Towards Autonomous Satellite Communications: An AI-based Framework to Address System-level Challenges

Juan Jose Garau-Luis*, Skylar Eiskowitz*, Nils Pachler*, Edward Crawley, Bruce Cameron

Engineering Systems Laboratory
Massachusetts Institute of Technology
{garau, eiskowit, pachler, crawley, bcameron}@mit.edu

Abstract

The next generation of satellite constellations is designed to better address the future needs of our connected society: highly-variable data demand, mobile connectivity, and reaching more under-served regions. Artificial Intelligence (AI) and learning-based methods are expected to become key players in the industry, given the poor scalability and slow reaction time of current resource allocation mechanisms. While AI frameworks have been validated for isolated communication tasks or subproblems, there is still not a clear path to achieve fully-autonomous satellite systems. Part of this issue results from the focus on subproblems when designing models, instead of the necessary system-level perspective. In this paper we try to bridge this gap by characterizing the system-level needs that must be met to increase satellite autonomy, and introduce three AI-based components (Demand Estimator, Offline Planner, and Real Time Engine) that jointly address them. We first do a broad literature review on the different subproblems and identify the missing links to the system-level goals. In response to these gaps, we outline the three necessary components and highlight their interactions. We also discuss how current models can be incorporated into the framework and possible directions of future work.

1 Introduction

The satellite communications landscape has been undergoing significant changes during the last few years. The growth of streaming platforms and other data-intensive services has pushed a television/video-centered industry into a data-dominated market characterized by higher throughput demands (Northern Sky Research 2019). In addition, new ground stations in unserved communities where terrestrial networks are not an affordable option (Reuter 2020), as well as new mobile terminals in almost every plane, make user bases larger and more complex. In response, highly-flexible megaconstellations are being deployed, which will flood the market in the upcoming years. What once was a static process from a resource control point of view is becoming increasingly dynamic and more challenging.

Satellite operators have been able to incorporate the necessary hardware improvements to adapt to this new dynamic context. Beamforming capabilities capture more complex terrestrial landscapes, digital payloads allow for real time power and bandwidth control (Balty, Gayrard, and Agnieray 2009; Angeletti, De Gaudenzi, and Lisi 2008), and the industry is entering an era marked by mega constellations that add new degrees of freedom (Vidal et al. 2021). However, current resource management policies are human-driven, which is no longer sufficient given the high-dimensionality and flexibility of upcoming systems. Therefore, operators are looking to outsource many of these human decisions to autonomous agents, as it will be the key to being competitive in this growing market (Coleman 2019).

Artificial Intelligence (AI) and learning-based solutions are prominent in satellite communications research as a promising way to control these complex space systems. From finding optimal policies to steer antennas (Abbas, Nasser, and Ahmad 2015), to leveraging time series data to identify contingency scenarios (Ackerman 2020), current research suggests that satellite constellations can achieve increasing levels of automation.

While AI is capable of matching the state-of-the-art performance in many of these specific tasks (Luis et al. 2020; Aravanis et al. 2015; Paris et al. 2019; Lei and Vázquez-Castro 2010; Durand and Abrão 2017; Ferreira et al. 2018; Luis et al. 2019), there is not a clear path to move the focus from solving specific tasks to the design of autonomous satellites able to achieve system-level goals. Part of the hardware components studied in recent papers such as antennas or power amplifiers are not isolated, but are part of larger entities that interact, as thoroughly studied by the space systems community (Angeletti, De Gaudenzi, and Lisi 2008). In that sense, there is little consideration on how the proposed AI models will fit into the whole constellation architecture.

Although scarce, some papers do attempt to bridge that gap. This is the case of (Guerster et al. 2019), where a Dynamic Resource Management (DRM) system concept is presented to specifically address the adapted process to control different resources such as power and bandwidth on a satellite. Authors divide their system into a deployed physical network and a short-term and long-term planners that make resource allocation decisions based on demand predictions. While the overall goal of this paper is well aligned with the industry needs, there is little emphasis on how the AI and learning-based technologies on a subsystem level connect to

*These authors contributed equally.

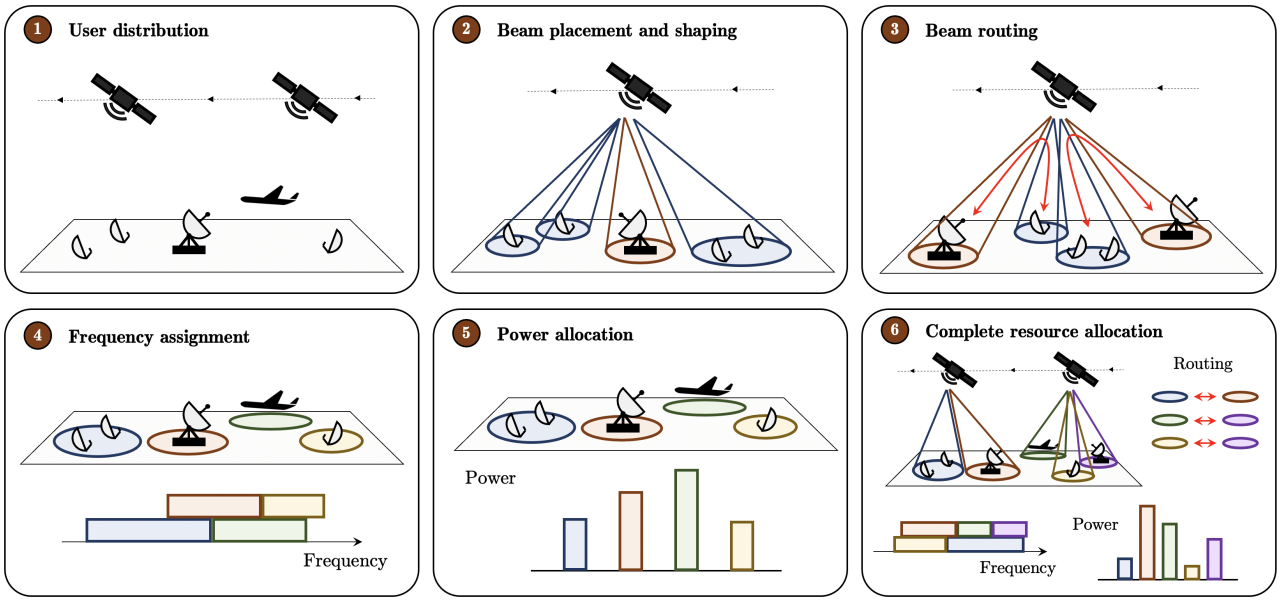

Figure 1: Resource allocation problem in satellite communications divided into a sequence of subproblems. From a set of fixed and mobile users with certain demand requirements (step 1), operators first decide how many beams to use, as well as their location and shape (step 2). Then, each beam is routed to a gateway (step 3) and a certain amount of bandwidth within the available frequency spectrum is allocated (step 4). The final subproblem involves powering each beam (step 5), resulting in a complete resource allocation (step 6).

their role on the system level and vice versa.

