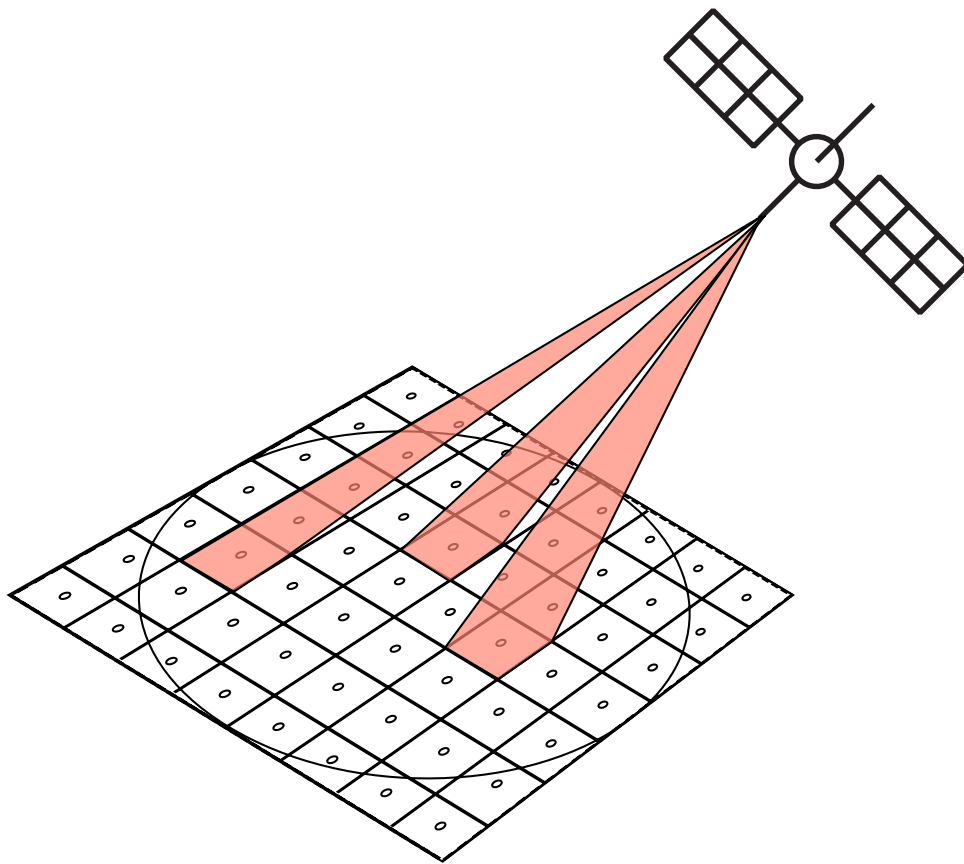# Dynamic Demand Modelling and Optimisation for Satellite Beam Hopping Systems

Masters Thesis

10. Semester ES

David Vanggaard Bedholm Johansen

June 4, 2025

MASTER'S THESIS

Title:

Dynamic Demand Modelling and Optimisation for Satellite Beam Hopping Systems
 Masters Thesis

Theme:
Space communication systems, Optimisation, Beam Hopping

Project Period:
Spring semester 2025

Project Group:
ES10

Participants:

David Vanggaard Bedholm Johansen
      david@vajo.dk
      David V B Johansen

Supervisors and contact:
Israel Leyva-Mayorga

Copies: 1

Page Numbers: 48

Date of Completion:
June 4, 2025

Abstract:

This thesis investigates beam hopping (BH) in satellite networks as a method for serving dynamic and non-uniform user demand. A Markov-based model is used to simulate user activity, and an optimisation problem is formulated to maximise the minimum capacity-to-demand ratio using predictive demand estimates. Simulation results indicate that demand-aware allocation generally improves demand satisfaction compared to naive time-split strategies, but the benefits of predictive optimisation are limited and highly dependent on user dynamics. The work highlights scenario-specific performance variation and discusses the challenges of computational complexity and user modelling. Future directions include more realistic user behaviour, advanced prediction, and scalable solution methods for practical BH deployment.

Preface

I'd like to thank my parents for their continued support in my education, this work stands because of them.

This report would not have been the same without the guidance of Israel Leyva-Mayorga, and for his supervision, I am grateful.

I am also thankful for the support of my friends.

My thanks also goes to Petar Popovski for allowing me to use their stock image collection [1].

Contents

List of Figures

List of Tables

1 | Introduction

The satellite industry is experiencing a boom of interest, with an increase in the number of space objects rising from 1000 objects in space to 10000 objects in space in 2024 [2]. This is due to an unprecedented number of launches as seen in Figure 1.1. The primary category of the launched satellites are telecommunication satellites [2]. Of the telecommunications satellites, the video broadcast has been a predominant business case; however, as users' habits change, the desire for broadband connection has risen [2].
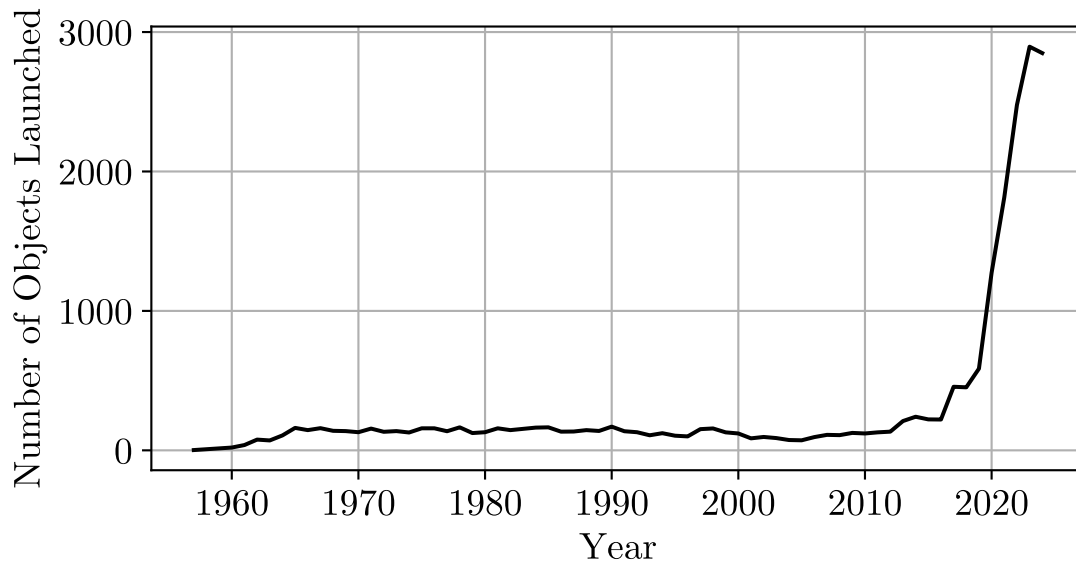
Figure 1.1: The yearly number of launched vehicles from 1957 to 2024 [3].

The transition from video broadcast to broadband connection has led to a focus away from geostationary orbit (GEO) to low Earth orbit (LEO) [4]. GEO is ideal for video broadcast, as relative to the Earth, a GEO satellite is stationary. So TV satellites can be simply pointed at the same place in the sky to achieve a connection [4]. However, broadband connection users have demands for lower latency and higher throughput, as multiple users need different network connections [4]. Opposed to a video broadcast which is only in downlink and sends the same information to all receivers. This made the use of LEO satellites the primary choice for broadband connection, as they have a lower orbit and are also cheaper to launch.

This increased demand for broadband internet over satellite communication is also marked by the development of 5G Non-Terrestrial Network (NTN) and future 6G NTN [5].

One of the most promising technologies to enhance satellite communication connectivity is beam hopping (BH) [6]. BH leverages the adaptability of modern antenna arrays to steer the antenna beams to illuminate different parts of the globe for a certain amount of time before hopping onto a different area of the Earth [7]. This is beneficial for the non-uniform distribution of users, to optimise frequency reuse, and lowering the cost of the launch by having fewer radio channels on the satellite, resulting in a lower mass, among other things [7] [8].

Satellites with the digital adaptability necessary for beam hopping (BH) already exist, namely, EUTELSAT Quantum, SES-17, and Spaceway. These platforms possess the required techni-

cal capabilities but, in the case of Spaceway, have already implemented BH as part of their communication protocol. It has also been included in the DVB-S2X standard [9].

The question of how the BH illumination plan is determined is one of the fundamental questions in incorporating BH into satellite networks and the subject of research papers [8] [10]. One less covered aspect of the illumination planning is ensuring resilience in demand matching systems, in the case that there is no demand, which would lead to longer outage periods.

Furthermore, the illumination plan also needs to deal with not just the spatial distribution of the demands, but also the temporal uncertainty of traffic across the geographic area. An example of the uncertainty of temporal demand is the 5G scheduling requests (SR) to the gNB sent by user equipment (UE) to initiate an uplink transmission [11] [12]. In the non-terrestrial context, due to inherent latency and protocol limitations, the uplink feedback can be unavailable or delayed. As a result, a BH illumination plan must rely on predictive or resilient scheduling strategies to mitigate dynamic user behaviour.

This thesis will investigate the following problem statement:

> How can beam hopping be leveraged to serve non-uniform and time-varying user demand in direct-to-cell satellite communication networks?

This is done by having Chapter 2 analyse and present the theory necessary behind cells, satellite communication, beam hopping, and satellite resources. Chapter 3 introduces the system model and defines the optimisation problem that will be studied, including user behaviour, demand modelling, and the mathematical formulation of beam allocation. Chapter 4 details the implementation of the simulation framework and the computational methods used to evaluate different allocation strategies. Chapter 5 presents the results of these evaluations, comparing the performance of various approaches under different scenarios. Chapter 6 discusses the implications of these results, including their limitations and relevance to practical system design Chapter 7 concludes the thesis and suggests directions for future work.

2 | Problem Analysis

This chapter aims to analyse the parallels between terrestrial cellular networks and NTNs, focusing on the transition from the basic concept of terrestrial networks towards NTNs. Further, the chapter presents dynamic coverage of LEO satellites, including the presentation of beamforming and BH. Finally, presenting the limited satellite resources and Quality of Service (QoS).

2.1 Terrestrial Cellular Networks: A Baseline for NTN

A cellular network connection is taken as a given today, with base stations covering large swaths of the Earth. The concept of cellular connectivity dates back to the developments at Bell Labs and the early telephone voice connections [13]. The concept of frequency reuse is a cornerstone in the development of cellular networks. Frequency reuse refers to the fact that a base station can reuse the same radio frequency spectrum as long as the areas in which they are used are a sufficient distance away from one another to where the co-interference is negligible [13]. The area in which a base station's transmitters are the most likely to serve users in the area is called a cell. The shape of a given cell is visualised as hexagonal for system-level evaluations and descriptions [13]. The reason for the name is that the principles of the shaping of the areas are similar to those of cells [13].
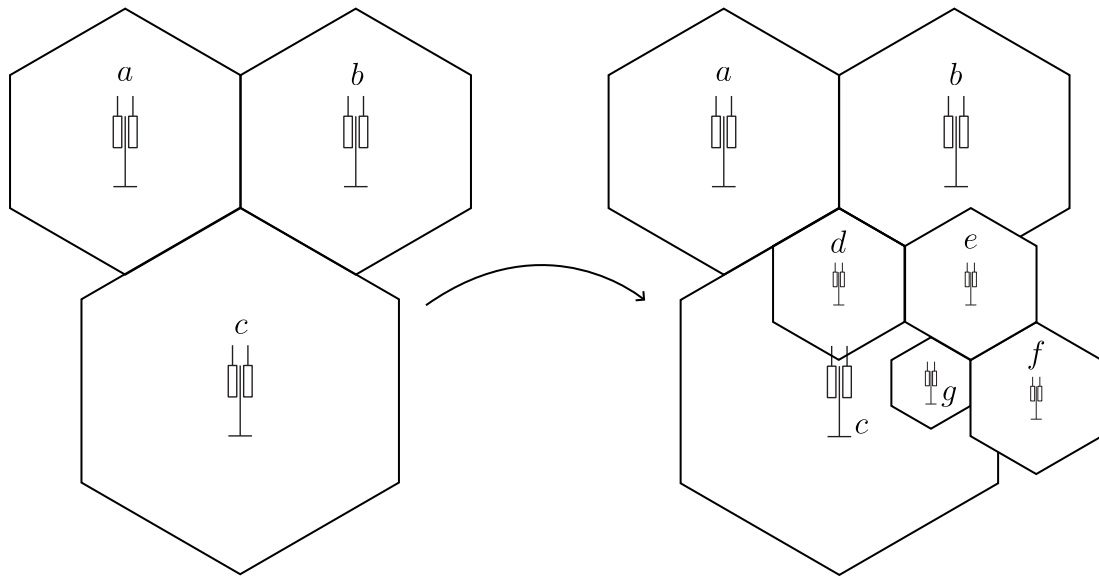
Figure 2.1: Three cells and splitting of cells in case of increased demand.

Another important concept developed at the time was cell splitting. If the number of users in a given cell grows larger than the set of channels available, the cell can be split into a new set of cells with smaller coverage areas. By splitting the cell, the number of users in the former area that can be served increases as the number of base stations and available radio channels increases [13].

Early base stations typically used a single omnidirectional antenna for simplicity, as it provided uniform coverage Figure 2.2a. However, omnidirectional antennas have severe limitations in capacity and interference management [13]. This evolved into fixed sector antennas, which are more directional to separate the cell into multiple subcells, also called sectors, see Figure 2.2b.

Fixed sector antenna increases service availability without requiring additional base stations. They also reduce the co-channel interference of different cells, as the spread of a single channel would be limited [13].

(a) Base station with an omnidirectional antenna.   (b) Base station with several sector antennas
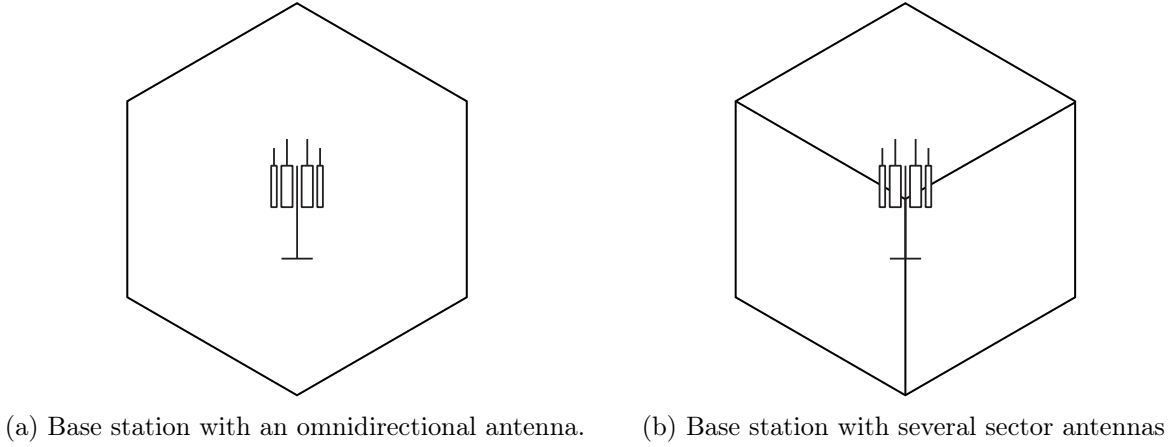
Figure 2.2: Example of sectorisation approach to achieving greater cellular connection.

Basestations' Antenna uses have since then evolved into more adaptive methods with beamforming and multiple-in-multiple-out (MIMO) systems that are used in 5G. These enable the sector antennas to adjust to the experienced network load dynamically. Instead of maintaining a fixed angle, modern base stations can dynamically adjust their beamwidth based on network demand. Massive MIMO (Ma-MIMO) allows for a single base station to achieve a high degree of spatial multiplexing, meaning that the cell is split into even more complex parts and achieves a higher degree of spectrum reuse.

These principles of frequency reuse, cell splitting, and advanced antenna technologies have been instrumental in the development of terrestrial networks. With the advent of NTNs, the principles from terrestrial networks serve as the basis for NTNs. NTNs offer new possibilities in terms of coverage in remote and maritime locations [5], and building on the development of dynamic methods in the terrestrial network, future NTN networks will likely also use dynamic methods to increase coverage [5].

