mobile communications, delivering enhanced mobile broadband, ultra-reliable low-latency communication (URLLC), and massive machine-type communications (mMTC) [8, 9]. Despite these advances, their dependence on dense terrestrial infrastructure limits coverage in remote, rural, maritime, and aerial environments, leaving large segments of the globe underserved. The transition toward 6G networks shifts the focus from localized performance enhancements to the pursuit of ubiquitous, global connectivity [10].

To achieve this vision, NTNs—encompassing satellites, HAPS, and UAVs—have emerged as indispensable complements to TNs [11]. While TNs provide high spectral efficiency and established infrastructure, NTNs offer wide-area, flexible, and resilient coverage, enabling service continuity where terrestrial systems alone fall short. The convergence of TNs and NTNs is expected to create a three-dimensional (3D) networking fabric, seamlessly linking users across space, air, ground, and sea [12]. In doing so, the evolution from 5G to 6G emphasizes not only higher data rates and lower latency but also expanded coverage, sustainable operation, and intelligent service delivery. Taken together, these requirements position TN–NTN integration as a foundational pillar of the 6G communication ecosystem, as illustrated in Fig. 1.

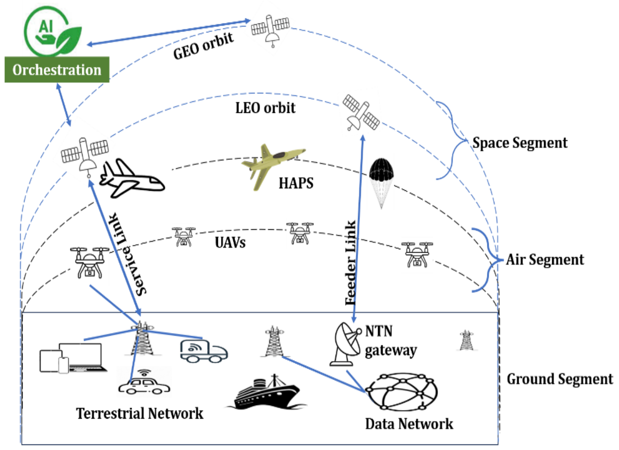

Fig. 1. AI-enabled orchestration of integrated terrestrial and non-terrestrial networks (TN–NTN) for global 6G connectivity.

B. Standards, Enablers, and Integration Drivers

The vision of integrated TN–NTN systems has been progressively shaped by standardization efforts, technological enablers, and real-world application demands [10]. On the standards side, 3GPP has laid important foundations [4]. Release 17 formally introduced NTN support in 5G, addressing issues such as mobility management and service continuity with satellite systems. This momentum is carried forward in Release 18 (5G-advanced), which strengthens interoperability by enabling direct-to-device satellite access and refining handover procedures between terrestrial and non-terrestrial segments. Looking ahead, Release 20 and beyond are expected to define the deeper convergence mechanisms required for 6G, with a focus on AI-native networking, ultra-reliable low-latency services, and energy-efficient operations. These activities are reinforced by international bodies such as the international telecommunication union, through its 2030 framework, and the Next Generation Mobile Networks alliance (NGMN), both of which highlight NTN integration as essential for achieving global inclusiveness and sustainability [13].

Complementing standardization are a set of technological enablers that make integration feasible. Multi-connectivity and dual connectivity provide users with simultaneous links across TN and NTN, ensuring seamless service continuity even under dynamic conditions. Integrated Access and Backhaul (IAB) architectures extend terrestrial footprints through satellites and high-altitude platforms. Reconfigurable Intelligent Surfaces (RIS) reshape propagation to recover link budget and improve spectral and energy efficiency. In parallel, Next-Generation Multiple Access (NGMA) techniques, such as Non-Orthogonal Multiple Access (NOMA) and Rate-Splitting Multiple Access (RSMA), increases user density and throughput [14].

C. Key Characteristics of 6G

Recent advances in 6G research outline a coherent set of characteristics that collectively mark a significant evolution in communication system design, capability, and operational logic beyond 5G. A dominant theme in the literature concerns the pursuit of extreme performance targets, with 6G expected to deliver terabit-per-second peak rates, ultra-low latency on the order of sub-milliseconds, and unprecedented levels of reliability to support emerging use cases such as holographic telepresence, large-scale autonomous mobility, and real-time distributed intelligence [10, 13]. These performance requirements are situated within a broader architectural shift toward 3D ubiquitous connectivity, where terrestrial, aerial, and satellite infrastructures are orchestrated as a unified communication fabric. Studies on TN–NTN integration consistently argue that such multi-layered spatial connectivity is essential for achieving seamless global coverage, robust service continuity, and resilience across diverse environments including sparsely populated, maritime, and disaster-prone regions [11, 12].

In parallel, the literature increasingly characterizes 6G as an AI-native network, moving beyond the auxiliary use of ML toward embedding intelligence directly into control, management, and orchestration layers. This perspective highlights the role of learning-enabled mechanisms in addressing the complexity of large-scale mobility, dense device ecosystems, spectrum coexistence, and dynamic multi-layer resource allocation across space–air–ground domains [15, 16].

Complementing this intelligence-driven orientation is the growing recognition of sustainability, security, and systemic resilience as intrinsic characteristics of 6G. Research underscores the need for energy- and carbon-aware network operation, motivated by the power constraints of NTN platforms and increasing global emphasis on environmentally aligned technologies [17]. At the same time, expanded multi-layer connectivity introduces new security and privacy challenges, prompting calls for integrated, cross-domain protection and fault-tolerant architectures capable of withstanding adversarial and environmental disruptions [18].

Taken together, these characteristics portray 6G as an ultra-performant, spatially pervasive, intelligence-driven, and sustainability-oriented communication paradigm. Within this framework, TN–NTN integration is not a peripheral enhancement but a structural requirement for realizing the comprehensive technical and societal aspirations articulated across the 6G research landscape.

III. Related Works

Research on the integration of TNs and NTNs has matured significantly in recent years, with multiple surveys and position papers addressing this topic from different perspectives. However, a close comparison reveals that existing reviews often emphasize one dimension disproportionately such as architectural frameworks, control-plane convergence, AI methods, virtualization, or standardization while leaving other aspects underexplored. The following synthesis compares the most relevant contributions and highlights how they differ, thereby contextualizing the unique positioning of this work.

The survey by Xu *et al*. [3] provides a broad discussion of integrated satellite–terrestrial network architectures for 6G, focusing on high-level reference designs, multi-layer compositions, and potential integration patterns. While it establishes an architectural foundation, it pays limited attention to orchestration mechanisms and energy/carbon-aware strategies. Similarly, the work in [12] frames integration as a 3D (space–air–ground) challenge, detailing issues such as doppler, synchronization, and multi-connectivity. Although valuable in capturing the physical constraints of TN–NTN integration, it stops short of proposing a systematic taxonomy that bridges architecture with AI-driven control or sustainability concerns.

In contrast, Kafle *et al*. [16] examine control-plane convergence through an integrated network control architecture. It emphasizes orchestration, performance monitoring, and policy enforcement across heterogeneous domains, thus offering insights into interoperability. However, its user-plane and sustainability perspectives remain limited. Extending this control-oriented view, the study of Ammar *et al*. [17] presents an in-depth survey of virtualization technologies. While it provides one of the most detailed accounts of orchestration stacks, it abstracts away radio-access fabric specifics and does not explicitly link orchestration strategies with carbon-aware site or gateway selection.

Several surveys have explored TN–NTN integration from an application or vertical perspective. Sultan and Chaudhary [19] investigate the integration of IoT into TN–NTN systems, offering insights into lightweight protocols, device constraints, and IoT traffic classes. This IoT-centric focus, however, treats AI orchestration and sustainability only tangentially, often equating sustainability with device-level energy saving rather than system-level carbon efficiency.

Standardization-focused reviews, such as [4], map ongoing research activities to 3GPP, and NGMN initiatives. These contributions are strong in identifying requirements and gaps from a standards perspective but remain limited in addressing the operationalization of AI-enabled closed-loop control or in linking standard features to sustainability levers.

AI and ML perspectives dominate the works in [6, 15]. The former highlights AI/ML as enablers for integration, discussing adaptive resource allocation, intelligent routing, and autonomous operation. However, it does not situate these techniques explicitly within different levels of TN–NTN coupling. The latter classifies ML approaches (supervised, unsupervised, reinforcement learning) and deployment styles (centralized, distributed, federated), mapping them to 3GPP integration scenarios with illustrative case studies. While this offers a methods-oriented taxonomy, it does not synthesize AI roles across architectural or sustainability dimensions.

An examination of the reviewed studies within the broader system context introduced earlier in Fig. 1, reveals that existing literature addresses only isolated components of TN-NTN integration. Some focus on the physical or architectural structure of TN–NTN systems, while others emphasize AI-driven mechanisms or specific application domains. However, very few consider how these technological elements relate to the operational intelligence and sustainability requirements that are intrinsic to 6G. Using Fig. 1 as a reference point makes this fragmentation evident and highlights the need for a more integrated synthesis that connects architectural design with AI-enabled operation and sustainability considerations.

Taken together, these surveys underscore the breadth of scholarship in TN–NTN integration. Yet, most adopt a unidimensional focus: architectural frameworks ([3], [12]), control/virtualization [17] IoT applications [19], standardization ([4]), or AI/ML methods ([6, 15]). None provide a cross-cutting perspective that systematically connects architectural design, AI-driven optimization, and sustainability strategies. This gap establishes the need for a comprehensive review that unifies these three dimensions, offering a balanced and integrative roadmap to guide future research and practice in building resilient, intelligent, and sustainable 6G networks. A comparative synthesis is presented in Table I, which highlights the contributions and gaps of existing reviews.

