

RESEARCH ARTICLE

Stochastic Modeling of Dependability Attributes for a Multilayer Satellite-HAP-Terrestrial System

Raina Raj¹  | S. Dharmaraja¹  | Vladimir Vishnevsky² | Dmitry Kozyrev^{2,3} 

¹Bharti School of Telecommunications Technology and Management, Indian Institute of Technology Delhi, New Delhi, India | ²V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia | ³RUDN University, Moscow, Russia

Correspondence: Raina Raj (rainaraj.curaj@gmail.com)

Received: 10 February 2025 | **Revised:** 26 May 2025 | **Accepted:** 22 July 2025

Funding: This work was supported by DST-RSF research project no. 22-49-02023 (RSF) and research project no. 64800 (DST).

Keywords: FSO/RF | mmWave communications | reliability | satellite-HAP-terrestrial systems | semi-Markov process | simulation

ABSTRACT

High-altitude platforms (HAPs) are an attractive solution for delivering long-term and stable relay services in communication systems, thanks to their superior capabilities such as extensive coverage, high mobility, and high-quality communication channels. The growing interest in deploying HAPs as relay nodes to bridge terrestrial and satellite networks has contributed to the advancement of integrated satellite-aerial-terrestrial (SAT) networks. To assess the dependability of such networks, a stochastic model is developed that incorporates the essential characteristics of satellite, HAP, and terrestrial components. A semi-Markov process (SMP) is employed to capture the non-Markovian behavior of the system states and to facilitate the evaluation of the model. The underlying study provides steady-state and time-dependent analysis of the system. The SMP framework enables the calculation of steady-state and transient probabilities, which are then used to determine key performance metrics such as system uptime and downtime probabilities, as well as the mean time to failure (MTTF) and reliability of the system. The effects of various component failure parameters on overall system performance are illustrated using graphical analysis. Finally, the analytical results derived from the model are validated through discrete-event simulation, ensuring the reliability of the proposed approach.

1 | Introduction

In recent times, the significance of land-based networks has significantly risen owing to the emergence of cutting-edge technologies like the fifth/sixth-generation (5G/6G) technologies and the Internet of Things (IoT). These networks do, however, have a number of limitations such as restricted coverage in isolated and rural areas, essentially because of the high costs associated with their implementation, and diminished network dependability as a result of environmental variables and natural disasters. The architecture of wireless networks for 6G is under development with the aim of enhancing coverage and reliability. The objective is to integrate nonterrestrial aerial networks like unmanned aerial vehicles (UAV), aircraft, and space networks with terrestrial aerial networks like cellular networks [1]. The combination of these networks leads to the integration of

the space-aerial-terrestrial (SAT) network [2]. The deployment of satellite communications is essential in numerous military, emergency, internet service, and commercial applications. The substantial growth of low earth orbit (LEO) satellites that may offer SAT network services has led to the establishment of worldwide connectivity through thousands of these types of satellites.

High-altitude platforms (HAPs), a crucial part of the SAT network, may provide efficient coverage in isolated areas and operate as a relay station between terrestrial and satellite to enhance the performance of the entire system. A HAP is a type of aircraft situated in the lower stratosphere, usually at an altitude of approximately 17 to 25 km [3]. It functions as a pseudo-satellite, offering services similar to those provided by traditional satellites. If a HAP is connected with a tether that supplies power from the ground, it is called a tethered high-altitude platform (tHAP) that

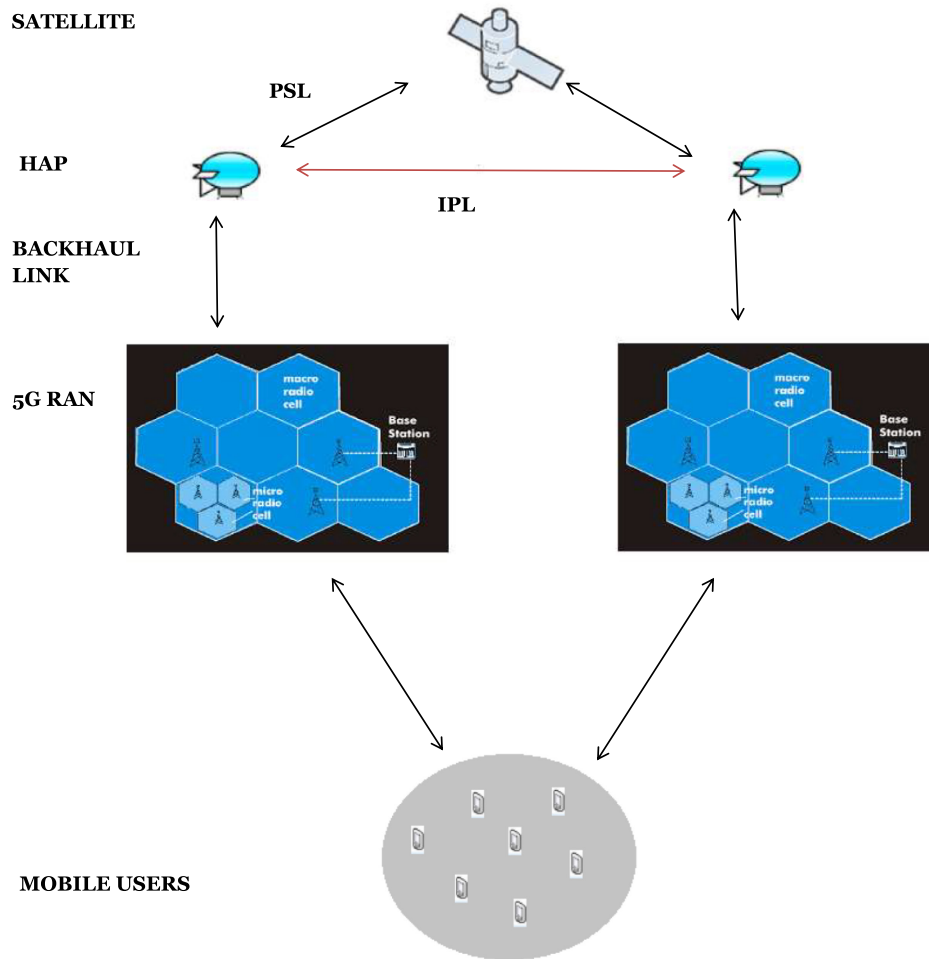


FIGURE 1 | Satellite-HAP-terrestrial network.

is established at the height of 200 m. HAP/tHAPs have a lot more potential than conventional satellites for monitoring natural disasters, observing agriculture and the atmosphere, tracking the weather, and relaying communications. They are also easier to deploy and maintain. Figure 1 represents an integrated SAT network where satellites and HAP/tHAPs are connected with each other through platform satellite links (PSL). HAP/tHAPs are connected with each other through the inter-platform link (IPL).

Furthermore, HAP/tHAPs are also connected to ground base stations, which serve as the primary interface for delivering services to end users. This results in a complex, multitier architecture involving terrestrial, aerial, and space-based components. Initially, HAPs were not considered part of this structure; however, their integration has become crucial for improving the quality of service (QoS), particularly in remote or underserved areas. Ensuring continuous and reliable service in such a heterogeneous SAT network is a major challenge, as it requires the coordinated operation of all components, including satellites, HAPs, and ground stations. One critical aspect of addressing this challenge is maintaining system dependability—minimizing service disruptions caused by component failures. This can be partially mitigated through design strategies such as redundancy and rapid component replacement. In the dependability analysis, reliability and availability are two key performance indicators for SAT networks, especially when deployed in

mission-critical applications such as disaster recovery, military communication, and remote internet access.

While numerous studies have focused on the architectural and design aspects of SAT networks incorporating HAP/tHAPs, there remains a significant gap in the literature regarding performance modeling and quantitative dependability analysis of such systems. Without a thorough modeling-based evaluation, it is difficult to understand how these networks perform under various failure scenarios or design conditions. Therefore, this research aims to fill that gap by providing a detailed dependability analysis of SAT networks integrated with HAP/tHAPs, using stochastic modeling techniques. By quantifying system reliability and availability through analytical or simulation-based approaches, this research aims to support more resilient network architectures in the deployment of SAT-HAP integrated systems.

1.1 | Relevant Work

Recent advancements in space technology and sophisticated space-based equipment have ushered in a new era of hybrid free-space optics (FSO)/radio frequency (RF) satellite communication. The demand for high data rates and robust satellite communication (SATCOM) connectivity is becoming increasingly vital, particularly for critical missions and disaster monitoring. In response,

several initiatives have proposed hybrid SAT FSO/RF systems, which aim to integrate satellite networks with terrestrial and aerial systems using both FSO and RF communication links [4].

Wideband communication in the millimeter wave (mmWave) bands is considered a cornerstone of 5G and is expected to play an even more significant role in 6G networks. The need for ultra-high data rates and the development of intelligent environments in 6G are effectively addressed by mmWave technologies. The technical potential of mmWave-enabled HAP communication systems is especially promising, offering a broad range of applications. The feasibility of using mmWave for HAPs to provide coverage in underserved areas has been explored in [5].

In recent years, research has increasingly focused on satellite, HAP, and mobile user integration [6–9]. A hybrid satellite-aerial-terrestrial network architecture for emergency scenarios was proposed in [10], where the authors compared terrestrial networks, low- and high-altitude platforms, and satellite systems; while also highlighting potential network technologies.

Satellite resource management has also garnered attention in recent studies [11–14]. For instance, [11] examined spectrum sharing between satellites and terrestrial networks; [12] investigated massive access via HAPs and LEO satellite backhaul solutions for rural regions lacking ground base stations; [13] proposed a hierarchical resource allocation strategy to manage the complexity of multilayer HAP-LEO architectures; and [14] analyzed asymmetry in optical wireless communication involving HAPs. Additionally, [15] proposed a unified space-air-ground IoT framework integrating HAPs and satellites, while [16] developed a constellation design methodology for large-scale HAP networks to ensure quality of service (QoS) and meet user demand. Dong et al. [17] presented a stochastic geometry-based approach for modeling the heterogeneous nonterrestrial networks (NTNs) by using the classical binomial point process and introducing a novel point process called the Matern hard-core cluster process (MHCCP) which possesses both properties of exclusivity and clustering. They validated the results through simulations. Dong et al. [18] proposed a stochastic geometry-based innovative model to characterize the impact of the finite-size distribution region of terrestrial terminals in the CSATN by jointly using a binomial point process (BPP) and a type-II Matern hard-core point process (MHCPP). The accuracy of the theoretical derivations was confirmed by Monte Carlo simulations.

To evaluate the reliability of HAP-enabled systems, several studies have proposed queueing models [19–21]. These works analyzed the reliability of various HAP components, such as tethers and rotors, using the k -out-of- n :F reliability model. However, their reliance on exponential failure and repair time distributions limits the applicability of their results to real-world scenarios. Recent contributions have advanced system modeling and performance evaluation. Choudhary et al. [22] proposed a stochastic model for HAP-5G integration. Guo et al. [23] studied the co-optimized performance of a multireconfigurable intelligent surface (RIS)-assisted integrated satellite–unmanned aerial vehicle–terrestrial network (IS-UAV-TN) serving multiple vehicle users. Bliss et al. [24] developed a HAP-assisted satellite relaying model for uplink data transmission, using Markov processes for performance evaluation and validating their results through simulation.

While several previous studies have contributed to the understanding of SAT network reliability, they often exhibit limitations, particularly in addressing multicomponent architectures involving HAPs. To the best of the authors' knowledge, there is a lack of reported work that comprehensively models the dependability of SAT networks by explicitly representing the interdependent behavior of multiple components within the system. Furthermore, most existing approaches rely on simplified assumptions such as exponentially distributed failure times, which may not accurately reflect the true operational characteristics of satellite components.

Motivated by these gaps, the objective of this study is to develop a comprehensive and realistic dependability model for SAT networks that

- incorporates a detailed multicomponent architecture,
- accounts for nonexponential failure behavior using general time distributions, and
- derives analytically tractable reliability and availability metrics to support system evaluation and optimization.

By doing so, this study aims to offer a more accurate representation of SAT network performance, especially in hybrid satellite-HAP architectures, where dependability is critical for ensuring service continuity in remote, mobile, or disaster-affected areas. A table consisting of the comparison of the existing literature with the underlying study is provided below in Table 1.