2.2   NTN dynamic coverage & demands

NTN is slated to play a key role in 5G Advanced and 6G systems to provide global connectivity [5]. Currently, several internet connectivity services provide communication through satellite networks, such as Telenet, OneWeb, and Starlink. Global connectivity is enabled due to their altitude and the non-stationary nature of the satellites [5]. The satellite orbits the Earth at a high altitude, resulting in a large moving coverage area. The upper bound of the coverage area is the area of the Earth visible to a satellite, given by [14]

$$S = 2\pi R_E^2 \left(1 - \left(\frac{R_E}{h + R_E} \right) \right) \tag{2.1}$$

This is the upper bound, as this is at the edge, it is the same as having an elevation angle of 0 deg as the edge of the coverage. Geostationary orbit (GEO) satellites were favoured in the past as their orbit progressed at the same speed as the rotation of the Earth and their high altitude of 35 786 km [4]. A GEO satellite's coverage would, therefore, not move, however, it would be at
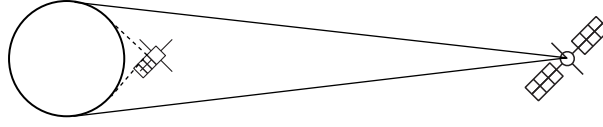
Figure 2.3: Sketch of comparison between the visibility of GEO and LEO.

a fixed point in the sky. The distance needed for GEO is also advantageous, as a single satellite could cover about 42.4 % of the Earth's surface, see Figure 2.3. A GEO satellite constellation would only need about four satellites to provide coverage for the whole globe. However, with the lowering cost of launches, non-geostationary satellite orbits (NGSO) are being favoured as they can provide a much lower latency and a higher bandwidth.

NGSOs are usually in the low Earth orbit (LEO) to medium Earth orbit (MEO). This results in a relative movement with the ground and much smaller coverage areas than traditional GEO constellations. Most of the Starlink satellites at the time of writing orbit at a height of 550 km, resulting in a visible area of 8 % of the globe [15]. Global coverage is then established using large constellations of satellites in various orbits to cover the entire globe continuously [5]. The number of satellites needed depends on a variety of factors, such as the orbital heights used, the inclination of the orbit, and the performance needed [5].

The antenna used to establish a connection with Earth terminals in the visible area has many configuration possibilities. The performance of an antenna is commonly described using Gain patterns, beamwidth and Effective Isotropic Radiated Power (EIRP) [16]. The EIRP and Gain Pattern are related by

$$EIRP(\theta, \phi) = P_{tx} \cdot G(\theta, \phi) \tag{2.2}$$

Where $EIRP(\theta, \phi)$ describes the density of the power in a given direction (θ, ϕ) of the antenna, P_{tx} describes the conducted power and gain $G(\theta, \phi)$ describe the ratio of power transmitted in a direction (θ, ϕ) relative to an isotropic antenna. Beamwidth is related to the gain parameter in the description of the angular width of the antenna. The half-power beamwidth (HPBW) is commonly used to describe the angle of coverage of an antenna's main lobe [16],

$$HPBW = \theta_2 - \theta_1 s.t \qquad\qquad G(\theta_1) = G(\theta_2) = \frac{1}{2}G_{max} \tag{2.3}$$

The shaping of the gain patterns is, along the path loss, the primary determinant of the coverage area.

The antenna systems in satellite communication have evolved similarly to terrestrial systems, seeking to deal with increased data demands, manage interference and maximally use the limited spectrum available by using more adaptive antenna systems and processing techniques [17].

2.2.1   Beamforming and multi-beam system

Beamforming is a technique that shapes the radiation pattern of an antenna using an array of individual elements [16]. A phased array is a type of antenna that uses the phase delay between the elements, resulting in the superimposing of the signals constructively in some directions and destructively in others [16], see Figure 2.4. The reciprocity holds for both transmission and reception. The implementation of phase delay between the separate elements can be implemented with fixed phase delay elements from the radio frequency source out to the elements [16]. Having fixed elements does not allow for dynamic beamforming, as the delay can not be changed. The ability to steer the beam electronically can be accomplished by digitising part of the RF-chain,
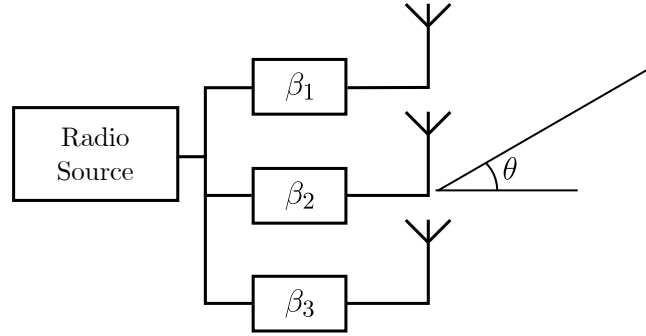
Figure 2.4: Example of a linear phased array combining the signal in a direction.

enabling dynamic phase delay between elements [18]. There are a variety of other types of antenna arrays, however, these rely on the passive reaction of the antenna, so they cannot be steered other than mechanically [18].

In more advanced configurations, such as multi-beam systems (MBS), multiple independently steered beams can be formed using separate RF chains [19] [20]. MBS, while adding complexity, enhances the area in which a single satellite can provide coverage similar to what using multiple arrays in terrestrial systems would accomplish by sectorisation.

The array performance is typically described by the Array Factor (AF), which describes the gain attributed to the antenna performance given identical antenna elements [16]. However, in the system-level evaluation employed in this report, the gain is assumed to be able to be steered arbitrarily through the range of the visible area of the satellite.

2.2.2 Beam Hopping

Beam hopping (BH) is a technique that covers multiple cells by illuminating the cells with an antenna beam for a certain amount of time before hopping [21]. BH leverages the increased adaptability of antenna systems, which allows the tracking of different cells on the ground at different times to possibly reach more cells within the visible area than there are beams available [21].

BH is promising because of the possible benefit of reducing the number of radio chains needed on the satellite to cover a larger area [21] [8]. This has the benefit of reducing the needed mass of the spacecraft, and possibly also the power budget [7]. The high degree of flexibility of BH also improves the future-proofing of the telecom satellite [21]. A telecom satellite is a large investment and will likely stay in operation for a long duration; adding BH is likely to extend the useful operational lifespan of the satellite.

The Spaceway satellite constellation has provided broadband reception to multiple users in North America in the Ka-band using the BH method [21]. There are also other satellites which could possibly implement it, such as the EUTELSAT Quantum and the SES-17 as these have almost fully digital communication hardware [7]. In the newer standard DVB-S2X, BH is also possible and is one of the first standards to directly specify it [7].

Several studies into how BH can be optimised and used have also been performed since around the 2010s [21]. In [8], the joint of the power allocation and scheduling is optimised to reach minimum rates by using various techniques to rewrite the coupled power scheduling problem. The issue of optimising the frequency reuse is also of major importance and was discussed in [7], with the use of selective precoding in clustered beams. The role of deep learning and RL has

also been discussed in various papers such as [22] and [10]. In [22], the learning is part of a two-step approach. The neural network is used to limit the combinatorial space of an optimisation problem [22]. The paper [10] considered the joint problem of allocation of bandwidth and beam illumination by utilising cooperative multi-agent deep reinforcement learning. [10] is one of the papers that highlight the aspect of time dynamics in user demand for beam allocation.

Several approaches have addressed beam allocation problems in BH, but relatively few studies have investigated the impact of dynamic, time-varying user demands. An understanding of how these dynamics affect the performance and optimisation of beam hopping is a crucial aspect of study in the development of BH technologies. This gap motivates the focus of the master's thesis on evaluating BH systems under time-dependent user demand scenarios.

2.2.3 Limited Satellite Resources

The downlink from a satellite will transmit a signal with a centre frequency, bandwidth and power. The power capable of being transmitted by satellite is limited due to size and weight limits imposed by both cost and launch capabilities. The centre frequency and bandwidth are limited by the regulations and standards applicable to the communication protocol.

Frequency Bands

5G NTN has two different frequency ranges one in the Ka-band and one in the S-band some are seen in Table 2.1. N256 is in the S-band and N512 is in the Ka-band [23].

Band	Uplink [MHz]	Downlink [MHz]	Channel Bandwidth [MHz]
n256	1980 - 2010	2170 - 2200	5,10,15,20
n512	27500-30000	17300 - 20200	50, 100, 200, 400

Table 2.1: Two frequency bands for 5G NTN [23].

Full duplex is achieved through frequency division duplexing, so there is an uplink and downlink frequency range [23]. There are other bandwidths as well, however, these have the largest bandwidth available and are applicable in the EU. N512 offers a higher bandwidth than N256, which allows for a higher communication channel capacity. However, the effects of the channel in the N256 and N512 are different, whereas N512 is generally worse off in regards to rain attenuation [24].

2.2.4 QOS

Quality of service (QoS) is an important topic in communication systems [25]. This term describes the desired End-to-end throughput performance of the service provided [25]. The parameters that describe quality can be many, but sticking to measurable terms, there are latency, reliability, and throughput [25]. Latency describes either the amount of throughput time it takes to receive a response or for the single to end where it needs to be. The reliability describes the ratio of transmitted packages to those received, conditioned on the network being available. The throughput is the rate of the data being sent, usually the average rate over a given period due to the variability of the instantaneous throughput [25].

These three different terms can be used to describe the type of service that you may provide for a given use case. The use case could be mobile streaming in which both reliability and data rate

need to be high, however, the latency is not of much concern. Meanwhile, an IoT device might want a low data rate, but a high reliability in case it is on a piece of critical infrastructure.

Traditionally, to achieve this, different protocols have to be implemented with a limited scope that allows for a specific target use case and QoS [26]. Network slicing is an alternative approach where multiple different QoS is achieved over the same physical layer [26]. This would allow the three primary services eMBB, URLLC and mMTC to coexist on the same physical layer [26]. Network slicing was implemented in 5G Core Network, through breakthroughs in software-defined networking (SDN) that allowed for the implementation of separate pipelines that can enable guaranteeing different QoS [26].

Services	Latency	Reliability	Throughput
eMBB	4 ms	Not Specified	$50\,\frac{\text{Mbit}}{\text{s}}$
URLLC	1 ms	99.999%	Low/moderate($1\,\frac{\text{Mbit}}{\text{s}}$)
mMTC Flexible (Seconds)	Depends on use case	Very low ($\frac{\text{kbit}}{\text{s}}$)	

Table 2.2: Overview of the QoS of the three different services [25] .

For parameters that are not explicitly specified, it is because values depend heavily on specific applications or deployments. For instance, eMBB generally provides moderate reliability adequate for common internet usage, while URLLC prioritizes ultra-low latency and reliability over throughput, typically handling smaller payloads [25]. mMTC services target scenarios with infrequent, low-data-rate communication, and therefore latency and reliability requirements vary significantly based on specific application contexts [25].

In this context, the demand or desired throughput is investigated to accomplish demand-matching. Achieving reliable QoS in NTN environments depends not only on the underlying satellite resources and protocols but also on the system's ability to adapt to time-varying and spatially inhomogeneous user demands. This necessitates defining a system that allows for the modelling of the user demand, and also the BH illumination plan solution.

3 | System Model Definition and Optimisation Problem Formulation

This chapter introduces the system model and the associated optimisation problem that will be used to evaluate beam hopping performance under user demand variability.

The BH system that is being covered is one in which the demand of the users in an area of visibility of a single satellite. This is done as the primary focus of this problem is the connection between a satellite and the users within the serviceable area beneath it, as seen in Figure 3.1.
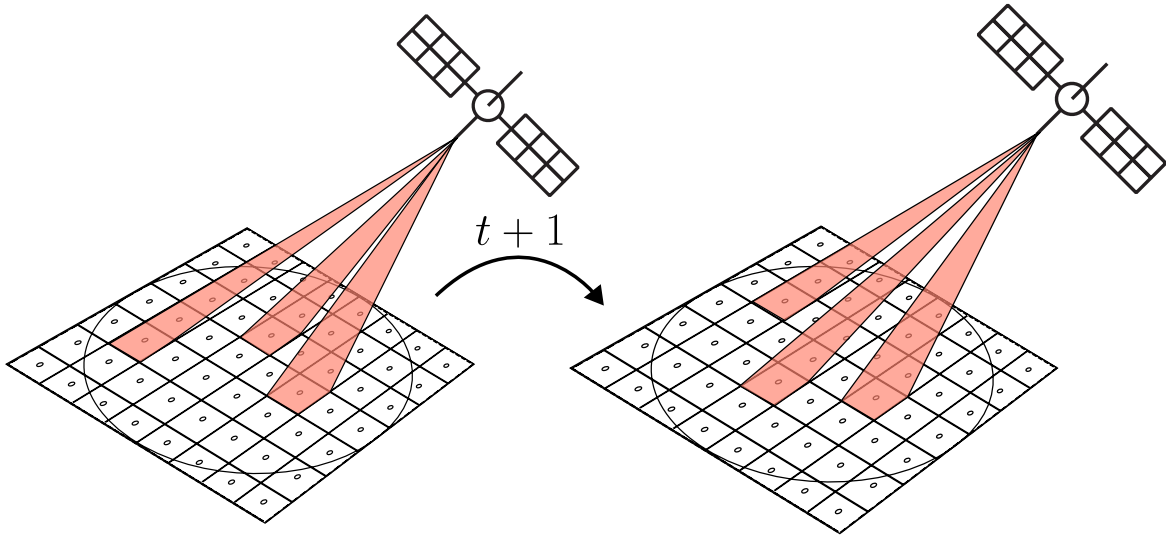
Figure 3.1: Illustration of the single multi-beam satellite beam hopping system.

To describe this system, it can be split into a few different system parts. There is the user model, their demand and the cells they are placed within. These cells are connected to a satellite which has to organise the beams. These different parts of the beam hopping system will be described in the following sections. An overview of the symbols can be found in Chapter B.

3.1 User Model and Demand

The users that the satellite serves can be described by the user model. The user state is defined by a tuple,

$$u = (\lambda, \phi, D)$$

The state consists of the user's location on the Earth given λ, ϕ as respectively the latitude and longitude. D describes the demanded throughput that the user needs over the configuration period T_R. The configuration period T_R is the duration over which the beams are allocated.

For simplicity, the users are placed at the centre point of their respective cells. This is done as the difference in capacity when dispersed over the area is negligible and not as relevant for the problem evaluation [7].

The user activity is modelled as either on or off. Upon initialisation of the individual user, the

On/Off state is determined by a Bernoulli Distribution [27],

$$f(k; p) = \begin{cases} p_{on} & \text{if } k = 1, \\ p_{off} = 1 - p_{on} & \text{if } k = 0. \end{cases}$$

where $k = 1$ is On and $k = 0$ is off. The demand that the user needs over the period is $D = k \cdot D_{on}$.

The user activity over time is modelled by a discrete Markov process to model the state change given a discrete time step. The probability of the state changes is described by the state change matrix [27].

$$A = \begin{bmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{bmatrix}$$

Where the probability of transitioning from off to on is described by α and the probability of going from an On state to an Off state is β.

Over the period, the variability of the channel largely depends on the number of state changes in the demand that are made, and the number of state changes in the system is called O.

The state of the user converges to a steady state probability,π, when O approaches ∞, given the Markov property [28]. The steady state is denoted as π. The steady-state distribution can be calculated by finding the eigenvalues of the transition probability matrix. Given that it is a two-state distribution, there exists a formula for calculating the steady-state

$$\pi = \begin{bmatrix} \pi_1 \\ \pi_2 \end{bmatrix} = \begin{bmatrix} \frac{\beta}{\alpha+\beta} \\ \frac{\alpha}{\alpha+\beta} \end{bmatrix} \tag{3.1}$$

This means that if the on or off state of the user has not been observed for several changes, the probability that it will be in either state converges to the same distribution regardless of the initial state.

This is important regarding the period, T_r, because if the number of changes O within the period is large, the state at the end of the period will converge to the steady state distribution.

3.2 Cells

The users are located in cells defined as the area in which a single beam serves this collection of users. The cells are described by the tuple,

$$c_k = (\lambda_k, \phi_k, \mathbf{U}) \tag{3.2}$$

where (λ_k, ϕ_k) are the location of the centre point of the cell. The centre points are distributed evenly along the longitudinal and latitudinal axes. The total number of visible cells is K given by their size and the minimum observation angle, θ_{min}, of the user.

The cells are distributed evenly across the angles in a way that results in a fixed midway point distance between each cell centre.

$\mathbf{U}$ is a tuple consisting of the states of users within the cell,

$$\mathbf{U} = (u_1, \ldots, u_{M_k})$$

with cell k having M_k users as a function of the cell. The location of all users is at the centre of each cell.