TABLE I: COMPARISON OF EXISTING TN–NTN REVIEW STUDIES AND IDENTIFIED GAPS

Ref.	Focus Area	Distinct Contribution	Key Gap Identified
[3]	Integrated satellite–terrestrial architectures for 6G	High-level architectural frameworks, multi-layer compositions	Limited focus on orchestration and sustainability
[4]	Research and standardization (3GPP, ITU-R, NGMN)	Strong mapping of requirements to standards	Lacks operationalization of AI-enabled control and sustainability considerations
[6]	AI/ML as enablers of TN–NTN integration	AI/ML functions: adaptive allocation, intelligent routing, autonomous operation	Does not situate AI within architectural coupling or sustainability frameworks
[12]	3D integration challenges (space–air–ground)	Physical-layer issues such as Doppler, synchronization, and multi-connectivity	No taxonomy linking architecture, AI, and sustainability
[16]	Network control-plane convergence	Control and orchestration functions across TN–NTN	Neglects user-plane integration and sustainability
[17]	Virtualization technologies (SDN, NFV, slicing)	Comprehensive review of orchestration stacks	Ignores radio-access fabric and carbon-aware resource orchestration
[19]	IoT integration into TN–NTN	Protocols, device constraints, IoT traffic optimization	Narrow IoT focus; lacks AI orchestration and system-level sustainability
[15]	ML-driven integration for 6G connectivity	Classification of ML approaches and architectures with case studies	Methods-oriented; limited synthesis across architecture and sustainability

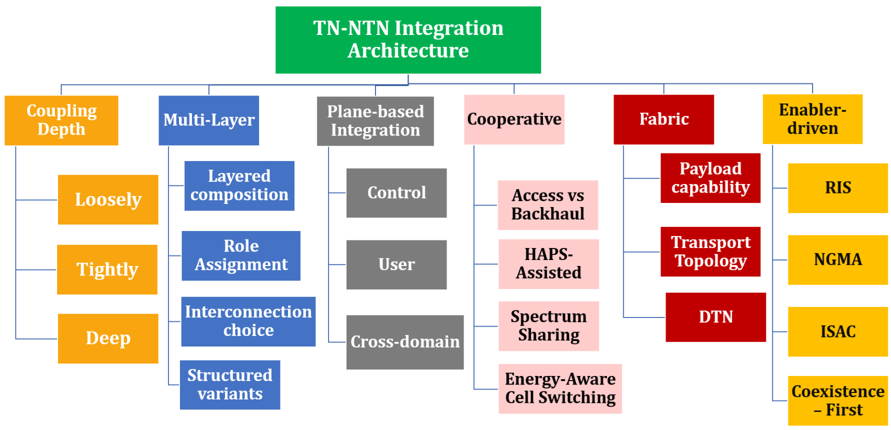

Fig. 2. Taxonomy architectural approaches for TN–NTN integration.

IV. ARCHITECTURAL TAXONOMY FOR TN-NTN INTEGRATION

The integration of terrestrial and non-terrestrial systems can be examined through multiple architectural perspectives. To ensure clarity and a structured analysis, this review classifies existing approaches into distinct taxonomic dimensions. Each taxonomy captures a specific aspect of integration, ranging from coupling depth to multi-layer composition and functional orchestration. The following subsections discuss these perspectives in detail, highlighting their design principles, inherent trade-offs, and implications for advancing 6G networks. The conceptual taxonomy is shown in Fig. 2, while detailed architectural illustrations of 6G systems can be found in existing surveys such as [20].

A. *Architectural Coupling*

In this review, *coupling* denotes the structural degree of TN–NTN integration—how access and core functions are placed, which standardized interfaces are used, and how identity, addressing, and policy are realized across the end-to-end protocol stack [16, 21].

1) *Loose coupling*

Loose coupling keeps terrestrial and non-terrestrial domains architecturally autonomous. Each runs its own access and core networks, and interconnection occurs at the IP layer through gateway peering or application proxies. Without a shared mobility anchor in the cellular stack, any continuity across domains is provided by transport or application mechanisms rather than network-native handover. Identities, addressing, and policy remain disjoint. This arrangement is easy to deploy and vendor-agnostic, yet duplicated control logic, inconsistent policies, and break-before-make mobility limit end-to-end quality and predictability.

2) Tight coupling

Tight coupling introduces architectural continuity by terminating non-terrestrial access on the same 5G core that serves the terrestrial domain through standardized interfaces. A shared mobility context enables coordinated handover and simultaneous connectivity across domains, so interruptions are bounded and measurable. Identities and policies are partly aligned typically a common subscriber space with segment-specific rules while operations and management can remain separate. The outcome is near-seamless service continuity and more efficient use of spectrum and transport, at the cost of policy reconciliation and integration complexity at the interworking boundary.

3) Ultra-tight coupling

Ultra-tight coupling treats terrestrial and non-terrestrial segments as co-equal parts of a single architecture. Access functions operate as peers, and the core is cloud-native and distributed, with user-plane processing placed at terrestrial edges and, where feasible, on high-altitude platforms or regenerative satellites. End-to-end slices span both domains under a unified policy and identity space, allowing resource pooling across spectrum, compute, and transport and enabling deterministic quality-of-service guarantees. Lateral meshes between non-terrestrial platforms reduce reliance on ground gateways and can lower end-to-end latency. Realizing this level of integration requires mature interoperability, fine-grained observability, and aligned security and exposure models among stakeholders.

A pragmatic evolution moves from IP-level interconnection (loose) to core-anchored interworking (tight) and ultimately to a unified access and core with cross-domain slice continuity (deep). Hallmark milestones include bringing non-terrestrial access under the same core control as the terrestrial network, enabling simultaneous connectivity with make-before-break inter-tier handover, distributing user-plane processing toward edges and aerial or space nodes to reduce latency and backhaul load, and extending slices so that policy, identity, and assurance apply uniformly across both segments.

B. Multilayer-Based Architectural Integration

A multi-layer approach specifies which physical strata are present, the role each plays along the service path, and how strata interconnect to form end-to-end routes. Two standard compositions are recognized in the literature: the Space–Air–Ground Integrated Network (SAGIN), in which the space tier (LEO/MEO/GEO) provides global reach and wide-area dissemination, the air tier (HAPS/UAV) supplies rapidly deployable, reconfigurable coverage and targeted capacity, and the ground tier (terrestrial access and core) delivers high-throughput, low-latency service while anchoring compute and storage; and its maritime extension, Space–Air–Ground–Sea (SAGS), which adds a sea tier (vessels, offshore platforms, buoys, coastal gateways) to address sparse infrastructure and long radio horizons offshore, shifting more access and aggregation to air and space [22].

Within either SAGIN or SAGS, performance hinges on role assignment and interconnection. Common patterns include space-centric aggregation with ground-based access, air-assisted coverage for restoration or demand hotspots, and mixed designs that stage lightweight functions aloft—aggregation, caching, header adaptation—to shorten paths and stabilize noisy links. Interconnection choices resolve to gateway-centric routes that ascend to space and return to ground gateways, mixed routes that combine gateways with lateral aloft segments (ISL/IHL) [23] to shorten paths and diversify failure modes, and air–ground aggregation paths where aerial nodes collect traffic from devices or small cells before forwarding to space or terrestrial backhaul. These choices determine where buffering and prioritization reside and thus the attainable envelopes for latency, resilience, and loss.

Deployments adapt these architectural patterns to their specific context. High-altitude platforms may be configured as super-macro canopies to stabilize wide areas and accelerate post-disaster restoration. Lateral aloft meshes are introduced where gateway density is limited or latency requirements are stringent. Integrated access-and-aggregation on air or space tiers is adopted where terrestrial build-out is constrained. The choice among these patterns is shaped by geography, demand dynamics, and the feasibility of gateway siting and backhaul.

Designing a credible multi-layer system requires careful balancing of interdependent factors [18]. The latency–coverage trade-off arises because space and high-altitude tiers extend reach but increase delay and timing dynamics, while dense ground deployments achieve lower latency at the cost of higher infrastructure investment. Capacity placement and gateway topology also shape performance, as inserting capacity aloft improves agility and restoration yet faces power and thermal limits, whereas concentrating capacity at gateways simplifies operations but elongates paths and risks bottlenecks; denser gateway grids, in turn, reduce route length but demand greater spectrum, siting, and backhaul resources. Energy posture and regulatory constraints further condition feasibility, with UAV endurance, satellite and HAPS power budgets, and terrestrial cell-sleeping policies requiring joint optimization, while spectrum allocations and coexistence requirements must be treated as primary architectural inputs.

C. Plane-Based Architectural Integration

Plane-based integration specifies how control, user, and management functions are distributed across terrestrial and non-terrestrial domains, and how their coordination ensures consistent policy, mobility, Quality-of-Service (QoS), and assurance. It provides the logical skeleton that binds heterogeneous strata into an end-to-end system.

1) Control-plane integration

The control plane governs registration, mobility, and policy enforcement. Loose integration maintains independent control domains with only gateway-level signaling, limiting handover efficiency [16]. Tight integration anchors terrestrial and non-terrestrial access on the same core functions, supporting shared subscriber identity, make-before-break handover, and coordinated dual connectivity. Deep integration extends this to a unified policy and intent namespace distributed across ground, aerial, and orbital nodes, enabling globally consistent but locally responsive control decisions.

2) User-plane integration

The user plane delivers forwarding, steering, and QoS treatment. Tight integration allows bearer- or flow-level splitting across domains while maintaining slice semantics. Deep integration distributes forwarding and buffering closer to access, including on high-altitude or regenerative satellites, to reduce path length and backhaul load. Key design issues include mobility anchor relocation without excessive state churn, buffering strategies for heterogeneous delays, and robust mapping of radio-layer QoS into transport schedulers.

3) Cross-plane co-design

Deterministic services such as time-sensitive control, holographic media, or joint sensing–communication demand coordinated plane behavior. Control commits resources and pre-computes feasible paths, the user plane ensures continuity through replication or fast switchover, and management validates conformance and initiates recovery under defined policy [24]. Time synchronization, admission control with budget tracking, and explicit failure semantics are prerequisites for making such guarantees verifiable.

4) Management and orchestration

The management plane spans configuration, telemetry, and assurance. In tight integration, management is federated with harmonized data models but domain-local operations [25]. Deep integration enables cross-domain orchestration in which a single service definition drives placement, scaling, and healing across layers, supported by closed-loop assurance. Achieving this requires normalized telemetry, synchronized timing across strata, and bounded monitoring overhead.

D. Cooperative-Driven Integration

Cooperative patterns capture how TN-NTN collaborate at run time once the overall layer composition and plane placement are defined. They determine how roles are allocated across tiers and how coordination sustains coverage, capacity, coexistence, and efficiency.

1) Access- versus backhaul-integrated operation

In access-integrated mode, the non-terrestrial stratum provides direct service to end devices, extending reach into maritime, remote, or emergency contexts. In backhaul-integrated mode, satellites or aerial platforms primarily transport traffic for terrestrial access nodes, stabilizing device requirements and reusing mature access technologies. The choice is deployment-specific [26]: access-integration offers coverage agility, while backhaul-integration supports dense access under fiber or microwave constraints. Tighter coupling enhances both, enabling seamless mobility in the former and coordinated traffic engineering in the latter.

2) HAPS-assisted cooperation

High-altitude platforms function as wide-area canopies and regional aggregation points. They absorb demand surges, mitigate terrain-induced coverage gaps, and collect traffic from UAV relays or ground cells before forwarding to gateways or satellites [27]. In disruptions, HAPS provide rapid baseline restoration, while in normal operation they support energy savings by allowing ground cells to enter sleep states. Their effectiveness depends on altitude, footprint, and payload budgets, and benefits can be amplified by predictive, AI-driven orchestration [28].

3) Spectrum sharing and coexistence

Where terrestrial and non-terrestrial domains operate in adjacent or overlapping bands, three strategies recur. Partitioned sharing assigns separate frequency, time, or spatial domains, yielding robustness but low efficiency [29]. Coordinated sharing exchanges intent and interference budgets across domains, increasing efficiency at the cost of tighter synchronization and trusted interfaces. Opportunistic sharing exploits idle spectrum windows, suitable for sparse deployments but dependent on rapid sensing and conservative protection. Deployment environments dictate the choice—coastal and urban edges often benefit from coordination, while remote areas can tolerate opportunistic operation.