1.2 | Contribution

Satellites offer a significant advantage in global coverage compared with terrestrial cellular systems. They are capable of delivering sufficient capacity in remote and nonland areas and play a vital role in emergency communications. Additional benefits include lower launch costs, shorter deployment delays, and the potential to establish satellite networks using inter-satellite links (ISLs) [30]. Despite their wide coverage and rapid movement, these satellites face notable limitations. Most satellite beams cover areas with radii exceeding 100 km, making them inadequate for extremely dense regions that require high capacity. Moreover, their high speed complicates resource management using conventional radio methods, which affects the quality of service (QoS) for mobile and IoT users [31]. The substantial path-loss attenuation often limits traditional satellite communication systems due to weak links between satellites and ground stations. To address these limitations, a multilayer architecture involving HAPs was introduced. HAPs are situated above the ground, providing strong connections to both satellites and ground stations [9]. This setup enhances system performance in terms of throughput and energy efficiency. HAPs function as relay points between ground-based cellular networks and user devices, helping to expand coverage in remote or infrastructure-deficient areas. HAP systems employ a range of wireless technologies, such as LTE and emerging 5G/6G standards, to enable communication between the platforms and terrestrial networks. They hold promise for narrowing the digital divide by delivering connectivity to underserved regions and supporting emergency communications during crises. Positioned in the stratosphere,

TABLE 1 | Comparison of existing literature with underlying study.

Reference	Type of network	Type of analysis	Transient analysis	Simulation results
Want et al. [15]	SAT network including UAV	Lagrange dual decomposition and concave-convex procedure method	No	Yes
Dong et al. [16]	HAP broadband networks	Multiple HAP coverage is analyzed with link geometry	No	Yes
Dong et al. [17]	UAV	Stochastic geometry based approach for modeling the heterogeneous nonterrestrial networks	No	Yes
Dong et al. [18]	SAT network with UAV	Stochastic geometry based innovative model to characterize the impact of the finite-size distribution region of terrestrial terminals	No	Yes
Vishnevsky et al. [19]	Only HAP	k -out-of- n : F model for reliability analysis of HAP	No	No
Vishnevsky et al. [20]	Only HAP	Heterogeneous k -out-of- n system for reliability analysis of HAP	No	No
Choudhary et al. [22]	HAP VoIP network	Dependability analysis using semi-Markov process	No	Yes
Bliss et al. [24]	SAT network with HAP-LEO	Steady-state Markov chain analysis of the UE-HAP	No	Yes
Raj [25]	SAT network including HAP	Performance analysis using Markov regenerative process	No	Yes
Raj and Dharmaraja [26]	SAT network including HAP	Performance analysis using semi-Markov process	No	Yes
Liu et al. [27]	SAT network with UAV	Analytical expression for the outage probability	No	No
Singya and Alouini [28]	SAT network with UAV	Analytical expressions of outage probability, asymptotic outage probability, ergodic capacity, effective capacity	No	Yes
Raj et al. [29]	SAT network including HAP	Performance analysis using SMP	No	Yes
Proposed study	SAT network including HAP	Dependability analysis using semi-Markov process	Yes	Yes

HAPs provide aerial cellular coverage that complements traditional infrastructure. In contrast, terrestrial cellular systems depend on land-based towers and antennas to serve densely populated areas with widespread connectivity.

The presented study significantly contributes to the field by developing a comprehensive dependability stochastic model tailored specifically for the architecture of SAT networks.

- The proposed study adopts a multicomponent architecture that models a wide array of SAT network elements—including HAPs, base stations, power subsystems, antennas, radio access networks, and so on—each with independent failure behaviors (realistic failure assumption for HAP), rather than assuming a simplified exponential decay behavior of components or aggregated system structure.

- The underlying study employs a SMP framework that captures nonexponentially distributed sojourn times for some of the stated proposed architecture. The SMP modeling framework is deliberately chosen in this study due to its superior ability to represent systems where component failure and repair times deviate from the exponential distribution. Unlike conventional continuous-time Markov chains (CTMCs), which are restricted by the memoryless property and assume exponentially distributed sojourn times, SMPs allow for general, nonexponential time distributions between state transitions. This flexibility is essential for modeling SAT networks, where various components exhibit diverse and often complex failure behaviors. For instance, some components experience delayed or phased failures, where the probability of failure changes over time, violating the constant-rate assumption inherent to exponential

models. By adopting the SMP approach, the model accurately captures these heterogeneous and time-dependent reliability characteristics, leading to more realistic and insightful dependability metrics. This makes SMP a robust and rational choice for analyzing the operational dynamics of SAT networks and for supporting reliability-based design and optimization decisions.

- The model is analyzed time dependent and time independent in both scenarios, that is, to analyze the system's performance metrics, steady-state and transient solutions are derived for the SMP. Subsequently, availability and reliability metrics are derived for the proposed model, serving as critical benchmarks for performance evaluation. The model's analytical results are rigorously validated against simulation data, ensuring both accuracy and robustness.

The findings of this study have several important implications for the design, deployment, and operation of real-world satellite communication systems. By providing detailed availability and reliability metrics derived from a component-specific stochastic model, the study equips system architects and network engineers with data-driven insights for designing more resilient SAT network architectures. This can lead to more informed decisions regarding redundancy strategies and component selection. The ability to model and analyze nonexponential failure behaviors using SMP may allow operators to better predict component lifecycles and schedule preventive maintenance or upgrades, ultimately reducing unexpected downtimes and maintenance costs. By quantifying reliability and availability at a micro level, the proposed model may aid in identifying potential weak links within the system.

Despite the valuable contributions of this study in developing a comprehensive dependability stochastic model for SAT networks, there are a few points that limited the study. The proposed model incorporates certain simple assumptions to manage the complexity of SAT network behavior, for example, the number of components involved in the SAT network, and consideration of exponential failure behavior for those components. These simplifications, while necessary, may limit the model's ability to capture all real-world operational dynamics, such as intricate interdependencies or rare failure modes. Also, the analysis is based on a fixed network architecture. Real satellite networks, however, often exhibit dynamic behavior due to changing orbital paths, variable traffic patterns, or network reconfigurations, which may not be fully reflected in the model. These challenges can be a further topic of research for the underlying model.

1.3 | Paper Organization

The proposed study is divided into the following sections. Section 2 consists of the network topology and a detailed description of the dependability model for the underlying study. The analysis of the dependability model using SMP is presented in Section 3 along with the model's performance measures. Section 4 provides in-depth presentations of important performance measures with respect to the several parameters of the system, while Section 5 offers detailed graphical explanations of the performance metrics of the system with respect to the parameters of the system. Finally, Section 6 concludes with suggestions for further research.

2 | Model Specification

2.1 | Components of the Model

Modeling every component of an integrated SAT network is a challenging task owing to the intricate complexities involved. To tackle this issue, we propose a stochastic model that incorporates only the essential and fundamental components required for packet service execution. The following section provides a detailed discussion of these components.

- *LEO satellite*: LEO satellites have the capacity to provide ample connectivity in remote and nonland locations and play a vital role in emergency communication. Additionally, LEO satellites have several advantages, including lower launch costs, shorter launch delays, and greater feasibility for constructing satellite networks by leveraging inter-satellite links (ISLs). The satellite system is connected to the aerial network through a free space optics (FSO) or radio frequency (RF) link. If the FSO link is interrupted due to some technical or hardware fault, the RF link will start working as a backup link; a repair process for FSO will be initiated. The RF link is also subject to failure and may fail during the service procedure. A repair process will be initiated to repair the failed RF link.
- *HAP*: A HAP is a type of aeronautical platform that operates in the stratosphere, situated at an altitude between 17 and 20 km. HAPs can take on various forms, including airships or planes, and are primarily used to transmit signals to terrestrial systems, which in turn provide services to users. Due to potential technical, software, or hardware faults, HAPs may encounter failures that disrupt signal transmission. To mitigate such disruptions, a proposed model includes the use of a redundant or backup HAP.
- *mmWave beamforming*: The HAP's ability to support more demanding application scenarios is limited by its communication capacity, which has become a bottleneck. The sub-6 GHz frequency band, which is heavily occupied, is insufficient to meet the requirements for ultra-high data traffic. However, a promising solution for HAP communications is the use of millimeter-wave (mmWave) frequency bands. With large antenna arrays packed into a small area on the HAP, three-dimensional (3D) beamforming can be performed. In addition, HAPs can serve as aerial access points or relays to significantly improve the coverage and QoS for terrestrial mmWave cellular networks. Due to the position and structure of antennas, the beamforming might stop functioning; subsequently, the connection to the terrestrial network will be disrupted. A repair facility is immediately initiated to make the mmWave beamforming work.
- *Terrestrial system*: The terrestrial network is composed of base stations (BS), macro stations (MS), and small cells (SC). The radio access network (RAN) is a network design utilized by telecommunication service providers to transmit signals to cellular users. At the core of this network lies the BS, which acts as a stable transmission point. The BS is connected to multiple antennae, known as an MS, that facilitate the transfer of signals to the cellular

users. To improve the QoS, the MS is further connected to small antennae, known as SC. Together, the BS, MS, and SC form the RAN. The signals transmitted by HAP are received by these components of the terrestrial system and further transferred to the user. If, due to some hardware or software fault, any of these components stop functioning, the terrestrial system will not be able to provide service.

- **Server:** In a communication network, a server is a crucial resource that not only facilitates packet processing but also responds to client requests. However, as the volume of incoming packets rises, the server's resources may become degraded and eventually lead to failure.
- **Power Supply:** Any disruption to the power supply (UPS) can result in the loss of packet reception; ultimately impacting network performance.

2.2 | Functioning of the Model

Figure 2 shows the state-transition diagram of the proposed model. In this study, it is assumed that only one component fails at a time, which is a common simplification in reliability modeling. This assumption enables the application of a SMP framework with manageable state-space complexity, facilitating analytical derivation of availability and reliability metrics. However, this does not fully reflect real-world scenarios where multiple components may fail concurrently due to shared environmental disruptions (e.g., solar flares, thermal cycling, or cascading faults). The current model also assumes a single repair facility that serves all components sequentially. In such a case, resource contention arises when multiple components require repair at the same time. To address this, future enhancements to the model could incorporate repair scheduling strategies, such as priority queueing, preemptive repair, or parallel repair facilities as per the system architecture.

As a result, the model may overestimate system availability and reliability, particularly in high-risk operational environments. The exclusion of multifailure scenarios also omits the queuing and prioritization challenges that arise when limited repair resources must handle multiple failures simultaneously. Although empirical validation using real satellite failure data would strengthen the model, such datasets are not publicly accessible due to operational security and proprietary constraints. We acknowledge this limitation and identify it as a key area for future model extension—potentially through integration with queueing theory or simulation-based multifailure frameworks.

Now, the detailed transitions between various states of the dependability model are outlined below.

- The initial state of the system is State 0, in which all components are functioning correctly. If a fault occurs in the FSO link between the satellite and HAP, the backup RF link will activate. The repair process for the FSO link will commence simultaneously in State 1. In the event of an additional failure in the RF link, the entire system will enter a breakdown state, leading to service termination in State 1_D .
- In the event that the HAP ceases to function, it will transition to State 2, and the backup HAP will take over. If the backup HAP also malfunctions, the system will enter State 2_D , and a repair process for the backup HAP will be initiated. If both HAPs are in the breakdown state, the service process will come to a halt. In state 2_D , both the primary HAP (HAP1) and the redundant HAP (HAP2) are nonoperational simultaneously. Our model currently assumes nonpreemptive, independent repair processes, meaning that repair of a component, once initiated, continues uninterrupted regardless of subsequent failures or system transitions. This aligns with practical considerations, where the primary component critical to service restoration is given precedence. This is implemented by setting repair