The demand of the cell is therefore an aggregate of several independent Markov processes, resulting in a bursty traffic model. All users have the same state transition matrix.

The number of users in a cell M_k is chosen to be modelled as a Bernoulli allocation with a high and low density of users. p_{high} is the probability that $M_k = M_{high}$ and p_{low} is the probability that $M_k = M_{low}$. This is to achieve a non-homogeneous demand across the cells instead of having only Markovian process determine the non-homogeneous demand.

The demand of the cell, D_k, is the aggregate of the demand of the users, which can be calculated using the 1-norm of all the user-states in a cell and multiplying by D_{on}.

The demand of the cell, D_k, is the aggregate demand of all users in the cell. This is expressed as the L1-Norm of the individual demand vector of the cell,

$$D_k = || \left[D_{k,1}, D_{k,2}, \ldots, D_{k,M_k} \right] ||_1 = \sum_{n=1}^{M_k} D_{k,n} \tag{3.3}$$

This results in an aggregate demand that is based on the individual user behaviour as seen in Figure 3.2.
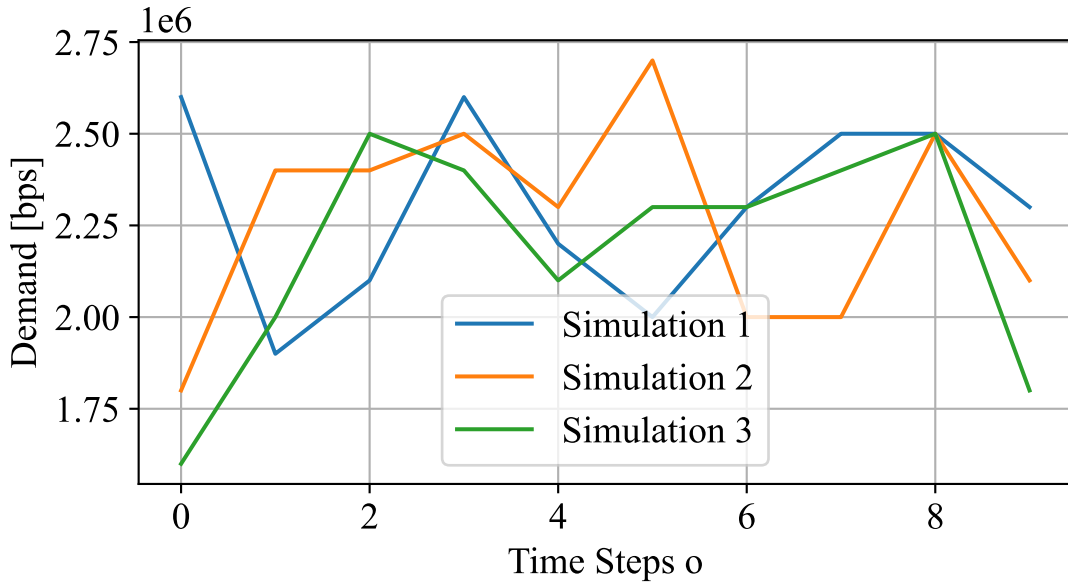
Figure 3.2: Plot of three realisation of thirty users with $100 \frac{\text{kbit}}{\text{s}}$ each.

3.3   Satellite and Allocation

To evaluate the beam hopping allocation across the cells, the satellite is simulated at a fixed point over the Earth, h_{sat}. The satellite has a number of beams, N_b, which can be allocated to the cells at different times. The beams will be sent with a fixed power, P_b, and the beams are configured by a beamforming mechanism, resulting in some gain G_{beam}. The beams can be steered independently towards the cells to serve the users within the cells. Only one beam can be active within a cell at any given time.

To describe which cells the satellite covers at any given time, a matrix can be defined, $X \in \mathbb{R}^{T \times K}$, with T being the number of slots in the configuration period and K being the total number of

slots [7]. The allocation grid can be seen in Figure 3.3. The individual elements can be described
by $x_{i,k}$, which indicates whether a cell has an active beam in its location, k at time i [10]. $x_{i,k}$ is
a binary value, $x(i,k) = \{0,1\}$, as only one beam is active in one cell at a time i.
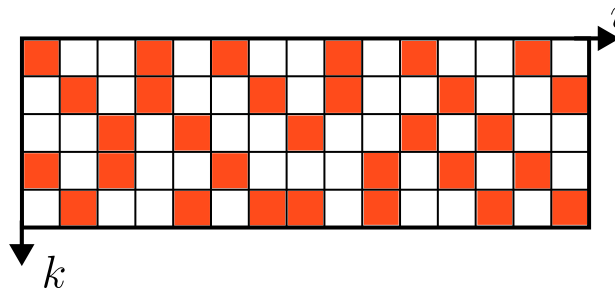
Figure 3.3: BH allocation grid with individual elements illustrated with a maximum of two beams
at any given time.

This description is useful as it allows us to set constraints on the allocation concerning the other
parameters.

$$N_b \geq \sum_{k=1}^{K} x_{i,k} \quad \forall i \tag{3.4}$$

which describes the constraint of the allocation and the number of beams active at any time.
To assign the user a throughput, a means to calculate the throughput is necessary, which is
described in the radio link section.

Each i in the allocation will have a duration of T_s, which will result in there being $\frac{T_r}{T_s} = S$.
The number of time slots S is not the same as the number of state changes O. The number S
describes essentially how many times over the period the beam can be allocated to a different cell.
While O describes the number of times that the user can change behaviour within the period.
This results in $S \gg O$.

3.4 Radio link

The capacity of the radio link between the satellite and the users in a cell, k, can be described
by the Shannon capacity [29].

$$R_k = B \cdot \log_2(1 + \text{SNR}) \tag{3.5}$$

The capacity given by Shannon is the upper limit of achievable throughput over a noisy channel
subject to additive white Gaussian noise (AWGN) [29]. The signal-to-noise ratio (SNR) can be
calculated by the,

$$SNR = \frac{P_r}{N} = P_{r,db} - N_{db} \tag{3.6}$$

which depends on the received power and the noise at the receiver. The noise is the thermal
noise as [24]

$$N = k_b TB \tag{3.7}$$

where k_b is the Boltzmann constant, T is the temperature and B is the bandwidth. The received
power can be calculated by cascading the effects of transmission, the path of propagation and
the reception [24].

$$P_{r,db} = G_{tx,db} + P_{tx,db} + G_{rx} - PL \tag{3.8}$$

The path loss is calculated by using the systems-level path loss model of [23]. $G_{tx,db}$ is the gain
of the transmitter antenna, which is the gain resulting from the beam forming in the antenna.
The $P_{tx,db}$ is the transmitted power in dB. The G_{rx} is the gain at the receiver. The PL is the
path loss consisting of the free space path loss (FSPL), atmospheric gas absorption, scintillation
and shadow fading. The FSPL is described by the equation,

$$FSPL(d, f_c) = 32.45 + 20\log_{10}(f_c) + 20\log_{10}(d) \tag{3.9}$$

d is the distance between the cell centre and the satellite in meters, and f_c is the transmitted
carrier as seen in figure Figure 3.4.
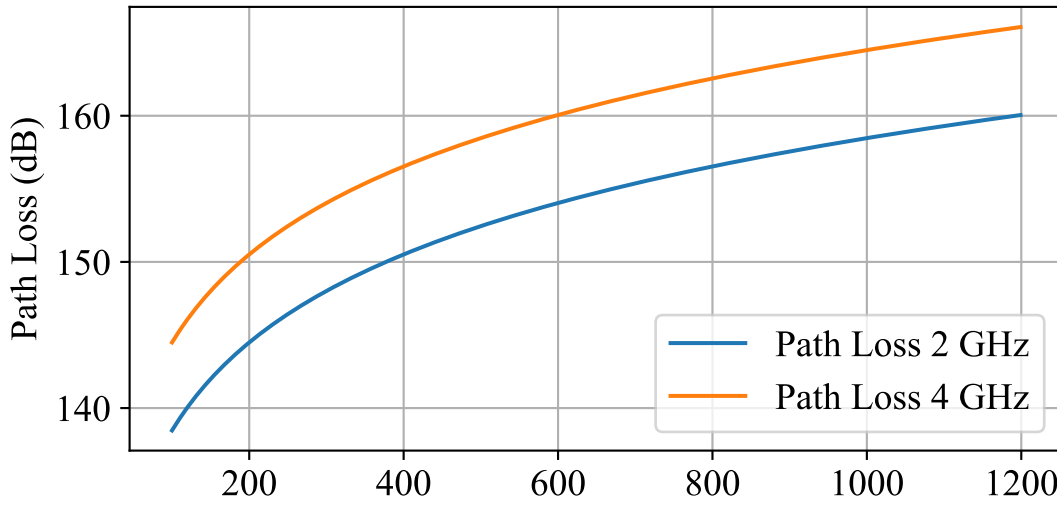
Figure 3.4: Free Space Path Loss over distance given $2\,\mathrm{GHz}$ based on Equation (3.9).

Given the channel model, the capacity of the radio link can be computed.

The capacity met for a single cell k using the allocation notation given can be formulated using
the following equation

$$C_k = \frac{1}{T_r} \sum_{i}^{S} R_k \cdot x_{i,k} \cdot T_s$$

$$C_k = \frac{1}{S} \cdot R_k \sum_{i}^{S} x_{i,k} \tag{3.10}$$

This represents the average demand met over the period T_r.

3.5   Prediction of demands

Given that the user demand is likely to change with the state, given the Markov property. Being
able to have some estimate of the changes that are happening over the period T_r might be
beneficial to ensure more consistent coverage.

As it is a sum of Markov chains, the probabilities of the following states after a certain number
of steps o can be calculated using the Poisson Binomial Distribution (PBD) [30]. The PBD is

a more general form of the Binomial distribution, it is the sum of independent Bernoulli trials
that can have different outcome probabilities [30]. The effect on the probability distribution of
different initial states is seen in Figure 3.5.
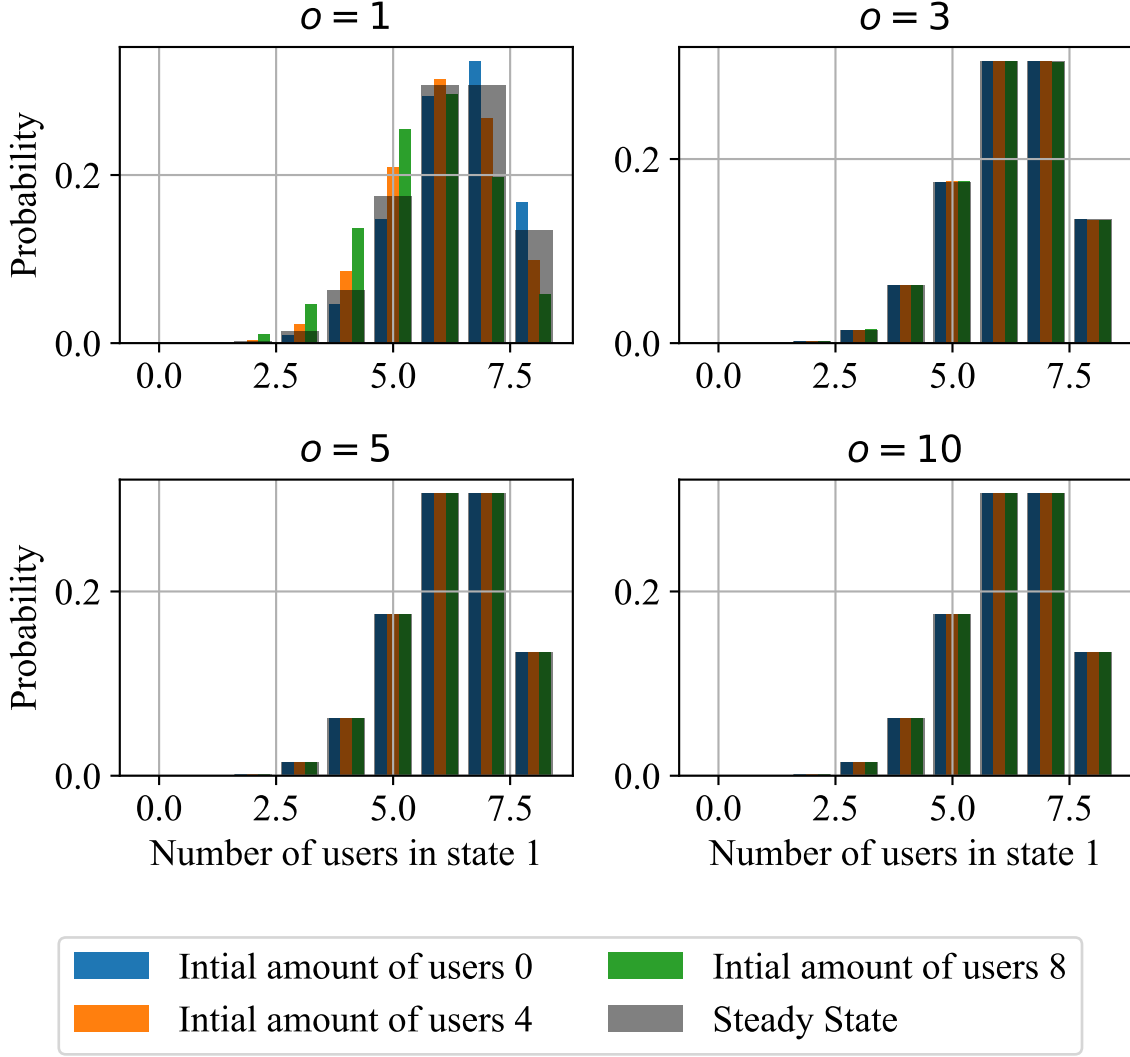
Figure 3.5: Three different initial states over different times, showcasing the convergence towards
steady state as o increases.

The demand of users in an arbitrary cell, k, after o steps can be defined as a PBD is defined as
it is a sum of multiple independent Markov processes given in Equation (3.3). For notation sake
the demand at step o of cell is given by $D_k^{(o)}$

$$D_k^{(o)} = \sum_{n=1}^{M} D_{k,n}^{(o)} \tag{3.11}$$

The probability of $D_{k,n}^{(o)}$ for the user n being in the ON state depends on the previous state $o-1$.
Henceforth the k is omitted as this generalizes to all cells in k. As an example the probability of
being on after one step from $o = 0$ to $o = 1$ is equal to

$$\Pr(D_n^{(1)} = 1 | D_n^{(0)} = 1) = 1 - \beta \qquad\qquad \Pr(D_n^{(1)} = 1 | D_n^{(0)} = 0) = \alpha \tag{3.12}$$

This is where the two different Bernoulli trial probabilities come from, and hence why it is a
PBD. The probability of it being in an ON state after o changes can be computed based on the
initial state of the user $D_n^{(o)}$ where $o = 0$, and the state change matrix, A

$$\Pr(D_n^{(\mathbf{o})} = 1 | D_n^{(\mathbf{0})} = 0) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} A^o \tag{3.13}$$

$$\Pr(D_n^{(\mathbf{o})} = 1 | D_n^{(\mathbf{0})} = 1) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} A^o \tag{3.14}$$

The probability mass function for the sum of demand after o steps can be written as the PBD
formulation [30].

$$\Pr\big(D^{(o)} = z\big) = \sum_{A \in F_z} \left(\prod_{i \in A} p_i \right) \left(\prod_{j \notin A} (1 - p_j) \right). \tag{3.15}$$

Where F_z is the set of subsets that can be selected from A with z elements. Where A is the set
of values $\{1, \ldots M\}$. As the users are subject to the same Markov Transition Matrix, the sum
of demand, $D_k^{(o)}$, can instead be defined as the sum of two binomial distributions.

$$D_n^{(o)} = X_1 + X_2 \quad X_1 \sim \text{Bin}(M_1, p_1^{(o)}) \quad X_2 \sim \text{Bin}(M_2, p_2^{(o)}) \tag{3.16}$$

Where M_1 is the number of users active at $o = 0$ and M_2 is the number of users that were inactive
in state $o = 0$, and $M_1 + M_2 = M$. The probability $p_1^{(o)}$ is equal to Equation (3.14) and the
probability $p_2^{(o)}$ is equal to Equation (3.13). The PMF of the sum of two binomial distributions
can be calculated using the binomial convolution [30]

$$\Pr(D_n^{(o)} = z) = \sum_{j=0}^{z} \Pr(X_1 = j) \cdot \Pr(X_2 = z - j)$$

$$= \sum_{j=0}^{z} \binom{M_1}{j} p_1^j (1 - p_1)^{M_1 - j} \cdot \binom{M_2}{z - j} p_2^{z-j} (1 - p_2)^{M_2 - (z-j)} \tag{3.17}$$