In this paper we try to close this research gap by proposing a similar resource management framework to the one presented in (Guerster et al. 2019), highlighting the necessary components in an architecture aimed at solving the complete DRM problem in satellite communications. We identify three interconnected subsystems, namely the Demand Estimator, Offline Planner, and Real Time Engine, that jointly affect the performance of the whole system on different resource management tasks such as power allocation, beam routing, and frequency assignment. Our goal is to analyze the potential value of different AI algorithms on both the system and subsystem levels and understand their trade-offs. Finally, we outline prevailing challenges to achieve better autonomous constellations and areas of future work for each task.

2 Dynamic Resource Management in Multibeam Satellite Constellations

The DRM problem in a satellite communications constellation consists of finding a feasible distribution of the constellation's resources that satisfies the changing demand needs of a user network. Fully solving this problem entails addressing a sequence of complex subproblems, each related to one of the resources to allocate, as illustrated in Figure 1. Overall, it is a time-sensitive problem, as the dynamic behavior of the users' demand requires updating parts of the solution to the subproblems in real time, as pictured in Figure 2. Given these time constraints and the high-dimensionality context of upcoming constellations, AI and learning-based algorithms are seen as essential to keep up with the constantly-changing demand in a way that won't require operators to severely over-provision resources.

2.1 User terminals

A user (represented by a small antenna or a plane in Figure 1) is defined as any entity connected to the constellation that expects certain throughput demand needs to be met (e.g., a person watching a movie through an online streaming platform). Two factors characterize each user: demand and position. In this context, we assume that the user's demand is flexible and may vary over time depending on the user's behaviour. For example, users navigating the web tend to generate highly-fluctuating demand, while backhauling customers tend to be more steady. The position of a user can either be fixed (e.g., common households antennas) or change over time (e.g., planes and ships).

2.2 Constellation's resources

To satisfy the users' demand, operators need to resolve four subproblems in order to distribute the constellation's resources: beam placement and shaping, gateway routing, frequency assignment, and power allocation.

Beam placement and beam shaping. Shown in the second step in Figure 1, the beam placement problem consists of achieving a feasible mapping of users to satellite beams. A beam is a communication channel between one or more users and a satellite. Different users can be assigned to the same beam only if they are all on the beam's terrestrial footprint, which is influenced by the beam's shape. There-

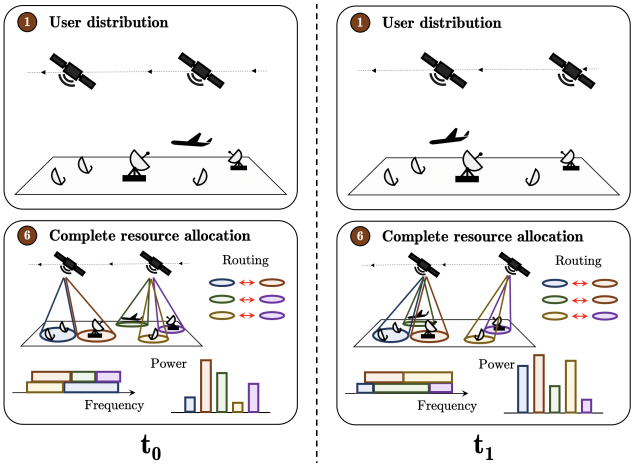

Figure 2: User distributions and resource allocations changing over time. A change in the user distribution from time t_0 to time t_1 prompts a change in the allocation of the constellation's resources.

fore, this process entails grouping users together. Also represented in the second step in Figure 1, the beam shaping problem consists of deciding the shapes and sizes of the beams, and therefore their footprints. Depending on the context, operators might prefer to use few beams with little overlap (which reduces interference) or a higher number of beams with simpler shapes (which is more energy-efficient). The beam shaping problem is not independent from the beam placement problem, so they need to be solved simultaneously.

Gateway routing. Depicted in the third step in Figure 1, the gateway routing problem consists of mapping beams to gateways. A gateway, shown as a bigger antenna in the figure, is an operator's ground station that gathers the traffic from many beams to the Internet and vice versa. There are multiple gateways connected to the network. How beams are mapped to gateways might be constrained by users' characteristics or preferences, and has a potential effect on the quality of the service.

Frequency assignment. Represented by the fourth step in Figure 1, the frequency assignment problem consists of assigning a central frequency and an amount of bandwidth to each beam. This goes beyond simply partitioning the frequency spectrum, since two beams can use overlapping frequencies due to several frequency reuse mechanisms available in each satellite. The extent to which two beams can use overlapping frequencies depends on how close their terrestrial footprints are.

Power allocation. Finally, the fifth step of the process shown in Figure 1 is power allocation, which sets the transmitted radio-frequency (RF) power to every beam. Generally, satellites are power-constrained, but there can be additional restrictions that limit the amount of power available to each beam. Operators can then reduce the necessary power of a specific beam by increasing its bandwidth, for instance.

2.3 System-level challenges

From a system's perspective, the emerging challenges when considering the complete DRM problem are threefold:

- First, each of the presented subproblems is not isolated; the decisions made at a certain step of the process impact the steps to come. For example, placing fewer beams (step 2 in Figure 1) might make routing easier (step 3) but requires higher power consumption (step 5).
- Another challenge comes from the interactions between different satellites in the constellation. One example is the handover process, where a beam serving a user switches from satellite to satellite due to constellation and user dynamics (as shown in Figure 2 for the plane).
- Finally, given the dynamic nature of the environment (e.g., changes in user demand patterns, moving beams, new users, or system failures), operators constantly reevaluate resource management decisions in real time. This is shown in Figure 2, where a change in the user distribution from time t_0 to t_1 requires a reconsideration of the resource allocation decisions. From an algorithmic perspective it is challenging to recompute an allocation given these time constraints and the high-dimensionality of the constellations.

In response to these challenges, AI and learning-based algorithms have the potential to play a substantial role in solving the resource allocation problem. This has started to be explored mainly at the subproblem level, mostly with a focus on brute performance. In the following section we discuss the specific AI models that have been proposed to address each subproblem.

3 Related Work

Multiple studies have addressed solving specific resource allocation subproblems by means of AI and learning-based methods. Most of the literature focuses on metaheuristic algorithms given their straightforward frameworks, but in the recent years authors have started to rely more on neural networks and Reinforcement Learning agents. A summary of the literature can be found in Table 1; we now discuss each subproblem separately.

