Received: 14 October 2023 | Revised: 9 May 2024 | Accepted: 24 May 2024

DOI: 10.1002/sat.1528

ORIGINAL PAPER

WILEY

Artificial intelligence-based data path control in low Earth orbit satellites-driven optical communications

Andrea Wrona | Andrea Tantucci

Department of Computer, Control and Management Engineering Antonio Ruberti, Sapienza University of Rome, Rome, Italy

Correspondence
Andrea Wrona, Department of Computer, Control and Management Engineering Antonio Ruberti, Sapienza University of Rome, Rome, Italy.
Email: wrona@diag.uniroma1.it

Summary

Free space optical communication has emerged as a promising technology for high-speed and secure data transmission between ground stations on Earth and orbiting satellites. However, this communication technology suffers from signal attenuation due to atmospheric turbulence and beam alignment precision. Low Earth orbit satellites play a pivotal role in optical communication due to their low altitude over the Earth surface, which mitigates the atmospheric precipitation effects. This paper introduces a novel data path control law for satellite optical communication exploiting artificial intelligence-based predictive weather forecasting and a node selection mechanism based on reinforcement learning. Extensive simulations on three case studies demonstrate that the proposed control technique achieves remarkable gains in terms of link availability with respect to other state-of-the-art solutions.

KEYWORDS
artificial intelligence, low Earth orbit, optical communication, reinforcement learning

1 | INTRODUCTION

In today's ever-connected world, the demand for high-speed, reliable and secure data transmission has never been greater. The proliferation of data-intensive applications, such as streaming video,[1] cloud computing[2] and the Internet of Things (IoT),[3] continues to place unprecedented strain on traditional communication networks. To meet these growing demands, the development of innovative communication technologies has become imperative. Among the innovative communication technologies, free space optics (FSO) has emerged as a promising solution to address these challenges.

1.1 | Free-space optical communication

In the telecommunications domain, FSO indicates all those wireless communications which, instead of making use of radio carriers in the form of a radio communication, make use of electromagnetic carriers belonging to the range of optical or infrared frequencies or wavelengths, aimed at transporting information between a transmitter and a receiver.[4]

FSO offers several advantages over traditional radio frequency (RF) communication systems when used in space-based applications. In particular, the main benefit of FSO communication relies on data transfer rates. Indeed, an FSO system based on lasers can achieve much higher data transfer rates with respect to RF communications. This is due to the fact that laser light has a much shorter wavelength than RF waves: the wavelength of laser light falls within the optical spectrum, typically in the range of 400 to 700 nanometres (nm), while RF waves can have much longer wavelengths, ranging from millimetres to meters.[5] The shorter wavelength of laser light allows for higher frequency modulation, which means that data can be encoded onto the carrier signal at much higher frequencies. This enables a more significant number of data bits to be transmitted per unit of time.

Moreover, laser communication systems can exploit a larger portion of the electromagnetic spectrum, including multiple wavelength channels, to transmit data simultaneously. This multiplexing capability increases the total data capacity of the communication link.

Other advantages of FSO systems over the RF counterpart rely on (i) the smaller divergence of the laser beam, which enables a higher concentration of optical power,[6] (ii) lower interference thanks to a point-to-point communication with a direct line of sight,[7] (iii) lower latency over longer distances[8] and (iv) more robust security due to the inherent difficulty to intercept FSO signals without being located precisely in the path of the beam.[9]

However, FSO communication systems come also with drawbacks and limitations related to atmospheric turbulence and bad weather conditions (thick clouds, rain, or snow) in the proximity of the optical ground stations (OGSs). These situations often lead to cloud and rain attenuation of the transmitted signal, causing most of the time (i) scintillation of the received optical signal,[10] (ii) beam wander[11] and (iii) beam divergence.[12]

In general, the effects of adverse weather conditions on FSO systems become more pronounced as the distance between the transmitter and the receiver increases. This is why, in the satellite telecommunications domain, an FSO link between an OGS and a satellite is typically operated by means of low Earth orbit (LEO) satellites.