The probability mass function can be used to describe the distribution of the possibilities of
the demand. A verification has been performed to verify the binomial convolution as a fitting
theoretical description, see Figure 3.6, it is compared with the Monte-Carlo Estimation of the
probability distribution.
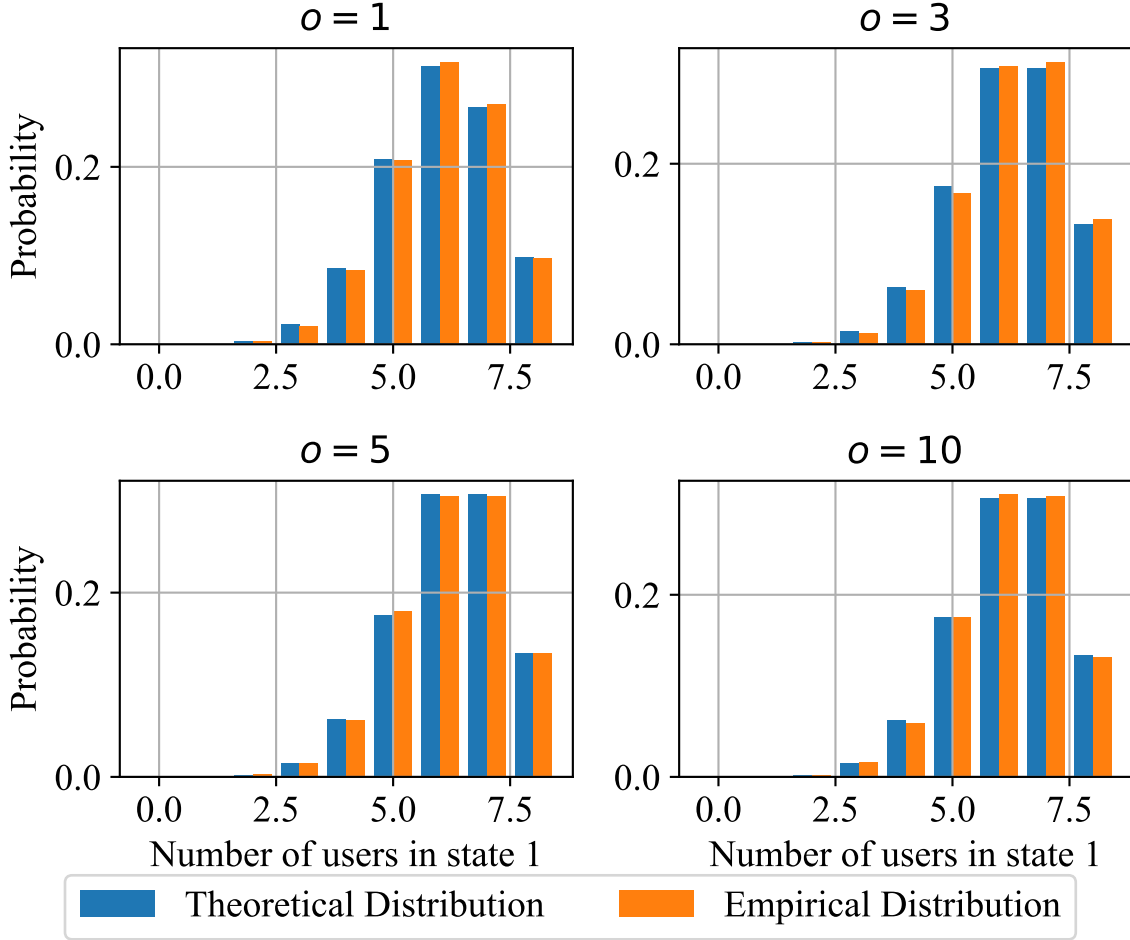
Figure 3.6: Comparison of Monte-Carlo Estimation of PMF with 10000 realizations and theoretical binomial convolution calculation of PMF.

The divergence between the theoretical and the empirical distribution achieve KL-Divergence of: 0,0003, 0.0005, 0.0001, 0.0003.

Given the PMF some predictive measure of how the demand is changing over time can be defined.

The measures that are evaluated are the average value of demand over the O steps.

$$\hat{D} = \mathbf{E}[\frac{1}{O} \sum_{o=0}^{O} \mathbf{D}^{(o)}] = \frac{1}{O} \sum_{o=0}^{O} \mathbf{E}[\mathbf{D}^{(o)}] \tag{3.18}$$

The expected value can be calculated using the probability mass function (PMF), given in Equation (3.17) [30], by

$$\mathbf{E}[\mathbf{D}^{(o)}] = \sum_{n=1}^{M} \Pr(\mathbf{D}^{(o)} = n) \cdot n \tag{3.19}$$

The average value of the demand over the time steps would limit the possibility of assigning too low a demand to each cell before optimisation.

Rather than the average demand over time, the average mode of the PMF could also be used

across time steps O.

$$\bar{D} = \frac{1}{O} \sum_{o=0}^{O} \text{Mode}(\text{Pr}(\mathbf{D}^{(o)} = n)) \tag{3.20}$$

The Mode is the largest value within the probability distribution. Another method would be to use the CDF of the process.

3.6 Optimisation Problem

The matter of the objective to meet can vary, as seen in the description of other studies into BH. In this thesis a demand-matching approach is used, where the maximisation of the minimum of the capacity to demand ratio, $\frac{C}{D}$, is chosen. This leads to the following objective,

$$\max \min_{1 \leq k \leq K} \frac{C_k}{D_k} \tag{3.21}$$

Where the max of the minimum of the capacity to demand ratio across the cells is chosen.

This is different from simply formulating a feasibility problem, as that sets hard bounds on whether it is possible to fully cover the demand. The reason for not simply going for $\max \min_k C_k$ is due to the fact that it would merely spread out the capacity evenly amongst the cells, neglecting the demands of the users themselves.

The capacity is calculated as in Equation (3.10). The demand is some measure of the demand of the cell. There are two constraints. The first can be seen in Equation (3.4) that at any given time slot, the sum of active beams can not be higher than the number of beams. The second is given that $x_{i,k}$ is a binary value $\{0, 1\}$. This can be formulated as the following problem.

$$
\begin{aligned}
\max \min_{1 \leq k \leq K} \quad & \frac{\frac{1}{S} \cdot R_k \sum x_{i,k}}{D_k} \\
\text{subject to:} \quad & \sum_{k=1}^{K} x_{i,k} \leq N_b, \forall i \\
& x_{i,k} \in \{0, 1\}, \quad \forall i, k
\end{aligned}
\tag{3.22}
$$

This problem is also only worth investigating in the case the max min does not just result in zero this requires, $K < N_b \cdot S$. The max-min statement is not solvable by most common solvers as it is non-smooth [31]. This can be mitigated by introducing an auxiliary variable, t, as a new objective. t can be defined as [31],

$$
\begin{aligned}
t &\leq \frac{\frac{1}{S} \cdot R_k \sum_i^{S} x_{i,k}}{D_k} \quad \forall k \\
D_k \cdot t - \frac{\frac{1}{S} \cdot R_k \sum_i^{S} x_{i,k}}{D_k} &\geq 0 \quad \forall k
\end{aligned}
\tag{3.23}
$$

where D_k and C_k are moved to the left side of the equation. Removing D_k from the denominator ensures numerical stability, in the case D_k should be zero.

The problem can then be reformulated as the maximisation of t,

$$\max t$$

subject to:

$$D_k \cdot t - \frac{\frac{1}{S} \cdot R_k \sum_i^S x_{i,k}}{D_k} \geq 0 \quad \forall k$$

$$\sum_{k=1}^{K} x_{i,k} \leq N_b, \quad \forall i$$

$$x_{i,k} \in \{0,1\}, \quad \forall i,k$$

(3.24)

If the demand D_k were equal to 0, the first constraint for that cell k would be fulfilled by default. This would mean that in the optimal case, it is unlikely for the cell to be allocated at any point in time. For specific types of communication, such as in emergencies, it is very undesirable not to be served for an entire period T_r. A bias can be added to the constraint, or another constraint can be added. The bias, R_{min} is added by adding a small value to D_k, if $D_k = 0$. The minimum requirement results in this cell at a minimum being served once during the period. This results in the following optimisation problem,

$$\max t$$

subject to:

$$\text{maximum}(D_k, \ R_{min}) \cdot t - \frac{\frac{1}{S} \cdot R_k \sum_i^S x_{i,k}}{D_k} \geq 0 \quad \forall k$$

$$\sum_{k=1}^{K} x_{i,k} \leq N_b, \quad \forall i$$

$$x_{i,k} \in \{0,1\}, \quad \forall i,k$$

(3.25)

Which is an Mixed-Integer Linear Programming (MILP) problem [32].

The measure of D_k used is evaluated across the initial state of the cell at $o = 0$, the predictive measure by using the average, $\hat{D}_k$, Equation (3.18), or the average mode over time, $\bar{D}_k$ Equation (3.20). To investigate the performance of the various demand predictions the system is implemented and evaluation parameters are described in the following implementation chapter. Including the user of naive solution

4 | Implementation and Evaluation

To implement the system an environment needs to be created that can compute the values associated with the optimisation. An overview of the flow of evaluation is in Figure 4.1.
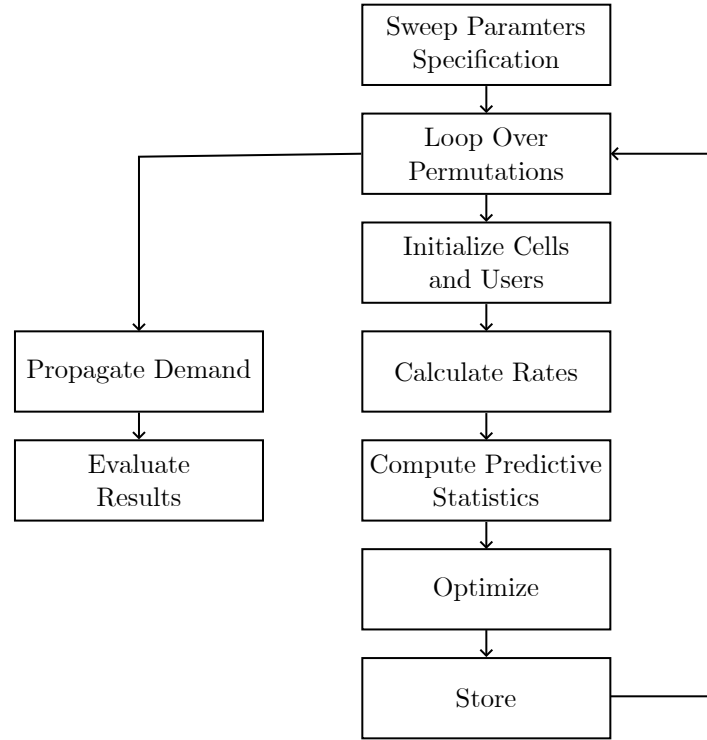
Figure 4.1: Flowchart for system evaluation.

To this end, a small Python library is implemented to organise and test the required functionality. The functionalities that are needed are implemented using JAX to achieve inherent hardware acceleration regardless of the device. JAX needs pure functions to enable the hardware acceleration achieved by using XLA [33]. A pure function is a function that has the same output given the same input, with no side effects [33]. This means that no global variables are used. The optimisation is computed using the Gurobi Optimiser with its Python interface.

To evaluate the system, it needs to be specified which parameters are used and are swept over. These parameters are then looped over its permutations. Given the parameters of a single permutation, the cells and users are initialised. Then the rates are calculated based on the distances, and the parameters used in the permutation, such as the transmit power and the gain of the satellite. Then the Predictive Statistics are computed. These are then used to optimise the problem, and the results are then stored. The demand is then propagated based on the Markov models used in the permutations. This is looped through until all permutations are evaluated and visualised in the results.

The individual blocks will be presented in the first part of the chapter afterwards, the parameters for evaluation are listed.

4.1 Implementation

The flowchart is implemented in the script **single_satellite_optimisation_script.py**. The following subsections provide an overview of the implementation.

4.1.1 Sweep Parameters Specification

The sweep parameters specification is simply a set of variables that are set to simulate the single-satellite situation. These can be split into two different sets of parameters, those that are fixed and those that are changed for the different permutations.

Fixed values include, for example, the frequency and bandwidth, as these are not relevant for assessing either the prediction or the resilience. The non-fixed parameters would, for example, be the minimum rate R_{min}, so as to evaluate the effect of different minimums.

The implementation of this is a list of simple variable assignments that are done at the top of each script. The variables used for the sweep are stored in a dictionary, so that they may be saved with the results of the individual simulations. This is saved using the pickle module. The pickle module can save in layers, meaning multiple dictionaries can be pickled sequentially in the same file.

4.1.2 Loop Over Permutations

The loop over permutations is implemented through a series of for loops, with an inner loop running for the number of iterations for each given configuration. The randomisation in JAX is an explicit pseudo-random method, meaning that the randomness is based on a key that needs to be rolled manually to ascertain randomness. This is done by setting keys based on a seed, which then, after each use, needs to be split to obtain a new pseudorandom number generation (PRNG) key [33].

The loop over the permutations calls the function in the script that optimises the beam allocation, optimise_allocation_of_beams and optimise_allocation_of_beams_with_prediction.

4.1.3 Initialise Cells and Users

The cells are defined by the tuple described in Equation (3.2). This is implemented as a named tuple in Python in the Project library to have explicit states [33]. The initialisation is done in two steps,

1. The cell's centre position is computed based on the visible area of the satellite.

2. The cells are populated according to the probability distributions described in the Chapter 3.

To determine the cell centre, the satellite position and the minimum observation angle are used to calculate the half-angle value of the cone that defines the range of latitudes on the ground. A slice of the cone along the longitude is seen in Figure 4.2.
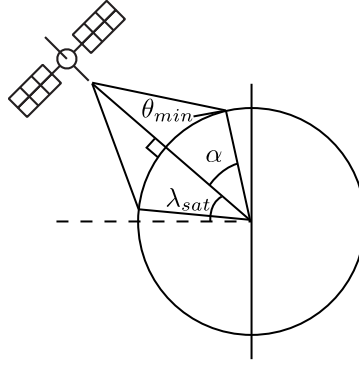
Figure 4.2: Slice of earth along a longitude axis, with satellite to indicate α.

For simplicity, the value, α, is used to generate a set of points that define the cell centres, within a square of the Earth with an angular width of α along each side. α depends on the satellite height h_{sat}.

The set of points is defined by a mesh grid that is computed by using linspace along the latitude and the longitude. The number of points along each axis is calculated based on the desired size of the cells on the ground.

$$N_{cells} = \text{int} \left(\frac{\alpha}{\frac{\text{cell\_size}}{\bar{R}_E}} \right) \tag{4.1}$$

N_{cells} is the number of cells along a singular axis. The placement of the cell centres along one axis should have an even spacing between one another, and to the edge of coverage. This results in the linspace needing to create $N_{cells} \cdot 2 + 1$ points along each axis. Every uneven value is then the centre point of the cells, and the even values are the borders. The cell mesh can then be computed by using the list of cell centres along each axis.

This square results in more cells being generated than are within the visibility area of the satellite, given θ_{min}. A mask is created of the generated area with the cells that are within the area being true. This can be computed by checking whether the angle between the vector pointing towards the satellite and the vector pointing towards the cell centre, called the central angle, $\Delta\sigma$ [14].

$$\Delta\sigma \leq \alpha \tag{4.2}$$

To speed up computation, instead of calculating the angles themselves, the cosine value is calculated instead. This results in the reversal of the inequality as cos(x) is strictly decreasing in the range $0 \leq x \leq \pi$.

$$\cos(\Delta\sigma) \geq \alpha \tag{4.3}$$

$\cos(\Delta\sigma)$ is calculated using spherical trigonometry given [14]

$$\cos(\sigma) = \sin(\phi_s)\sin(\phi_c) + \cos(\phi_s)\cos(\phi_c)\cos(\lambda_c - \lambda_s) \tag{4.4}$$

This is computed in a vectorised and JIT-compiled manner.