3.1 AI and learning-based methods for resource allocation subproblems

Beam placement and beam shaping. Most studies that approach the beam placement problem in the literature focus on beam scheduling, i.e., defining start and end times for each beam (Wang et al. 2020a; Deng et al. 2018; He, Jia, and Zhong 2017; Anzalchi et al. 2010; Angeletti, Fernandez Prim, and Rinaldo 2006). All these works show efficient Genetic Algorithms (GA) that optimize based on user satisfaction, total capacity, and/or power usage in the context of satellites with tens of beams. Some authors (Hu et al. 2020b) have also proposed Reinforcement Learning (RL) techniques, specifically Q-learning algorithms, to deal with the added complexity when moving to the thousand-beam range. Regarding other beam placement-specific tasks, researchers have recommended Simulated Annealing (SA)

Method	Algorithm	Beam plac. & shape	Routing	Freq. assign.	Power alloc.
Metaheuristics	**Genetic Algorithm**	2020a; 2018; 2017; 2010; 2006	2011	2017; 2015; 2006; 2019	2015; 2010; 2019
	Particle Swarm Opt.			2020	2017; 2020
	Simulated Annealing	2014		2020; 2018; 2014	2020; 2018
Reinforcement Learning	**Q-learning**	2020b	2020	2020; 2020; 2018; 2018	2020; 2020; 2018
	Policy Gradient				2019; 2020
Supervised learn.	**Neural architectures**			1997; 2005	

Table 1: AI methods and algorithms applied to each resource allocation subproblem in the literature. Studies might focus on the complete subproblem or solve it partially. Some studies address parts of two different subproblems.

(Camino et al. 2014) to specifically solve the question of how many beams to place and where. Although the beam shaping subproblem has been addressed by means of heuristics and mathematical programming frameworks (Kyrgiazos, Evans, and Thompson 2013; Camino et al. 2016), we could not find AI-based proposals for this task.

Routing. The scarce literature on beam and gateway routing mainly focuses on how to route users to gateways in order to minimize delay. While most studies suggest non-AI load balancing solutions (Torkzaban et al. 2020; Yang et al. 2016), some authors have proposed GA (Rao and Wang 2011) and Q-Learning (Gong et al. 2020) frameworks in the context of highly-varying environments.

Frequency assignment. From the four subproblems that constitute the complete resource allocation process, the frequency assignment is the most studied in both depth and breadth. In general terms, when all beams use the same bandwidth, this problem can be understood as assigning the central frequency for each beam. Both metaheuristic (GA, (He, Jia, and Zhong 2017; Tirmizi, Mishra, and Zadgaonkar 2015), SA, (Vidal, Legay, and Goussetis 2020)) and RL-based (Q-learning, (Hu et al. 2018)) formulations have proven successful in solving this task. Some studies do consider flexible bandwidth and address it with both metaheuristic frameworks (GA, (Angeletti, Fernandez Prim, and Rinaldo 2006; Paris et al. 2019), SA (Cocco et al. 2018)) and learning-based models (Q-Learning, (Liao et al. 2020; Zheng et al. 2020), supervised learning with neural networks, (Funabiki and Nishikawa 1997; Salcedo-Sanz and Bousoño-Calzón 2005)). While metaheuristics are better suited for scenarios with fewer beams, the high-dimensionality of real operations makes learning-based techniques a competitive solution for time-constrained contexts. Specific instances of the independent bandwidth allocation problem (i.e., choosing bandwidth fractions) have also been studied through an AI lens (PSO, (Pachler et al. 2020), Q-Learning, (Ferreira et al. 2018)). Finally, although most models consider a single pool of frequency, some research has focused on the effects of frequency reuse and multiple polarizations (SA, (Camino et al. 2014)).

Power allocation. Finally, a large number of studies address the power allocation problem by means of AI and learning-based methods, where the goal is to autonomously choose how much power to provide to each beam while respecting beam- and satellite-wise maximum power constraints. Metaheuristic algorithms (GA (Aravanis et al. 2015; Anzalchi et al. 2010; Paris et al. 2019), PSO (Durand and Abrão 2017; Pachler et al. 2020), and SA (Vidal, Legay, and Goussetis 2020; Cocco et al. 2018)) prove to be effective ways to meet these constraints. However, it is hard for them to scale up to thousands of beams and still converge in the range of seconds. Consequently, authors have also looked into RL-based methods, specifically Q-Learning (Zhang et al. 2020; Liao et al. 2020; Ferreira et al. 2018) and Policy Gradient (Luis et al. 2019, 2020).

As seen in Table 1, some models address decisions concerning more than one subproblem. For example, the Genetic Algorithm proposed in (Angeletti, Fernandez Prim, and Rinaldo 2006) encompasses both beam placement and frequency assignment decisions – but it does not completely address either of them. In most studies, specific subproblems or use cases have been simplified in order to achieve convergence within reasonable time. It remains unclear the best way in which AI and learning-based techniques can expand to encompass multiple subproblems in high-dimensional and time-restricted scenarios.

3.2 System-level gaps in current approaches

Despite the progress on the specific subproblems, gaps in the literature are evident at the system level, where challenges outlined in Section 2.3 remain untouched. First, studies often analyze isolated subproblems, but do not consider the dependence on the next subproblem. For example, authors in (Camino et al. 2014) need to rely on a demand-related metric instead of a user satisfaction metric because their model does not consider power allocation (last step of the resource allocation process). It remains unclear how both metrics relate without solving the latter subproblem. As discussed in the previous section, when more than one subproblem is addressed, most of the time those are not completely solved, and simplifications are applied.

Next, studies fail to address the satellite as a part of a larger constellation network, and so they avoid the added complexity of this higher-dimensional problem. For example, most frequency assignment algorithms (e.g., (Vidal, Legay, and Goussetis 2020; Liao et al. 2020)) overlook that,

given the constellation dynamics, frequency must be assigned to beams that are constantly changing. An allocation that is valid for a specific point in time may have severe limitations at subsequent times, as different beams are handed over to a satellite following routing decisions.

Finally, any DRM problem framework must also situate the subproblems and their interactions in the context of realistic operating scenarios. Most studies assume a given future user distribution (Aravanis et al. 2015; Cocco et al. 2018), which fails to capture the reality that in many cases users' future demand and location are unknown. Forecasting user demand for a short upcoming horizon leads to less uncertainty, but also less time to react and utilize this new knowledge. Challenges like this are overlooked when studies don't simulate the user distribution of the problem. Furthermore, studies omit the increasing uncertainty that directly comes from having higher dimensionality, leading to optimistic results and solutions that might actually be infeasible due to the uncaptured uncertainty.