1.2 | LEO satellites

LEO satellites play a pivotal role in the advancement of FSO communication. Situated at altitudes between 180 and 2000 km above the Earth's surface,[13] LEO satellites have relatively short orbital periods, typically completing one orbit around the Earth in 90–120 min[14]: This frequent orbiting allows for better coverage and faster data transmission.

Due to their relatively low altitude, atmospheric effects, such as signal attenuation due to rain fade and atmospheric turbulence, have a reduced impact on an FSO system compared with GEO satellites. This results in more reliable and consistent FSO communication links, characterised by low latency, high-data throughput and improved signal strength, making them a preferred choice for real-time, high-data-rate FSO applications.[15]

Moreover, this type of satellite can be deployed in large constellations,[16,17] which provide continuous global coverage and improve the overall reliability of the FSO communication. This aspect is crucial for applications that require uninterrupted connectivity, such as satellite-based internet services.

LEO satellites are commonly used in the space industry for scientific research and Earth observation purposes and for military operations, as well as for communication, navigation and remote sensing applications.[18]

On the other hand, LEO satellites present two main disadvantages compared with the GEO ones. The first one deals with their shorter lifespan, requiring periodic orbital adjustments to maintain their position in the orbit or replacements of units within the fleet.[19,20] The second one is related with visibility issues: Since LEO satellites' rotation speed is much higher than the Earth's rotational speed, FSO terrestrial signals have to be handed over to another satellite within the fleet. A satellite handover is performed when the serving satellite is below a minimum elevation angle relative to the corresponding OGS: This may have a significant impact on the communication quality, because of communication loss during the handover process.[21]

Despite these downsides, LEO satellites represent the ideal technology for optical satellite communications, problems related to laser signal attenuation in the presence of adverse atmospheric conditions remain. To address these technological difficulties, various methodologies have been proposed by the scientific community.

1.3 | State of the art

Researchers and engineers have created a variety of strategies and technologies to solve the problems of power attenuation in FSO systems due to atmospheric fading. A standard procedure rely on adaptive optics.[22,23] Systems implementing this technology correct for turbulence-induced distortions by changing the geometry of optical components like mirrors or deformable lenses based on real-time observations of air turbulence. This technique aids in optical beam stabilisation and minimises scintillation effects.

Other techniques rely on filtering and error correction, in which proper filters and modulation methods try to filter out noise coming from the interference of fog or clouds.[24] Since in many situations it is not possible to filter out the noise, it is possible to employ broader laser beams to reduce the effects of beam spreading brought on by turbulence.[12] This strategy, nevertheless, could result in slower data speeds.[25] Eventually, another common standard approach is to implement redundant FSO lines, equipping satellites or OGSs with more than one laser communication terminals (LCTs)[26], or FSO/RF hybrid systems,[27] in order to enhance the communication system reliability.

The aforementioned fading mitigation techniques intervene at the hardware level on the individual receiver or transmitter but do not take into consideration any changes to the architecture or topology of the communication system.

In order to limit bad weather effects, it is possible to intervene at architecture level linking in a wired fashion two or more OGSs within a same network. In this way, if the signal suffers some degradation in an area, the other OGSs, located in areas where the weather is favourable, may compensate said attenuation. The communication loss is mitigated by continuously forwarding the signal to the OGS(s) under untoward weather conditions, at least until the latter improve. This technique is called site diversity.[28]

The site diversity has proven to be a disruptive approach for the reliability of FSO communications, since it (i) enables geographical diversity (usually GSs are located more than 100 km away one from each other[28]) to reduce the likelihood of simultaneous signal degradation at all sites,[28] (ii) realises spatial separation to ensure that the OGSs locations are subject to different weather patterns and atmospheric conditions[29] and (iii) involves using multiple antennas at each site, pointing in different directions or at different elevation angles. This configuration allows the system to quickly switch between antennas to find the clearest signal path, thus improving the link availability and reducing the number of outages or dead times.[30]

Although the site diversity technique adds complexity and high cost to the infrastructure, the benefits in terms of improved reliability and availability often justify its implementation, particularly for mission-critical applications.