Having now the mask of the cells that are within the visible area, the cells can be populated with users. The cell's positions and the visibility mask are looped over and the cells are populated. Firstly the number of users that are within a cell is determined by draw of jax.random.choice() that chooses between either the high-density cell type or the low-density cell type. The high and low-density amount of probability is set in parameter specification.

For each user, a named tuple is generated, which is added to a list that is put into the cell tuple. For each user, a choice is made whether the user is on or off, as is specified in the Chapter 3.

4.1.4 Calculate Rates

The calculation of achievable data rates per cell is essential for evaluating the capacity constraints in the optimisation framework. These rates depend primarily on the satellite–cell link distance, which affects the path loss and consequently the signal-to-noise ratio (SNR). The other parameters are constant values, such as gains and power transmitted, that are fixed across cells.

Cell centres are transformed from geodetic (topographic) coordinates to Cartesian coordinates using the following equations [14]:

$$x = R_E \cdot \cos(\lambda) \cdot \cos(\phi)$$
$$y = R_E \cdot \cos(\phi) \cdot \sin(\lambda)$$
$$z = R_E \cdot \sin(\phi)$$

where R_E is the Earth radius plus the satellite altitude, λ is the longitude, and ϕ is the latitude. This transformation facilitates straightforward computation of Euclidean distances, required for determining the free-space path loss (FSPL).

For simplicity, only free-space path loss (FSPL), is considered in this implementation. Other channel impairments such as atmospheric gas absorption, scintillation, or log-normal shadowing are not included. This choice is to focus on the structural effects of beam allocation and demand prediction, rather than precise channel modelling. However, the implementation remains modular, allowing for future inclusion of additional loss components if required.

The SNR computation and the Shannon capacity are then computed using the expressions in Section 3.4. The distance is computed using the JIT-Compiled functions and the SNR computation, and Shannon capacity is computed using vectorization of pure Functions to speed up computation.

4.1.5 Compute Predictive Statistics

The predictive statistics need to be computed before the optimisation step. The predictive statistics would involve the calculation of the binomial convolution to estimate the PMF. The number of users in each binomial distribution, i.e. the number of people that are either on or off at the initial draw, is important. If this number is very large the convolution could become infeasible, however, if the number of users in each cell is small, it can be computed by standard convolution, using the jnp.convolve and then normalizing it. If the number of users was larger, it may be beneficial to perform FFT-Convolution instead [34].

Some care has to be taken when implementing the computation of the PMF in regards to being compatible with XLA compilation. The number of total users in the cell is the same, however, the number of users that are on or off is subject to changes.

Given the computed PMF, the predictive values can be readily computed for each cell and then the two predicted values can be used in the following optimisation.

To make computation more efficient, the transition probabilities after o steps can be computed once for all the steps until O is reached. This can be done as the transition probabilities for all users are the same. The changing value is the total number of users in the cell.

4.1.6 Optimise

The optimisation step is implemented using the Gurobi optimiser interface in Python. The Gurobi optimiser is chosen as it has good performance for MILP problems and is available

through the university's academic license.

The Gurobi interface functions by the definition of optimisation variables, setting constraints by using conditional inequalities and setting an objective. Optimisation of MILP can be difficult, so the specification of parameters is crucial. To speed up computation, the threads parameter can be set to 0 to use all available threads. The MIPGap is a useful parameter to stop the optimisation early. The MIPGap describes the objective bound and the objective value achieved [35].

$$\frac{B - V}{V} \tag{4.5}$$

The bound, B, is an estimate of the upper bound of performance, and V is the value of the optimisation. The MIP gap decreases as the solution approaches the bound. The boundary and value evolve over time, converging towards optimality, if solvable. The setting of a MIP gap sets a near-optimality proof boundary, as the MIP gap can still be large even if the optimal solution is found.

4.1.7   Store

Care must be taken to ensure that the data is not stored in a manner that is excessively time-consuming to load. Only ndarrays and the string labels are stored. The cell state is troublesome to store due to it being a named tuple and a recursive definition, and so it is not stored in its entirety. Only the number of active users and the maximum number of users are stored. The data is stored in chunks for each iteration and prediction type for simplicity in the evaluation process.

4.1.8   Propagate Demand

The demand is propagated before evaluation, as it does not need to be computed before the optimisation. The optimisation is performed independently as the probability distributions used are based on the initial distribution.

This means that for each iteration of the optimisation, the demand can be propagated to evaluate the performance over independent realizations of the user behaviour, $N_{realisations}$.

The propagation of the demand is done with a vectorised scanning function to implement the user transitions in an efficient manner. The scanning function is over the state transitions of O, and then this is vectorised. This is imperative as the number of users and the number of cells is large. The number of users active for each cell is saved in the storage step.

4.2   Metrics

Some metrics need to be defined to evaluate the implementation. The implementation of the metrics themselves is also important, as some of the metrics need to be implemented in a way that requires less computation. As the optimisation problem of MILP already will take a significant time to compute, it is desired for the surrounding functions to take less time, as there is time to gain. The results of the optimisation are also part of the results.

4.2.1   Disconnection time

To describe the downtime of the connection, the disconnect time is computed by summing up the amount of time slots between each illumination time within a single cell. This is to see how

often the cell is illuminated, and also in some ways to ascertain resilience. If the cell is never illuminated, the device in that area will not be able to establish a satellite connection.

The disconnect time can be computed as the number of consecutive slots where no beam is assigned to the slot, $x[i,k] = 0$, until the cell is illuminated $x[i,k] = 1$, for each cell. If a cell is never illuminated within the scheduling period, the disconnect time is equal to the entire configuration period, T_r, which indicates a complete lack of coverage.

This can readily be vectorized, as the disconnect time can be computed for cells independently and then collected. The disconnect time metric function has a schedule with the shape defined in (S, K).

4.2.2 Capacity to Demand Ratio

The capacity over the demand ratio is used as a metric as it is an indicator of how well demand is met. This can show a metric of the underallocation of the solution to the optimisation problem.

The allocation method is based on optimising for average performance across the entire period T_R. This makes an underallocation of demand for instances during the period unsuitable, as it is not indicative of the objective of the optimisation or takes account of the periods in which the cell is allocated and has a much higher rate than the instantaneous demand at that point in time.

The demand will transition to a variety of states O times over T_R instead of using the value of the demand for each time period the dynamic demand is averaged, $\tilde{D}$, across the period. The C/D metric used in the results section is the ratio between the achieved capacity given the schedule and the average propagated demand over the period $\tilde{D}$ realized $N_{realisations}$ times.

4.2.3 Naive Solution

To serve as a baseline, against the demand-meeting optimisation problem a naive solution is used. The naive solution is one in which the beams are split evenly across all cells regardless of demands. The solution can be defined as the share of time a single cell will be illuminated given by the ratio,

$$f = \frac{N_b}{T_s \cdot K} \tag{4.6}$$

This number is multiplied by the Shannon capacity of the cells themselves to provide the achieved capacity and average across the period T_r. The achieved capacity is then used to compute the C/D to evaluate the performance of performing the optimisation and the need for considering the demands.

4.3 Parameters for evaluation

The parameters used for evaluation are in the following table Table 4.1.

Parameter	Value	Unit/Note
T_r	1	s (Reconfiguration period)
T_s	0.001	s (Time step)
O	{2, 5}	Number of transitions (demand steps)
R_{min}	{0, 100, 200}	kbit/s
D_{on}	100	kbit/s
α	{0.7, 0.07}	ON transition probability
β	{0.3, 0.03}	OFF transition probability
p_{high}	0.2	High density prob.
p_{low}	0.8	Low density prob.
M_{high}	400	Users per high density cell
M_{low}	80	Users per low density cell
N_b	19	Number of beams
h	600	km (Satellite altitude)
L_{cell}	50	km (Cell size)
P_{tx}	75.4	W (Per beam)
G_{tx}	30	dBi
G_{rx}	0	dBi
T	290	K
B	30	MHz
f_c	2	GHz
θ_{min}	40	deg (Min obs. angle)
$start_{ON}$	0.6	Initial ON-state prob.
N_{iter}	15	Iterations per sweep
$N_{realisations}$	100	Number of demand realisations
MIPGap	1	%(Objective value bound gap)
R_E	6378	km

Table 4.1: Parameters used in the system evaluation. Braces indicate that multiple values are swept in the parameter study. See text for details of permutation and dependencies.

The allocation period T_r is chosen as 1 s, to both limit the amount of time slots and also to correspond with the time in which the edge cells are likely to be within the area. The slot duration is chosen to be 1 ms this is to satisfy the requirement that the number of time slots should be sufficient to achieve a non-zero solution to the optimisation problem. The D_{on} value is from a SatNEx V presentation [36].

The transition probabilities are chosen to be representatives of a highly variable case and one that has more inertia. The values are scaled to have the same steady-state probabilities as seen in

$$\frac{\alpha}{\alpha + \beta} = \frac{0.7}{0.3 + 0.7} = \frac{0.07}{0.03 + 0.07} \tag{4.7}$$

The combination of steady-states and the table parameters responsible for the amount of users in each cell is chosen iteratively based on having an initial demand that is slightly above one before the variable user channel is introduced.

The gain of the transmitter and receiver are chosen in accordance with 3GPP, [12]. The transmit power is chosen to be the same as in 3GPP [12]. The frequency range chosen is the lower frequency range of the NTN in revision 17, N256 downlink. The number of beams 19 is chosen as in [12]. The height of the satellite corresponds to a LEO satellite in [12].

The R_{min}, O, α and β are swept N_{iter} times in each permutation. It is to be noted that the α and β are bound together, so there is no combination of $\alpha = 0.7$ and $\beta = 0.03$ as this would change the steady-state of the Markov chain.

5 | Results

The results of the optimisation have produced in 450 different optimised schedules given the different prediction methods, R_{min}, iterations, and O. his produces a large set of allocation schedules and makes the computation of the variable demand across all cells excessively time-consuming. This means that the cell closest to the nadir and the cell furthest from the nadir are evaluated.

The average result of the optimisation relative to the initial demand is in Figure 5.1.
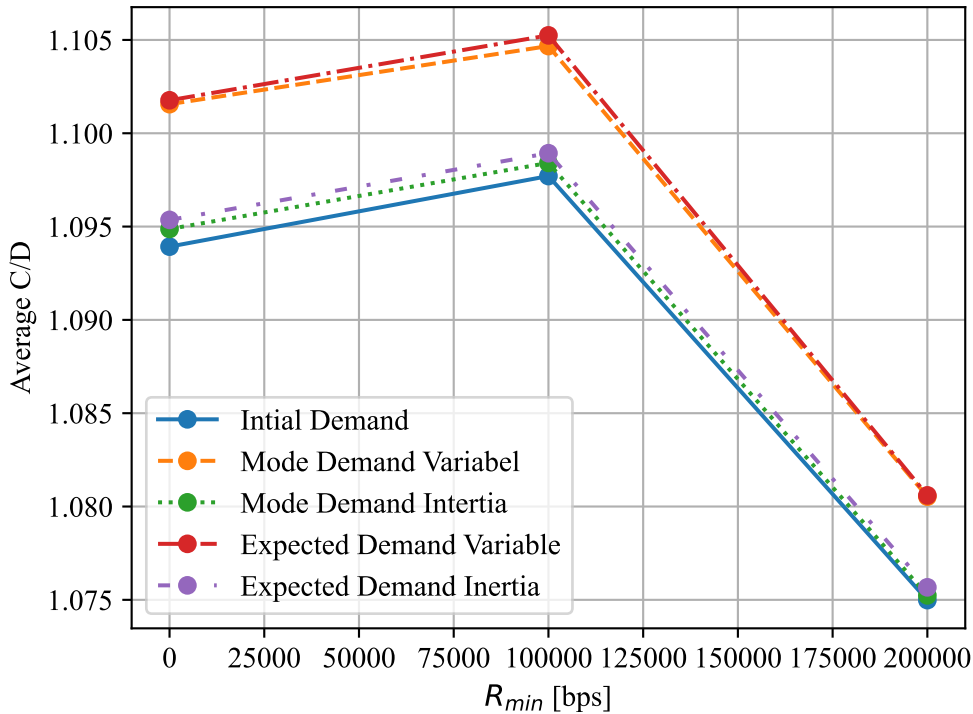

Figure 5.1: Average results of the optimisation.

The C/D compared with optimising with respect to the initial demand is larger than the others. This can be expected as the predictions are in most cases larger than the initial prediction. There is a fall-off when the R_{min} is increased beyond the $100 \frac{\text{kbit}}{\text{s}}$. In the following, only results corresponding to $R_{min} = 0$ are shown in the subsequent figures and tables, unless otherwise stated. The trends observed for other values of R_{min} were found to be similar and do not provide additional insight. The plots can be found in the source code, Chapter A, under the folder **visualizers/rmin200000/** and **visualizers//rmin100000/**.

The disconnect time metric of the resulting optimised schedule is in Figure 5.2.
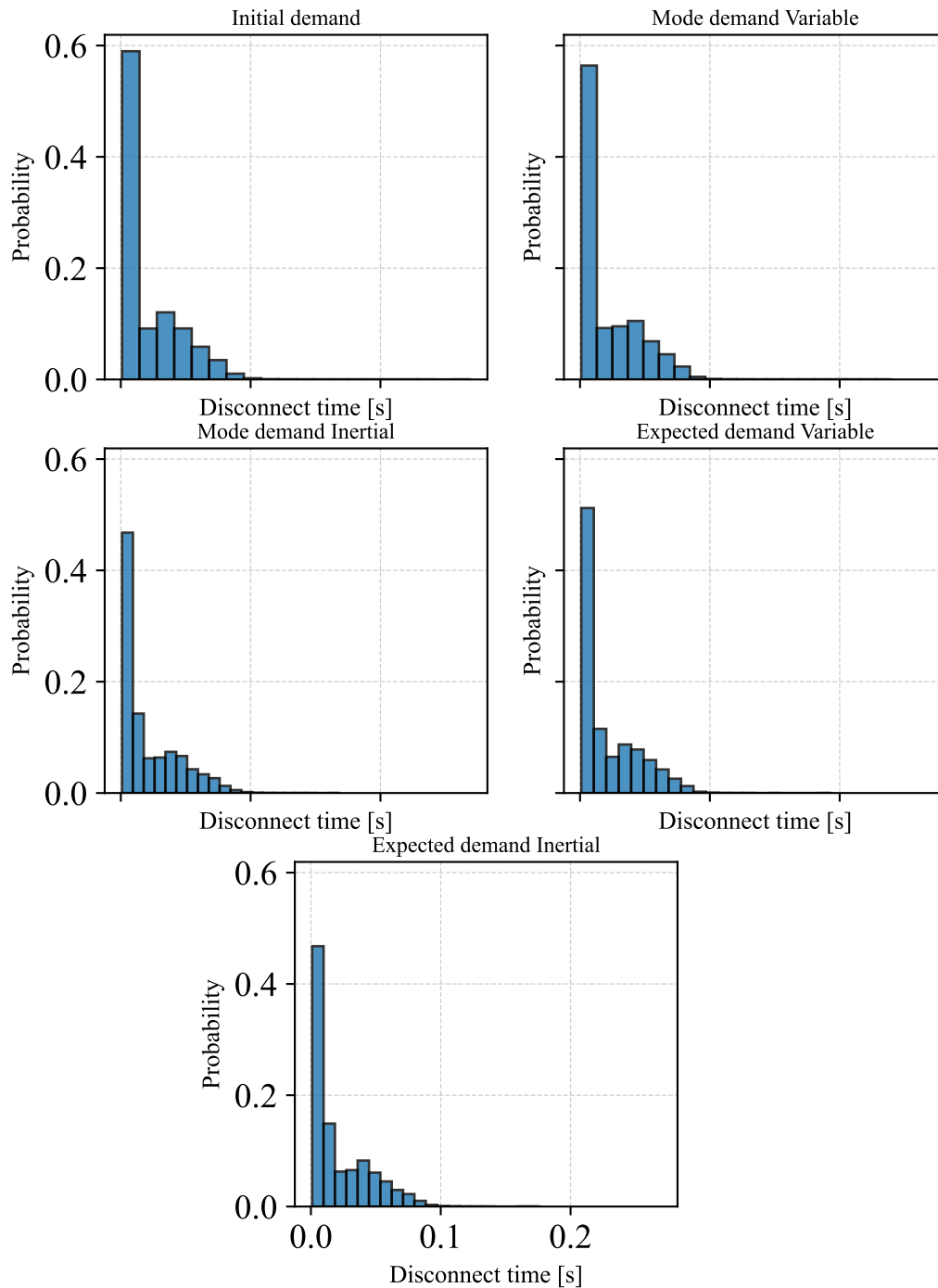
Figure 5.2: Plot of the disconnect time across the five different configurations.