4 Addressing system-level complexity with an AI-based framework

In this section we introduce the necessary elements of a resource management system and explore the role of AI and learning-based methods to address the system-level complexities. The proposed elements are similar to the ones introduced in (Guerster et al. 2019), and correspond to three components: the Demand Estimator (DE), responsible for forecasting the user distribution; the Offline Planner (OP), responsible for long-term proactive decisions; and the Real Time Engine (RTE), responsible for short-term reactive decisions. These three subsystems, as depicted in Figure 3, jointly address the four resource allocation subproblems. We also discuss how these systems should interact to meet all the constraints imposed by real-time operation and large constellations.

4.1 System goals

As described in Section 2.3, three challenges remain unaddressed from a high-level problem perspective: 1) subproblems need to be considered simultaneously, 2) satellites need to be considered integrated in a larger entity, the constellation, and 3) the environment is dynamic and high-dimensional.

From a real operation point of view, the last point requires operators to rely on user distribution forecasts instead of perfect information to decide on the resource allocation. Operators can decide to predict future distributions with a specific level of uncertainty depending on how far the prediction horizon is. While a lower uncertainty given by shorter horizons is best for finding better allocations, it can substantially reduce the time windows available to compute them. These two features (uncertainty and computing time) define the challenges a DRM framework must operate under. Any architecture for this resource allocation system must be able to address different levels of uncertainty and computing windows.

4.2 DRM Framework overview

Our framework, depicted in Figure 3, consists of three components that jointly aim to solve the complete resource allocation problem described in Figure 1 and address the necessary real-time changes. This process starts with the Demand Estimator forecasting the user distribution (which includes the demand and location of users) at future horizons. The long-term predictions are passed to the Offline Planner, and the short-time predictions are sent to the Real Time Engine.

In conjunction, both the OP and RTE are responsible for the four subproblems. While they both can act on the same set of decisions, the objective and responsibilities of each subsystem are significantly different. The goal of the OP is to compute complete allocations from scratch ahead of time based on user distribution estimates. Given the problem's high-dimensionality and complexity, the time needed to run this process is not negligible, which limits the OP to long-term decisions where computing time is not a constraint. However, since the long-term demand predictions are uncertain, the real and estimated demand may significantly differ. The objective of the RTE is then to tune the resource allocation originally computed by the OP to resolve these discrepancies based on more detailed knowledge of the short-term user distribution. To that end, the RTE should first decide which specific resources it will change, as there might be more than one possible alternative. The RTE iteratively repeats that process until the next OP update.

4.3 AI-based decision-making in the subsystems

Demand Estimator. The DE is tasked with producing forecasts that predict changes in user distributions as accurately as possible, so that the OP and RTE avoid under- and over-provisioning resources. Due to the increasingly complex user demand patterns, it may no longer be sufficient to rely on simple statistical methods that are unable to model complex, nonlinear behaviours (Suganthi and Samuel 2012). Following similar trends in other industries (Fildes and Kumar 2002; Zhao and Magoulès 2012; Wang and Srinivasan 2017; Panapakidis and Dagoumas 2016; Sagheer and Kotb 2019) time series forecasting models like recurrent neural networks and attention-based mechanisms (Vaswani et al. 2017) are essential to accurately predict these patterns. These methods take advantage of large amounts of data to constantly improve their performance and capture both short- and long-term patterns (Vaswani et al. 2017). Furthermore, novel meta learning algorithms may be able to leverage little data coming from new and out-of-distribution users (Panigrahi, Nanda, and Swarnkar 2021).

Offline Planner. The OP is responsible for creating a feasible long-term resource distribution plan that satisfies the forecasted users' demand. Given that this problem is high-dimensional and complex, conventional resource allocation techniques such as Lagrangian-based binary search for frequency assignment (Park et al. 2012) or heuristic-based power allocation (Destounis and Panagopoulos 2011) are unlikely to succeed, even on a subproblem level. In this line, metaheuristics have proven successful in solving each of the individual subproblems (Deng et al. 2018; Rao and

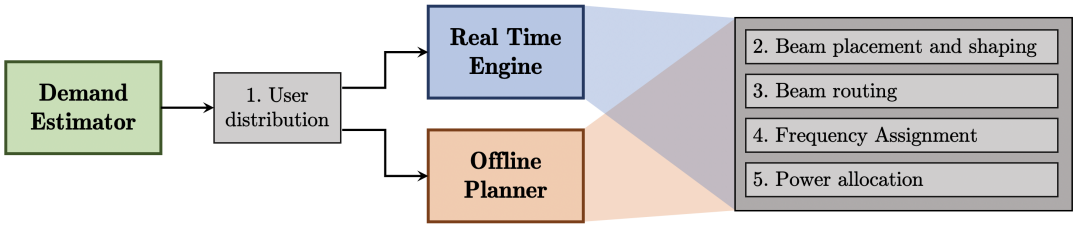

Figure 3: Necessary components in a resource management system. The *Demand Estimator* is responsible for predicting changes in the user distribution (demand and location) at different time horizons. The *Offline Planner* then addresses long-term resource allocation tasks, while the *Real Time Engine* works on a time scale of seconds to minutes, fixing any service failure and leveraging resource-efficient opportunities.

Wang 2011; Tirmizi, Mishra, and Zadgaonkar 2015; Durand and Abrão 2017), and show promising results when dealing with multiple subproblems (Camino et al. 2014; Pachler et al. 2020; He, Jia, and Zhong 2017). In cases where the dimensionality is so high that these techniques are no longer practical, learning-based alternatives have been explored (Hu et al. 2020b; Gong et al. 2020; Wang et al. 2020a; Luis et al. 2019).

Real-Time Engine The RTE has to refine the resource allocation proposed by the OP based on the discrepancies between the real and forecasted user distributions. The goal of the RTE is not to allocate on all possible variables, but rather to a lower-dimensional subset that does not compromise a fast convergence time. In that sense, the Reinforcement Learning-based agents presented in the literature could be better suited to reallocate power (Zhang et al. 2020; Liao et al. 2020; Ferreira et al. 2018; Luis et al. 2019, 2020), reassign frequencies (Liao et al. 2020; Zheng et al. 2020; Ferreira et al. 2018; Hu et al. 2018), reroute (Gong et al. 2020), or rearrange the beam placement (Hu et al. 2020b). In the case of routing and beam placement tasks, the lack of more studies also opens the door to consider metaheuristic-based alternatives, such as (Wang et al. 2020a; Deng et al. 2018; He, Jia, and Zhong 2017; Anzalchi et al. 2010; Angeletti, Fernandez Prim, and Rinaldo 2006; Rao and Wang 2011).