Said technique employs sophisticated control and switching mechanisms to monitor the quality of signals received at different sites in real-time. When one site experiences signal degradation, the system automatically switches to an alternate site with better signal quality. These switching mechanisms were initially manual and human-driven, while nowadays are usually based on statistical analysis of weather forecasts,[31–34] with the switching system being controlled by intelligent algorithms.

The choice of which OGSs to point at or to transmit from is driven by a series of key performance indicators (KPIs) and follows an optimal routing/resource allocation logic.[35,36] The most important KPI for any satellite communication system is the link availability, but other design drivers for the multi-station site diversity algorithms may include (i) the total system throughput, which corresponds to the successful transmission rate over a certain time interval and is strictly related to the link availability, (ii) the energy consumption for the movement/re-pointing of LCTs, which impacts on the total power budget for the on-board payload, (iii) the topology of the OGSs network (i.e. specific OGSs network topologies may prevent the possibility to re-route the user traffic from one OGS to the others, thus limiting the subset for choosing the second OGS to be the second LCT pointed towards), (iv) user plane latency and jitter and (v) on-board switching capabilities.

An additional KPI to consider is the information freshness, also called Age of Information (AoI). This parameter is of particular importance in applications like the tracking of of ships and cargo[37] and in guaranteeing a high quality of experience in the context of massive multiple access networks.[38]

Since the installation of redundant OGSs may represent a waste of investment for the network operator, several works focused on the minimisation of the number of required OGSs to guarantee a minimum given system performance.[39–41] A different optimisation approach relies on the hypothesis that OGSs have been already positioned, and the problem focuses on how to choose the set of OGSs to connect to in order to maximise the availability. Fuchs and Moll[42] calculate the correlated and uncorrelated availability for OGS networks in the scenario of space-to-ground optical communication links with GEO satellites. An efficient optimisation algorithm is presented, in order to choose the best OGS starting from 5 years of cloud data. It is shown how many OGSs deployed in a very wide area can guarantee a network availability near to 100%. A complementary optimisation approach is proposed in Lyras et al,[43] with the selection of the minimum number of ground stations (GSs) satisfying the monthly availability requirements of the total network, so minimising service and maintenance costs. Eventually, in the scenario presented in Erdogan et al,[44] the optimisation process consists in selecting the best GS among several candidates, trying to provide a reliable connectivity through large-scale site diversity. Results show that the optimal choice mostly depends on the altitude and the zenith angle of the set of GSs.

Recent advancements in artificial intelligence (AI), particularly in the field of reinforcement learning (RL), have opened up new possibilities for optimising satellite communication strategies. As an example, the work in Zhu et al[45] proposes a Deep RL load-balancing strategy for 6G regional satellite communication systems. The approach makes use of virtual network embedding (VNE) in order to deal with the limited availability of resources a small satellite network can offer. Wrona et al[26] propose an AI-based predictive handover strategy for optical communications between a GEO satellite equipped with two LCTs and an OGS network, making use of weather forecasts. Other works making use of AI and RL focused on the resource allocation and traffic splitting for RF satellite communications,[46–49] shifting attention from the OGS network level to the one of the LEO constellation.

However, none of the above-mentioned works have tackled the issue of defining an intelligent handover and path planning procedure for an FSO-based point-to-point communication system between terrestrial OGS networks and a LEO fleet.

In this paper, a mixed deep learning (DL) for weather forecasts and RL approach for intelligent signal routing is proposed, in order to tackle the problem of minimising the outage probability on a multi-agent scenario. The main innovations of the present work are as follows:

- multihop data routing between two OGS networks that cannot communicate directly but only passing through a LEO satellite fleet;
- weather predictions over the OGSs areas via supervised learning exploiting historical hourly weather data;
- a centralised control law realised through an intelligent agent exploiting the RL framework with an intrinsic optimisation of the link availability.