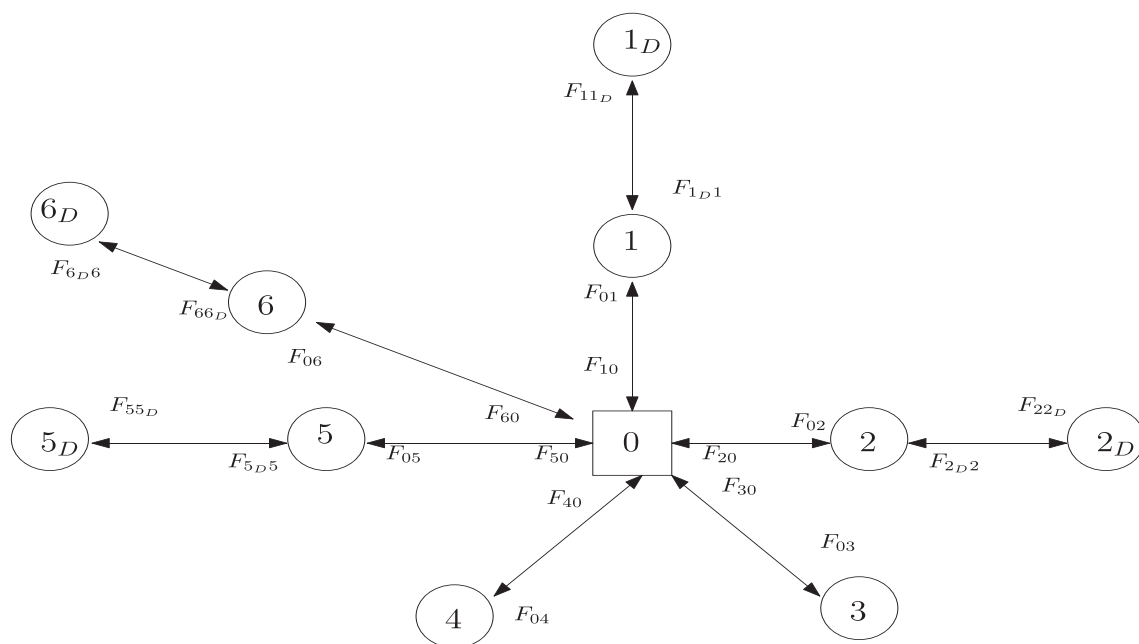


FIGURE 2 | State transition diagram for the underlying model.

transition rates from state 2_D to the state where only HAP1 is restored (e.g., state 2), rather than both being repaired simultaneously or HAP2 being repaired first. Since each state transition in an SMP marks a renewal point, the sojourn time distribution resets upon entering state 2_D . However, our model captures the deterministic knowledge that the repair of HAP1 had already been in progress by adjusting the transition rates out of 2_D accordingly. That is, we reduce the expected remaining repair time for HAP1 in 2_D to reflect that its repair was already partially completed in the

TABLE 2 | Description of states.

States	Description
State 0	In this state, all the components of the system are in operable condition.
State 1	In this state, the FSO link stops working and in place, the RF link starts working. Also, the repair facility is initiated for the FSO link.
State 1_D	In this state, both FSO and RF links are interrupted, therefore, the system stops operating. The repair facility, started in state 1 continues to perform in this state.
State 2	In this state, the operation of the HAP is interrupted due to some error; hence, a standby HAP starts working in place of the main HAP. Also, the repair process for the main HAP is initiated.
State 2_D	Here, both the HAPs stop functioning, resulting in the breakdown of the whole system. The repair process for the functioning of HAP continues in this state as well.
State 3	In this state, the beam-forming of mmWave antenna is disrupted due to some technical or external factor. A repair facility is initiated for the proper functioning of the antenna.
State 4	In this state, the BS/MS/SC are in a down state and the system stops functioning. A repair process is started for the functioning of BS/MS/SC.
State 5	In this state, the main server fails due to some technical fault, and the system is run by the standby server. Also, the repair process is performed on the main server.
State 5_D	Here, both the servers break down and the whole system stops working. Simultaneously, the repair process of State 5 carries on.
State 6	In this state, the power supply of the system is interrupted and a backup UPS provides the power. Also, the repair process for the main UPS start functioning.
State 6_D	In this state, both the power supplies are in downstate, and consequently, the whole system stops working. The repair process is initiated.

previous state (state 2). This ensures consistency with SMP properties while maintaining realistic continuity in repair operations.

- In the event of hardware or software faults in the mmWave beamforming antennas on the HAP, the link between the HAP and terrestrial network will be disrupted, rendering the system unable to provide service. Consequently, the system will transition to State 3, and a repair process will be initiated.
- In the event of hardware, software, or technical errors in any component of the terrestrial network, the connection to the HAP will be disrupted; resulting in the termination of user services. Therefore, the system will transition to State 4, and a repair process will be initiated.
- In the event of a system fault, if the server fails to function, a backup server will take its place to continue providing service. Thus, the system will transition to State 5. But, when the backup server also stops working, the system will move to State 5_D , and a repair process will be initiated.
- In case of power supply failure, the system will transition to State 6, and the backup power supply will take over. If the backup power supply also fails, the entire system will transition to a breakdown state 6_D .

The description of states is provided in Table 2.

The proposed model permits the assumption that the maintenance and component failure times will be exponentially distributed. However, the lifetime of HAP may be simulated using the gamma distribution [32]. For the suggested model, the distributions of sojourn times for each state are shown in Table 3.

3 | Steady-State Analysis

This section will outline the suggested model's steady-state analysis. Let the stochastic process for the underlying system be $\{X(t), t \geq 0\}$. A SMP can accurately simulate the underlying stochastic process because of the gamma distribution for the sojourn time of a single state and the Markov characteristics of the time epochs [33].

The proposed stochastic model's state space is designated as

$$\Omega = \{0, 1, 1_D, 2, 2_D, 3, 4, 5, 5_D, 6, 6_D\}.$$

The SMP is characterized by its kernel matrix $K(t)$. The kernel matrix $K(t)$ is given by $K_{m,n}(t) = P\{Y_1 = n, T_1 \leq t | Y_0 = m\}$, where (Y_n, T_n) , $n \geq 0$, represents a Markov renewal sequence of random variables, with $m, n \in \Omega$. The function $K_{m,n}(t)$ describes the probability that the SMP, upon entering state m , transitions to state n within time t . In contrast, it can be viewed as the probability that the next state is n , conditional on the transition occurring within time t while the SMP is in state m . Moreover, let $F_{m,n}(t)$ be the cumulative distribution function (CDF) of the sojourn time in state m and associated with the transition from state m to state n . The elements $K_{m,n}(t)$ can be defined based on the reachability from state m as follows:

TABLE 3 | Distributions of sojourn times for the underlying model.

CDF	Distribution	Parameter description	CDF	Distribution	Parameter description
F_{01}	Exp (λ_{01})	λ_{01} : failure rate FSO link	F_{10}	Exp (λ_{10})	λ_{10} : repair rate FSO link
F_{11_D}	Exp (λ_{11_D})	λ_{11_D} : failure rate RF link	F_{1_D1}	Exp (λ_{1_D1})	λ_{10} : repair rate of FSO/RF link
F_{02}	Gamma (β, λ_{02})	β, λ_{02} : scale and shape parameter for failure of primary HAP	F_{20}	Exp (λ_{20})	λ_{20} : repair rate of primary HAP
F_{22_D}	Gamma (β, λ_{22_D})	β, λ_{22_D} : scale and shape parameter for failure of backup HAP	F_{03}	Exp (λ_{03})	λ_{03} : failure rate of beamforming antenna
F_{2_D2}	Exp (λ_{2_D2})	λ_{2_D2} : repair rate of HAP	F_{30}	Exp (λ_{30})	λ_{30} : repair rate of beamforming antenna
F_{04}	Exp (λ_{04})	λ_{04} : failure rate of RAN network	F_{40}	Exp (λ_{40})	λ_{40} : repair rate of RAN network
F_{05}	Exp (λ_{05})	λ_{05} : failure rate of main server	F_{50}	Exp (λ_{50})	λ_{50} : repair rate of main server
F_{55_D}	Exp (λ_{55_D})	λ_{55_D} : failure rate of backup server	F_{5_D5}	Exp (λ_{5_D5})	λ_{5_D5} : repair rate of server
F_{06}	Exp (λ_{06})	λ_{06} : failure rate of main power supply	F_{60}	Exp (λ_{60})	λ_{60} : repair rate of main power supply
F_{66_D}	Exp (λ_{66_D})	λ_{66_D} : failure rate of backup power supply	F_{6_D6}	Exp (λ_{6_D6})	λ_{6_D6} : repair rate of power supply

$$K_{m,n}(t) = \begin{cases} 0, & \text{if state } n \text{ is not reachable from state } m \text{ within time } t, \\ F_{m,n}(t), & \text{if only state } n \text{ is reachable from state } m \text{ within time } t, \\ \int_0^t \prod_{\substack{k \in \Omega \\ k \neq n}} [1 - F_{m,k}(x)] dF_{m,n}(x), & \text{if multiple states are reachable from state } m \text{ within time } t, \end{cases}$$

By employing the two-stage method, the underlying SMP model is analyzed. As per this method, in the first stage, the one-step transition probabilities of the embedded Markov chain are derived using a kernel matrix $K(t)$. And, in the second state, using the mean sojourn time of each state, the steady-state probabilities of the SMP are computed. The entries of $K(t)$ can be further determined using the CDFs provided in Table 3 as follows.

Let $A_1 = \lambda_{01} + \lambda_{03} + \lambda_{04} + \lambda_{05} + \lambda_{06}$, $A_2 = A_1 + \lambda_{02}$.

$$\begin{aligned} K_{01}(t) &= \int_0^t \overline{F_{02}}(x) \overline{F_{03}}(x) \overline{F_{04}}(x) \overline{F_{05}}(x) \overline{F_{06}}(x) dF_{01}(x) \\ &= \frac{\lambda_{01}}{(\beta-1)!} \frac{1-e^{-(A_1)t}}{A_2} - \left[\frac{1-e^{-(A_1)t}}{A_1} - \frac{1-e^{-(A_2)t}}{A_2} \right] \\ &\quad [\lambda_{01}t^{\beta-1} - \lambda_{02}\lambda_{01}(a^{\beta-1} - t^{\beta-1})], \end{aligned} \quad (1)$$

Here, in general, $\overline{F_{mn}}(t) = 1 - F_{mn}(t)$ for all m and n . Similarly, the expressions of $K_{02}(t)$, $K_{03}(t)$, $K_{04}(t)$, $K_{05}(t)$, $K_{06}(t)$ are computed.

$$\begin{aligned} K_{02}(t) &= \int_0^t \overline{F_{01}}(x) \overline{F_{03}}(x) \overline{F_{04}}(x) \overline{F_{05}}(x) \overline{F_{06}}(x) dF_{02}(x) \\ &= \frac{\lambda_{02} [1 - e^{-(A_2)t}]}{(\beta-1)! (A_2)} [a^{\beta-1} - t^{\beta-1}], \end{aligned} \quad (2)$$

$$K_{03}(t) = \int_0^t \overline{F_{01}}(x) \overline{F_{02}}(x) \overline{F_{04}}(x) \overline{F_{05}}(x) \overline{F_{06}}(x) dF_{03}(x), \quad (3)$$

$$\begin{aligned} &= \frac{\lambda_{03}}{(\beta-1)!} \frac{1-e^{-(A_1)t}}{A_2} - \left[\frac{1-e^{-(A_1)t}}{A_1} - \frac{1-e^{-(A_2)t}}{A_2} \right] \\ &\quad [\lambda_{03}t^{\beta-1} - \lambda_{02}\lambda_{03}(a^{\beta-1} - t^{\beta-1})], \end{aligned} \quad (4)$$

$$K_{04}(t) = \int_0^t \overline{F_{01}}(x) \overline{F_{03}}(x) \overline{F_{02}}(x) \overline{F_{05}}(x) \overline{F_{06}}(x) dF_{02}(x), \quad (5)$$

$$\begin{aligned} &= \frac{\lambda_{04}}{(\beta-1)!} \frac{1-e^{-(A_1)t}}{A_2} - \left[\frac{1-e^{-(A_1)t}}{A_1} - \frac{1-e^{-(A_2)t}}{A_2} \right] \\ &\quad [\lambda_{04}t^{\beta-1} - \lambda_{02}\lambda_{04}(a^{\beta-1} - t^{\beta-1})], \end{aligned} \quad (6)$$

$$K_{05}(t) = \int_0^t \overline{F_{01}}(x) \overline{F_{03}}(x) \overline{F_{04}}(x) \overline{F_{02}}(x) \overline{F_{06}}(x) dF_{05}(x), \quad (7)$$

$$\begin{aligned} &= \frac{\lambda_{05}}{(\beta-1)!} \frac{1-e^{-(A_1)t}}{A_2} - \left[\frac{1-e^{-(A_1)t}}{A_1} - \frac{1-e^{-(A_2)t}}{A_2} \right] \\ &\quad [\lambda_{05}t^{\beta-1} - \lambda_{02}\lambda_{05}(a^{\beta-1} - t^{\beta-1})], \end{aligned} \quad (8)$$

$$K_{06}(t) = \int_0^t \overline{F_{01}(x)} \overline{F_{03}(x)} \overline{F_{04}(x)} \overline{F_{05}(x)} \overline{F_{02}(x)} dF_{06}(x), \quad (9)$$