4.4 System interactions

The proposed DRM framework hinges on the distribution of functionalities across the three subsystems, and, as a result, it is important to study how the DE, OP, and RTE interact. Here, we discuss these interactions across the pairs and how AI can handle them.

Demand Estimator and OP/RTE. The DE interacts with both the OP and RTE by feeding in forecasts of the future user distribution. The key challenge the DE must deal with is uncertainty. To quantify uncertainty, common metrics (e.g., MAPE, MAE, and MSE) measure how well a model behaves on a test dataset, but more important to our application are measures of under- and over-predictions in demand. Satellite operators typically overprovision, because an underestimation of future user demand is unacceptable in serving their customers. The DE itself may account for this preference by training a model with an asymmetric loss function that places a heavier penalty on underestimations. Furthermore, the forecasts being probabilistic instead of point estimates gives insight into the amount of uncertainty at the prediction level. This allows the OP to better plan for worst-case scenarios and reduce the number of times the RTE needs to take action. To avoid producing overly-conservative allocations, the OP and RTE subsystems could also incorporate robustness mechanisms against uncertainty, as explored in the literature for both metaheuristic frameworks (Keane 1995; Al-Aomar 2006) and RL (Bellemare, Dabney, and Munos 2017; Dabney et al. 2018).

OP and RTE. The interaction between OP and RTE is more subtle since it is a consequence of the shared control of the constellation's resources. While the OP does a long-term recurrent allocation of all resources, the RTE should only react when nominal demand predictions are out of distribution. We envision this happening under six circumstances: 1) demand spike, when the real demand is significantly higher than expected; 2) demand diminish, when real demand is significantly lower than expected; 3) new user, when a new user connects to the system using a new (non-allocated) beam; 4) mobility re-route, when a ship or plane follows a different route than the one expected; 5) gateway failure, when a gateway antenna temporarily or permanently fails; and 6) satellite failure, when a satellite temporarily or permanently fails. Both the RTE's type of action and time to react significantly vary depending on the event. Nevertheless, all these situations can be considered as out-of-distribution occurrences, which have traditionally been a challenge to AI and learning-based models as a form of non-stationarity or high stochasticity. Recent studies (Peng et al. 2018; Mankowitz et al. 2019; Ghosh et al. 2017) propose generalization mechanisms which might be able to address these scenarios.

The challenges emerging from system interactions pose additional problems to AI and have been traditionally understudied in satellite communications literature. We consider these issues as one of the remaining open challenges when it comes to the Dynamic Resource Management problem and a necessary direction of future work. In the next section we extend on this idea and other open challenges and future work directions.

5 Open Challenges

Different challenges remain regarding the use of AI and learning-based technologies for the Dynamic Resource Management problem. These can be analyzed from a subproblem perspective, a system-level perspective considering interactions, and from the perspective of autonomous systems. In this section we cover all three and propose directions of future work, highlighting those areas where AI is still not present.

5.1 Limitations on subproblem performance

We first address the performance limitations that subproblem-oriented AI algorithms face and which recently-proposed frameworks could help mitigate them.

- **High-dimensionality**. All four subproblems covered in this work (beam placement and shaping, gateway routing, frequency assignment, and power allocation) have been traditionally under-studied in high-dimensional scenarios (Luis et al. 2020), such as those dealing with tens of thousands of users or thousands of beams. Although some works do try to address this issue (Hu et al. 2020b; Luis et al. 2019) by simply running tests on high-dimensional scenarios, recently-proposed mechanisms against high-dimensionality are still to be explored in the satellite communications context. This is the case, for instance, of specific RL frameworks (de Wiele et al. 2020; Zahavy et al. 2018).
- **Time limitations**. Time convergence is often an overlooked factor in current approaches, and can be especially critical for some of the metaheuristic-based algorithms presented. While most works focus on a static picture for which all resources can be allocated without a time horizon, future studies should incorporate new findings on speeding up the convergence time (Nia and Alipouri 2009). Learning-based approaches are more robust against this issue given their quick online evaluations and offline training frameworks.
- **Constellation Dynamics**. Constraints imposed by constellation dynamics are usually neglected. Most works focus on single satellites, disregarding dependencies between the different satellites of the constellation and how they change over time. Constraint-based frameworks have been studied for both metaheuristic algorithms (Homaifar, Qi, and Lai 1994) and RL (Garcia and Fernández 2015; Dalal et al. 2018; Bohez et al. 2019).
- **Out-of-distribution events.** As highlighted in Section 4.4, most events requiring the RTE involvement occur due to out-of-distribution events such as demand spikes or system failures. This issue has traditionally been a challenge for AI and learning-based methods, although recent works propose different methods to increase environment diversity (Ghosh et al. 2017; Lee et al. 2019; Tobin et al. 2017) and consequently the agents' robustness.

5.2 Limitations due to interactions

When the DE, OP, and RTE act together, additional limitations might emerge. These are mainly related with forecasting quality and the OP-RTE interaction.

- **Long-horizon forecasting**. The DE is faced with challenging prediction windows when a delay exists between the current and the target times. Uncertainty in these situations might be too large using traditional methods. Newer models that use attention mechanisms that learn complex patterns in data like (Li et al. 2019) or (Wu et al. 2020) could be explored.
- **Multiuser prediction**. To forecast future user distributions and demands, the DE can do it on a per-user basis or rely on multivariate output models (Charkraborty Kishan Mehrotra, Mohan, and Ranka 1990; Wan et al. 2019). The latter is not sufficiently-explored and can be the basis of frameworks to leverage low amounts of data (e.g., new users, non-stationary patterns) (Yu, Rao, and Dhillon 2016).
- **Search space complexity.** As a consequence of the OP-RTE interaction, once the OP sets a long-term resource allocation, the RTE will amend some parts of it as real-time operations take place. While maybe no more than tens of beams will be part of this process during a RTE cycle, the related search space will be notably more complex, as the RTE will be required to simultaneously make decisions on multiple variable types (e.g., beam center, bandwidth, power) instead of following a sequential process.

5.3 Areas of future work

We finally address unexplored research directions that are related to more advanced problems in the context of autonomous constellations.