The remainder of this paper is organised as follows: Section 2 provides general background on RL strategies within the context of multi-agent systems. This section establishes the necessary foundation and background information for understanding the subsequent discussions. Section 3 presents the mathematical modelling of LEO satellites' and Earth's dynamical motion. In Section 4, extensive simulations show the effectiveness of the proposed approach with respect to the system KPIs, and, eventually, Section 5 sums up the carried out research, pointing out the achieved results, describing as well its limitations and outlining future research directions.

2 | PRELIMINARIES

In this section, the mathematical foundations of recurrent neural networks (RNNs) and RL control will be presented.

2.1 | Long short-term memory (LSTM) neural networks

The first stage of the technique described here utilises a collection of region-specific DL weather forecasting models. The supervised AI system feeds historical numerical weather data that describe each OGS geographical zone as input. With a temporal resolution T_R that must be in accordance with the time constants characterising the rotation and switching operations of the LCTs, each piece of data is labelled with the real weather conditions that prevailed at that precise moment. By reviewing the data from the previous days, one can attempt to identify patterns between the weather state and the accessible features (atmospheric pressure, temperature, etc.).

The LSTM, first suggested in Hochreiter and Schmidhuber[50] to address the well-known problem of vanishing gradient that characterises RNN, is the most promising neural network topology to accomplish this goal. Because LSTM gains from feedback connections between neurons, it differs significantly from classical feedforward neural networks in that it can analyse complete data sequences in addition to individual data points. The LSTM structure is therefore ideally adapted to handle time-series data, such as the one we are addressing in our work. By specifying a certain time window $\mathcal{T}_p$ of length T_W, the AI model tries to predict weather at time $k+1$ by looking at the actual weather encountered in the T_W previous time instants, that is,

$$\mathcal{T}_p = \{k, k-1, \ldots, k-T_W-1\}. \tag{1}$$

At their core, LSTMs are composed of memory cells that enable them to store and manipulate information over extended sequences. These memory cells have three crucial components:

1. Cell state: it is like a conveyor belt that runs through the entire LSTM network. It can transport information across time steps without much modification. The cell state can be updated, allowing it to capture relevant information and discard irrelevant details.
2. Hidden state: also known as the output state, carries information from previous time steps to the current one. It acts as a working memory that helps LSTMs remember past information that is crucial for making predictions or decisions.
3. Gates: LSTMs employ three types of gates to control the flow of information:
   - Forget gate. This gate decides what information from the cell state should be discarded or kept. It takes as input the previous hidden state and the current input and outputs a value between 0 and 1 for each component of the cell state, where 0 means 'forget' and 1 means 'keep'.
   - Input gate. This gate determines what new information should be added to the cell state. It computes a candidate cell state and decides which parts of it should be added to the current cell state.
   - Output gate. The output gate controls what information should be output as the hidden state. It takes the current cell state and the input, and it generates the new hidden state.

Like any other machine learning procedure, LSTM can handle in an efficient way large-scale datasets even when launched on workstations with limited hardware resources, thanks to the mini-batch strategy.[51] Instead of processing the entire dataset at once, the training samples are divided into mini-batches, thus applying the gradient descent algorithm several times within each epoch.

Moreover, it is also important to stress the fact that LSTM neural networks are automatically implemented with the gradient flipping feature. Gradients that are above a certain safe threshold are prevented from becoming so by rescaling them. This keeps the gradients from blowing up and still permits efficient learning. In LSTM, gradient clipping is usually applied globally to all gradients with respect to all tunable weights.[50]

In this work, each LSTM network is trained on local OGS meteorological data, because if not so it would be difficult for a single predictor to generalise across the various climates of the OGSs' geographic regions, which can actually be located at very different latitudes.