4) Energy-aware cooperation

Multi-tier systems can shift users and flows across strata to minimize energy consumption while preserving QoS. Low-load ground cells may transfer users to aerial or LEO overlays, while high-power coverage may yield to terrestrial small cells when demand intensifies locally [30]. Decisions balance link budgets, trajectory visibility, switching costs, and carbon intensity of the underlying power source. Guarding against excessive switching and aligning policies with sustainability targets are key. AI-based prediction and policy learning further stabilize such strategies under fluctuating demand and channel conditions.

E. Fabric-Based Architecture

Fabric-based architecture defines the structural choices that determine which functions are executed aloft, how platforms interconnect with terrestrial gateways, and how the system sustains operation under intermittency. Unlike plane-based or cooperative perspectives, the focus here is on payload design and the transport fabric.

1) Payload capability

Payloads are categorized as transparent or regenerative [31]. Transparent payloads simply forward waveforms, offering low power use and simpler certification, but shift all adaptation tasks to terrestrial gateways, increasing latency and creating bottlenecks. Regenerative payloads re-emit signals at higher layers, enabling header adaptation, retransmission, caching, and lightweight user-plane functions. These reduce latency and improve traffic steering over noisy links but demand higher power, thermal capacity, and validation. Transparent payloads suit gateway-centric hub-and-spoke designs, while regenerative payloads support distributed processing and deeper TN–NTN integration.

2) Transport topology

Three interconnection patterns dominate. Gateway-centric routing uses terrestrial gateways, leveraging existing infrastructure but concentrating delay and resilience issues. Lateral inter-platform links create mesh topologies that shorten routes and add path diversity, though they require precise pointing and timing. Integrated IAB reuses spectrum for backhaul through satellites or HAPS, rapidly extending coverage in fiber-sparse areas [32]. Transparent payloads centralize QoS at gateways, whereas regenerative payloads enable on-board shaping to reduce jitter and blocking. In practice, hybrid approaches combine these patterns, balancing gateway density, mesh connectivity, and resource partitioning to optimize latency, resilience, and capacity.

3) Intermittency and DTN

In disrupted environments such as polar routes, oceans, or sparse gateway regions, delay-/Disruption-Tolerant Networking (DTN) supports store–carry–forward with time-aware contact plans and custody transfer [33]. DTN separates delay-tolerant from delay-critical flows, protecting real-time traffic during short visibility windows. Transparent payloads treat aloft segments as variable trunks, while regenerative payloads enable selective retransmission and aggregation. Properly designed, DTN transforms intermittency from a limitation into a predictable performance parameter, ensuring reliable operation where continuous end-to-end paths are unavailable.

F. Enabler-Driven Architecture

Enabling technologies are reshaping integrated TN–NTN architectures by influencing the placement of functions, the interfaces exercised, and the performance envelopes attainable. Three enablers stand out - RIS, NGMA, and ISAC alongside coexistence-first design principles that treat spectrum sharing as a primary constraint rather than a secondary optimization.

1) Reconfigurable Intelligent Surfaces (RIS)

RIS panels are engineered radio surfaces that manipulate propagation without full transceivers [34]. In TN–NTN systems, they can be deployed on terrestrial sites to mitigate blockage, mounted on high-altitude platforms to act as reconfigurable reflectors, or installed on maritime platforms to stabilize links near the horizon. Their architectural role lies in shaping coverage and recovering link budgets at stratum boundaries, thereby reducing the need for dense gateways or additional relays. Design choices concern placement, aperture, and the specific interfaces supported (service or feeder links). Properly integrated, RIS reduces energy per delivered bit and extends system agility, but requires dedicated control and calibration resources.

2) Next-Generation Multiple Access (NGMA)

NGMA methods, such as NOMA and RSMA [2, 35], shift contention from connection establishment to receiver processing. This enables uplink aggregation at HAPS or UAVs, where partially separated streams can be forwarded efficiently, or regenerative processing on satellites, where precoding or stream separation reduces feeder load. Architecturally, the main variables are the locus of processing (ground, air, or orbital) and the mapping of user groups to transport trunks. NGMA reduces signaling overhead and random-access contention but introduces additional processing and validation demands, particularly when functions are moved aloft.

3) Integrated Sensing and Communication (ISAC)

ISAC jointly designs waveforms and scheduling to serve communication and sensing tasks. In SAG and SAGS deployments, altitude and geometry provide advantageous baselines for localization and environmental monitoring. ISAC makes timing determinism and line-of-sight preservation first-order design constraints, leading to architectural choices [1] such as stable HAPS anchors, prioritized lateral aloft links, and compute placement on HAPS or regenerative satellites for fast fusion and feedback. These features favor deeper coupling, where shared identity and time bases can support both sensing and communications.

4) Coexistence-first blueprints

In scenarios, where TN and NTN share spectrum, coexistence becomes an architectural baseline [29]. Four patterns are common. Partitioned coexistence separates resources by time, frequency, or space, offering robustness at the cost of spectral efficiency. Coordinated coexistence exchanges scheduling and interference budgets across domains, enabling efficient use in contested areas but requiring trusted interfaces. Opportunistic coexistence exploits temporal or spatial gaps, suiting sparse deployments with bursty traffic. Reverse-pairing duplexing assigns opposite directions to TN and NTN in shared bands, reducing interference along feeder arcs and coastal cells. These strategies shape gateway siting, beam geometry, and telemetry needs, directly constraining feasible topologies.

These enablers are complementary but compete for power, spectrum, and control budgets. RIS can offset gateway density but requires physical siting and calibration. NGMA improves spectrum use but increases on-platform processing demands. ISAC strengthens sensing capabilities but narrows flexibility in path diversity and energy-saving policies. Coexistence-first designs cap peak throughput but can recover efficiency through RIS or NGMA if supported by sufficient observability. Collectively, these enablers are not peripheral but central to architectural design, determining where capacity is inserted, how paths are conditioned, and which service guarantees can be credibly sustained at scale. Table II summarizes how the architectural approaches for TN–NTN integration, as discussed above, have been addressed in the existing literature.

TABLE II: REVIEW OF ARCHITECTURES APPROACHES FOR TN–NTN INTEGRATION

Ref.	Focus	Architectural Approach	Contribution
[22]	Multi-layer integration of TN–NTN for 6G	3GPP roadmap + modular waveform + cross-domain management	Proposed a multi-layer approach combining terrestrial, aerial, and satellite; addressed seamless connectivity, resource optimization, and FSO management.
[23]	Hierarchical NTN with LAPs, HAPs, and satellites	Comparative analysis of multilayer configs	Analyzed different multilayer configurations; provided design guidelines for balancing flexibility, coverage, and latency vs. stand-alone systems.
[18]	NTN architecture under realistic hardware limitations	RIS + NGMA integration + impairment-aware architecture	Proposed multi-layer NTN including HAPS-SMBS, RIS-enabled UAVs, and NOMA; evaluated hardware impairments; suggested future directions in RF mitigation, UAV power
[16]	Integrated control plane for TN–NTN convergence	Control-plane architecture for orchestration & monitoring	Designed TN–NTN control plane for interoperability and QoS assurance; enabled performance monitoring and end-to-end orchestration across heterogeneous domains.
[24]	Cross-domain SDN for multi-layer space–terrestrial integrated networks	Domain-split SDN (satellite, aerial, terrestrial)	Proposed SDN-based MLSTIN architecture; improved reconfigurability and decision-making efficiency; identified challenges in managing heterogeneous devices.
[25]	Controller placement in SDN-based satellite–terrestrial networks	Distributed multi-layer hierarchical controller model	Proposed super/master/slave controller hierarchy; introduced time-slot stabilization and inter-layer strategies; improved scalability, stability, and cost-efficiency.
[26]	Cooperative NTN–TN architecture for extreme 6G coverage	Cooperative HAPS integration, shared spectrum, dual connectivity	Proposed cooperative NTN–TN with spectrum/resource sharing; evaluated HAPS performance in 6G simulator, enabling ubiquitous, cost-efficient extreme coverage.
[27]	Cooperative multilayer edge caching in satellite–terrestrial networks	Three-layer cooperative caching (BS, satellite, gateway) + iterative algorithms	Proposed cooperative caching strategies to minimize content retrieval delay; introduced cache hit probability analysis and optimal placement algorithms.
[28]	TaNTIN – Collaborative technologies for B5G/6G TN–NTN	Conceptual framework (AI, blockchain, MEC, tactile Internet, AR/VR)	Reviewed collaborative technologies in TaNTIN; highlighted their role in enabling QoS for telemedicine, e-education, gaming, and business applications.
[29]	Spectrum coexistence in S-band for TN–NTN	Stochastic geometry modeling of coexistence scenarios	Derived analytical models for coverage and data rate; identified conditions where TN–NTN coexistence is feasible; provided insights for spectrum sharing.
[30]	Space–Air–Ground–Sea (SAGS) cooperative integration system	AI-driven situational awareness (RL, GCN, multimodal fusion), robust transmission	Designed a SAGS cooperative architecture for global coverage; improved convergence, latency, and throughput via AI-enhanced awareness, reliability, and scheduling.
[31]	Dynamic topology & routing for STNs	Dyna-STN framework + OSPF overlay routing	Proposed dynamic discrete topology model with virtual nodes; improved routing, packet forwarding, and service continuity under time-varying STN topology.
[32]	Integrating TN–NTN via IAB technology	System-level simulation tool for IAB + satellite backhaul	Analyzed IAB with LEO/GEO satellites; evaluated registration time, link capacity, latency; showed feasibility with constraints from satellite links.
[33]	Delay & Disruption Tolerant Networking (DTN) for terrestrial & TCP/IP	Systematic Literature Review (SLR)	Surveyed DTN applications for terrestrial/space; classified studies by architecture, routing, performance; provided research gaps via color-coded matrix.
[2]	3GPP-driven TN–NTN integration	3GPP guidelines + AI-assisted beamforming	Illustrated TN–NTN integration roadmap via 3GPP; highlighted AI/beamforming role; validated effectiveness through numerical analysis.
[1]	AI-assisted maritime communications (SeaX-G architecture)	Review + AI/ML-based optimization scenarios	Proposed TN–NTN architecture for maritime; mapped 6G enablers (AI, spectrum, offloading) to use cases (fleet coordination, logistics, emergency response).

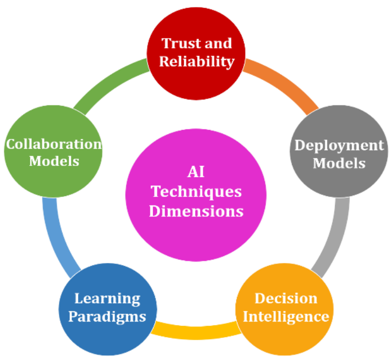

Fig. 3. Key dimensions of AI techniques for TN–NTN integration.