- **Transfer learning for satellite architectures**. All models and algorithms so far have been designed in the context of specific satellite architectures. This assumption limits the application of one model to multiple satellite architectures. In the future, it is expected that satellites will have flexibility in their own configurations (De Weck, Scialom, and Siddiqi 2004), to better address operators' needs. Being able to successfully transfer the models to these new configurations will be essential to maintain service quality.
- **New prediction models**. Historical data of current user bases might be rich enough to forecast using simple neural architectures. However, it is possible that this is not always the case, especially if new users choose on-demand services more frequently. In those cases, more complex neural architectures including Graph Neural Networks to leverage spatiotemporal data (Wang et al. 2020b), attention mechanisms to improve performance (Vaswani et al. 2017), or transfer learning frameworks capable of few-shot-learning (Panigrahi, Nanda, and Swarnkar 2021) might be better suited.
- **Orbits as resources.** In this paper, like most studies on satellite communications, we assume all satellites are located on the same orbit. This might not always be the case, especially for megaconstellations. In that sense, we could consider allocating orbital resources as the fifth

subproblem from a constellation perspective. It is likely that AI and learning-based methods can also be applied in that context.

- **Multiagent systems.** Finally, we want to highlight the possibility that the resource control is decentralized as opposed to a centralized DE, OP, and RTE. These scenarios align with the literature on multiagent systems (Foerster et al. 2018; Rashid et al. 2018), and satellite constellations are a specific use case that is starting to be explored (He, Wang, and Wang 2020; Hu et al. 2020a).

6 Conclusion

In this work, we propose an AI-based framework to tackle the system-level challenges in the Dynamic Resource Management problem for satellite communications. This framework incorporates three necessary components (Demand Estimator, Offline Planner, and Real-Time Engine) to address the sequence of subproblems (beam placement and shaping, gateway routing, frequency assignment, and power allocation) that lead to the complete resource allocation. We identify potential component interactions that are often overlooked by current approaches and discuss why AI and learning-based methods are well suited to handle them. In this context, we examine the benefits of applying AI to solve the complete resource allocation problem, highlighting prevailing performance limitations that AI could overcome, as well as unexplored areas of future work.

Acknowledgments

This work was supported by SES. The authors would like to thank SES for their input to this paper and their financial support. The project that gave rise to these results also received the support of a fellowship from "la Caixa" Foundation (ID 100010434). The fellowship code is LCF/BQ/AA19/11720036.

References

Abbas, N.; Nasser, Y.; and Ahmad, K. E. 2015. Recent advances on artificial intelligence and learning techniques in cognitive radio networks. *Eurasip Journal on Wireless Communications and Networking*, 2015(1): 174.

Ackerman, D. 2020. Method finds hidden warning signals in measurements collected over time.

Al-Aomar, R. 2006. Incorporating robustness into genetic algorithm search of stochastic simulation outputs. *Simulation modelling practice and theory*, 14(3): 201–223.

Angeletti, P.; De Gaudenzi, R.; and Lisi, M. 2008. From "Bent Pipes" to "Software Defined Payloads": Evolution and Trends of Satellite Communications Systems.

Angeletti, P.; Fernandez Prim, D.; and Rinaldo, R. 2006. Beam Hopping in Multi-Beam Broadband Satellite Systems: System Performance and Payload Architecture Analysis. In *24th AIAA International Communications Satellite Systems Conference*. Reston, Virigina: American Institute of Aeronautics and Astronautics. ISBN 978-1-62410-027-7.

Anzalchi, J.; Couchman, A.; Gabellini, P.; Gallinaro, G.; D'Agristina, L.; Alagha, N.; and Angeletti, P. 2010. Beam hopping in multi-beam broadband satellite systems: System simulation and performance comparison with non-hopped systems. In *2010 5th Advanced Satellite Multimedia Systems Conference and the 11th Signal Processing for Space Communications Workshop*, 248–255. IEEE. ISBN 978-1-4244-6831-7.

Aravanis, A. I.; Shankar M. R., B.; Arapoglou, P. D.; Danoy, G.; Cottis, P. G.; and Ottersten, B. 2015. Power allocation in multibeam satellite systems: A two-stage multi-objective optimization. *IEEE Transactions on Wireless Communications*, 14(6): 3171–3182.

Balty, C.; Gayrard, J. D.; and Agnieray, P. 2009. Communication satellites to enter a new age of flexibility. *Acta Astronautica*, 65(1-2): 75–81.

Bellemare, M. G.; Dabney, W.; and Munos, R. 2017. A distributional perspective on reinforcement learning. In *International Conference on Machine Learning*, 449–458. PMLR.

Bohez, S.; Abdolmaleki, A.; Neunert, M.; Buchli, J.; Heess, N.; and Hadsell, R. 2019. Value constrained model-free continuous control.

Camino, J.-T.; Artigues, C.; Houssin, L.; and Mourgues, S. 2016. Mixed-integer linear programming for multibeam satellite systems design: Application to the beam layout optimization. In *2016 Annual IEEE Systems Conference (SysCon)*, 1–6. IEEE.

Camino, J.-T.; Mourgues, S.; Artigues, C.; and Houssin, L. 2014. A greedy approach combined with graph coloring for non-uniform beam layouts under antenna constraints in multibeam satellite systems. In *2014 7th Advanced Satellite Multimedia Systems Conference and the 13th Signal Processing for Space Communications Workshop (ASM-S/SPSC)*, 374–381. IEEE. ISBN 978-1-4799-5893-1.

Charkraborty Kishan Mehrotra, K.; Mohan, C. K.; and Ranka, S. 1990. Forecasting the Behavior of Multivariate Time Series using Neural Networks. Technical report.

Cocco, G.; De Cola, T.; Angelone, M.; Katona, Z.; and Erl, S. 2018. Radio Resource Management Optimization of Flexible Satellite Payloads for DVB-S2 Systems. *IEEE Transactions on Broadcasting*, 64(2): 266–280.

Coleman, M. 2019. Is AI key to the survival of satcoms? - SpaceNews.

Dabney, W.; Ostrovski, G.; Silver, D.; and Munos, R. 2018. Implicit quantile networks for distributional reinforcement learning. In *International conference on machine learning*, 1096–1105. PMLR.

Dalal, G.; Dvijotham, K.; Vecerik, M.; Hester, T.; Paduraru, C.; and Tassa, Y. 2018. Safe exploration in continuous action spaces. *arXiv preprint arXiv:1801.08757*.

De Weck, O. L.; Scialom, U.; and Siddiqi, A. 2004. Optimal reconfiguration of satellite constellations with the auction algorithm. In *A Collection of the 22nd AIAA International Communications Satellite Systems Conference and Exhibit Technical Papers*, volume 1, 457–473. American Institute of Aeronautics and Astronautics Inc. ISBN 1563477122.

de Wiele, T.; Warde-Farley, D.; Mnih, A.; and Mnih, V. 2020. Q-Learning in enormous action spaces via amortized approximate maximization. *arXiv preprint arXiv:2001.08116*.

Deng, B.; Jiang, C.; Kuang, L.; Guo, S.; Lu, J.; and Zhao, S. 2018. Two-Phase Task Scheduling in Data Relay Satellite Systems. *IEEE Transactions on Vehicular Technology*, 67(2): 1782–1793.