$$= \frac{\lambda_{06}}{(\beta-1)!} \frac{1-e^{-(A_1)t}}{A_2} - \left[\frac{1-e^{-(A_1)t}}{A_1} - \frac{1-e^{-(A_2)t}}{A_2} \right] \left[\lambda_{06} t^{\beta-1} - \lambda_{02} \lambda_{06} (x^{\beta-1} - t^{\beta-1}) \right], \quad (10)$$

$$K_{10}(t) = \int_0^t \overline{F_{11D}(x)} dF_{10}(x) = \frac{\lambda_{10}}{\lambda_{10} + \lambda_{11D}} \left(1 - e^{-(\lambda_{10} + \lambda_{11D})t} \right), \quad (11)$$

$$= \int_0^t e^{-\lambda_{11}x} \cdot \lambda_{10} e^{-\lambda_{10}x} dx = \lambda_{10} \int_0^t e^{-(\lambda_{10} + \lambda_{11})x} dx = \lambda_{10} \cdot \left[\frac{1 - e^{-(\lambda_{10} + \lambda_{11})t}}{\lambda_{10} + \lambda_{11}} \right] \quad (12)$$

$$K_{55D}(t) = \int_0^t \overline{F_{50}(x)} dF_{55D}(x) = \frac{\lambda_{55D}}{\lambda_{50} + \lambda_{55D}} \left[1 - e^{-(\lambda_{50} + \lambda_{55D})t} \right], \quad (13)$$

$$K_{66D}(t) = \int_0^t \overline{F_{60}(x)} dF_{66D}(x) = \frac{\lambda_{66D}}{\lambda_{60} + \lambda_{66D}} \left[1 - e^{-(\lambda_{60} + \lambda_{66D})t} \right],$$

$$\begin{aligned} K_{50}(t) &= \frac{\lambda_{50}}{\lambda_{50} + \lambda_{55D}} \left[1 - e^{-(\lambda_{50} + \lambda_{55D})t} \right], \\ K_{60}(t) &= \frac{\lambda_{60}}{\lambda_{60} + \lambda_{66D}} \left[1 - e^{-(\lambda_{60} + \lambda_{66D})t} \right], \\ K_{11D}(t) &= \frac{\lambda_{11D}}{\lambda_{10} + \lambda_{11D}} \left[1 - e^{-(\lambda_{10} + \lambda_{11D})t} \right]. \end{aligned} \quad (14)$$

Similarly, $K_{2D2}(t), K_{5D5}(t), K_{6D6}(t), K_{30}(t), K_{40}(t)$ can be computed following the same method.

$$K_{2D2}(t) = 1 - e^{-\lambda_{2D2}t}, K_{6D6}(t) = F_{6D6}(t) = 1 - e^{-\lambda_{6D6}t}, K_{5D5}(t) = F_{5D5}(t) = 1 - e^{-\lambda_{5D5}t}, K_{30}(t) = F_{30}(t) = 1 - e^{-\lambda_{30}t},$$

$$K_{40}(t) = \int_0^t dF_{40}(x) = 1 - e^{-\lambda_{40}t}, K_{1D1}(t) = F_{1D1}(t) = 1 - e^{-\lambda_{1D1}t},$$

$$K_{22D}(t) = \int_0^t \overline{F_{20}(x)} dF_{22D}(x) = \lambda_{20}^{\beta} \left(\frac{1 - e^{-(\lambda_{20} + \lambda_{22D})t}}{\lambda_{20} + \lambda_{22D}} \right) (a^{\beta-1} - t^{\beta-1}),$$

$$K_{55D}(t) = \int_0^t \overline{F_{50}(x)} dF_{55D}(x) = \frac{\lambda_{55D}}{\lambda_{50} + \lambda_{55D}} \left[1 - e^{-(\lambda_{50} + \lambda_{55D})t} \right], K_{66D}(t) = \int_0^t \overline{F_{60}(x)} dF_{66D}(x) = \frac{\lambda_{66D}}{\lambda_{60} + \lambda_{66D}} \left[1 - e^{-(\lambda_{60} + \lambda_{66D})t} \right].$$

$$K_{20}(t) = \int_0^t \overline{F_{22D}(x)} dF_{20}(x) = \frac{1}{(\beta-1)!} \left[(\beta-1)! (1 - e^{-\lambda_{20}t}) t^{\beta-1} \left[(1 - e^{-\lambda_{20}t}) - \lambda_{20} \left(\frac{1 - e^{-(\lambda_{20} + \lambda_{22D})t}}{\lambda_{20} + \lambda_{22D}} \right) \right] - \lambda_{22D} (a^{\beta-1} - t^{\beta-1}) \right] \left[(1 - e^{-\lambda_{20}t}) + \lambda_{20} \left(\frac{1 - e^{-(\lambda_{20} + \lambda_{22D})t}}{\lambda_{20} + \lambda_{22D}} \right) \right].$$

Define

$$K_{m,n}(\infty) = \lim_{t \rightarrow \infty} K_{m,n}(t),$$

where $m, n \in \Omega$. The matrix $K(\infty)$ represents the transition probabilities (K_{mn}) for the underlying model whose row sum is 1.

$$K(\infty) = \begin{matrix} & \begin{matrix} 0 \\ 1 \\ 1_D \\ 2 \\ 2_D \\ 3 \\ 4 \\ 5 \\ 5_D \\ 6 \\ 6_D \end{matrix} & \begin{bmatrix} 0 & K_{01} & 0 & K_{02} & 0 & K_{03} & K_{04} & K_{05} & 0 & K_{06} & 0 \\ K_{10} & 0 & K_{1D1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & K_{11D} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ K_{20} & 0 & 0 & 0 & K_{2D2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & K_{22D} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ K_{30} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ K_{40} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ K_{50} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & K_{5D5} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & K_{55D} & 0 & 0 & 0 \\ K_{60} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & K_{6D6} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & K_{66D} & 0 \end{bmatrix} \end{matrix}.$$

For the embedded Markov chain (EMC) of the SMP, the steady-state probabilities (v_i) are computed as follows

$$v = vK(\infty), \sum_{i \in \Omega} v_i = 1. \quad (15)$$

TABLE 4 | Parameter values (per unit time) from Reference [22].

λ_{01}	0.08	λ_{02}	0.6	λ_{22_D}	0.5
λ_{03}	1.5	λ_{04}	1.3	a, β	1.1
λ_{05}	0.8	λ_{06}	0.9	λ_{6_D6}	1.4
λ_{10}	0.33	λ_{20}	0.67	λ_{66_D}	1.3
λ_{30}	1.1	λ_{40}	1.5	λ_{5_D5}	0.2
λ_{50}	0.02	λ_{60}	0.02	λ_{55_D}	0.1
λ_{11_D}	0.2	λ_{1_D1}	0.4	λ_{2_D2}	0.5

On solving the set of equations 15, the steady-state probabilities, v_i are obtained. Therefore, the following formula is applied to compute the steady-state probabilities.

$$\pi_i = \frac{v_i h_i}{\sum_{j \in \Omega} v_j h_j}, i \in \Omega, \quad (16)$$

where v_i are the steady-state probabilities obtained from Equation (15).

In Equation (16), h_i denotes the expected sojourn time, spent by the process at each state i . These h_i 's are computed as follows.

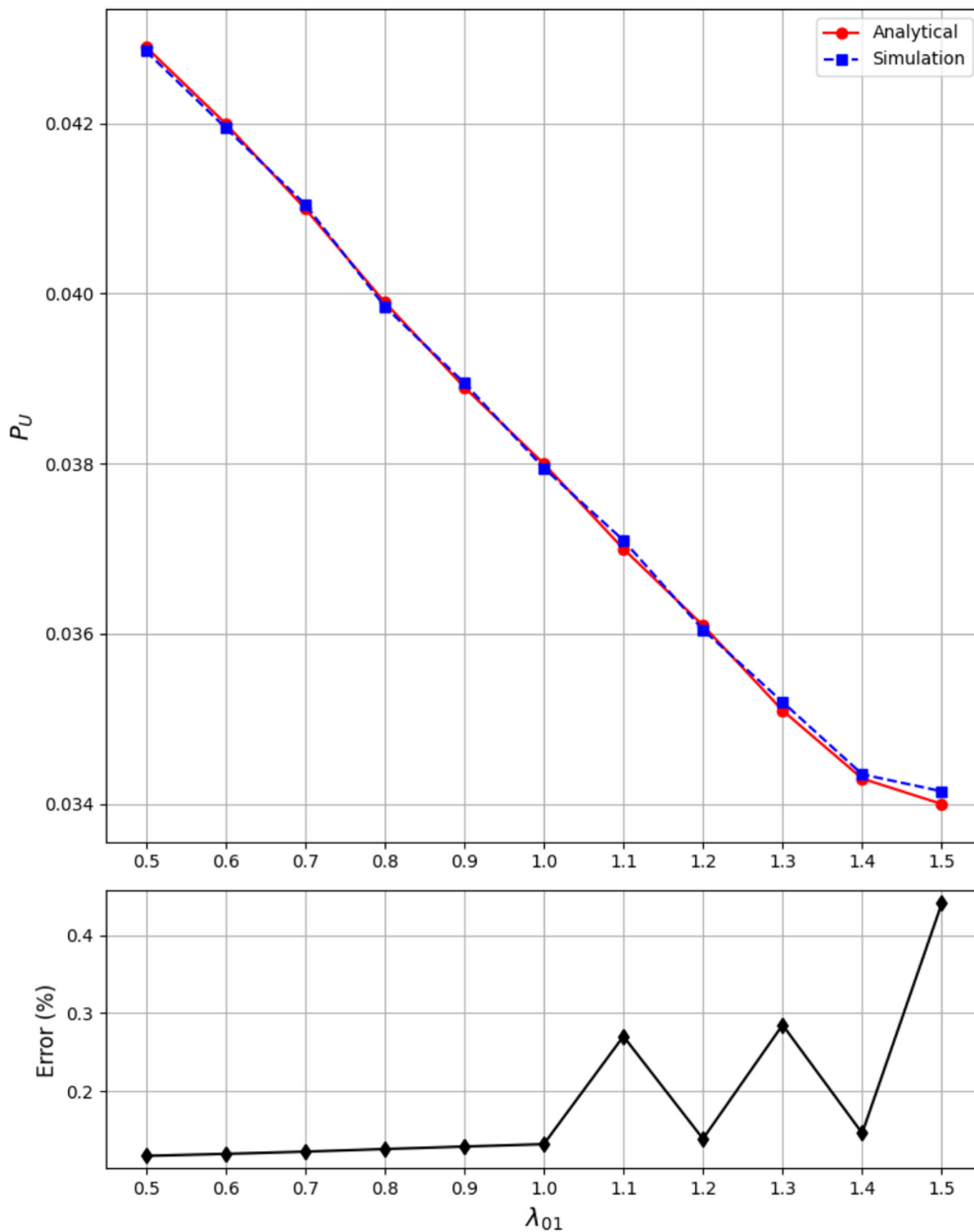


FIGURE 3 | P_U versus failure rate of FSO link per unit time (λ_{01}).

$$\begin{aligned}
 h_0 &= \int_0^{\infty} \bar{F}_{01}(x) \cdot \bar{F}_{02}(x) \cdot \bar{F}_{03}(x) \cdot \bar{F}_{04}(x) \cdot \bar{F}_{05}(x) \cdot \bar{F}_{06}(x) dx \\
 &= \frac{1}{(\beta-1)!} \left[\frac{(\beta-1)!}{\lambda_{01} + \lambda_{03} + \lambda_{04} + \lambda_{05} + \lambda_{06}} - a^{\beta-1} \left(\frac{1}{\lambda_{01} + \lambda_{03} + \lambda_{04} + \lambda_{05} + \lambda_{06}} - \frac{1}{\lambda_{01} + \lambda_{02} + \lambda_{03} + \lambda_{04} + \lambda_{05} + \lambda_{06}} \right) \right] \\
 h_1 &= \int_0^{\infty} \bar{F}_{10}(x) \cdot \bar{F}_{11_D}(x) dx = \frac{\lambda_{10}}{\lambda_{10} + \lambda_{11_D}}, \quad h_{1_D} = \int_0^{\infty} \bar{F}_{11_D}(x) dx = \frac{1}{\lambda_{1_D1}}, \\
 h_2 &= \int_0^{\infty} \bar{F}_{20}(x) \cdot \bar{F}_{22_D}(x) dx = \frac{1}{(\beta-1)!} \left[\frac{(\beta-1)!}{\lambda_{20}} - a^{\beta-1} \left(\frac{1}{\lambda_{20}} - \frac{1}{\lambda_{20} + \lambda_{22_D}} \right) \right], \quad h_{2_D} = \int_0^{\infty} \bar{F}_{22_D}(x) dx = \frac{1}{\lambda_{2_D2}}, \\
 h_3 &= \int_0^{\infty} \bar{F}_{30}(x) dx = \frac{1}{\lambda_{30}}, \quad h_4 = \int_0^{\infty} \bar{F}_{40}(x) dx = \frac{1}{\lambda_{40}}, \quad h_{6_D} = \int_0^{\infty} \bar{F}_{66_D}(x) dx = \frac{1}{\lambda_{6_D6}}, \\
 h_5 &= \int_0^{\infty} \bar{F}_{50}(x) \cdot \bar{F}_{55_D}(x) dx = \frac{\lambda_{50}}{\lambda_{50} + \lambda_{55_D}}, \quad h_6 = \int_0^{\infty} \bar{F}_{60}(x) \cdot \bar{F}_{66_D}(x) dx = \frac{\lambda_{60}}{\lambda_{60} + \lambda_{66_D}}, \quad h_{5_D} = \int_0^{\infty} \bar{F}_{55_D}(x) dx = \frac{1}{\lambda_{5_D5}}.
 \end{aligned}$$