V. AI-DRIVEN TECHNIQUES FOR 6G TN-NTN SYSTEMS

AI has emerged as a cornerstone in shaping TN–NTN convergence, providing an overarching intelligence layer that influences resource management, mobility, scheduling, and orchestration. To capture this breadth, the following subsections organize AI contributions into learning paradigms, functional domains, and application-specific strategies. This structure enables a holistic view of how AI shapes integrated architectures into adaptive and sustainable communication systems. Fig. 3 illustrates the key dimensions of AI techniques for TN–NTN integration.

A. AI Learning Paradigm

AI has emerged as a central enabler for TN–NTN integration, governing both the optimization of network functions and the orchestration of heterogeneous resources. The architectural challenge lies in selecting appropriate learning paradigms for different decision loops ranging from fine-grained radio resource management to system-wide policy control.

Supervised Learning (SL) and Unsupervised Learning (UL): SL is applied where labeled datasets exist, for example in channel prediction or mobility classification, enabling accurate regression or classification of radio conditions [14]. UL is suited for anomaly detection, clustering of traffic patterns, and unsupervised beam selection in highly dynamic NTN scenarios, particularly where labeled data is scarce [15].

Reinforcement Learning (RL) and Deep RL (DRL): RL enables agents to adapt policies based on feedback from the environment, making it particularly relevant for spectrum sharing, inter-tier handover, and resource scheduling under uncertainty. DRL scales this paradigm to high-dimensional state spaces, allowing UAV trajectory optimization, adaptive satellite beamforming, and dynamic slice admission to be jointly optimized with latency and energy constraints [36, 37].

Federated Learning (FL): FL distributes model training across devices, gateways, or even satellites, aggregating model updates rather than raw data. This paradigm respects privacy and reduces backhaul load, while enabling adaptation to region-specific conditions. Architecturally, FL aligns with deep coupling, where common models are maintained but personalized updates reflect local context [38].

Self-supervised learning leverages unlabeled traffic and environmental data to pre-train representations that reduce the need for curated datasets [39]. This is particularly valuable in NTN environments where labeling costs are prohibitive and channel conditions evolve rapidly. Pre-trained models can then be fine-tuned for tasks such as interference prediction, anomaly detection, or beam alignment.

Large Language Models (LLMs) introduce a new layer of capability at the level of intent interpretation and policy synthesis. By processing high-level service descriptions or operator intents, LLMs can translate requirements into admissible network policies, interfacing with orchestration frameworks via standardized northbound APIs [40]. While not replacing domain-specific optimization, LLMs augment human operators by providing explainable reasoning, rapid policy adaptation, and multi-domain coordination across terrestrial and non-terrestrial segments.

Taken together, these paradigms define a layered AI toolkit. SL and UL provide foundational perception and clustering capabilities; RL and DRL deliver adaptive decision-making in dynamic environments; FL extends learning across distributed nodes; self-supervision improves scalability where labels are sparse; and LLM-assisted control elevates AI's role to high-level policy management. Their integration transforms TN–NTN systems from static infrastructures into adaptive fabrics capable of sustaining 6G-scale inclusiveness, resilience, and sustainability.

B. Radio Resource Management (RRM)

RRM is a critical function in the integration of TN–NTN, as it governs spectrum allocation, beam scheduling, power control, and interference management across heterogeneous strata. Architecturally, RRM decisions determine how resources are mapped onto multi-layer stacks, coordinated across coupling depths, and executed through enablers such as RIS and NGMA. AI has increasingly been adopted to address these challenges, offering adaptive decision-making under the uncertainty of dynamic topology, intermittent connectivity, and diverse service requirements.

1) Beamforming and beam-hopping.

Satellite and high-altitude systems rely heavily on multi-beam coverage and beam-hopping to manage spatial and temporal variability. Conventional static scheduling often leads to inefficient utilization under bursty traffic or mobility-induced demand. AI techniques, particularly RL and DRL, have been shown to optimize beam-hopping patterns by learning demand distributions and minimizing service outage [41], while SL models exploit historical propagation data to calibrate beam pointing and reduce sidelobe interference. In deeply coupled architectures, AI-driven beam management can be integrated with terrestrial schedulers to ensure end-to-end QoS continuity across space–air–ground domains.

2) Power allocation

Power allocation is a fundamental design concern in cooperative TN–NTN systems [42]. AI-driven methods enable adaptive and distributed power control that accounts for Doppler shifts, variable channel states, and platform-specific energy budgets. Multi-agent RL frameworks have demonstrated superior performance in balancing transmit power across heterogeneous nodes, while FL has been proposed for UAVs and LEO satellites to collaboratively optimize power without centralizing sensitive channel state data [2]. Such approaches align with sustainability goals by minimizing energy per delivered bit and supporting carbon-aware switching policies.

3) NGMA schemes

NGMA schemes such as NOMA and RSMA, rely on intelligent user grouping to achieve high spectral efficiency. Clustering methods based on unsupervised learning have been applied to partition users according to channel and QoS characteristics, while DRL-based schedulers dynamically adapt grouping to maintain stable throughput under gateway-sparse or interference-prone conditions [43, 44]. Architecturally, AI-enabled grouping decisions interact closely with the payload and transport fabric, determining whether grouping is realized at the terrestrial edge, the air tier, or regenerative satellites.

4) RIS control

RIS control offer an additional lever for shaping propagation in TN–NTN systems. However, the high dimensionality of RIS configuration makes manual optimization impractical. AI has been widely investigated as a solution: SL techniques map channel conditions to RIS states [34], while DRL frameworks optimize RIS configuration jointly with beamforming and power allocation, reducing outage probability and enhancing spectral efficiency [45]. In tight and deep coupling regimes, AI-driven RIS orchestration ensures consistency with end-to-end slice policies and facilitates cross-plane coordination.

Across these domains, AI converts RRM from a static configuration problem into an adaptive control loop that exploits real-time telemetry and historical data. By embedding learning paradigms—SL, RL/DRL, FL, and self-supervision—into beam management, power allocation, NGMA grouping, and RIS control, integrated TN–NTN systems achieve higher resilience, efficiency, and scalability. Importantly, these AI-driven strategies not only optimize performance but also reinforce the architectural goals of inclusiveness, sustainability, and service intelligence envisioned for 6G.

C. Mobility and Connectivity Management

Mobility and multi-connectivity are central challenges in the integration of TN–NTN. Architecturally, they determine how user sessions are maintained across heterogeneous strata, how anchors are relocated in multi-layer stacks, and how continuity is preserved under different coupling depths. Conventional handover and connectivity procedures originally designed for terrestrial domains struggle in integrated settings due to long propagation delays, Doppler shifts, and intermittent link visibility. AI offers mechanisms to predict mobility patterns, orchestrate make-before-break transitions, and optimize multi-connectivity policies in real time.

1) Connectivity management

In loosely coupled deployments, mobility events across TN and NTN are typically handled as disjoint domain transitions, often leading to service interruption. AI-driven approaches mitigate this by predicting handover triggers in advance, using trajectory data, historical mobility traces, and environmental context. RL and DRL have been employed to select optimal target cells in dynamic aerial and satellite scenarios, reducing handover failures and packet loss [15, 46]. FL extends these capabilities by enabling localized mobility prediction models at UAVs, HAPS, or satellites without centralizing user data, thereby aligning with privacy and latency constraints. In deeply coupled architectures, AI allows distributed mobility anchors to coordinate anchor relocation seamlessly, ensuring continuity even in gateway-sparse or high-mobility environments.

2) Multi-connectivity

Simultaneous connectivity across TN and NTN is critical for resilience, throughput aggregation, and URLLC. AI enhances multi-connectivity by dynamically selecting and weighting active links across space, air, and ground domains based on instantaneous channel conditions, queue states, and service-level objectives. Multi-agent RL has been applied to balance traffic across heterogeneous interfaces, improving reliability under fast-varying satellite topologies [37]. SL and UL methods assist in clustering and prioritizing link combinations, optimizing scheduling across terrestrial, aerial, and orbital segments. Moreover, AI-driven slice-aware policies can assign different flows to different links, latency-critical data through terrestrial or aerial paths, while delay-tolerant traffic is offloaded to satellite links, thereby harmonizing efficiency and QoS [36, 47].

By embedding AI into mobility and multi-connectivity functions, integrated TN–NTN architectures evolve from reactive handover and static link aggregation toward proactive, predictive, and policy-driven connectivity. Learning paradigms such as DRL, FL, and self-supervised models enable adaptation to dynamic mobility regimes, ephemeris dynamics, and traffic heterogeneity [38]. These approaches not only stabilize user experience across domains but also reduce signaling overhead, energy consumption, and outage probability, reinforcing the architectural objectives of seamless global service continuity and sustainable 6G operations.

D. Scheduling and Slicing

Scheduling and slicing are central to service differentiation in integrated TN–NTN architectures. They determine how heterogeneous resources are partitioned across layers, how priorities are enforced under different coupling depths, and how service-level objectives are preserved across the control, user, and management planes. AI augments these processes by enabling predictive, adaptive, and cross-domain coordination of resources.

1) Scheduling

Traditional schedulers rely on static heuristics that fail to adapt to fast-varying satellite visibility, mobility, or bursty demand [48]. AI techniques transform scheduling into a predictive and adaptive task. RL and DRL models dynamically assign time–frequency–power resources across strata, optimizing latency and throughput under non-stationary conditions [48]. Multi-agent RL extends this by coordinating scheduling between terrestrial base stations, UAV relays, and satellites, thereby balancing load while reducing inter-tier interference. SL models trained on historical traffic traces anticipate demand surges and pre-allocate capacity, while unsupervised clustering groups flows with similar QoS requirements for efficient batch scheduling.

2) Cross-domain slicing

Slice admission and orchestration in TN–NTN settings require policies that adapt to heterogeneous constraints such as orbital dynamics, gateway density, and backhaul availability [49]. AI-based decision engines predict slice demand, reserve resources proactively, and harmonize scheduler decisions across domains. For example, DRL-based frameworks map traffic classes to slices while respecting latency budgets, whereas FL allows distributed slice controllers on satellites or HAPS [50] to adapt policies locally and share model updates for global consistency. Self-supervised approaches are emerging to detect slice performance drifts without labeled fault data, allowing early adaptation before SLA violations occur.

3) SLA assurance

End-to-end SLA assurance requires that scheduling and slicing decisions remain consistent across multiple planes. AI facilitates closed-loop assurance by correlating telemetry from ground, air, and space nodes into unified performance views. Graph Neural Networks (GNNs) and anomaly detection models identify bottlenecks in multi-layer paths, triggering adaptive re-scheduling or slice reconfiguration [45]. AI-driven controllers also implement pre-emption and priority enforcement, ensuring that critical flows maintain service continuity even during congestion or link outages.