Destounis, A.; and Panagopoulos, A. D. 2011. Dynamic power allocation for broadband multi-beam satellite communication networks. *IEEE Communications letters*, 15(4): 380–382.

Durand, F. R.; and Abrão, T. 2017. Power allocation in multibeam satellites based on particle swarm optimization. *AEU - International Journal of Electronics and Communications*, 78: 124–133.

Ferreira, P. V. R.; Paffenroth, R.; Wyglinski, A. M.; Hackett, T. M.; Bilen, S. G.; Reinhart, R. C.; and Mortensen, D. J. 2018. Multiobjective Reinforcement Learning for Cognitive Satellite Communications Using Deep Neural Network Ensembles. *IEEE Journal on Selected Areas in Communications*, 36(5): 1030–1041.

Fildes, R.; and Kumar, V. 2002. Telecommunications demand forecasting - A review. *International Journal of Forecasting*, 18(4): 489–522.

Foerster, J.; Farquhar, G.; Afouras, T.; Nardelli, N.; and Whiteson, S. 2018. Counterfactual multi-agent policy gradients. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32.

Funabiki, N.; and Nishikawa, S. 1997. A gradual neural-network approach for frequency assignment in satellite communication systems. *IEEE Transactions on Neural Networks*, 8(6): 1359–1370.

Garcia, J.; and Fernández, F. 2015. A comprehensive survey on safe reinforcement learning. *Journal of Machine Learning Research*, 16(1): 1437–1480.

Ghosh, D.; Singh, A.; Rajeswaran, A.; Kumar, V.; and Levine, S. 2017. Divide-and-Conquer Reinforcement Learning.

Gong, X.; Sun, L.; Zhou, J.; Wang, J.; and Xiao, F. 2020. Adaptive Routing Strategy Based on Improved Q-learning for Satellite Internet of Things. In *International Conference on Security, Privacy and Anonymity in Computation, Communication and Storage*, 161–172. Springer.

Guerster, M.; Jose Garau Luis, J.; Crawley, E.; and Cameron, B. 2019. Problem representation of dynamic resource allocation for flexible high throughput satellities. In *IEEE Aerospace Conference Proceedings*, volume 2019-March. IEEE Computer Society. ISBN 9781538668542.

He, S.; Wang, T.; and Wang, S. 2020. Load-Aware Satellite Handover Strategy Based on Multi-Agent Reinforcement Learning. In *GLOBECOM 2020-2020 IEEE Global Communications Conference*, 1–6. IEEE.

He, Y.; Jia, Y.; and Zhong, X. 2017. A traffic-awareness dynamic resource allocation scheme based on multi-objective optimization in multi-beam mobile satellite communication systems. *International Journal of Distributed Sensor Networks*, 13(8): 155014771772355.

Homaifar, A.; Qi, C. X.; and Lai, S. H. 1994. Constrained optimization via genetic algorithms. *Simulation*, 62(4): 242–253.

Hu, X.; Liao, X.; Liu, Z.; Liu, S.; Ding, X.; Helaoui, M.; Wang, W.; and Ghannouchi, F. M. 2020a. Multi-agent deep reinforcement learning-based flexible satellite payload for mobile terminals. *IEEE Transactions on Vehicular Technology*, 69(9): 9849–9865.

Hu, X.; Liu, S.; Chen, R.; Wang, W.; and Wang, C. 2018. A Deep Reinforcement Learning-Based Framework for Dynamic Resource Allocation in Multibeam Satellite Systems. *IEEE Communications Letters*, 22(8): 1612–1615.

Hu, X.; Zhang, Y.; Liao, X.; Liu, Z.; Wang, W.; and Ghannouchi, F. M. 2020b. Dynamic beam hopping method based on multi-objective deep reinforcement learning for next generation satellite broadband systems. *IEEE Transactions on Broadcasting*, 66(3): 630–646.

Keane, A. J. 1995. Genetic algorithm optimization of multi-peak problems: studies in convergence and robustness. *Artificial Intelligence in Engineering*, 9(2): 75–83.

Kyrgiazos, A.; Evans, B.; and Thompson, P. 2013. Irregular beam sizes and non-uniform bandwidth allocation in HTS multibeam satellite systems. In *31st AIAA International Communications Satellite Systems Conference (ICSSC)*.

Lee, K.; Lee, K.; Shin, J.; and Lee, H. 2019. Network Randomization: A Simple Technique for Generalization in Deep Reinforcement Learning.

Lei, J.; and Vázquez-Castro, M. A. 2010. Joint power and carrier allocation for the multibeam satellite downlink with individual SINR constraints. In *IEEE International Conference on Communications*. ISBN 9781424464043.

Li, S.; Jin, X.; Xuan, Y.; Zhou, X.; Chen, W.; Wang, Y.-X.; and Yan, X. 2019. Enhancing the Locality and Breaking the Memory Bottleneck of Transformer on Time Series Forecasting. (NeurIPS).

Liao, X.; Hu, X.; Liu, Z.; Ma, S.; Xu, L.; Li, X.; Wang, W.; and Ghannouchi, F. M. 2020. Distributed Intelligence: A Verification for Multi-Agent DRL-Based Multibeam Satellite Resource Allocation. *IEEE Communications Letters*, 24(12): 2785–2789.

Luis, J. J. G.; Guerster, M.; Del Portillo, I.; Crawley, E.; and Cameron, B. 2019. Deep reinforcement learning for continuous power allocation in flexible high throughput satellites. In *2019 IEEE Cognitive Communications for Aerospace Applications Workshop, CCAAW 2019*. Institute of Electrical and Electronics Engineers Inc. ISBN 9781728100487.

Luis, J. J. G.; Pachler, N.; Guerster, M.; Del Portillo, I.; Crawley, E.; and Cameron, B. 2020. Artificial Intelligence Algorithms for Power Allocation in High Throughput Satellites: A Comparison. In *IEEE Aerospace Conference Proceedings*. IEEE Computer Society. ISBN 9781728127347.

Mankowitz, D. J.; Levine, N.; Jeong, R.; Shi, Y.; Kay, J.; Abdolmaleki, A.; Springenberg, J. T.; Mann, T.; Hester, T.; and Riedmiller, M. 2019. Robust Reinforcement Learning for Continuous Control with Model Misspecification.

Nia, M. B.; and Alipouri, Y. 2009. Speeding up the genetic algorithm convergence using sequential mutation and circular gene methods. In *2009 Ninth international conference on intelligent systems design and applications*, 31–36. IEEE.

Northern Sky Research. 2019. VSAT and Broadband Satellite Markets. Technical report.