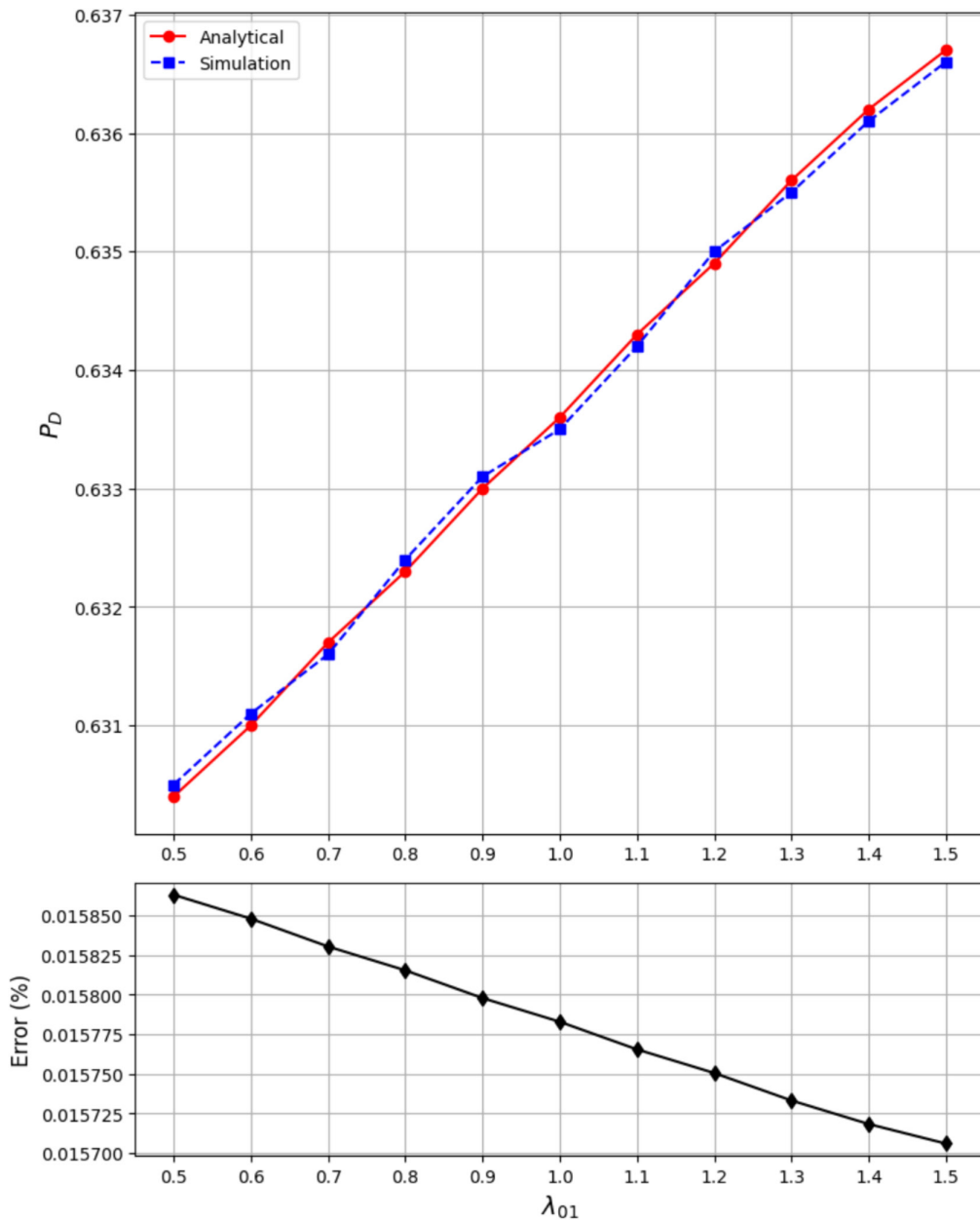


FIGURE 4 | P_D versus failure rate of FSO link per unit time (λ_{01}).

4 | Performance Measures

After computing the steady-state distribution, the relevant performance metrics for the proposed system are determined.

1. The probability of the system being in the upstate is given by

$$P_U = \pi_0.$$

2. The probability of the system being in downstate is given by

$$P_D = \pi_{1_D} + \pi_{2_D} + \pi_3 + \pi_4 + \pi_{5_D} + \pi_{6_D}.$$

3. The mean time to failure (*MTTF*) is the reliability measure that is most frequently utilized. To compute the *MTTF*,

we assume that the failed states $A = \{1_D, 2_D, 5_D, 6_D, 3, 4\}$ are absorbing state and $T = \{0, 1, 2, 5, 6\}$ are transient states. Hence, the modified transition probability matrix, denoted by P is defined as

$$P = \begin{pmatrix} N & R \\ O & M \end{pmatrix}$$

where N represents the transition probabilities of transient states, M denotes the transitions of absorption states, and R, O are the transition probabilities from transient states to absorption and absorption to transient states, respectively. The expected sojourn time of each state i , is calculated by $h_i n_i = -y^T(0)$, $i \in T$, where $y(0)$ is the initial probability vector of size 11 and $y^T(0) =$

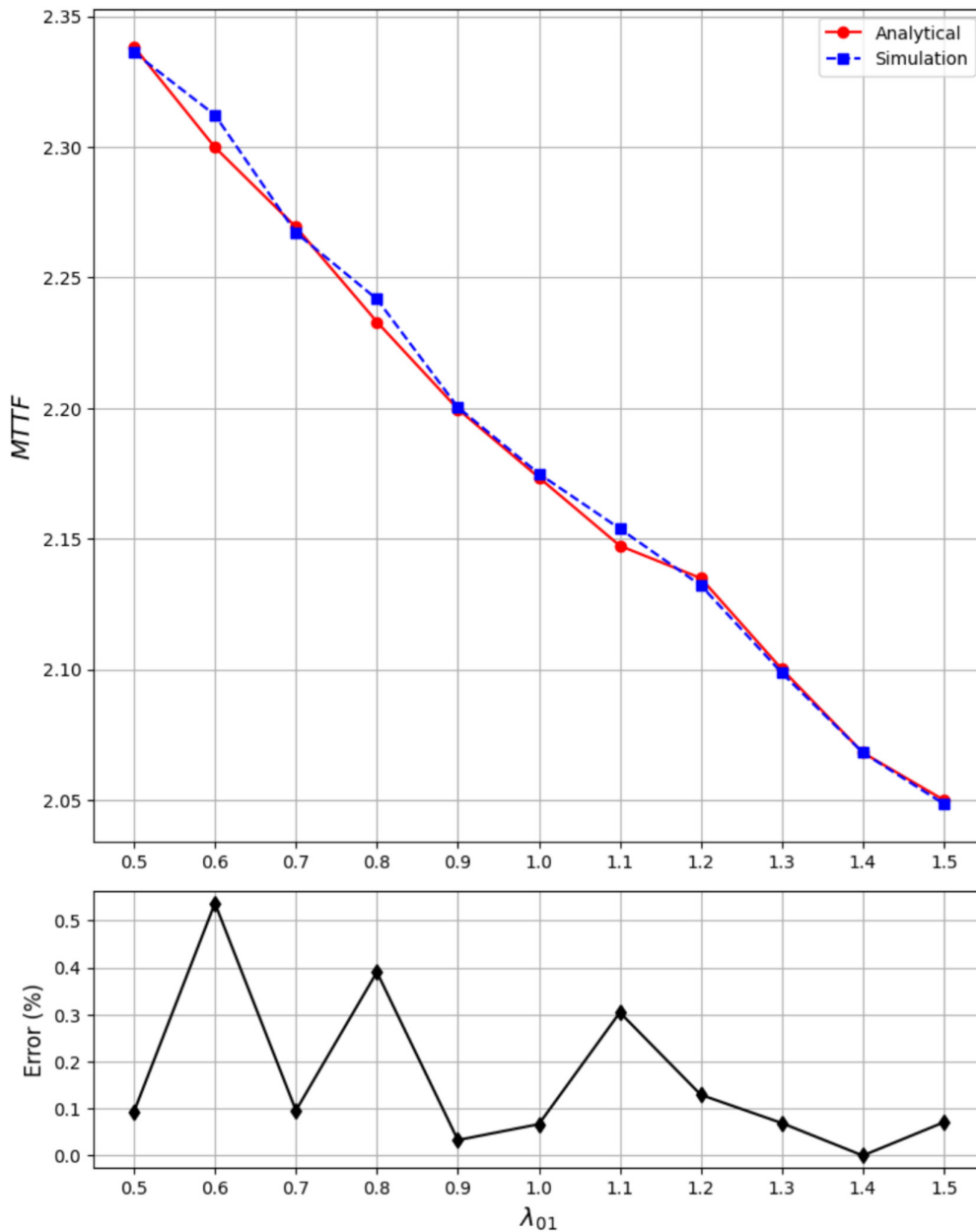


FIGURE 5 | *MTTF* (unit time) versus failure rate of FSO link per unit time (λ_{01}).

is a subvector for the transient states. Let $y^T(0) = [1, 0, 0, 0, 0]$. Therefore, the *MTTF* is computed as follows

$$MTTF = \sum_{i \in T} n_i h_i$$

where n_i is the mean number of times state i is visited until the absorption. The following equations provide the value of n_i :

$$n_i = \pi_i + \sum_{j \in T} n_j p_{ij}, i, j \in T$$

where π_i is the probability that the process initiates from state i and p_{ij} represents the (i, j) th component of the matrix $P = [p_{ij}]$. Here, it is assumed that $\pi_i = [1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]$.

5 | Numerical Illustration

In this section, numerical analyses are conducted to assess the feasibility of the proposed model in achieving steady-state availability and reliability results. To address the lack of precise parameterization, the sensitivity of the model's characteristics to slight variations in specific parameter values is examined, considering that parameter estimation falls beyond the scope of this paper. Table 4 presents the parameter values necessary for generating the results. It is important to note that these values are chosen solely for numerical demonstration purposes. Some of the parameters in Table 4 are referred from the work [22]. Due to the unavailability of the existing literature, the rest of the parameters has been assumed in order to carry out the analysis.