E. Routing and Transport Service

Routing and transport functions in integrated TN–NTN define how traffic traverses multi-layer paths, interacts with payload and topology choices, and adapts to coupling depth. Conventional routing protocols, designed for static terrestrial domains, are challenged by orbital dynamics, intermittent visibility, and heterogeneous link characteristics. AI enhances routing and transport by embedding predictive, adaptive, and data-driven decision-making into path selection, congestion control, and reliability management.

1) Path selection

Dynamic topologies in LEO constellations, UAV relays, and HAPS networks render static routing tables inadequate. DRL has been applied to learn optimal next-hop policies under changing connectivity graphs, reducing latency and improving delivery ratio. GNNs further enable AI-driven routing by embedding the time-varying network graph into a feature space where shortest-delay or energy-efficient paths can be inferred in real time [14]. In deep coupling architectures, AI-based controllers can coordinate routing jointly across space, air, and ground tiers, ensuring end-to-end path continuity and slice compliance.

2) Congestion and load balancing

Transport efficiency depends on how flows are balanced across heterogeneous routes and gateways. MARL frameworks have demonstrated the ability to distribute traffic across lateral inter-satellite links (ISLs), gateways, and HAPS relays, thereby avoiding bottlenecks and stabilizing latency variance [51]. SL and UL approaches leverage historical traffic patterns to anticipate congestion and pre-emptively reroute traffic before overload occurs. Such AI-driven balancing is particularly relevant to backhaul-integrated operation, where gateway capacity and feeder link utilization dictate performance.

3) DTN

In contexts where intermittency is intrinsic, AI augments DTN protocols by optimizing store–carry–forward operations. Predictive models exploit orbital ephemeris, mobility trajectories, and contact history to refine contact plans, custody transfer, and buffer prioritization [33]. RL agents dynamically schedule transmissions across contact opportunities, improving delivery ratios and reducing wasted retransmissions under uncertain link availability.

4) Transport-layer adaptation

End-to-end quality depends not only on routing but also on adaptive transport. AI-enhanced congestion control mechanisms, including DRL-based TCP variants, dynamically tune window sizes and retransmission timers to heterogeneous round-trip times in TN–NTN links [52]. Similarly, FL has been explored for adaptive coding and error control across distributed aerial and space nodes, enabling localized adaptation without centralizing telemetry. These approaches ensure that transport semantics (throughput, fairness, reliability) are preserved across strata.

F. Computation Offloading and Service Placement

Computation offloading and function placement are central to the design of integrated TN–NTN since they determine where tasks such as inference, caching, and analytics are executed across ground, aerial, and orbital resources. Architecturally, offloading decisions intersect with multi-layer compositions, cooperative patterns, and payload capabilities. Traditional offloading frameworks assume stable terrestrial links and homogeneous compute resources, but TN–NTN integration introduces additional uncertainties including intermittent visibility, heterogeneous energy budgets, and highly variable latency. AI has therefore emerged as an essential enabler for adaptive and context-aware offloading strategies.

1) Dynamic offloading decisions

AI techniques determine when and where tasks should be executed under dynamic connectivity conditions. RL and DRL have been applied to optimize task partitioning between terrestrial edge nodes, UAV relays, and LEO satellites, balancing delay against system load [53]. FL extends these schemes by enabling local models on UAVs and high-altitude platforms to collaboratively learn offloading policies without centralizing user data, which aligns with privacy and latency constraints [54]. Predictive supervised learning models have also been used to anticipate link intermittency, such as LEO handover windows, and proactively migrate tasks to stable anchors.

2) Service placement across strata

The location of compute functions, whether at ground edges, aerial nodes, or regenerative payloads, directly shapes service quality. DRL-based orchestration dynamically allocates Virtual Network Functions (VNFs) across tiers to minimize latency and reduce backhaul load [55]. Self-supervised learning methods have also been explored to infer optimal placement under sparse telemetry conditions, particularly in maritime or polar contexts [56]. Architecturally, these placement strategies align with programmable fabrics and cross-plane orchestration.

AI transforms computation offloading and placement from static heuristics into adaptive processes that account for mobility, intermittency, and heterogeneous resource availability. Table III summarizes how the AI techniques for TN–NTN integration, as discussed above, have been addressed in the existing literature.

TABLE III: REVIEW OF AI TECHNIQUES FOR 6G TN–NTN SYSTEMS

Study	Focus	AI Techniques	Contribution
[41]	Predictive beamforming for RSMA	Transformer + CNN (TranCN)	Introduced TranCN for predictive precoder design using historical CSI, improving WESR and reducing feedback in dynamic NTN channels.
[2, 42, 43, 44]	AI-enabled spectrum sharing	HDRL + AI-assisted frameworks	Proposed HDRL and AI-based spectrum sharing; positioned TN–NTN integration within 3GPP guidelines with validated numerical insights.
[34, 45]	AI for spectrum and beamforming management	AI/ML (RIS beamforming, SDN)	Reviewed RIS and SDN–AI frameworks for spectrum/beamforming optimization, improving energy efficiency and adaptability
[14,46,47]	AI for TN–NTN integration and connectivity	ML (supervised, RL, DRL) + AI frameworks	Surveys and frameworks on ML/AI-enabled TN–NTN integration; covers beamforming, spectrum sharing, Doppler mitigation, and AI-empowered NTN strategies.
[48]	User scheduling in satellite–HAPS–ground networks	Ensemble DNNs	ML-based scheduling to balance load in heterogeneous networks, enabling real-time optimization under high complexity.
[37, 49, 50]	AI-based network slicing & scheduling (vehicular, power grid, IoT resilience)	RL (PPO, A2C) + AI orchestration	Joint slicing & scheduling with RL, resilient slicing designs; reduced costs, higher slice satisfaction, and fault-tolerant STECN operations.
[13, 51, 52]	AI-enhanced routing & load balancing in TN–NTN	MADDPG, Federated RL, Fuzzy logic, CNN	Proposed AoI-aware routing, fuzzy-logic load balancing, and hybrid CNN–fuzzy routing; improved QoE, reduced overload, and optimized inter-satellite traffic.
[53, 55]	RL-based task offloading (vehicular & general)	RL (Q-learning, DQL, DDQN, DDPG)	RL-driven offloading strategies for MEC-enabled TN–NTN, reducing system overhead and energy consumption by up to 55%.
[54]	Privacy-preserving task offloading	DRL-based optimization	Introduced privacy-aware task offloading balancing completion time, energy, and user privacy.
[56]	Traffic offloading in hybrid satellite–terrestrial networks	Recurrent Neuro-Fuzzy + SDN	Dynamic offloading with RNFM + SDN; ~99% accuracy in resource allocation, outperforming conventional prediction algorithms.

G. *Efficacy and Progress of AI*

Existing research provides growing evidence that AI has already produced tangible performance gains across several core functions of 6G TN–NTN systems [2]. Reinforcement learning, Transformer-based prediction, and graph neural networks have demonstrated improved beam management, more accurate spectrum coordination, and more adaptive routing and congestion control in dynamic LEO–HAPS–terrestrial environments [6]. These techniques consistently outperform traditional rule-based and heuristic approaches, particularly under conditions of fast-changing topology, heterogeneous link quality, and intermittent visibility. AI has also proven effective in mobility prediction, handover optimization, task offloading, and service placement, reducing latency, energy consumption, and computation load in UAV- and HAPS-assisted MEC architectures [14, 15]. Collectively, these results indicate that AI is not merely a conceptual enabler but a practical driver of improved efficiency, resilience, and adaptability in integrated space–air–ground networks.

Despite this progress, the maturity of AI technologies within TN–NTN systems is uneven. Techniques such as RL-based scheduling, supervised channel prediction, and clustering-driven NGMA user grouping are relatively well validated and approaching deployment readiness. By contrast, federated learning for distributed NTN environments, DRL for mobility prediction, and AI-driven spectrum sharing require further evaluation under realistic propagation conditions, hardware limitations, and multi-operator settings. More emergent directions—including self-supervised learning for sparse NTN datasets and LLMs for intent-driven orchestration [40]—remain exploratory, constrained by limited datasets, insufficient explainability, and the computational constraints of satellites, HAPS, and UAV platforms. Furthermore, the efficacy of AI is strongly conditioned by architectural context, with tightly coupled TN–NTN systems offering the unified observability and shared control needed for reliable end-to-end optimization [47]. Overall, the literature shows that while AI has achieved clear and measurable gains, significant methodological, architectural, and operational challenges must still be addressed before large-scale, real-world deployment becomes feasible.

VI. SUSTAINABILITY STRATEGIES FOR GREEN TN-NTN

Beyond performance optimization, integrated TN–NTN systems must align with long-term sustainability goals. Energy efficiency, spectrum-aware operations, carbon-sensitive orchestration, and rigorous energy modeling collectively define how future 6G infrastructures can remain environmentally responsible while meeting service requirements. To reflect this, the sustainability discussion is organized into multiple dimensions, each addressing a specific lever of green operation—from optimizing day-to-day network functions to embedding carbon awareness into orchestration frameworks. Fig. 4 illustrates the key sustainability strategies for TN–NTN integration. The following subsections outline these strategies in detail.

Fig. 4. Key sustainability strategies for TN–NTN integration.

A. Network Operations

AI has become a critical enabler of energy-aware operations in integrated TN–NTN, allowing networks to balance sustainability objectives with service performance. By leveraging predictive analytics, adaptive control, and distributed learning, AI optimizes how energy is consumed across terrestrial, aerial, and orbital domains. Traditional schedulers often over-provision resources, leading to idle power waste [57–59]. AI-based scheduling frameworks predict demand patterns using historical traffic and mobility traces, enabling proactive activation of cells, beams, or payloads only when required. RL and DRL agents have shown effectiveness in minimizing idle consumption while preserving latency and QoS [60].

AI supports dynamic redistribution of traffic across ground, aerial, and space tiers based on energy efficiency. For example, low-demand periods allow UAVs or HAPS to offload users to terrestrial small cells entering low-power states, while during surges, satellite overlays absorb excess demand [61]. MARL enables cooperative policies that jointly consider endurance limits, link budgets, and QoS requirements, achieving balanced energy use without compromising service continuity [56, 62]. Energy efficiency also depends on reducing signaling and retransmission overhead.

AI-driven RIS control allows surfaces to adapt their reflection states in real time, improving link budgets and lowering transmit power requirements. Similarly, clustering and grouping strategies in NGMA schemes can be optimized through unsupervised learning and DRL, ensuring efficient user grouping that minimizes control rounds and retransmissions. AI-enabled green operations transform energy management from reactive policies into predictive and adaptive orchestration loops. By integrating AI into scheduling, load shifting, and enabler control, TN–NTN systems can significantly reduce energy per delivered bit while sustaining service quality, reinforcing both sustainability and scalability objectives for 6G networks.