2.2 | Markov decision process (MDP) and RL

RL is one of the branches of machine learning: It deals with intelligent agents performing actions over an environment and then observing its state and the reward function.[52] The agents' goal is to find a policy which maximises the expected cumulative reward, without being aware of the dynamical equations governing the environment (data-driven control). Usually, an RL problem is formalised through a MDP, defined through a tuple $(\mathcal{S},\mathcal{A},P,R,\gamma)$, where we have the following:

- $\mathcal{S}$ is the state space
- $\mathcal{A}$ is the action space
- $P(s'|s,a)$ is the probability that the environment transitions from state s to state s' when the agent chooses the action a.
- $R(s,a,s')$ is the immediate reward the agent gets when transitioning from state s to s' after taking action a
- $\gamma \in [0, 1)$ is the discount factor, represent the agent's preference for immediate rewards $(\gamma \approx 0)$ over the future ones $(\gamma \approx 1)$.

The objective of the agent in an MDP is to find a policy $\pi : S \rightarrow A$ that maximises the cumulative function of the rewards over a (potentially) infinite horizon.[52] Usually, the said function is defined as follows:

$$G_k = \sum_{i=0}^{\infty} \gamma^k R_{k+i+1}, \tag{2}$$

where $0 \leq \gamma \leq 1$ is the discount rate.

When the transition probability $P(\cdot)$ is not known, it is possible to rely on RL techniques, in which the agent learns the optimal policy through the experience. All MDPs handled through RL require the estimation of a value function or an action-value function $Q_\pi(\cdot)$ for a given policy $\pi(\cdot)$[52]:

$$Q_\pi(s,a) = \mathbb{E}_\pi[G_k|s_k = s, a_k = a], \forall s \in \mathcal{S}, \forall a \in \mathcal{A}. \tag{3}$$

Action-value functions satisfy recursive relationships through the Bellman equation, which expresses a link between the action-value function of a state with the action-value function of the next state

$$\begin{aligned} Q_\pi(s,a) &= \mathbb{E}_\pi[G_k|s_k = s, a_k = a] \\ &= \mathbb{E}_\pi[R_{k+1} + \gamma G_{k+1}|s_k = s, a_k = a] \\ &= \sum_{s'} P(s'|s,a)(R + \gamma \sum_{a'} \pi(a'|s') Q_\pi(s',a')), \end{aligned} \tag{4}$$

where (s',a') is the next state-action couple with respect to (s,a). Hence, solving an MDP through RL means finding the optimal action-value function $Q^*(s,a) = \max_\pi Q_\pi(s,a)$ for which it holds the Bellman principle of optimality[52]:

$$Q^*(s,a) = \sum_{s'} P(s'|s,a)(R + \gamma \max_{a'} Q^*(s',a')). \tag{5}$$

One of the most popular and used RL algorithm to estimate $Q^*(s,a)$ is the Q-learning[53]; said algorithm updates the Q-values according to the following law:

$$Q(s_k,a_k) \leftarrow (1-\alpha)Q(s_k,a_k) + \alpha[R_{k+1} + \gamma \max_a Q(s_{k+1},a)], \tag{6}$$

where α is the so-called learning rate. A popular choice for the policy $\pi(\cdot)$ is the ε-greedy policy through which the agent selects with probability ε a random action and with probability $1-\varepsilon$ the action associated to the maximum value in the Q-table. For finite MDPs, it has been proven that the Q-Learning algorithm is able to converge to the optimal Q-function if the Q-Learning update rule given by (6) is used and if the learning rate α satisfies the following conditions:

$$\sum_k \alpha_k = \infty, \sum_k \alpha_k^2 < \infty. \tag{7}$$

This condition requires that $\alpha \in [0;1)$ which translates in the fact that each state-action pair is visited infinitely often. By adopting an ε-greedy policy, this condition can be stochastically satisfied. The pseudocode of the algorithm has been reported in Algorithm 1.