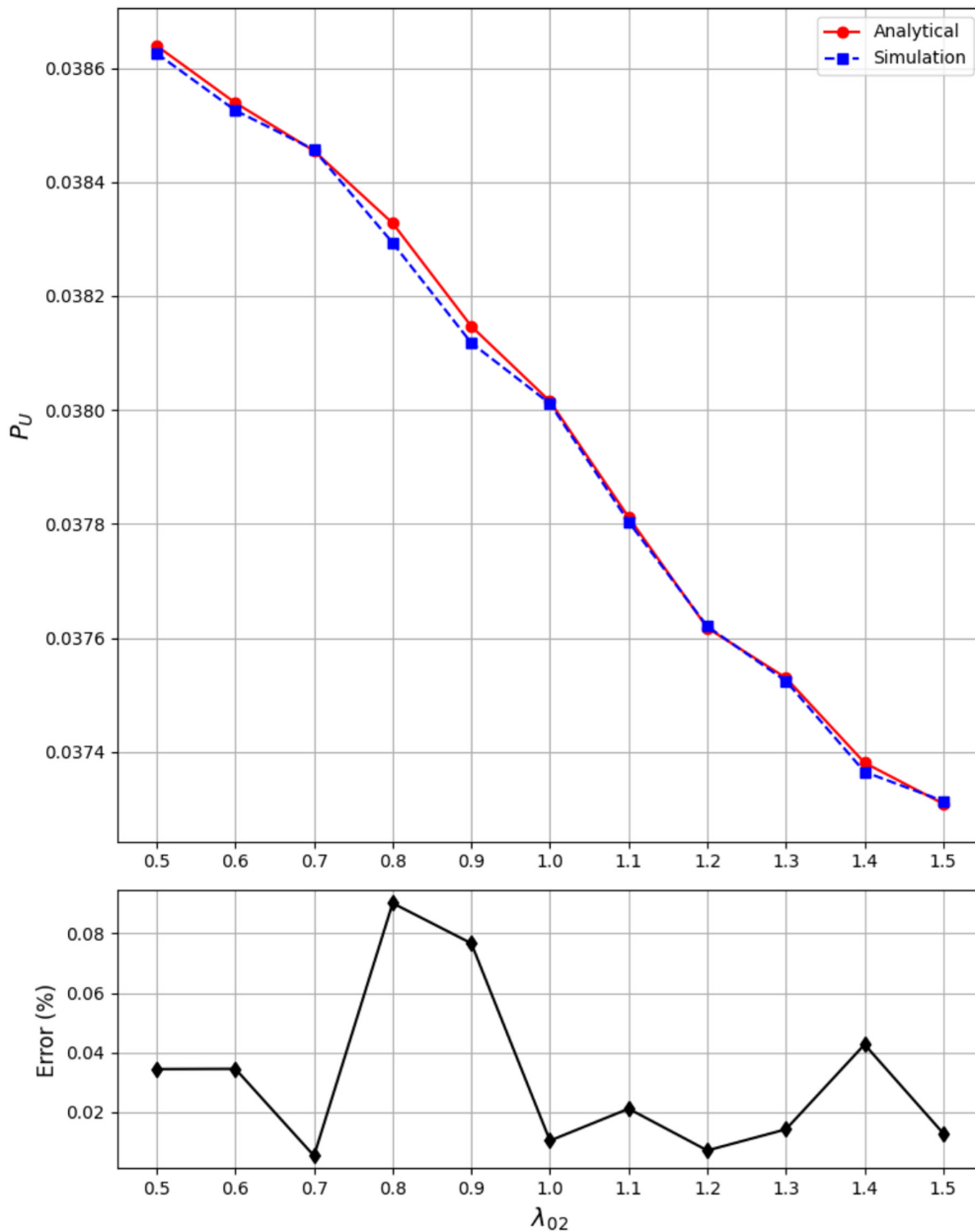


FIGURE 6 | P_U versus failure rate of primary HAP per unit time (λ_{02}).

The behavior of the system's attributes P_U , P_D , and $MTTF$ in relation to the failure rate (λ_{01}) of the satellite link break is illustrated in Figures 3–5. As the value of λ_{01} increases, the likelihood of the system being in the operational state decreases, resulting in a decrease in P_U . Conversely, the probability of the system being in the downstate increases, leading to an increase in P_D . Since $MTTF$ represents the duration for which the system remains in an operational state before failure, an increase in λ_{01} results in a decrease in $MTTF$ value.

Figures 6–8 depicts the system dynamics concerning the failure rate of the HAP (λ_{02}). The graphs indicate that an increment in λ_{02} leads to a decrement in both P_U and $MTTF$ as the likelihood of the system being in the available state decreases. Conversely, an increase in λ_{02} results in an increase in P_D , representing the probability of the system being in the downstate.

5.1 | Transient Analysis

This section presents the transient analysis of dependability attributes for a SAT network integrated with a HAP. Due to the non-Markovian nature of sojourn time distributions in the proposed dependability model, deriving closed-form expressions for transient behavior is infeasible. Hence, a simulation-based approach is adopted to evaluate the transient availability of the network. The simulations are performed using a MATLAB program, and the accuracy of the results is validated through a t-distribution test. Transient availability is obtained for various parameter settings, specifically focusing on the failure rates of the HAP and the FSO link, with results presented within a 95% confidence interval.

Figures 9 and 10 illustrate these findings. Figure 9 shows how the probability of the system being in the operational (up) state

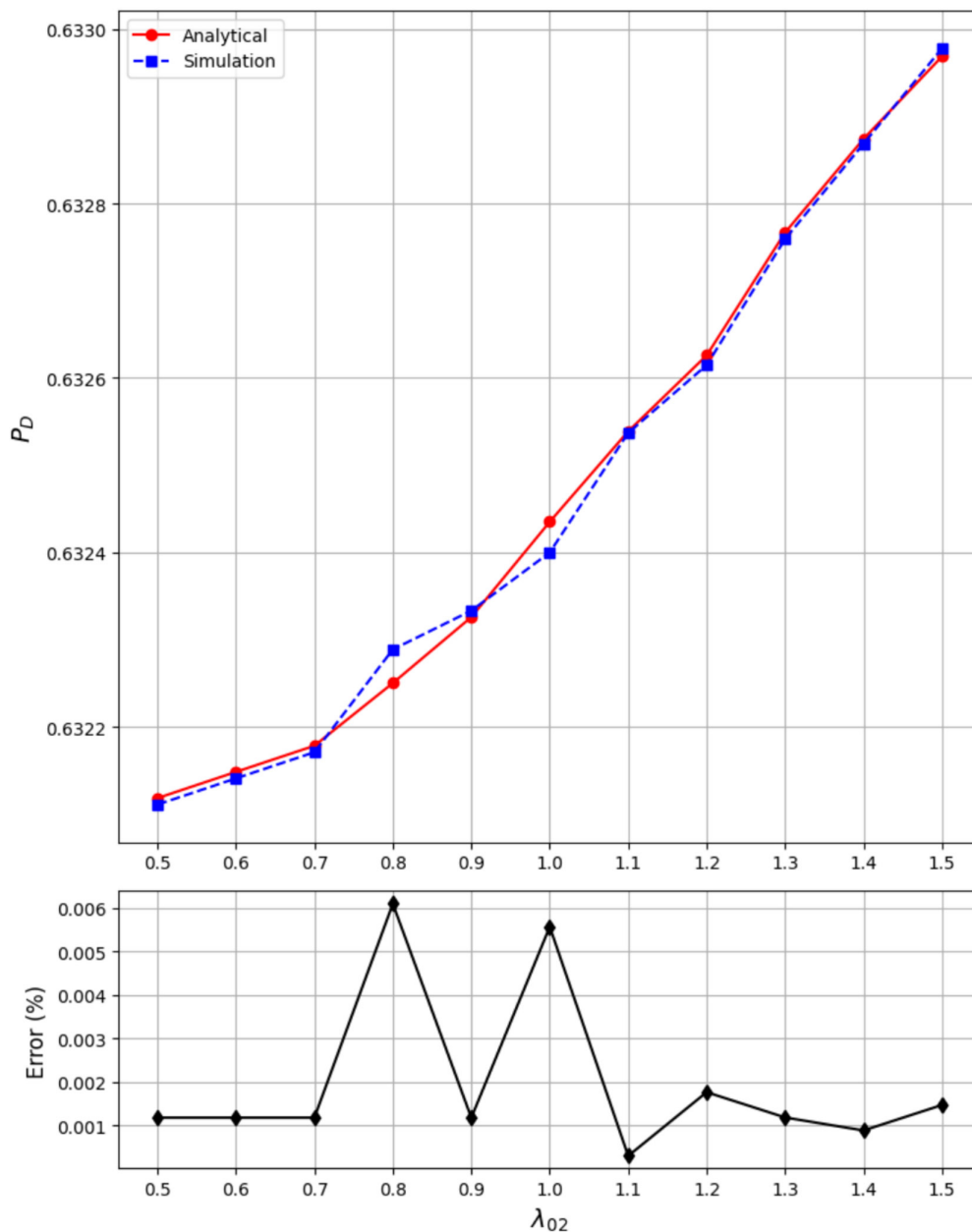


FIGURE 7 | P_D versus failure rate of primary HAP per unit time (λ_{02}).

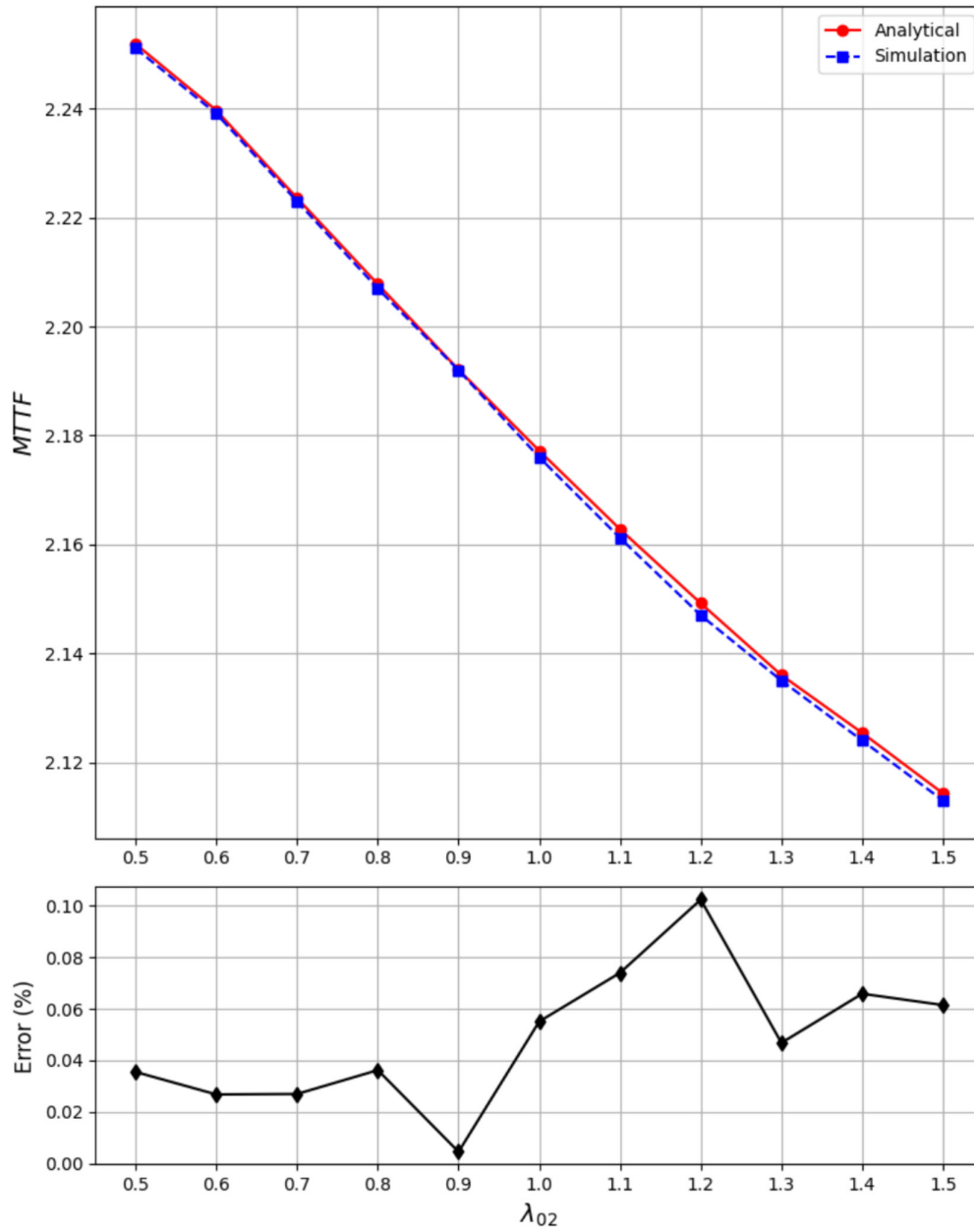


FIGURE 8 | *MTTF* (unit time) versus failure rate of primary HAP per unit time (λ_{02}).

evolves over time for different values of the FSO link failure rate (λ_{01}). The graph indicates that availability decreases as λ_{01} increases. Furthermore, availability trends toward a steady-state value over time, suggesting convergence to long-run behavior. Figure 10 depicts the availability of the network over time for various HAP failure rates (λ_{02}). Similar to the previous case, increasing λ_{02} results in a decline in availability. Again, the network eventually reaches a steady-state, as indicated by the flattening of the availability curves over time.