B. Spectrum Management

Spectrum management in integrated TN–NTN architectures is closely tied to energy performance. Coexistence strategies that govern how terrestrial and non-terrestrial segments share spectrum—partitioned, coordinated, or opportunistic—carry distinct energy implications. Partitioned coexistence provides robust interference isolation [58] but requires excess spectrum allocation, leading to inefficient energy-per-bit outcomes [35]. Coordinated coexistence improves spectral efficiency through cross-domain scheduling and interference alignment but introduces control signaling overhead and tight synchronization requirements that raise energy cost. Opportunistic coexistence reduces active spectrum usage by exploiting idle gaps, yet frequent sensing and reconfiguration may offset energy savings in dense deployments.

Critical applications such as time-sensitive control or safety services demand deterministic spectrum use with strict latency [63] and reliability bounds, often at the expense of energy flexibility. In contrast, delay-tolerant or non-critical traffic can exploit elastic coexistence policies, allowing the system to trade off guaranteed capacity for lower energy use. The architectural challenge lies in jointly scheduling deterministic and elastic flows without compromising service assurance.

AI offers a means to balance coexistence efficiency and energy consumption. Predictive AI models also anticipate traffic and interference patterns, enabling proactive mode-switching that preserves service quality while minimizing unnecessary energy expenditure

C. Carbon-Aware Orchestration

Carbon-aware orchestration extends energy efficiency by explicitly considering the carbon intensity of power sources that sustain terrestrial, aerial, and orbital infrastructure. Rather than focusing solely on energy per delivered bit, this approach aligns workload placement and routing with the availability of renewable energy and low-carbon supply. By embedding real-time carbon-intensity maps into orchestration frameworks, routing and workload placement can favor gateways, edge sites, or compute clusters powered by renewable energy. AI-based predictors further anticipate regional variations in carbon availability, enabling proactive redirection of traffic or tasks before intensity peaks occur [64].

Carbon-aware strategies extend to gateway activation and lateral link utilization. RL and GNNs have been applied to select optimal gateway subsets and inter-platform links based not only on latency and capacity but also on renewable supply conditions [65].

This ensures that non-terrestrial segments dynamically align with green terrestrial entry points. Workload placement and backhaul routing can be continuously tuned to carbon constraints. AI-driven orchestration shifts compute-intensive functions toward sites with lower carbon footprints and defers delay-tolerant workloads to renewable-rich regions or times. Similarly, backhaul functions such as caching or aggregation can be reallocated across strata to minimize reliance on fossil-intensive gateways, with slice controllers enforcing service-level guarantees.

D. Energy Modeling and Efficiency

Energy modeling is fundamental to sustainable design in integrated TN–NTN. A unified framework must capture heterogeneous energy characteristics across ground, aerial, and orbital strata to ensure performance and sustainability. The primary efficiency metric is energy consumed per successfully delivered bit, encompassing transmit power, protocol overhead, and retransmissions [66]. This measure enables comparison across terrestrial, aerial, and satellite links and supports carbon-aware routing and resource allocation.

UAVs and HAPS are constrained by propulsion and payload endurance, while satellites rely on limited solar harvesting and thermal dissipation. These factors condition coverage duration, payload operation, and communication reliability, making accurate budget modeling essential. Efficiency metrics must integrate energy with service-level requirements such as latency and resilience. Composite indicators—linking energy per bit with delay or reliability trade-offs—guide orchestration decisions on routing, load balancing, and resource allocation across strata [67].

Standardized efficiency models that reflect link-level and platform-level budgets, combined with cross-stratum trade-off metrics, provide the foundation for sustainable TN–NTN operation and inform advanced AI-enabled orchestration strategies. Table IV summarizes how the sustainability strategies for TN–NTN integration, as discussed above, have been addressed in the existing literature.

TABLE IV: REVIEW OF SUSTAINABILITY STRATEGIES FOR GREEN TN–NTN ECOSYSTEMS

Study	Focus	Sustainability strategies	Contribution
[59–61]	Adaptive / cooperative transmissions in integrated STNs and UAV-assisted links	Trade-off among energy efficiency (EE), spectral efficiency (SE), and reliability (SER/latency)	Propose adaptive direct/cooperative transmission schemes; demonstrate cooperative relays and adaptive mode switching improve EE while sustaining QoS.
[68, 69]	Joint UAV offloading, edge processing, and satellite forwarding under uncertainty	Robust optimization with probabilistic constraints and online learning for uncertain UAV–satellite links	Weighted minimization of propulsion, computation, and transmission energy; robust designs improve sustainability under uncertain or dynamic NTN environments.
[70]	IoT task offloading via satellite–terrestrial terminals (TSTs)	Two-stage MEC offloading with sequential fractional & dual decomposition	E-CORA algorithm reduces IoT device energy by balancing ground/space offloading; lowers IMD energy consumption while maintaining service quality.
[62, 71, 72]	Dynamic scheduling and offloading under satellite energy constraints (LEO SEC)	Lyapunov-based dynamic optimization with energy budgets	Propose task scheduling and offloading strategies that minimize completion time while respecting LEO long-term energy limits; proven near-optimal guarantees.
[35, 58, 73]	Energy-aware spectrum/power allocation (RSMA, cognitive STN, MIMO-NOMA TSN)	EE-centric spectrum reuse with constrained DRL, interference-aware power allocation, and hybrid beamforming	Introduce constrained SAC, closed-form optimal power allocation, and clustering/beamforming to achieve higher EE under QoS and coexistence limits.
[63]	RIS/UAV-assisted IoT with satellite connectivity	STAR-RIS + UAV joint trajectory/power optimization	Propose ISRU framework with alternating optimization; achieves 40% energy savings and higher sum-rates vs. unoptimized schemes.
[5, 64, 65]	Multi-tier orchestration: TN–UAV–HAPS–Sat, maritime SAS-NTN, and UAV multicast orchestration	Cross-stratum sustainable orchestration (cell switching, power/trajectory optimization, multicast grouping)	Show improved EE, SE, and multicast delivery through latency-aware cell switching, maritime decomposition algorithms, and mobility-aware UAV scheduling.
[66, 67, 74]	Energy-aware edge computing & freshness for IoT (LEO SEC, AoI, MEC for STINs)	Energy-constrained MEC + AoI–EE trade-offs	Introduce Lyapunov and drift-plus-penalty schemes; balance energy vs. freshness; reduce IMD energy consumption with satellite-assisted MEC.

VII. HOLISTIC INTEGRATION FRAMEWORK FOR 6G

The integration of TN-NTN in 6G cannot be fully addressed by treating architectural taxonomies, artificial intelligence techniques, and sustainability strategies as independent dimensions. A comprehensive perspective is required to capture how these elements interact to form a coherent and adaptive communication fabric. To this end, we propose a holistic integration framework for 6G, which synthesizes structural, operational, and ecological perspectives into a unified foundation for TN–NTN convergence.

From the architectural standpoint, the taxonomies outlined earlier provide the structural backbone of integration. Variations in coupling depth, multi-layer composition, and functional plane distribution determine how control, user, and management functions are allocated and interconnected across terrestrial, aerial, and space domains. These design choices establish the attainable performance envelope—covering latency, reliability, and coverage—while simultaneously shaping the extent to which orchestration and intelligence can be embedded into the network.

Built upon this structural backbone is the AI-enabled intelligence layer, which operationalizes the architecture by enabling dynamic optimization, predictive decision-making, and distributed adaptation. Advanced learning paradigms—including reinforcement learning, federated learning, and self-supervised models—equip the network to evolve beyond static configurations, ensuring that beams, slices, mobility anchors, and service placement can be continuously aligned with real-time conditions. This intelligence layer transforms the architectural design into a responsive and adaptive fabric capable of withstanding high mobility, intermittent connectivity, and heterogeneous service demands.

The third pillar of the framework is defined by sustainability strategies, which embed ecological and societal imperatives into the integration process. Energy-aware operations, spectrum coexistence, carbon-sensitive orchestration, and unified efficiency modeling elevate sustainability from a secondary consideration to a primary design criterion. By integrating these strategies directly into architectural choices and AI-driven orchestration loops, the framework ensures that 6G deployments advance not only toward universal coverage and service intelligence but also toward climate resilience and environmentally responsible operation.

Collectively, Fig. 5 illustrates a holistic framework where architecture shapes the system's design, AI shapes its operational behavior, and sustainability shapes its long-term viability. Their convergence forms the cornerstone of TN–NTN integration in 6G, providing both a conceptual roadmap for standardization efforts and a practical guide for real-world deployment.

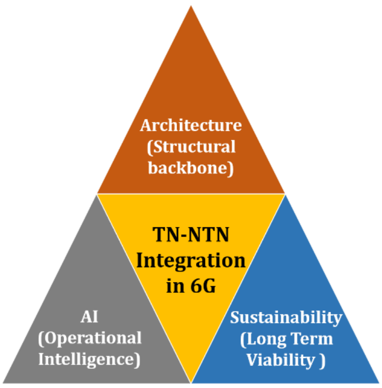

Fig. 5. Holistic framework for TN–NTN integration in 6G.

VIII. Challenges and Future Research Directions

Despite notable advances in TN–NTN research, several challenges continue to shape the trajectory of 6G development. Architecturally, integrating terrestrial, aerial, and satellite components remains difficult due to heterogeneous mobility, link characteristics, and protocol differences, and current standards provide only partial support for deep cross-segment coupling. These limitations highlight the need for architectural models that can accommodate multi-segment heterogeneity while enabling unified coordination.

AI-enabled operation introduces additional constraints. Existing models often rely on limited or unevenly distributed datasets and are rarely tested under the extreme dynamics characteristic of LEO constellations, UAV platforms, or time-varying propagation conditions. Ensuring robustness, reducing communication overhead for distributed learning, and designing resource-efficient models suitable for power-constrained airborne and satellite nodes remain open areas for investigation.

Sustainability considerations further complicate system design. The energy footprint of satellite constellations, HAPS platforms, and dense terrestrial deployments is substantial, yet systematic frameworks for carbon-aware routing, gateway selection, or lifecycle assessment are still underdeveloped. More work is needed to integrate energy and carbon modeling into architectural and operational decision-making.

Taken together, these challenges point to future research directions that require closer alignment between system architecture, AI-driven control, and sustainability principles. Promising avenues include integrated design frameworks that jointly optimize these dimensions, the use of digital-twin environments for predictive and cross-segment optimization, and sustainable-by-design approaches that embed environmental considerations early in 6G TN–NTN development.

IX. Conclusion

This review has examined TN–NTN integration through a unified perspective that treats architectural design, AI-enabled operation, and sustainability as interdependent elements of future 6G systems. The primary contribution of this work is the development of an integrated architecture–AI–sustainability triadic framework, which demonstrates how these dimensions collectively shape system behavior, design decisions, and long-term operational viability.

The architectural analysis provides a structured foundation for understanding the emerging integration models and the constraints inherent in multi-layer, multi-domain convergence. The synthesis of AI techniques shows how learning-driven mechanisms can enhance adaptability, autonomy, and coordination across dynamic space–air–ground environments. The sustainability discussion expands the scope by emphasizing the environmental implications of 6G deployments and demonstrating why energy- and carbon-efficient strategies must become core design principles rather than auxiliary considerations.