Algorithm 1 Q-learning

Inputs: learning rate $\alpha \in [0, 1)$; discount rate $\gamma \in [0; 1]$; small $\varepsilon > 0$

Output: $Q(s, a)$

Initialise $Q(s, a)$, $\forall s$, $\forall a$
for all episodes **do**
  reset s
  for each step of the episode **do**
    choose action a following ε-greedy policy
    perform action a and observe s', r'
    perform Q-Learning update rule over $Q(s, a)$
    $s \leftarrow s'$
  end for
end for

It is worth highlighting the fact that, as just seen, MDPs offer a mathematical framework for simulating decision-making issues in dynamic contexts where actions' results are unpredictable and possibly contingent on the state of the environment at the time. In fact, in the context of satellite communications, variables like satellite movement and signal interference can cause the topology of the communication network to change dynamically. It is possible to depict these dynamic changes using the MDP formulation, representing them as shifts between various states, each of which is associated with a distinct communication network configuration. Hence, the RL framework is as a matter of fact well-suited for satellite communications problems, which present a network topology changing with time.

3 | MODELLING

Let us consider an FSO-like communication system made by two OGS networks, one transmitting data from N_{tr} OGSs and the other one acting as receiver with N_{re} stations. Each of the two zones can be subject to different atmospheric conditions, going from sunny to cloudy to stormy, which affect the data delivery from the transmitting to the receiving zone. The two sets of OGSs cannot communicate using terrestrial wired or wireless technologies, but they must rely on a LEO constellation composed of N_{sat} satellites. The communication is a point-to-point one realised through laser beams. The system scenario is depicted in Figure 1.

In what follows, a detailed mathematical modelling of the overall communication system is presented, including the formulation of the orbiting LEO satellite equations of motion, the ground-to-satellite and intersatellite visibility assessment and the MDP characterisation.

3.1 | Satellite equations of motion

Low Earth orbit satellites are considered one of the best options for satellite communication due to their short orbital period, which provides wide coverage and a high service availability. Usually, constellations of LEO satellites are used since singularly; they have a small field of view and can communicate with only a small fraction of Earth at a time.

In order to define a LEO constellation of satellites, the orbit itself must be characterised. Given an inertial frame of reference and an arbitrary epoch (a specified point in time), exactly six parameters are necessary to unambiguously define an arbitrary and unperturbed orbit. These are the semi-major axis a, the eccentricity e, the inclination i, the argument of perigee ω, the longitude of the ascending node Ω, also denoted as the Right Ascension of the Ascending Node (RAAN) for geocentric orbits, and the true anomaly f[54].[55]

The orbital parameters can be used to compute, at every epoch, the position and velocity of the satellite around that orbit. To describe the motion of spacecraft, it is usually used a coordinate frame which is inertial and fixed with respect to the stars, namely, the Earth centred inertial (ECI) reference frame.[56] In particular, the x-y plane coincides with the equatorial plane of Earth. The x-axis is permanently fixed in a direction relative to the celestial sphere, which does not rotate as Earth does. The z-axis lies at a 90° angle to the equatorial plane and extends through the North Pole (see Figure 2).