5.2 | Simulation Approach

To validate the reliability aspects of the proposed model, a simulation approach is utilized. The parameters used in the analytical calculations (as listed in Table 4) are replicated in a

MATLAB program for this purpose. The simulation is iterated 100 times with the same parameter set, resulting in 100 metric values. The accuracy of the simulation results is assessed using the *t*-distribution test. “The *t*-distribution is used for smaller sample sizes to account for the uncertainty in the standard deviation of the sample”. Next, the mean and standard deviation of the 100 values are computed and a confidence interval (C.I.) for 95% accuracy of the results is obtained. The CI for the 95% accuracy of the results can be calculated by:

$$C.I. = \bar{X} \pm \frac{s^* t_{0.05, k-1}}{\sqrt{k}}$$

where $\bar{X}$ is sample mean, *s* is the standard deviation, $t_{0.05, k-1}$ is tabulated value for 95% CI and *k* – 1 degree of freedom, and *k* is

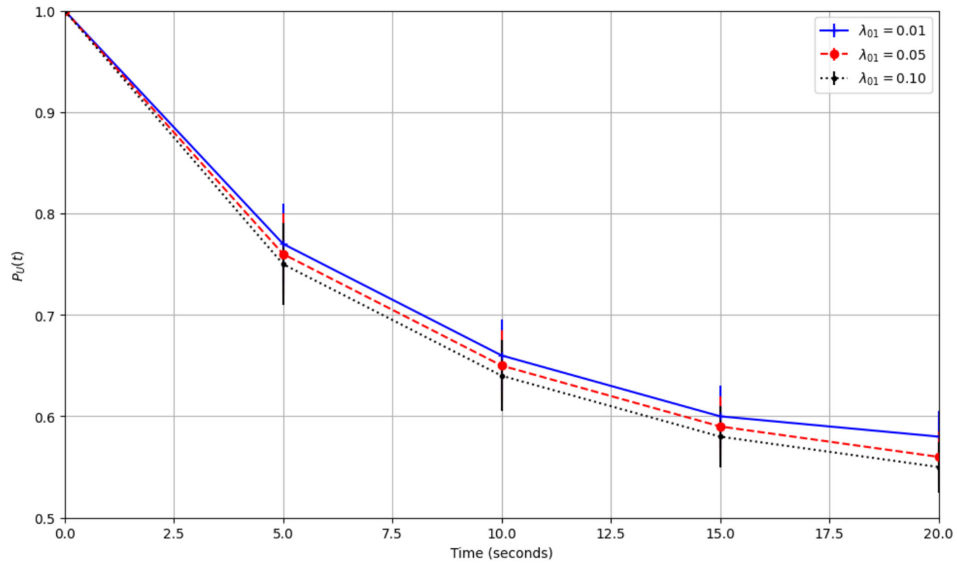


FIGURE 9 | P_U versus time for different failure rate of FSO link per unit time (λ_{01}).

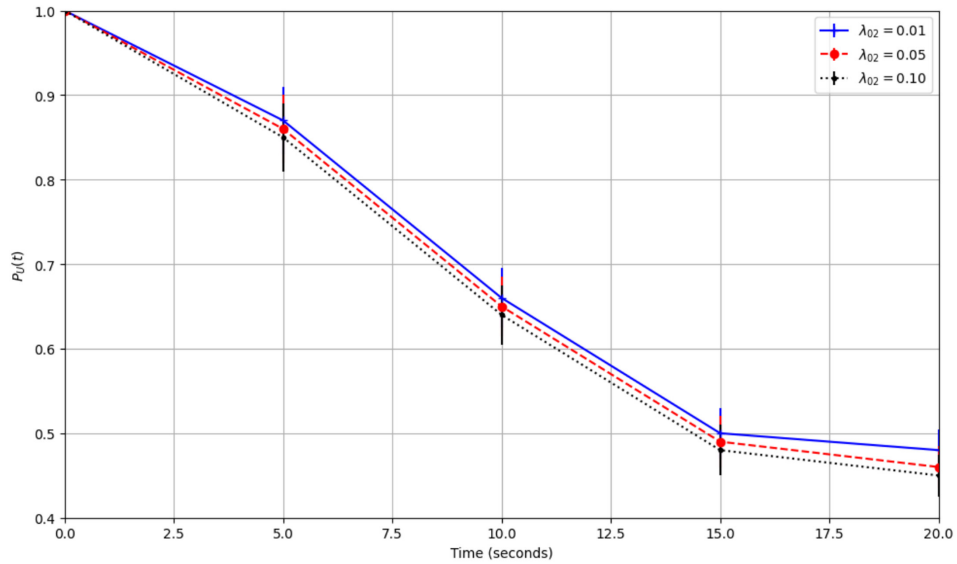


FIGURE 10 | P_U versus time for different failure rate of HAP per unit time (λ_{02}).

sample size. When the sample size k is large than 30, $t_{0.05,k-1}$ is replaced by $z_{0.05}$.

To quantitatively evaluate the deviation between the simulation and analytical results, we have computed the relative error (in percentage) at each data point using the formula:

$$\text{Relative Error} = \frac{|\text{Simulation} - \text{Analytical}|}{\text{Analytical}} \times 100\%$$

Theoretical and simulation results are compared for the system's upstate, downstate, and mean time to failure (*MTTF*) across various failure rates. The graphs illustrate the analysis results with solid lines, while the simulated outcomes are depicted with dotted lines. The relative errors across all points were found to be consistently below (0.5%), demonstrating a strong agreement. The alignment between the findings of both methods suggests

a close correspondence, affirming the approximate accuracy of the proposed model and validating the hypothesized framework.

6 | Conclusions and Future Works

The proposed stochastic model for the satellite-aerial-terrestrial network considers important components, including redundancy, and is analyzed using an SMP that accounts for the non-exponentially distributed sojourn times of various states of the integrated network. The steady-state probabilities enable for the computation of the system's upstate and downstate probabilities, as well as its *MTTF*. Numerical illustrations are incorporated to demonstrate the steady-state analysis of the model, and a sensitivity analysis is carried out to determine the impact of minor variations in the values of the network parameter on the model's reliability. Moreover, the proposed model undergoes verification

through discrete event simulation, yielding results that closely match the analytical findings. This alignment underscores the validity of the underlying model.

In the future, the proposed model can be modified by considering catastrophic scenarios and propagation failure of the components. A new architecture may be proposed to increase the reliability and availability of the space-air-ground networks. In our future research work, we aim to integrate the concept of software rejuvenation into our analysis. Additionally, we intend to explore the simultaneous failure of multiple components within the proposed architecture of space-air-ground networks.

Acknowledgments

The authors acknowledge all the competing authorities in preparing the manuscript.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

References

1. H. Guo, J. Li, J. Liu, N. Tian, and N. Kato, "A Survey on Space-Air-Ground-Sea Integrated Network Security in 6G," *IEEE Communications Surveys & Tutorials* 24, no. 1 (2021): 53–87.
2. J. Liu, Y. Shi, Z. M. Fadlullah, and N. Kato, "Space-Air-Ground Integrated Network: A Survey," *IEEE Communications Surveys and Tutorials* 20, no. 4 (2018): 2714–2741.
3. G. K. Kurt, M. G. Khoshkholgh, S. Alfattani, et al., "A Vision and Framework for the High Altitude Platform Station (HAPS) Networks of the Future," *IEEE Communications Surveys and Tutorials* 23, no. 2 (2021): 729–779.
4. S. Shah, M. Siddharth, N. Vishwakarma, R. Swaminathan, and A. Madhukumar, "Adaptive-Combining-Based Hybrid FSO/RF Satellite Communication With and Without HAPS," *IEEE Access* 9 (2021): 81492–81511.
5. Z. Xiao, L. Zhu, Y. Liu, et al., "A Survey on Millimeter-Wave Beamforming Enabled UAV Communications and Networking," *IEEE Communications Surveys and Tutorials* 24, no. 1 (2021): 557–610.
6. T. Pecorella, L. S. Ronga, F. Chiti, S. Jayousi, and L. Franck, "Emergency Satellite Communications: Research and Standardization Activities," *IEEE Communications Magazine* 53, no. 5 (2015): 170–177.
7. Z. Qu, G. Zhang, H. Cao, and J. Xie, "Leo Satellite Constellation for Internet of Things," *IEEE Access* 5 (2017): 18391–18401.
8. L. You, K. X. Li, J. Wang, X. Gao, X. G. Xia, and B. Ottersten, "Massive MIMO Transmission for LEO Satellite Communications," *IEEE Journal on Selected Areas in Communications* 38, no. 8 (2020): 1851–1865.
9. N. Zhang, S. Zhang, P. Yang, O. Alhussein, W. Zhuang, and X. S. Shen, "Software Defined Space-Air-Ground Integrated Vehicular Networks: Challenges and Solutions," *IEEE Communications Magazine* 55, no. 7 (2017): 101–109.
10. Y. Wang, Y. Xu, Y. Zhang, and P. Zhang, "Hybrid Satellite-Aerial-Terrestrial Networks in Emergency Scenarios: A Survey," *China Communications* 14, no. 7 (2017): 1–13.

11. W. Feng, N. Ge, and J. Lu, *Coordinated Satellite-Terrestrial Networks: A Robust Spectrum Sharing Perspective* (IEEE, 2017), 1–5.
12. Z. Jia, M. Sheng, J. Li, D. Zhou, and Z. Han, "Joint HAP Access and LEO Satellite Backhaul in 6G: Matching Game-Based Approaches," *IEEE Journal on Selected Areas in Communications* 39, no. 4 (2020): 1147–1159.
13. Y. Li, N. Deng, and W. Zhou, "A Hierarchical Approach to Resource Allocation in Extensible Multi-Layer LEO-MSS," *IEEE Access* 8 (2020): 18522–18537.
14. N. Tiwari, S. De, and D. Selvamuthu, "Modeling of Channel Asymmetry in HAP-Based Optical Wireless Communications," (2023).
15. J. Wang, C. Jiang, Z. Wei, C. Pan, H. Zhang, and Y. Ren, "Joint UAV Hovering Altitude and Power Control for Space-Air-Ground IoT Networks," *IEEE Internet of Things Journal* 6, no. 2 (2018): 1741–1753.
16. F. Dong, H. Han, X. Gong, J. Wang, and H. Li, "A Constellation Design Methodology Based on QoS and User Demand in High-Altitude Platform Broadband Networks," *IEEE Transactions on Multimedia* 18, no. 12 (2016): 2384–2397.
17. W. Y. Dong, S. Yang, and S. Chen, "Uplink Performance Analysis of Heterogeneous Non-Terrestrial Networks in Harsh Environments: A Novel Stochastic Geometry Model," *IEEE Transactions on Communications* (2025).
18. W. Y. Dong, S. Yang, P. Zhang, and S. Chen, "Stochastic Geometry Based Modelling and Analysis of Uplink Cooperative Satellite-Aerial-Terrestrial Networks for Nomadic Communications With Weak Satellite Coverage," *IEEE Journal on Selected Areas in Communications* 42 (2024): 3428–3444.
19. V. Vishnevsky, D. Selvamuthu, V. Rykov, D. Kozyrev, and N. Ivanova, "Reliability Modeling of a Flight Module of a Tethered High-Altitude Telecommunication Platform," in *IEEE* (2022), 1–6.
20. V. M. Vishnevsky, D. V. Kozyrev, V. V. Rykov, and Z. F. Nguyen, "Reliability Modeling of an Unmanned High-Altitude Module of a Tethered Telecommunication Platform," *Informatsionnye Tekhnologii i Vychislitel'nye Sistemy* 4 (2020): 26–38.
21. V. M. Vishnevsky, D. Selvamuthu, V. Rykov, D. V. Kozyrev, N. Ivanova, and A. Krishnamoorthy, *Reliability Assessment of Tethered High-Altitude Unmanned Telecommunication Platforms* (Springer Nature Singapore, 2024).
22. N. Choudhary and K. Khaitan nee Gupta, "High Altitude Aeronautical Platform for VoIP: Dependability Analysis," *Wireless Personal Communications* 125, no. 3 (2022): 2277–2303.
23. K. Guo, M. Wu, X. Li, H. Song, and N. Kumar, "Deep Reinforcement Learning and NOMA-Based Multi-Objective RIS-Assisted IS-UAV-TNs: Trajectory Optimization and Beamforming Design," *IEEE Transactions on Intelligent Transportation Systems* 24 (2023): 10197–10210.
24. M. Bliss, F. J. Block, T. C. Royster, and D. J. Love, "Uplink NOMA for Heterogeneous NTN With LEO Satellites and High-Altitude Platform Relays," in *IEEE* (2022), 172–177.
25. R. Raina and S. Dharmaraja, "Stochastic Modelling of Multi-Layer HAP-LEO Systems in 6G for Energy Saving: An Analytical Approach," *Computer Communications* 210 (2023): 22–34.
26. R. Raina, "Stochastic Modeling of Enhancing Energy Efficiency for HAP in HAP-LEO Systems Using Markov Regenerative Process," *Computer Networks* 242 (2024): 110253.
27. X. Liu, M. Lin, Q. Huang, J. Wang, and J. Ouyang, "Performance Analysis for Multi-User Integrated Satellite and UAV Cooperative Networks," *Physical Communication* 36 (2019): 100762.
28. P. K. Singya and M. S. Alouini, "Performance of UAV-Assisted Multiuser Terrestrial-Satellite Communication System Over Mixed FSO/RF Channels," *IEEE Transactions on Aerospace and Electronic Systems* 58, no. 2 (2021): 781–796.