Together, these insights underscore the importance of advancing TN–NTN research through a holistic lens that recognizes the mutual influence of structure, intelligence, and sustainability. The proposed triadic framework offers a coherent basis for identifying unresolved challenges—such as interoperability gaps, resource limitations, and the absence of unified sustainability metrics—and provides a forward-looking direction for developing resilient, intelligent, and environmentally responsible 6G TN–NTN ecosystems.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

The authors declare that they have equal contributions to this paper. All authors had approved the final version.

FUNDING

The author extends her appreciation to Prince Sattam bin Abdulaziz University for funding this research work through the project number (PSAU/2025/01/37302).

ACKNOWLEDGMENT

The author extends her appreciation to Prince Sattam bin Abdulaziz University for supporting this research work through its Deanship of Scientific Research.

REFERENCES

[1] Giannopoulos, P. Gkonis, A. Kalafatelis *et al.*, "From 6G to SeaX-G: Integrated 6G TN/NTN for AI-assisted maritime communications—architecture, enablers, and optimization problems," *J. Mar. Sci. Eng.*, vol. 13, no. 6, 1103, 2025.

[2] M. A. Jamshed, A. Khausik, M. Dajer *et al.*, "Non-terrestrial networks for 6G: Integrated, intelligent and ubiquitous connectivity," *IEEE Commun. Stand. Mag.*, vol.9, no. 3, pp. 86–93, 2025.

[3] H. Xu, S. Z. Chen, and M. Ai, "Integrated satellite and terrestrial network architecture for 6G," *ZTE Technology Journal*, vol. 29, no. 5, pp. 9–15, 2023.

[4] M. Harounabadi and T. Heyn, "Toward integration of 6G-NTN to terrestrial mobile networks: Research and standardization aspects," *IEEE Wirel. Commun.*, vol. 30, no. 6, pp. 20–26, 2023.

[5] M. Ozturk, M. Salamatmoghadasi, and H. Yanikomeroglu, "Integrating terrestrial and non-terrestrial networks for sustainable 6G operations: A latency-aware multi-tier cell-switching approach," *IEEE Netw.*, 2025. doi: 10.1109/MNET.2025.3554393

[6] M. Khalid, J. Ali, and B. Roh, "Artificial intelligence and machine learning technologies for integration of terrestrial in non-terrestrial networks," *IEEE IoTM*, vol. 7, no. 1, pp. 28–33, 2024.

[7] T. Vaiyapuri, E. Kariri, A. Abdelrahman, and R. Aldawood, "Enhancing YOLO with adversarial and transfer learning for UAV-based urban wildlife conservation," *Journal of Soft Computing and Data Mining*, vol. 6, no. 1, pp 278–288, 2025.

[8] M. M. Azari, S. Solanki, S. Chatzinotas *et al.*, "Evolution of non-terrestrial networks from 5G to 6G: A survey," *IEEE Commun. Surv. Tutor*, vol. 24, no. 4, pp. 2633–2672, 2022.

[9] A. Yuan, Z. Yang, and Z. Sun, "Evolution of satellite communication systems toward 5G/6G for 2030 and beyond," *Engineering*, 2025. doi: 10.1016/j.eng.2025.06.025

[10] M. A. Jamshed, A. Kaushik, S. Manzoor *et al.*, "A tutorial on non-terrestrial networks: Towards global and ubiquitous 6G connectivity," *Foundations and Trends in Networking*, vol. 14, no. 3, pp. 160–253, 2025.

[11] N. Chen, Y. Song, Y. Cao *et al.*, "Network-layer perspectives on satellite--terrestrial integrated networks in 6G: A comprehensive review," *Engineering*, 2025. doi: 10.1016/j.eng.2025.05.012

[12] G. Geraci, D. López-Pérez, M. Benzaghta, and S. Chatzinotas, "Integrating terrestrial and non-terrestrial networks: 3D opportunities and challenges," *IEEE Commun. Mag.*, vol. 61, no. 4, pp. 42–48, 2022.

[13] W. Chen, J. Lee, J. Montojo, and M. Shafi, "Recent development of 5G-advanced in 3GPP: Commercial needs and a bridge to 6G," *GetMobile Mob. Comput. Commun*, vol. 29, no. 1, pp. 9–13, 2025.

[14] M. Ansarifard, N. Mokari, M. R. Javan *et al.*, "AI-based radio and computing resource allocation and path planning in NOMA NTNs: AoI minimization under CSI uncertainty," *IEEE Trans. Veh. Technol.*, vol. 73, no. 10, pp. 14874–14887, 2024.

[15] M. A. Aygul, H. Turkmen, H. A. Cirpan, and H. Arslan, "Machine learning-driven integration of terrestrial and non-terrestrial networks for enhanced 6G connectivity," *Comput. Netw.*, vol. 255, 110875, 2024.

[16] V. P. Kafle, M. Sekiguchi, H. Asaeda, and H. Harai, "Integrated network control architecture for terrestrial and non-terrestrial network convergence," *IEEE Commun. Stand. Mag.*, vol. 8, no. 1, pp. 12–19, 2024.

[17] S. Ammar, C. P. Lau, and B. Shihada, "An in-depth survey on virtualization technologies in 6G integrated terrestrial and non-terrestrial networks," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 3690–3734, 2024. doi: 10.1109/OJCOMS.2024.3414622

[18] F. Khennoufa, K. Abdellatif, H. Yanikomeroglu *et al.*, "A multi-layer non-terrestrial networks architecture for 6G and beyond under realistic conditions and with practical limitations," *IEEE IoTM*, vol. 8, no. 5, pp. 136–143, 2025.

[19] A. Sultan and M. Chaudhary, "Integration of terrestrial IoT and non-terrestrial networks in 6G and beyond: A comprehensive review," *Authorea Preprints*, 2025, doi: 10.36227/techrxiv.173933032.20096926/v1

[20] V. K. Quy, A. Chehri, N. M. Quy, N. D. Han, and N.T Ban, "Innovative trends in the 6G era: A comprehensive survey of architecture, applications, technologies, and challenges," *IEEE Access*, vol. 11, pp. 39824–39844, 2023.

[21] S. Saleh, P. Zheng, X. Liu *et al.*, "Integrated 6G TN and NTN localization: Challenges, opportunities, and advancements," *IEEE Commun. Stand. Mag.*, vol. 9, no. 2, pp. 63–71, 2025.

[22] A. Rago, A. Guidotti, C. Amatetti *et al.*, "Innovative multi-layer approaches for 6G integrated terrestrial and non-terrestrial networks," *IEEE Commun. Stand. Mag.*, vol. 9, no. 2, pp. 39–47, 2025.

[23] D. Wang, M. Giordani, M.-S. Alouini, and M. Zorzi, "The potential of multilayered hierarchical nonterrestrial networks for 6G: A comparative analysis among networking architectures," *IEEE Veh. Technol. Mag.*, vol. 16, no. 3, pp. 99–107, 2021.

[24] Y. Shi, Y. Cao, J. Liu, and N. Kato, "A cross-domain SDN architecture for multi-layered space-terrestrial integrated networks," *IEEE Netw.*, vol. 33, no. 1, pp. 29–35, 2019.

[25] O. A. Ayofe, J. M. Mu'azu, A.M. Tekanyi *et al.*, "Reliable SDN controller architecture for integrated satellite-terrestrial network convergence: A review," in *Proc. 2023 2nd International Conference on Multidisciplinary Engineering and Applied Science*, 2023, pp. 1–6.

[26] Y. Hokazono, H. Kohara, Y. Kishiyama, and T. Asai, "Extreme coverage extension in 6G: Cooperative non-terrestrial network architecture integrating terrestrial networks," in *Proc. 2022 IEEE Wireless Commun and Netw Conference*, 2022, pp. 138–143.

[27] X. Zhu, C. Jiang, L. Kuang, and Z. Zhao, "Cooperative multilayer edge caching in integrated satellite-terrestrial networks," *IEEE Trans. Wirel. Commun.*, vol. 21, no. 5, pp. 2924–2937, 2021.

[28] M. W. Akhtar and S. A. Hassan, "TaNTIN: Terrestrial and non-terrestrial integrated networks—A collaborative technologies perspective for beyond 5G and 6G," *Internet Technol. Lett.*, vol. 7, no. 1, e274, 2024.

[29] N. Okati, A. N. Barreto, L. U. Garcia, and J. Wigard, "Co-existence of terrestrial and non-terrestrial networks in S-band," arXiv preprint, arXiv:2401.08453, 2024.

[30] X. Wang, B. Yi, S. Kumari *et al.*, "Bridging terrestrial and non-terrestrial networks: A novel architecture for space-air-ground-sea integration system," *IEEE Wirel. Commun.*, vol. 32, no. 3, pp. 20–27, 2025.

[31] S. Li, Q. Wu, and R. Wang, "Dynamic discrete topology design and routing for satellite-terrestrial integrated networks," *IEEE/ACM Trans. on Netw*, vol. 32, no. 5, pp. 3840–3853, 2024.

[32] D. Pugliese, M. Quadrini, D. Striccoli *et al.*, "Integrating terrestrial and non-terrestrial networks via IAB technology: System-level design and evaluation," *Comput. Netw.*, vol. 253, 110726, 2024.

[33] A. Castillo, C. Juiz, and B. Bermejo, "Delay and disruption tolerant networking for terrestrial and TCP/IP applications: A systematic literature review," *Network*, vol. 4, no. 3, pp. 237–259, 2024.

[34] C. E. Worka, F. A. Khan, Q. Z. Ahmed, P. Sureephong, and T. Alade, "Reconfigurable Intelligent Surface (RIS)-assisted Non-Terrestrial Network (NTN)-based 6G communications: A contemporary survey," *Sensors*, vol. 24, no. 21, 6958, 2024.

[35] X. Li, Y. Li, X. Song, L. Shao, and H. Li, "Energy efficient power allocation for MIMO-NOMA-based integrated terrestrial-satellite networks," *Int. J Satell Comm Netw*, vol.42, no.3, pp.181–199, 2024.