Pachler, N.; Garau Luis, J. J.; Guerster, M.; Crawley, E.; and Cameron, B. 2020. Allocating Power and Bandwidth in Multibeam Satellite Systems using Particle Swarm Optimization. In *2020 IEEE Aerospace Conference*, 1–11. IEEE. ISBN 978-1-7281-2734-7.

Panapakidis, I. P.; and Dagoumas, A. S. 2016. Day-ahead electricity price forecasting via the application of artificial neural network based models. *Applied Energy*, 172: 132–151.

Panigrahi, S.; Nanda, A.; and Swarnkar, T. 2021. A Survey on Transfer Learning. *Smart Innovation, Systems and Technologies*, 194(10): 781–789.

Paris, A.; Del Portillo, I.; Cameron, B.; and Crawley, E. 2019. A Genetic Algorithm for Joint Power and Bandwidth Allocation in Multibeam Satellite Systems. In *IEEE Aerospace Conference Proceedings*, volume 2019-March. IEEE Computer Society. ISBN 9781538668542.

Park, U.; Kim, H. W.; Oh, D. S.; and Ku, B.-J. 2012. A dynamic bandwidth allocation scheme for a multi-spot-beam satellite system. *Etri Journal*, 34(4): 613–616.

Peng, X. B.; Andrychowicz, M.; Zaremba, W.; and Abbeel, P. 2018. Sim-to-Real Transfer of Robotic Control with Dynamics Randomization. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, 3803–3810. IEEE. ISBN 978-1-5386-3081-5.

Rao, Y.; and Wang, R. 2011. Performance of QoS routing using genetic algorithm for Polar-orbit LEO satellite networks. *AEU-International Journal of Electronics and Communications*, 65(6): 530–538.

Rashid, T.; Samvelyan, M.; de Witt, C. S.; Farquhar, G.; Foerster, J.; and Whiteson, S. 2018. QMIX: Monotonic Value Function Factorisation for Deep Multi-Agent Reinforcement Learning.

Reuter, J. 2020. How satellite helps reach underserved areas around the globe.

Sagheer, A.; and Kotb, M. 2019. Time series forecasting of petroleum production using deep LSTM recurrent networks. *Neurocomputing*, 323: 203–213.

Salcedo-Sanz, S.; and Bousoño-Calzón, C. 2005. A Hybrid Neural-Genetic Algorithm for the Frequency Assignment Problem in Satellite Communications. *Applied Intelligence*, 22(3): 207–217.

Suganthi, L.; and Samuel, A. A. 2012. Energy models for demand forecasting - A review. *Renewable and Sustainable Energy Reviews*, 16(2): 1223–1240.

Tirmizi, A.; Mishra, R. S.; and Zadgaonkar, A. S. 2015. An efficient channel assignment strategy in cellular mobile network using hybrid genetic algorithm. *International Journal of Electrical, Electronics and Computer Engineering*, 4(2): 127.

Tobin, J.; Fong, R.; Ray, A.; Schneider, J.; Zaremba, W.; and Abbeel, P. 2017. Domain randomization for transferring deep neural networks from simulation to the real world. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 23–30. IEEE. ISBN 978-1-5386-2682-5.

Torkzaban, N.; Gholami, A.; Baras, J. S.; and Papagianni, C. 2020. Joint satellite gateway placement and routing for integrated satellite-terrestrial networks. In *ICC 2020-2020 IEEE International Conference on Communications (ICC)*, 1–6. IEEE.

Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, ; and Polosukhin, I. 2017. Transformer: Attention is all you need. *Advances in Neural Information Processing Systems 30*, (Nips): 5998–6008.

Vidal, F.; Legay, H.; and Goussetis, G. 2020. Joint Power, Frequency and Precoding Optimisation in a Satellite SDMA Communication System. In *2020 10th Advanced Satellite Multimedia Systems Conference and the 16th Signal Processing for Space Communications Workshop (ASM-S/SPSC)*, 1–8. IEEE.

Vidal, F.; Legay, H.; Goussetis, G.; Garcia Vigueras, M.; Tubau, S.; and Gayrard, J. 2021. A methodology to benchmark flexible payload architectures in a megaconstellation use case. *International Journal of Satellite Communications and Networking*, 39(1): 29–46.

Wan, R.; Mei, S.; Wang, J.; Liu, M.; and Yang, F. 2019. Multivariate Temporal Convolutional Network: A Deep Neural Networks Approach for Multivariate Time Series Forecasting. *Electronics*, 8(8): 876.

Wang, L.; Xu, S.; Hu, X.; Wang, W.; and Ma, S. 2020a. Dynamic Beam Hopping of Multi-beam Satellite Based on Genetic Algorithm. In *2020 IEEE International Conference on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom)*, 1364–1370. IEEE.

Wang, X.; Ma, Y.; Wang, Y.; Jin, W.; Wang, X.; Tang, J.; Jia, C.; and Yu, J. 2020b. Traffic flow prediction via spatial temporal graph neural network. In *Proceedings of The Web Conference 2020*, 1082–1092.

Wang, Z.; and Srinivasan, R. S. 2017. A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models. *Renewable and Sustainable Energy Reviews*, 75(November 2016): 796–808.

Wu, N.; Green, B.; Ben, X.; and O'Banion, S. 2020. Deep Transformer Models for Time Series Forecasting: The Influenza Prevalence Case. *arXiv*.

Yang, Y.; Xu, M.; Wang, D.; and Wang, Y. 2016. Towards energy-efficient routing in satellite networks. *IEEE Journal on Selected Areas in Communications*, 34(12): 3869–3886.

Yu, H.-F.; Rao, N.; and Dhillon, I. S. 2016. Temporal Regularized Matrix Factorization for High-dimensional Time Series Prediction. Technical report.

Zahavy, T.; Haroush, M.; Merlis, N.; Mankowitz, D. J.; and Mannor, S. 2018. Learn what not to learn: Action elimination with deep reinforcement learning. In *Advances in Neural Information Processing Systems*, 3562–3573.

Zhang, P.; Wang, X.; Ma, Z.; Liu, S.; and Song, J. 2020. An online power allocation algorithm based on deep reinforcement learning in multibeam satellite systems. *International Journal of Satellite Communications and Networking*, 38(5): 450–461.

Zhao, H. X.; and Magoulès, F. 2012. A review on the prediction of building energy consumption. *Renewable and Sustainable Energy Reviews*, 16(6): 3586–3592.

Zheng, F.; Pi, Z.; Zhou, Z.; and Wang, K. 2020. LEO Satellite Channel Allocation Scheme Based on Reinforcement Learning. *Mobile Information Systems*, 2020.