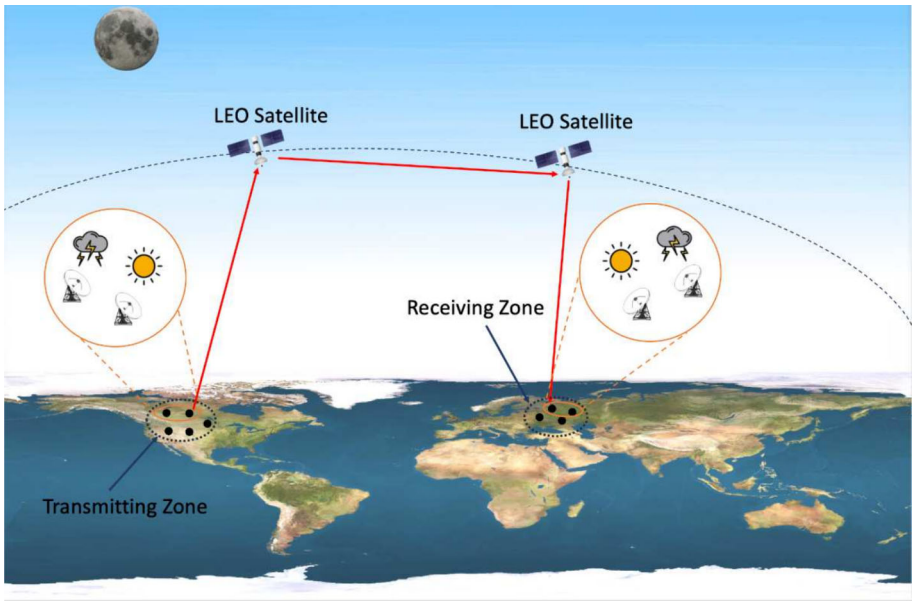

FIGURE 1 System scenario.

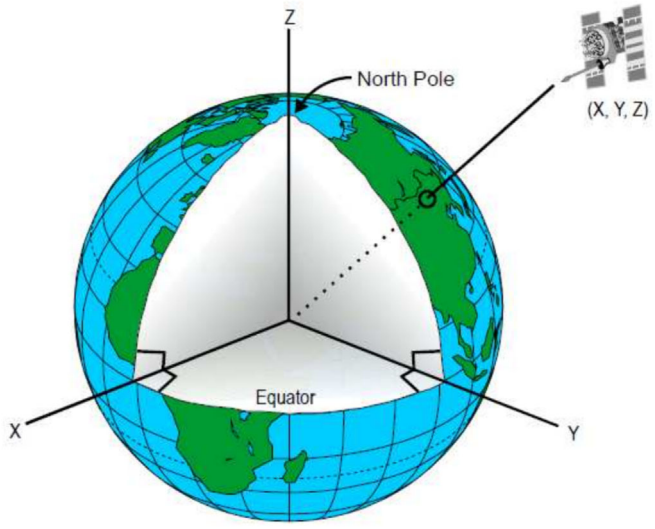

FIGURE 2 Earth centered inertial reference frame.

Let us define as $R_x(\phi), R_y(\eta)$ and $R_z(\psi)$ the standard rotation matrices:

$$R_x(\phi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos & \phi \sin \phi 0 - \sin \phi \cos \phi \end{bmatrix}, \tag{8}$$

$$R_y(\eta) = [\cos \eta 0 - \sin \eta 010 \sin \eta 0 \cos \eta], \tag{9}$$

$$R_z(\psi) = [\cos \psi \sin \psi 0 - \sin \psi \cos \psi 0001]. \tag{10}$$

Algorithm 2 shows how to pass from the orbital parameters to the satellite position and velocity in the ECI coordinates.

Algorithm 2 Orbital parameters to ECI coordinates

Inputs: $a, e, \Omega, i, \omega, f$

Parameters: $\mu = 3.986004418 \times 10^{14}[m^3/s^2]$

Outputs: r, v

$p = a(1 - e^2)$ ▷ semilatus rectum

$cf = \cos f, sf = \sin f$

$r = p/(1 + e(cf))$ ▷ Safe Division

$v = \sqrt{mi/p}$ ▷ Safe sqrt and safe division

Define a rotation matrix based on angles and axes

$\text{ang} = \begin{bmatrix} \omega & i & \Omega \end{bmatrix}^T$

$\text{axes} = \begin{bmatrix} 3 & 1 & 3 \end{bmatrix}^T$

$M = \text{ang2mat}(\text{ang}, \text{axes})$

Compute position and velocity in ECI

Transpose M

$r \leftarrow rM \begin{bmatrix} cf & sf & 0 \end{bmatrix}^T$

$v \leftarrow vM \begin{bmatrix} -sf & e + cf & 0 \end{bmatrix}^T$

At this point, it is possible to define the satellite equations of motion as a second-order differential equation which is dependent on the satellite position vector r:

$$\ddot{r} = \mu \frac{r}{||r||^3}, \tag{11}$$

where $||r||$ is the euclidean norm of the position vector and $\mu = 3.986004418 \times 10^{14}$ (m^3/s^2) is the geocentric gravitational constant.