[36] C. Zhu, Y. Bouteraa, M. Khishe, *et al*, "Enhancing unmanned marine vehicle path planning: A fractal-enhanced chaotic grey wolf and differential evolution approach," *Knowledge-Based Systems*, vol. 317, 113481, 2025. doi: 10.1016/j.knosys.2025.113481

[37] S. Y. Lien, and D. J Deng, "Autonomous non-terrestrial base station deployment for non-terrestrial networks: A reinforcement

learning approach," *IEEE Trans Veh Technol*, vol. 71, no. 10, pp. 10894-10909, 2022.
[38] V. K. Quy, N. M. Quy, T. T Hoai *et al*., "From federated learning to quantum federated learning for space-air-ground integrated networks," in *Proc. 2024 IEEE Conference on Standards for Communications and Networking,* pp. 402–407.
[39] T. Vaiyapuri and A. Binbusayyis, "Deep self-taught learning framework for intrusion detection in cloud computing environment," *IAES Int. J. Artif. Intell*, vol. 13, no. 1, p. 747. 2024.
[40] B. Rong and H. Rutagemwa, "Leveraging large language models for intelligent control of 6G integrated TN-NTN with IoT service," *IEEE Netw.*, vol. 38, no. 4, pp. 136–142, 2024.
[41] S. Zhang, S. Zhang, W. Yuan, and T. Q. S. Quek, "Transformer-empowered predictive beamforming for rate-splitting multiple access in non-terrestrial networks," *IEEE Trans. Wirel. Commun.*, vol. 23, no. 12, pp. 19776–19788, 2024.
[42] N. Heydarishahreza, T. Han, and N. Ansari, "Spectrum sharing and interference management for 6G LEO satellite-terrestrial network integration," *IEEE Commun Surv Tutor*, vol. 27, no. 5, pp. 2794–2825, 2024.
[43] M. Umer, M. A. Mohsin, A. A. Nasir, H. Abou-Zeid, and S. A. Hassan, "Intelligent spectrum sharing in integrated TN-NTNS: A hierarchical deep reinforcement learning approach," *IEEE Wirel. Commun.*, vol. 32, no. 3, pp. 64–71, 2025.
[44] Y.-C. Liang, J. Tan, H. Jia, J. Zhang, and L. Zhao, "Realizing intelligent spectrum management for integrated satellite & terrestrial networks," *J. Commn Inf Netw*, vol.6, no.1, pp. 32–43, 2021.
[45] G. Hua, G. Lin, Y. Zhang, and Y. Zhao, "Satellite-terrestrial integrated network slicing resource management based on reinforcement learning," in *Proc. Chinese Intelligent Systems Conference*, 2023, pp. 111–121.
[46] M. A. Jamshed, A. Kaushik, O. A. Dobre, Y. J. Zhang, and X. Lin, "Guest editorial: Unveiling the cosmic connection: AI's pivotal role in bridging terrestrial and non-terrestrial networks," *IEEE Wirel. Commun.*, vol. 32, no. 3, pp. 18–19, 2025.
[47] S. Mahboob and L. Liu, "Revolutionizing future connectivity: A contemporary survey on AI-empowered satellite-based non-terrestrial networks in 6G," *IEEE Commun Surv Tutor*, vol. 26, no. 2, pp. 1279–1321, 2024.
[48] H. Dahrouj, S. Liu, and M.-S. Alouini, "Machine learning-based user scheduling in integrated satellite-HAPS-ground networks," *IEEE Netw.*, vol. 37, no. 2, pp. 102–109, 2023.
[49] H. Wu, J. Chen, C. Zhou, J. Li, and X. Shen, "Learning-based joint resource slicing and scheduling in space-terrestrial integrated vehicular networks," *J. Commn Inf Netw*, vol. 6, no. 3, pp. 208–223, 2021.
[50] H. H. Esmat, B. Lorenzo, and W. Shi, "Toward resilient network slicing for satellite--terrestrial edge computing IoT," *IEEE Internet Things J.*, vol. 10, no. 16, pp. 14621–14645, 2023.
[51] Y. Gao, H. Yang, X. Wang *et al.*, "A fuzzy-logic-based load balancing scheme for a satellite-terrestrial integrated network," *Electronics (Basel)*, vol. 11, no. 17, p. 2752, 2022.
[52] F. Wang, D. Jiang, Z. Wang, Z. Lv, and S. Mumtaz, "Fuzzy-CNN based multi-task routing for integrated satellite-terrestrial networks," *IEEE Trans. Veh. Technol.*, vol. 71, no. 2, pp. 1913–1926, 2021.
[53] N. Waqar, S. A. Hassan, A. Mahmood, K. Dev, D.-T. Do, and M. Gidlund, "Computation offloading and resource allocation in MEC-enabled integrated aerial-terrestrial vehicular networks: A reinforcement learning approach," *IEEE Trans. on Intell Transp Syst*, vol. 23, no. 11, pp. 21478–21491, 2022.
[54] W. Lan, K. Chen, Y. Li, J. Cao, and Y. Sahni, "Deep reinforcement learning for privacy-preserving task offloading in integrated satellite-terrestrial networks," *IEEE Trans. Mob. Comput.*, vol. 23, no. 10, pp. 9678–9691, 2024.
[55] N. N. Ei, P. S. Aung, Z. Han, W. Saad, and C. S. Hong, "Deep reinforcement learning-based resource management for task offloading in integrated terrestrial and non-terrestrial networks," *IEEE Internet Things J.*, vol. 12, no. 1, pp. 11977–11993, 2024.
[56] M. K. Mishra and R. K. Mishra, "Traffic offloading and resource allocation enabled hybrid satellite-terrestrial networks in cognitive integrated systems using a recurrent neuro-fuzzy model," *Int J Satell Commun Netw*, vol. 42, no. 6, pp. 461–480, 2024.
[57] V. K. Quy, D. C. Nguyen, D. Van Anh, and N. M. Quy, "Federated learning for green and sustainable 6G IIoT applications," *Internet of Things*, vol. 25, 101061, 2024. doi: 10.1016/j.iot.2024.101061
[58] Y. Ruan, Y. Li, C.-X. Wang, R. Zhang, and H. Zhang, "Energy efficient power allocation for delay constrained cognitive satellite terrestrial networks under interference constraints," *IEEE Trans. Wirel. Commun.*, vol. 18, no. 10, pp. 4957–4969, 2019.
[59] S. Kandeepan, K. Gomez, L. Reynaud, and T. Rasheed, "Aerial-terrestrial communications: terrestrial cooperation and energy-efficient transmissions to aerial base stations," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 4, pp. 2715–2735, 2014.
[60] Y. Ruan, Y. Li, C.-X. Wang, and R. Zhang, "Energy efficient adaptive transmissions in integrated satellite-terrestrial networks with SER constraints," *IEEE Trans. Wirel. Commun.*, vol. 17, no. 1, pp. 210–222, 2017.
[61] D. Ron and J.-R. Lee, "Intelligent energy efficiency and service reliability optimization for UAV-aided terrestrial networks," *IEEE Trans. Green Commun. Netw.*, 2025. doi: 10.1109/TGCN.2025.3601729.
[62] J. Min, Z. Ximu, G. Xuemai, L. Xiaofeng, and G. Qing, "Joint UE location energy-efficient resource management in integrated satellite and terrestrial networks," *J. Commn. Inf. Net.*, vol. 3, no. 1, pp. 61–66, 2018.
[63] W. D Lukito, W. Xiang, P. Lai *et al.*, "Integrated STAR-RIS and UAV for satellite IoT communications: An energy-efficient approach," *IEEE Internet Things J.*, vol. 12, no. 9, pp. 11356–11371, 2024.
[64] S. S. Hassan, Y. K. Tun, N. H. Tran, W. Saad, C. S. Hong, and others, "Seamless and energy-efficient maritime coverage in coordinated 6G space-air-sea non-terrestrial networks," *IEEE Internet Things J.*, vol. 10, no. 6, pp. 4749–4769, 2022.
[65] F. Wang, H. Yang, S. Zhang, J. Shi, Z. Li, and T. Q. S. Quek, "Mobility-aware multicast orchestration for low-altitude UAVs with integrated terrestrial and non-terrestrial networks," *IEEE Trans. Commun*, vol. 73, no. 11, pp. 12854-12868, 2025.
[66] Q. Wang, X. Liang, H. Zhang, and L. Ge, "AoI-aware energy efficiency resource allocation for integrated satellite-terrestrial IoT networks," *IEEE Tran. Green Commun Netw*, vol. 9, no. 1, pp. 125–139, 2024.
[67] L. Cheng, G. Feng, Y. Sun, S. Qin, F. Wang, and T. Q. S. Quek, "Energy-constrained satellite edge computing for satellite-terrestrial integrated networks," *IEEE Trans. Veh. Technol.*, vol. 74, no. 2, pp. 3359–3374, 2024.
[68] X. Tang, Y. Jiang, R. Zhang, Q. Du, J. Liu, and N. Liu, "Energy-efficient integrated communication and computation via non-terrestrial networks with uncertainty awareness," *IEEE Internet Things J.*, vol. 12, no. 17, pp. 35165–35178, 2025.
[69] X. Gao, J. Wang, X. Huang, Q. Leng, Z. Shao, and Y. Yang, "Energy-constrained online scheduling for satellite-terrestrial integrated networks," *IEEE Trans. Mob. Comput.*, vol. 22, no. 4, pp. 2163–2176, 2021.
[70] Z. Ji, S. Wu, C. Jiang, D. Hu, and W. Wang, "Energy-efficient data offloading for multi-cell satellite-terrestrial networks," *IEEE Commun Lett*, vol. 24, no. 10, pp. 2265–2269, 2020.
[71] C. Peng, Y. He, S. Zhao, Y. Li, X. Wang, and B. Deng, "Energy efficiency optimization for uplink traffic offloading in the integrated satellite-terrestrial network," *Wireless Networks*, vol. 28, no. 3, pp. 1147–1161, 2022.
[72] J. Li, K. Xue, D. S. L. Wei, J. Liu, and Y. Zhang, "Energy efficiency and traffic offloading optimization in integrated satellite/terrestrial radio access networks," *IEEE Trans. Wirel. Commun*, vol. 19, no. 4, pp. 2367–2381, 2020.
[73] Q. Zhang, L. Zhu, Y. Chen, and S. Jiang, "Constrained DRL for energy efficiency optimization in RSMA-based integrated satellite terrestrial network," *Sensors*, vol. 23, no. 18, 7859, 2023.
[74] Z. Song, Y. Hao, Y. Liu, and X. Sun, "Energy-efficient multiaccess edge computing for terrestrial-satellite Internet of Things," IEEE Internet Things J., vol. 8, no. 18, pp. 14202–14218, 2021.

Thavavel Vaiyapuri is a faculty member College of Computer Engineering and Sciences, Prince Sattam bin Abdulaziz University, Saudi Arabia. She received her Ph.D. degree in computer science and has over a decade of academic and research experience. Her research interests include artificial intelligence of things (AIoT), sustainable and green computing, low-carbon AI frameworks with applications in healthcare. She is a member of the IEEE Computer Society and a Fellow of the Higher Education Academy (HEA), U.K. Dr. Vaiyapuri is dedicated to advancing interdisciplinary AI research that supports the UN SDGs.

Karthiyayini Murugesan received her M.Tech. degree in computer science and engineering from Sethu Institute of Technology, India, in 2019. She is currently preparing for doctoral research in AI with a focus on sustainable and intelligent systems. Her research interests include sustainable edge intelligence for healthcare. With a strong interest in interdisciplinary applications of AI, she aspires to develop socially impactful solutions aligned with UN SDGs.