All cells are connected at some point, as indicated by the axis of the figure. In the optimisation results it can be seen that for 0.6% of the cases, the disconnect time is only 2.5 ms. This shows lower disconnect time which is indicative of lower jitter across the period.

The distribution of the naive solution across the edge cell and nadir cell can be seen in the Figure 5.3 and its associated CDF in Figure 5.4.
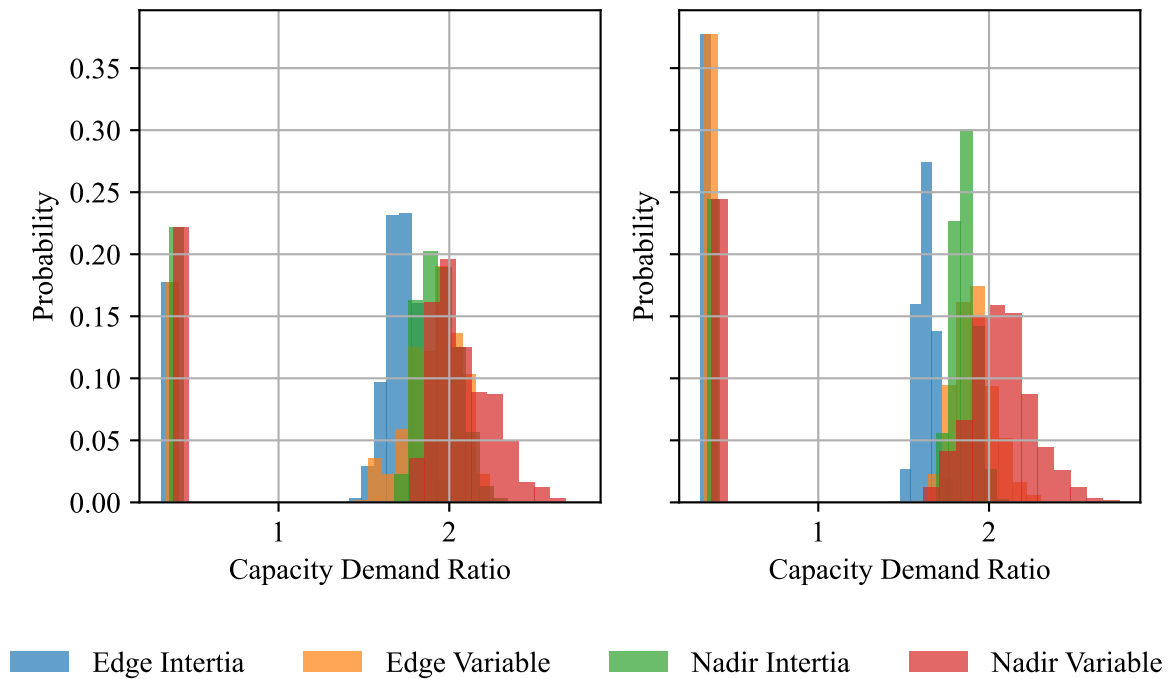
Figure 5.3: Histogram of the resulting C/D and their probabilities in the naive solution, the left and right plots are respectively variable and inertia case.

Figure 5.4: CDF of the C/D in the naive case, the left and right plots are respectively variable and inertia case.

The distribution can be seen to be bimodal, split between meeting and not meeting the demand of the cell. This is due to the fact that the cell is either in a low-density or high-density distribution.

This is apparent in the CDF, Figure 5.4, as the CDF plateaus around the probability of a cell being initialised with that density.

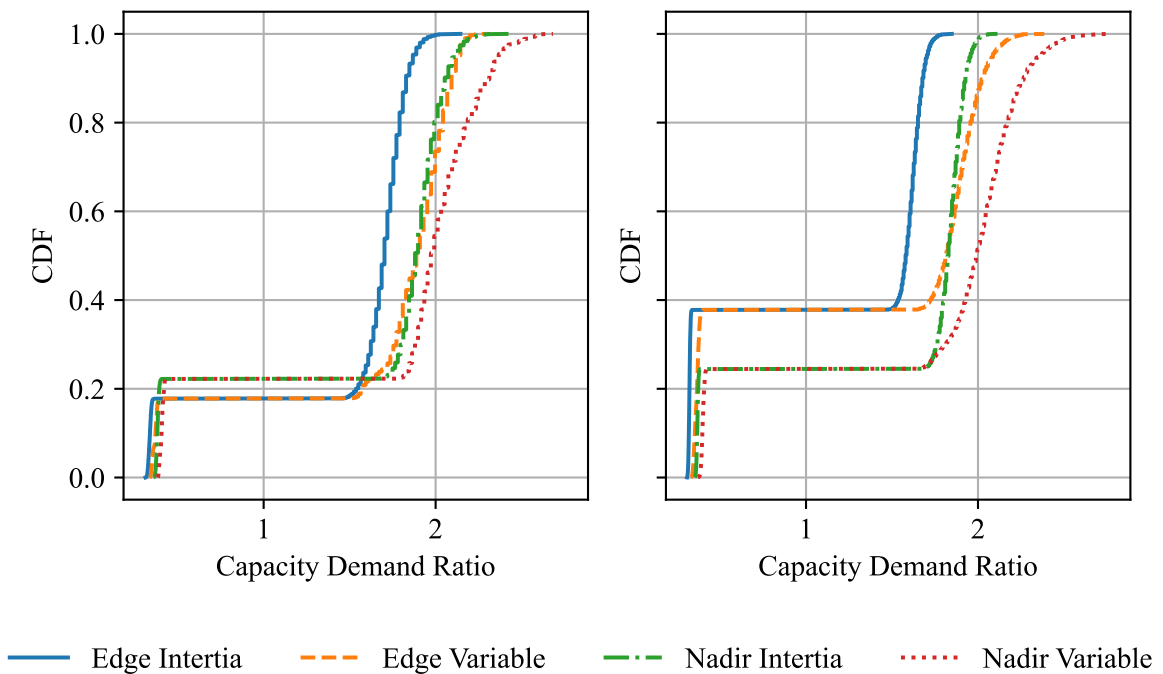The naive case is combined with the solutions of rmin $= 0$ and $O = 2$, in Figure 5.5 and Figure 5.6.

Figure 5.5: CDF of different predictions and the naive solution in the nadir cell, the left and right plots are respectively variable and inertia case.

Figure 5.6: CDF of different predictions and the naive solution in the edge cell, the left and right plots are respectively variable and inertia case.

When comparing the naive case to the C/D is closer to one and has no bimodal split but rather rises steadily. This means that regardless of the initialisation of a high or a low-density cell, the optimisation process brings the solution closer to C/D than the naive case.

When comparing the Figure 5.6 and Figure 5.5 there is no significant difference when comparing either the nadir cell or the edge cell performance when using the optimisation algorithm. However, there is a difference when comparing the inertia case and the variable case. The variable case performs worse than the inertia case in terms of having a higher C/D ratio. To highlight this the results are plotted without the naive case in Figure 5.9 and Figure 5.10.
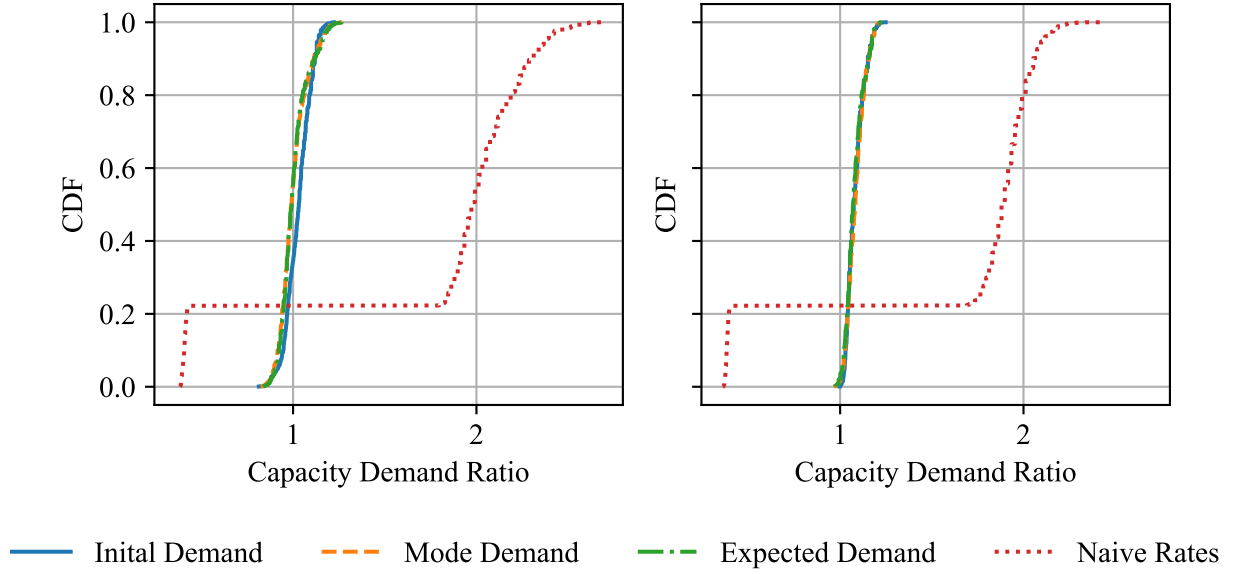

Figure 5.7: CDF of different predictions in the nadir cell, the left and right plots are respectively variable and inertia case.
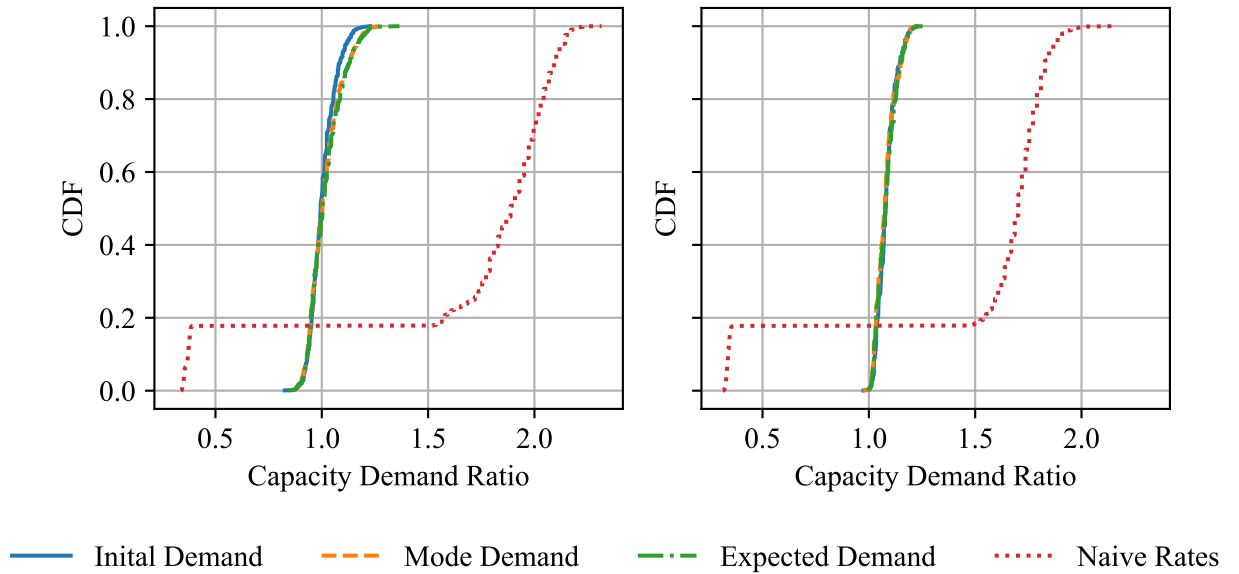

Figure 5.8: CDF of different predictions in the edge cell, the left and right plots are respectively variable and inertia case.

The variable case performs significantly worse than the inertia case. This is presented in Table 5.1 and Table 5.2.

Table 5.1: CDF at C/D = 1 for Nadir Case

Type	Variable CDF at C/D = 1	Interti CDF at C/D = 1
Initial Demand	0.3447	0.0013
Mode Demand	0.5507	0.013
Expected Demand	0.5553	0.0207

Table 5.2: CDF at C/D = 1 for Edge Case

Type	Variable CDF at $CD = 1$	Interti CDF at C/D = 1
Initial Demand	0.5307	0.0033
Mode Demand	0.4853	0.0047
Expected Demand	0.4693	0.0013

In the variable case for the nadir cell, the predictive measures do not perform better; the initial demand prediction actually performs better when $O = 2$. Next, the $O = 5$ case is presented, in which the demand changes more often. Here, the naive case is omitted, as it provides the same results as in the $O = 2$ case, where there is a bimodal split.
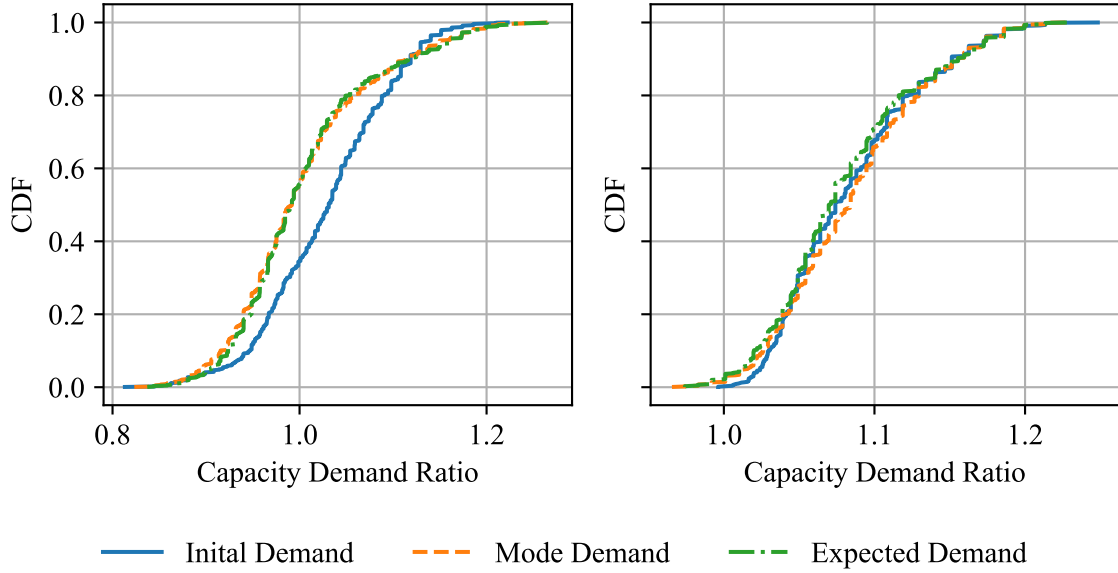
Figure 5.9: CDF of different predictions in the nadir cell, the left and right plots are respectively variable and inertia case.