29. R. Raj, D. Selvamuthu, and D. Mishra, "Traffic Control Analysis in High Altitude Platforms Using a Semi-Markov Process Approach," in *IEEE* (2024), 1–6.

30. D. Zhou, M. Sheng, J. Luo, R. Liu, J. Li, and Z. Han, "Collaborative Data Scheduling With Joint Forward and Backward Induction in Small Satellite Networks," *IEEE Transactions on Communications* 67, no. 5 (2019): 3443–3456.

31. Q. Qi, X. Chen, L. Lei, C. Zhong, and Z. Zhang, "Outage-Constrained Robust Design for Sustainable B5G Cellular Internet of Things," *IEEE Transactions on Wireless Communications* 18, no. 12 (2019): 5780–5790.

32. M. Sharma, D. Chadha, and V. Chandra, "High-Altitude Platform for Free-Space Optical Communication: Performance Evaluation and Reliability Analysis," *Journal of Optical Communications and Networking* 8, no. 8 (2016): 600–609.

33. L. B. Castañeda, V. Arunachalam, and S. Dharmaraja, *Introduction to Probability and Stochastic Processes With Applications* (John Wiley & Sons, 2012).

Appendix

Derivation of Equation (1)

We are interested in evaluating the integral

$$\int_0^t (1 - F_{02}(x)) dF_{01}(x)$$

where

- $F_{01}(x)$ is the CDF of a gamma distribution with shape β and rate λ_{01} .
- $F_{02}(x)$ is the CDF of an exponential distribution with rate λ_{02} .

We know

$$1 - F_{02}(x) = e^{-\lambda_{02}x}, \quad dF_{01}(x) = f_{01}(x)dx = \frac{\lambda_{01}^\beta x^{\beta-1} e^{-\lambda_{01}x}}{(\beta-1)!} dx.$$

Thus, the integral becomes

$$\begin{aligned} \int_0^t (1 - F_{02}(x)) dF_{01}(x) &= \int_0^t e^{-\lambda_{02}x} \cdot \frac{\lambda_{01}^\beta x^{\beta-1} e^{-\lambda_{01}x}}{(\beta-1)!} dx. \\ &= \frac{\lambda_{01}^\beta}{(\beta-1)!} \int_0^t x^{\beta-1} e^{-(\lambda_{01}+\lambda_{02})x} dx. \end{aligned}$$

This integral does not have a closed-form in elementary functions, and is usually expressed using the lower incomplete gamma function.

We approximate the given expression

$$x^{\beta-1} e^{-ax} \approx t^{\beta-1} e^{-ax}, \quad (\text{holding } x^{\beta-1} \text{ constant at } x = t).$$

Then

$$\int_0^t x^{\beta-1} e^{-ax} dx \approx t^{\beta-1} \int_0^t e^{-ax} dx = t^{\beta-1} \cdot \frac{1 - e^{-at}}{a}.$$

So

$$\int_0^t e^{-\lambda_{02}x} dF_{01}(x) \approx \frac{\lambda_{01}^\beta}{(\beta-1)!} \cdot t^{\beta-1} \cdot \frac{1 - e^{-(\lambda_{01}+\lambda_{02})t}}{\lambda_{01} + \lambda_{02}}.$$

The above approximation can be improved using moment matching. We consider the integral

$$I(t) = \int_0^t x^{\beta-1} e^{-ax} dx, \quad \text{where } a = \lambda_{01} + \lambda_{02}.$$

We have

$$1 - F_{02}(x) = e^{-\lambda_{02}x}, \quad dF_{01}(x) = f_{01}(x)dx = \frac{\lambda_{01}^\beta x^{\beta-1} e^{-\lambda_{01}x}}{(\beta-1)!} dx.$$

Thus, the integral simplifies to

$$\int_0^t e^{-\lambda_{02}x} f_{01}(x) dx = \frac{\lambda_{01}^\beta}{(\beta-1)!} \int_0^t x^{\beta-1} e^{-(\lambda_{01}+\lambda_{02})x} dx = \frac{\lambda_{01}^\beta}{(\beta-1)!} \cdot I(t).$$

A basic approximation is obtained by replacing the slowly-varying function $x^{\beta-1}$ with its value at the upper limit $x = t$

$$I(t) \approx t^{\beta-1} \int_0^t e^{-ax} dx = t^{\beta-1} \cdot \frac{1 - e^{-at}}{a}.$$

While this approximation is simple, it does not accurately capture the behavior of the integrand over the full interval $[0, t]$, especially when $x^{\beta-1}$ changes significantly. To improve accuracy, we apply moment matching.

We define

$$A(t) = \frac{1 - e^{-at}}{a}, \quad B(t) = \frac{1 - e^{-\lambda_{01}t}}{\lambda_{01}}.$$

These are the integrals of e^{-ax} and $e^{-\lambda_{01}x}$ over $[0, t]$, respectively.

The idea is to approximate $I(t)$ as

$$I(t) \approx t^{\beta-1} \cdot A(t) + \text{correction term}.$$

We observe that.

- $A(t)$ approximates the exponential decay with total rate $a = \lambda_{01} + \lambda_{02}$.
- $B(t)$ approximates the decay using only the gamma rate λ_{01} .

The difference $B(t) - A(t)$ reflects the discrepancy in decay rates between the gamma and exponential assumptions.

To capture the effect of the changing polynomial term $a^{\beta-1}$, we introduce a correction function

$$R(t) = \lambda_{01} t^{\beta-1} - \lambda_{02} \lambda_{01} (a^{\beta-1} - t^{\beta-1}).$$

The corrected approximation is given by

$$I(t) \approx t^{\beta-1} \cdot \frac{1 - e^{-at}}{a} - (B(t) - A(t)) \cdot R(t).$$

Substituting explicitly

$$\begin{aligned} I(t) &\approx t^{\beta-1} \cdot \frac{1 - e^{-at}}{a} - \left(\frac{1 - e^{-\lambda_{01}t}}{\lambda_{01}} - \frac{1 - e^{-at}}{a} \right) \cdot \\ &\quad [\lambda_{01} t^{\beta-1} - \lambda_{02} \lambda_{01} (a^{\beta-1} - t^{\beta-1})]. \end{aligned}$$

Let

$$\int_0^t a^{\beta-1} e^{-ax} dx \approx \left[\frac{1-e^{-at}}{a} \right] \cdot t^{\beta-1} - \left[\frac{1-e^{-\lambda_{01}t}}{\lambda_{01}} - \frac{1-e^{-at}}{a} \right] \cdot R(t),$$

where $R(t)$ is a correction function chosen to capture the difference between gamma and exponential behavior.

Let

$$R(t) = \lambda_{01} t^{\beta-1} - \lambda_{02} \lambda_{01} (a^{\beta-1} - t^{\beta-1}).$$

Combining all terms, the approximation becomes

$$\begin{aligned} \int_0^t (1-F_{02}(x)) dF_{01}(x) &\approx \frac{\lambda_{01}}{(\beta-1)!} \cdot \frac{1-e^{-\lambda_{01}t}}{\lambda_{01}+\lambda_{02}} \\ &- \left(\frac{1-e^{-\lambda_{01}t}}{\lambda_{01}} - \frac{1-e^{-(\lambda_{01}+\lambda_{02})t}}{\lambda_{01}+\lambda_{02}} \right) \\ &\cdot [\lambda_{01} t^{\beta-1} - \lambda_{02} \lambda_{01} (a^{\beta-1} - t^{\beta-1})]. \end{aligned}$$

Derivation of h_0 Values

We want to evaluate the integral

$$I(t) = \int_0^t (1-F_{01}(x))(1-F_{02}(x)) dx,$$

where

- $F_{01}(x) = 1 - e^{-\lambda_{01}x}$ is the CDF of an exponential distribution with rate λ_{01} ,
- $F_{02}(x)$ is the CDF of a gamma distribution with shape parameter $\beta \in \mathbb{N}$ and rate λ_{02} .

Then

$$1 - F_{01}(x) = e^{-\lambda_{01}x}, \quad 1 - F_{02}(x) = \sum_{k=0}^{\beta-1} \frac{(\lambda_{02}x)^k}{k!} e^{-\lambda_{02}x}.$$

Thus, the integral becomes

$$I(t) = \int_0^t e^{-\lambda_{01}x} \left(\sum_{k=0}^{\beta-1} \frac{(\lambda_{02}x)^k}{k!} e^{-\lambda_{02}x} \right) dx = \sum_{k=0}^{\beta-1} \frac{\lambda_{02}^k}{k!} \int_0^t x^k e^{-(\lambda_{01}+\lambda_{02})x} dx.$$

Now, for the last term in the sum (i.e., $k = \beta - 1$), we use the identity

$$\int_0^t x^n e^{-ax} dx = \frac{n!}{a^{n+1}} - \frac{t^n}{a} e^{-at} + (\text{higher order terms}),$$

for a compact closed-form, define

$$I_1(t) = \int_0^t (1-F_{01}(x)) dx = \int_0^t e^{-\lambda_{01}x} dx = \frac{1-e^{-\lambda_{01}t}}{\lambda_{01}}.$$

Next, we note that

$$F_{02}(t) = \frac{\gamma(\beta, \lambda_{02}t)}{(\beta-1)!}, \quad (1-F_{02}(t)) = \frac{1}{(\beta-1)!} \int_{\lambda_{02}t}^{\infty} u^{\beta-1} e^{-u} du.$$

Now substitute back into the integral

$$I(t) = \int_0^t e^{-\lambda_{01}x} \cdot \frac{1}{(\beta-1)!} \int_{\lambda_{02}x}^{\infty} u^{\beta-1} e^{-u} du dx.$$

Switch the order of integration:

$$I(t) = \frac{1}{(\beta-1)!} \int_0^{\infty} u^{\beta-1} e^{-u} \left(\int_0^{\min(x, u/\lambda_{02})} e^{-\lambda_{01}x} dx \right) du.$$

Now evaluate the inner integral:

$$\int_0^{\min(x, u/\lambda_{02})} e^{-\lambda_{01}t} dt = \begin{cases} \frac{1-e^{-\lambda_{01} \cdot (u/\lambda_{02})}}{\lambda_{01}}, & \text{if } u \leq \lambda_{02}x, \\ \frac{1-e^{-\lambda_{01}x}}{\lambda_{01}}, & \text{if } u > \lambda_{02}x. \end{cases}$$

Now split the integral at $u = \lambda_{02}x$ and combine the two parts

$$I(t) = \frac{1}{(\beta-1)!} \left[\frac{1}{\lambda_{01}} \int_0^{\infty} u^{\beta-1} e^{-u} du - \left(\frac{1}{\lambda_{01}} - \frac{1}{\lambda_{01}+\lambda_{02}} \right) a^{\beta-1} \right].$$

Using $\int_0^{\infty} u^{\beta-1} e^{-u} du = (\beta-1)!$, the expression simplifies to

$$I(t) = \frac{1}{(\beta-1)!} \left[\frac{(\beta-1)!}{\lambda_{01}} - a^{\beta-1} \left(\frac{1}{\lambda_{01}} - \frac{1}{\lambda_{01}+\lambda_{02}} \right) \right].$$

Final result

$$\begin{aligned} \int_0^t (1-F_{01}(x))(1-F_{02}(x)) dx &= \frac{1}{(\beta-1)!} \\ &\left[\frac{(\beta-1)!}{\lambda_{01}} - a^{\beta-1} \left(\frac{1}{\lambda_{01}} - \frac{1}{\lambda_{01}+\lambda_{02}} \right) \right]. \end{aligned}$$