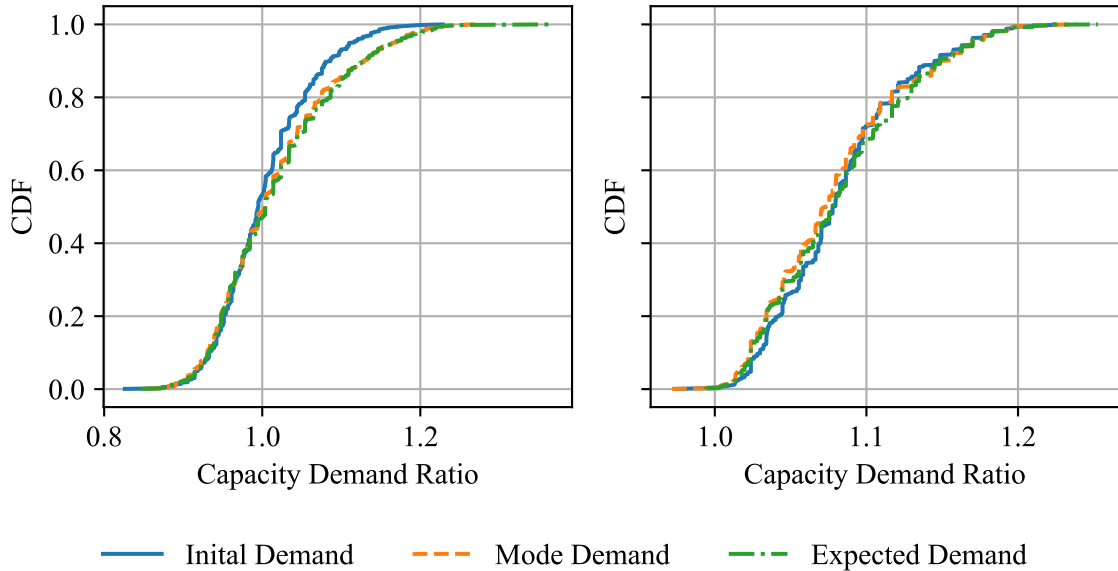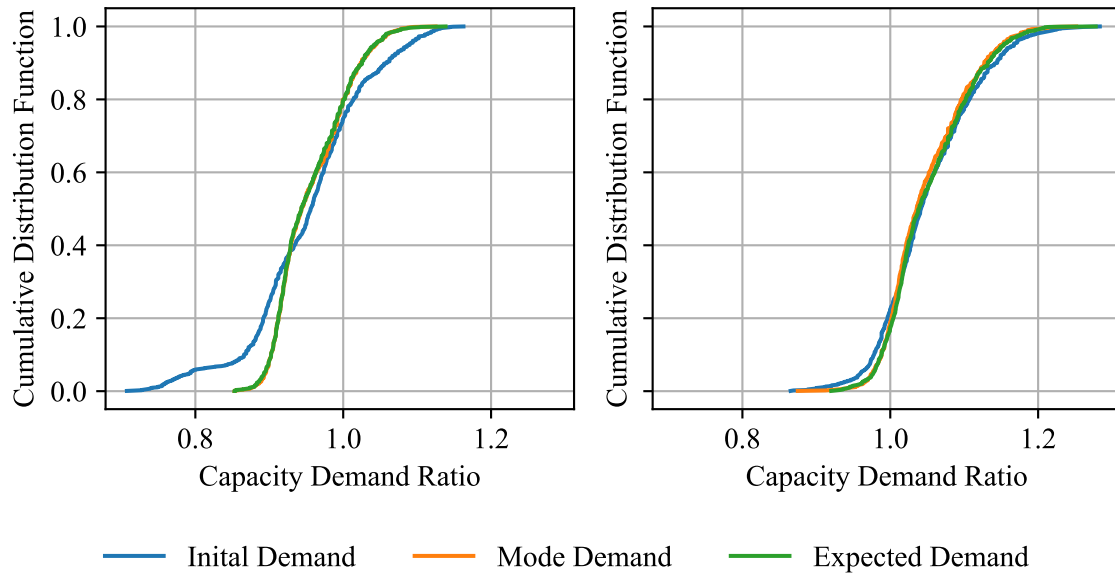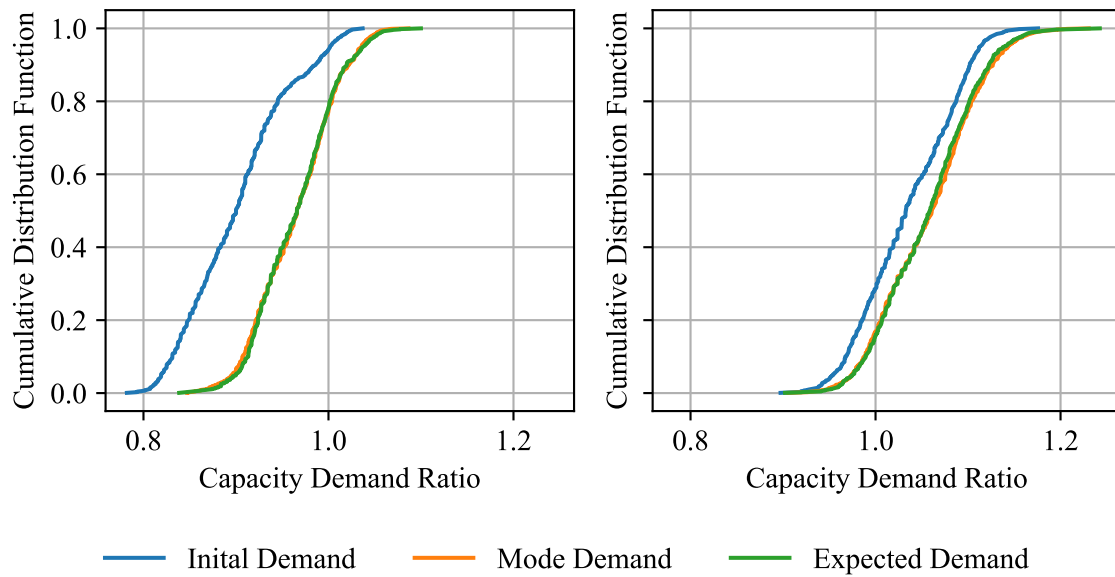Figure 5.10: CDF of different predictions in the edge cell with $O = 5$, the left and right plots are respectively variable and inertia case.

The CDF value when the C/D = 1 is higher than in the case with O = 2, as can be seen in the results Table 5.3 and Table 5.4.

Table 5.3: CDF at C/D = 1 for Nadir Case

Type	Variable CDF at C/D = 1	Inertial CDF at C/D = 1
Initial Demand	0.748	0.224
Mode Demand	0.800	0.187
Expected Demand	0.798	0.173

Table 5.4: CDF at C/D = 1 for Edge Case

Type	Variable CDF at C/D = 1	Inertial CDF at C/D = 1
Initial Demand	0.941	0.286
Mode Demand	0.774	0.168
Expected Demand	0.782	0.156

In the case of having a prediction, the edge cell case showcases a clearly improved performance compared to simply using the initial demand.

To summarise using the optimisation problem that optimises to meet demand results in low disconnect time regardless of it not being part of the optimisation problem. This signifies that the system is more resilient in terms of discovering new connection. The BH demand meeting optimisation is shown to be beneficial in the case of non-uniform dynamic demand compared to a fixed even time split. The effects of the variability can also be seen to be worsening the performance as CDFs indicate. The benefit of having a predictive measure is limited as can be seen from the figures, the only case that shows a clear benefit is the edge case with with $0 = 5$ steps. The difference between the performance of the Mode Demand prediction and Expected Demand prediction is small.

These results are compiled based on a model of user behaviour in which the satellite has total knowledge of the user behaviour probabilities. Modelling of the user behaviour itself more complex with users that may behave differently. This will be further discussed in the next chapter.

6 | Discussion

This chapter discusses the main findings from the previous sections and considers the broader context and limitations of the presented work. The evaluation of the demand-aware beam hopping framework relied on certain modelling choices and simplifying assumptions, particularly regarding user behaviour and system knowledge.

The discussion also addresses aspects such as computational practicality, potential directions for improving user modelling and prediction, and the relevance of these results for more realistic satellite communication systems. Finally, suggestions for possible system-level enhancements and future research directions are outlined.

6.1 Results

The results from the previous section indicates that changing the R_{min} has an effect on the C/D as seen in Figure 5.1. However, with chosen statistics R_{min} is unlikely to have any role, as the number of users is likely always be above 2. The reason for the number 2, is given the $R_{min} = 200 \frac{\text{kbit}}{\text{s}}$. As the results of the optimisation and the parameters going into it are stored the file and the D_k can be seen in the cases and plotted, as seen in Figure 6.1.
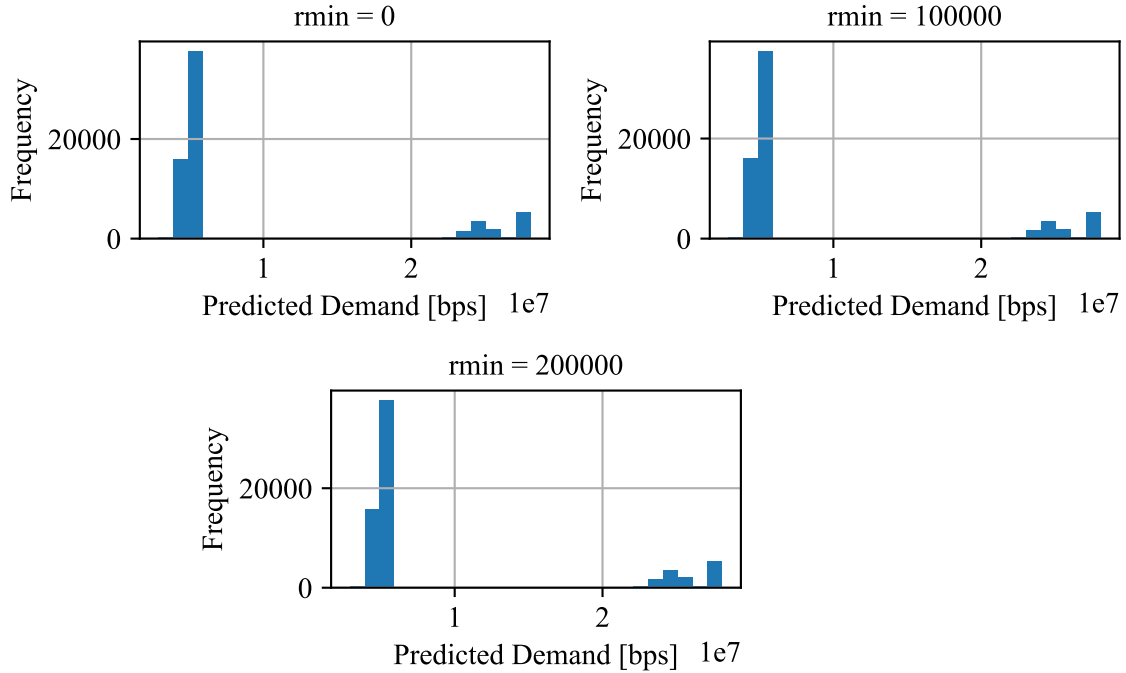
Figure 6.1: Three histograms of the predicted demand that is used for the optimisation problem.

There is not a single case of the D_k being low enough for the R_{min} to have had an effect, yet in the plot there is a difference that could indicate that there is something happening implementation wise that affects optimisation. This could be an effect of using the jax.numpy.maximum() call inside of the definition of the constraint when optimising the system. This highlights also the needs of further consideration of the underlying system parameters and the user model.

The C/D not being improved in most cases by utilising the predictive may be because the demand prediction is likely higher than the initial demand in all cases. This could lead to every cell essentially needing a higher demand and could possibly effect the BH allocation.

6.2 Digital Twinning

Digital twinning is methodology that is receiving more attention at the time of writing. Digital twin are a data based approach. Digital twin methodology is a technology bridging the gap between the physical data and the digital twin to enable informed or optimal choices based on the digital representation[37]. Digital twin may prove useful in the a real world case where the user behaviour is both uncertain and changes over time, and would aid in the assessment of uncertainty of the user model.

There are many attempts to make a unified methodology to define digital twins, one of them is the directional graph based Bayesian network definition seen in Figure 6.2.
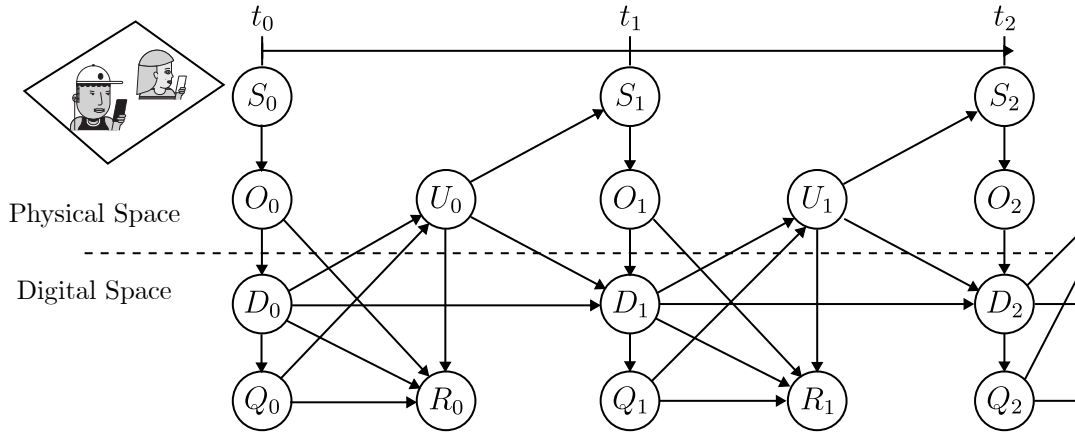

Figure 6.2: Directional digital twin graph [37].

The digital twin function by in the physical space some users are active and having certain network characteristics. The state S, this is then observed O before processing in to the digital space. The observation is converted into the digital state space, D. The digital state space is used to provide a control signal, U, that will affect the state in some manner for the next time step. Q is the quantities of interest that the operator may monitor and also affect the control signal U. R is the reward function of the digital twin system measures the performance, success or cost of the action, state or outcome [37].

The state could in the case of a beam hopping network be the users in the cell as seen in Figure 6.2. In this the user behaviour could be observed by their request for scheduling or their requests for resource allocation, to control the allocation of the demand. The different functions for each node can be at different times contain different functions. Initially the system before the deployment of the satellite could be set to some predicted value rather than the ones obtained during operations. Post deployment the digital twin would then be based on the operations of the satellite.

Digital twin of the user behaviour of the state could be based on hidden Markov models (HMM) [38]. A HMM approaches Markov chain by divorcing the Markov state from the observation of the state [38]. the In context of the thesis the transition probabilities of all users has been

known. This would be smaller transition as the model is already based on Markov models. A HMM approach would work in the digital twin paradigm as it accounts for the difference between physical reality, the hidden state, and the digital twin, the best estimate based on observations. Standard HMM assumes stationary probabilities of the hidden state transition, but there exist extensions that account so that model can account for seasonality of behaviour,[39], .i.e. night time day time difference in data demand. [40] extends the HMM from single Markov process to multiple coupled Markov process.

An alternate method of developing digital twins is with the use of Neural Network Models. Neural network can effectively model non-linear behaviour and predict future states based on historical or modelled data [41]. This would require possibly more complex models or real world data that could be applied to train the model [41].

6.3 User Model Behaviour and Cell initialisation

The cell initialisation is only split between the two cases of high and low density which is somewhat limiting when assessing diversity in user behaviour. This limitation becomes evident in the naive case presented in the results section. Due to the simplified distinction, the number of active users across cells that are of the same density is very similar. This means that the allocation algorithm in effect allocates a similar amount to each cell. Increasing the diversity may make the system more interesting to evaluate as this would also increase the benefit of predictions, if similar user behaviour model and parameters are used.

The current Markovian approach, relying solely on the immediate past state, does not capture time-dependent variations such as daily patterns or long-term correlations. The Markovian approach was in this case chosen for simplicity and computational efficiency. Including temporal diversity would significantly change the the demand dynamics, resulting in a change in the scope and effectiveness of prediction methods. As it could make a PBD unfitting of the underlying distribution. Including long-term correlation into the user behaviour through higher-order Markov process would allow for implementation of possibly more realistic models. Investigating Quasi-Markov Processes (QMP) is another promising direction, as QMPs allow non-stationary transition probabilities [42]. Additionally, exploring digital twin approaches that model time-varying probability distributions could further enhance prediction quality and system responsiveness.

6.4 Computational Speed

The time required for a single optimisation in this thesis exceeds the allocation period T_R itself. The resulting beam hopping allocation approach impractical to implement in real-time or near-real-time deployment scenarios. While reducing the computational complexity was not the focus of the thesis, it remains a challenge for implementing beam hopping strategies in operational satellite networks. [43] shows an improved method using continuos relaxation of binary variables coupled with a particle swarm optimisation (PSO) algorithm significantly improves computational speed. [22] proposes a mixed learning and optimisation method, that searches for a subset of solution before and optimisation routing determines the final schedule. The combination of learning and optimisation lessens the computational time dramatically.

Adopting approaches such as these that can lessen the computational time could significantly enhance scalability and practicality of demand matching BH.

6.5 System-Level Extensions

While this study focused on the average throughput to demand ratio, real communication systems operate on a frame or packet level. The recent study [43] proposes a method to jointly optimise the power and beams at a frame level, achieving improvements in power usage and capacity-demand matching. Incorporating frame-level scheduling beyond allocation of the illumination plan could provide a more nuanced of delay, jitter and resource contention, and would go beyond the feasibility of meeting some throughput demand.

The addition of a frame-level approach also opens up an expansion into considering not just the downlink, but also the uplink. The inclusion of the uplink opens up for more challenges regarding scheduling and feedback mechanisms. These mechanisms introduce additional protocol overhead, making the joint uplink-downlink optimisation both more complex and realistic.

7 | Conclusion

This work set out to investigate the developing field of beam hopping method of allocation in satellite communication. It did this in a context of providing throughput to a user base that is non uniform and dynamic in time.

Firstly a presentation of cellular connection theory is presented, moving from the terrestrial case towards the non-terrestrial case, and some of the benefits satellite communication can bring. This presentation is tied to the development in antenna systems that leads to the presentation of the what a beam hopping system is. BH benefits, current implementation of BH and studies into BH are presented. The limited satellite resources and QoS are briefly covered, as well.

A system model was developed using a Markov-based ON/OFF user model. The beam hopping allocation problem was formulated to maximise the minimum capacity-to-demand ratio across cells, and probabilistic demand estimation using the Poisson Binomial Distribution was introduced, based on the expected value across time and the average mode across time.

The system was implemented in a simulation framework, utilising JAX for effective computation of system parameters and Gurobi optimisation tools to evaluate performance across a range of user dynamics and allocation strategies based on the predictive measures.

The results show that the capacity to demand optimisation based on the initial value and the predictive measure performs better than the naive solution in which the beams are scheduled evenly in time. It also shows that the system has a low disconnect time relative to the optimisation period.

The results show that, except for specific scenarios with highly variable user behaviour and frequent transitions, predictive optimisation offers limited measurable benefits over simply using the initial demand. Notably, improved performance was observed only in edge cells with frequent state transitions.

These findings highlight the need of scenario-specific analysis when considering advanced predictive allocation methods. The thesis also discussed the limitations of the chosen user models and the computational demands of the optimisation approach. Future work should focus on incorporating more complex user dynamics, exploring the role of advanced prediction algorithms such as digital twins, expanding to higher levels of communication such as packet and frames, and improving computational efficiency to better inform the practical application of beam hoping in satellite networks.

Bibliography

[1] *Wireless Connectivity: An Intuitive and Fundamental Guide | Wiley eBooks | IEEE Xplore.* [Online]. Available: https://ieeexplore.ieee.org/book/9116608 (visited on 06/03/2025).

[2] A. Smith, *Space Connect: The rise of LEO satellite constellations*, en-US, Feb. 2025. [Online]. Available: https://www.itu.int/hub/2025/02/space-connect-the-rise-of-leo-satellite-constellations/ (visited on 05/15/2025).

[3] *Search OSOidx.* [Online]. Available: https://www.unoosa.org/oosa/osoindex/search-ng.jspx (visited on 05/15/2025).

[4] *What are the differences between LEO and GEO satellites? '23*, en-GB, Section: LEO Connectivity. [Online]. Available: https://ast-networks.com/insights/leo-connectivity/leo-and-geo-satellites/ (visited on 05/15/2025).

[5] A. Guidotti, A. Vanelli-Coralli, V. Schena, *et al.*, "The path to 5G-Advanced and 6G Non-Terrestrial Network systems," in *2022 11th Advanced Satellite Multimedia Systems Conference and the 17th Signal Processing for Space Communications Workshop (ASMS/SPSC)*, ISSN: 2326-5949, Sep. 2022, pp. 1–8. DOI: 10.1109/ASMS/SPSC55670.2022.9914764. [Online]. Available: https://ieeexplore.ieee.org/document/9914764 (visited on 05/15/2025).

[6] *(PDF) Beam Hopping in Multi-Beam Broadband Satellite Systems: System Performance and Payload Architecture Analysis.* [Online]. Available: https://www.researchgate.net/publication/236334523_Beam_Hopping_in_Multi-Beam_Broadband_Satellite_Systems_System_Performance_and_Payload_Architecture_Analysis (visited on 05/15/2025).

[7] L. Chen, Vu Nguyen Ha, L. Wu, S. Chatzinotas, and B. Ottersten, "The Next Generation of Beam Hopping Satellite Systems: Dynamic Beam Illumination With Selective Precoding," *IEEE Transactions on Wireless Communications*, vol. 22, no. 4, pp. 2666–2682, Apr. 2023, ISSN: 1558-2248. DOI: 10.1109/TWC.2022.3213418. [Online]. Available: https://ieeexplore.ieee.org/document/9923617 (visited on 04/17/2025).

[8] L. Chen, L. Wu, E. Lagunas, *et al.*, "Joint Power Allocation and Beam Scheduling in Beam-Hopping Satellites: A Two-Stage Framework With a Probabilistic Perspective," *IEEE Transactions on Wireless Communications*, vol. 23, no. 10, pp. 14 685–14 701, Oct. 2024, ISSN: 1558-2248. DOI: 10.1109/TWC.2024.3417707. [Online]. Available: https://ieeexplore.ieee.org/document/10582841 (visited on 04/17/2025).

[9] C. Rohde, R. Wansch, G. Mocker, *et al.*, "Beam-hopping over-the-air tests using DVB-S2X super-framing," in *36th International Communications Satellite Systems Conference (ICSSC 2018)*, Oct. 2018, pp. 1–7. DOI: 10.1049/cp.2018.1711. [Online]. Available: https://ieeexplore.ieee.org/document/8742519 (visited on 05/15/2025).

[10] Z. Lin, Z. Ni, L. Kuang, C. Jiang, and Z. Huang, "Dynamic Beam Pattern and Bandwidth Allocation Based on Multi-Agent Deep Reinforcement Learning for Beam Hopping Satellite Systems," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 4, pp. 3917–3930, Apr. 2022, ISSN: 1939-9359. DOI: 10.1109/TVT.2022.3145848. [Online]. Available: https://ieeexplore.ieee.org/document/9693289 (visited on 05/15/2025).

[11] E. Dahlman, S. Parkvall, and J. Skold, *5G NR: The Next Generation Wireless Access Technology*, 1st. USA: Academic Press, Inc., 2018, ISBN: 0128143231.

[12] *Specification # 38.321*. [Online]. Available: https://portal.3gpp.org/desktopmodules /Specifications/SpecificationDetails.aspx?specificationId=3194 (visited on 05/30/2025).

[13] V. H. M. Donald, "Advanced mobile phone service: The cellular concept," *The Bell System Technical Journal*, vol. 58, no. 1, pp. 15–41, 1979. DOI: 10.1002/j.1538-7305.1979.tb02209.x.

[14] H. D. Curtis, *Orbital Mechanics for Engineering Students*. Elsevier Butterworth-Heinemann, 2005.

[15] *Starlink Satellite Constellation - eoPortal*, en. [Online]. Available: https://www.eoportal.org/satellite-missions/starlink (visited on 06/03/2025).

[16] C. A. Balanis, *Antenna Theory: Analysis and Design*. USA: Wiley-Interscience, 2005, ISBN: 0471714623.

[17] G. Maral and M. Bousquet, *Satellite communications systems. Systems, techniques and technology*. Jan. 2009, Journal Abbreviation: Wiley Series in Communication and Distributed Systems, Chichester, New York: Wiley, 1993, 2nd ed. Publication Title: Wiley Series in Communication and Distributed Systems, Chichester, New York: Wiley, 1993, 2nd ed., ISBN: 978-0-470-83498-5. DOI: 10.1002/9780470834985.

[18] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 501–513, 2016. DOI: 10.1109/JSTSP.2016.2520912.

[19] P. Angeletti, F. Prim, and R. Rinaldo, "Beam hopping in multi-beam broadband satellite systems: System performance and payload architecture analysis," vol. 1, Jun. 2006. DOI: 10.2514/6.2006-5376.

[20] 3GPP, "3rd generation partnership project; technical specification group radio access network; solutions for nr to support non-terrestrial networks (ntn) (release 16)," 3GPP, Tech. Rep., 2023.

[21] *(PDF) Beam Hopping in Multi-Beam Broadband Satellite Systems: System Performance and Payload Architecture Analysis*. [Online]. Available: https://www.researchgate.net /publication/236334523_Beam_Hopping_in_Multi-Beam_Broadband_Satellite_Syste ms_System_Performance_and_Payload_Architecture_Analysis (visited on 05/15/2025).

[22] L. Lei, E. Lagunas, Y. Yuan, M. G. Kibria, S. Chatzinotas, and B. Ottersten, "Beam Illumination Pattern Design in Satellite Networks: Learning and Optimization for Efficient Beam Hopping," *IEEE Access*, vol. 8, pp. 136 655–136 667, 2020, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2020.3011746. [Online]. Available: https://ieeexplore.ieee.org/document/9146871 (visited on 04/21/2025).

[23] 3GPP, "3rd generation partnership project; technical specification group radio access network; nr; user equipment (ue) radio transmission and reception; part 5: Satellite access radio frequency (rf) and performance requirements (release 18)," 3rd Generation Partnership Project (3GPP), Tech. Rep., Dec. 2024.

[24] 3GPP, "3rd generation partnership project; technical specification group radio access network; study on new radio (nr) to support non-terrestrial networks (release 15)," 3GPP, Tech. Rep., 2020.

[25] S. Dahmen-Lhuissier, *5G*, en-gb. [Online]. Available: https://www.etsi.org/technologies/5g?tmpl=component&utm_source=chatgpt.com (visited on 06/02/2025).

[26] X. Foukas, G. Patounas, A. Elmokashfi, and M. K. Marina, "Network Slicing in 5G: Survey and Challenges," *IEEE Communications Magazine*, vol. 55, no. 5, pp. 94–100, May

2017, ISSN: 1558-1896. DOI: 10.1109/MCOM.2017.1600951. [Online]. Available:
https://ieeexplore.ieee.org/document/7926923 (visited on 06/04/2025).

[27] S. M. Ross, *Introduction to Probability Models*, 12th. Academic Press, 2019.

[28] S. M. Kay, "Markov Chains," en, in *Intuitive Probability and Random Processes Using MATLAB®*, S. M. Kay, Ed., Boston, MA: Springer US, 2006, pp. 739–775, ISBN: 978-0-387-24158-6. DOI: 10.1007/0-387-24158-2_22. [Online]. Available:
https://doi.org/10.1007/0-387-24158-2_22 (visited on 05/22/2025).

[29] A. Goldsmith, *Wireless Communications*, ISBN: 139781107385955. [Online]. Available: https://web.stanford.edu/class/ee359/doc/WirelessComm_Chp1-16_March32020.pdf (visited on 05/18/2025).

[30] William Feller, *An Introduction to Probability Theory and Its Applications, Volume 1, 3rd Edition | Wiley*, en. [Online]. Available:
https://www.wiley.com/en-us/An+Introduction+to+Probability+Theory+and+Its +Applications%2C+Volume+1%2C+3rd+Edition-p-9780471257080 (visited on 06/04/2025).

[31] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004. [Online]. Available:
https://web.stanford.edu/~boyd/cvxbook/.

[32] P. Bonami, M. Kilinç, and J. Linderoth, "Algorithms and Software for Convex Mixed Integer Nonlinear Programs," en, in *Mixed Integer Nonlinear Programming*, J. Lee and S. Leyffer, Eds., New York, NY: Springer, 2012, pp. 1–39, ISBN: 978-1-4614-1927-3. DOI: 10.1007/978-1-4614-1927-3_1.

[33] *User guides — JAX documentation*. [Online]. Available:
https://docs.jax.dev/en/latest/user_guides.html#user-guides (visited on 05/26/2025).

[34] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, 3rd. Pearson Education, 2010.

[35] *What is the MIPGap?* en-US, Nov. 2024. [Online]. Available: https://support.gurobi .com/hc/en-us/articles/8265539575953-What-is-the-MIPGap (visited on 05/28/2025).

[36] Value taken from satnexV Presentaion.

[37] M. G. Kapteyn, J. V. R. Pretorius, and K. E. Willcox, *A Probabilistic Graphical Model Foundation for Enabling Predictive Digital Twins at Scale*, arXiv:2012.05841 [cs], Apr. 2021. DOI: 10.48550/arXiv.2012.05841. [Online]. Available:
http://arxiv.org/abs/2012.05841 (visited on 02/11/2025).

[38] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Berlin, Heidelberg: Springer-Verlag, 2006, ISBN: 0387310738.

[39] O. Carr, *Oliver-carr/Time-Varying-Hidden-Markov-Model*, original-date: 2018-04-10T14:52:31Z, Jan. 2024. [Online]. Available:
https://github.com/oliver-carr/Time-Varying-Hidden-Markov-Model (visited on 06/03/2025).

[40] M. Brand, N. Oliver, and A. Pentland, "Coupled hidden Markov models for complex action recognition," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, ISSN: 1063-6919, Jun. 1997, pp. 994–999. DOI: 10.1109/CVPR.1997.609450. [Online]. Available:
https://ieeexplore.ieee.org/document/609450 (visited on 06/03/2025).

[41] T. Kreuzer, P. Papapetrou, and J. Zdravkovic, "Artificial intelligence in digital twins—A systematic literature review," *Data & Knowledge Engineering*, vol. 151, p. 102 304, May 2024, ISSN: 0169-023X. DOI: 10.1016/j.datak.2024.102304. [Online]. Available:

https://www.sciencedirect.com/science/article/pii/S0169023X24000284 (visited on 06/03/2025).

[42]   J. Janssen and R. Manca, "Applied Semi-Markov Processes," *Applied Semi-Markov Processes*, Jan. 2006, ISSN: 0-387-29547-X. DOI: 10.1007/0-387-29548-8.

[43]   L. Chen, A. Wang, E. Lagunas, L. Lei, and S. Chatzinotas, "Power Allocation and Beam Illumination Design for Time-Flexible Satellite Systems," in *2024 IEEE 29th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, ISSN: 2378-4873, Oct. 2024, pp. 01–06. DOI: 10.1109/CAMAD62243.2024.10942998. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10942998 (visited on 06/03/2025).

Appendices

A | Source Code

The source code for the project can be found using following the QR code Figure A.1,

Figure A.1: Link to the source code for project.

B | Overview of the symbols

An overview of the symbols used in the system is given in Table B.1.

Symbol	Description	Unit/Note
u	User state tuple (λ, ϕ, D)	–
λ	User/cell longitude	[deg]
ϕ	User/cell latitude	[deg]
D	User demand over T_R	[bit/s]
T_R	Configuration period	[s]
k	User activity state (on/off)	$k \in \{0, 1\}$
p_{on}, p_{off}	Probability user is on/off	–
α	Probability of transition off $\to$ on	–
β	Probability of transition on $\to$ off	–
A	State transition matrix (Markov)	–
O	Number of demand state changes per T_R	–
π	Steady-state probability vector	–
c_k	Cell state tuple $(\lambda_k, \phi_k, \mathbf{U})$	–
$\mathbf{U}$	List of user states in cell k	–
M_k	Number of users in cell k	–
p_{high}	Probability cell is high density	–
p_{low}	Probability cell is low density	–
M_{high}	Number of users in high density cell	–
M_{low}	Number of users in low density cell	–
D_k	Aggregate cell demand (sum over users)	[bit/s]
$D_k^{(o)}$	Aggregate users after o steps	[bit/s]
N_b	Number of satellite beams	–
h_{sat}	Satellite altitude	[km]
G_{beam}	Beam gain	[dBi]
P_b	Power per beam	[W]
$x_{i,k}$	Allocation: 1 if beam to cell k at time i, else 0	–
X	Allocation matrix $(T \times K)$	–
T_s	Duration of a single slot	[s]
S	Number of slots in T_R, $S = T_R/T_s$	–
θ_{min}	Minimum observation angle	[deg]
K	Total number of cells	–
B	Channel bandwidth	[Hz]
f_c	Carrier frequency	[Hz]
$G_{tx,db}$	Transmitter gain (dB)	[dB]
$P_{tx,db}$	Transmitter power (dB)	[dB]
G_{rx}	Receiver gain	[dB]
PL	Path loss	[dB]
$FSPL$	Free space path loss	[dB]
d	Distance (satellite to cell)	[m]
N	Noise power	[W]
k_b	Boltzmann constant	$[1.38 \times 10^{-23}$ J/K]
T	System noise temperature	[K]
R_k	Shannon capacity for cell k	[bit/s]
C_k	Average capacity allocated to cell k	[bit/s]
R_{min}	Minimum allocation requirement	[bit/s]
$\hat{D}_k$	Average predicted demand over O	[bit/s]
$\bar{D}_k$	Average mode of predicted demand over O	[bit/s]
t	Auxiliary optimisation variable (min C/D)	–

Table B.1: Overview of symbols used in the system model and optimisation problem.