3.2 | Visibility analysis

In order for the satellite to exchange information with a GS or with another satellite, there is a condition which needs to be analysed, the visibility.

In satellite communication, the visibility is a very important concept since it can determine if a certain information exchange can happen or not and how good is the communication channel in terms of noises. The former is related to the concept of geometric visibility which is named after the visibility definition for the Euclidean space: *Given a set of obstacles, two points in the space are said to be visible to each other, if the line segment that joins them does not intersect any obstacles*. In the space domain, this translates into the fact that the relative position vector between one satellite to the other does not have to intersect the Earth. The latter instead is related to the concept of electronic visibility which deals with analysing other parameters such as the elevation angle and carrier to noise ratio (C/N0). The angle of elevation is the angle between the horizontal line and the line of sight which is usually above the horizontal line. The C/N0 expresses how high is the noise component with respect to the information carrier: The lower the ratio is, the more the noise is prevalent and vice-versa.

Since the paper does not focus on the quality of the communication link, some assumptions have been made to simplify the following analysis.

1. The information exchanged between the satellite and the GS and between one satellite and another is always good with a negligible amount of noise;
2. The satellite is visible with respect to the GS if the elevation angle is greater than a certain threshold. This in order to exclude the case of interference of buildings in the vicinity of the GS;
3. The satellite is visible with respect to another one if the geometric visibility condition is satisfied, that is, if the relative position vector does not intersect the Earth surface.

In the following subsections, the implementation of the visibility algorithms related to assumptions 2 and 3 will be detailed.

3.2.1 | Ground station to satellite visibility

As already explained, a satellite is considered visible from a generic ground station if the elevation angle is above a certain threshold. The elevation angle is computed with respect to the horizontal plane of the GS, so the East-North-Up (ENU) coordinate frame has been considered, which is the reference frame of the ground station's antenna. This implies a change of coordinates of the satellite position and velocity vectors from the ECI reference frame to the ENU frame. This transformation can be performed by applying two rotations to the original coordinates: the first one to pass from the ECI to the Earth centered Earth fixed (ECEF) coordinates and the second one to pass from the ECEF to ENU coordinates.

Since the ECEF reference frame is noninertial and is rotating along with the Earth, a new dynamic equation must be introduced to take this rotation into account. So if we define θ as the angle of rotation of the Earth, the Earth rotational dynamics can be easily written as follows:

$$\dot{\theta} = \omega_E, \tag{12}$$

where $\omega_E = 2\pi/86400 \approx 7.29 \times 10^{-5}$ (rad/s) is the angular velocity of the Earth. In Algorithms 3 and 4 are detailed the steps to compute the two rotations. In the following, the notation x_{RF} with RF $\in$ {ECI, ECEF, ENU} denotes the reference frame of the generic vector x, while the notation $x_{\mathrm{RF},c}$ with $c \in \{x,y,z\}$ denotes the x, y and z components of the generic vector x expressed in the RF coordinates, respectively.

Algorithm 3 ECI to ECEF coordinates transformation

Inputs: $r_{\mathrm{ECI}}, v_{\mathrm{ECI}}, \theta$
Parameters: ω_E
Outputs: $r_{\mathrm{ECEF}}, v_{\mathrm{ECEF}}$

$R = R_z(\theta)$
$r_{\mathrm{ECEF}} = R r_{\mathrm{ECI}}$
$a = v_{\mathrm{ECI},x} + \omega_E r_{\mathrm{ECI},y}$
$b = v_{\mathrm{ECI},y} - \omega_E r_{\mathrm{ECI},x}$
$c = v_{\mathrm{ECI},z}$
$\tilde{v} = \begin{bmatrix} a & b & c \end{bmatrix}^T$
$v_{\mathrm{ECEF}} = R\tilde{v}$

Algorithm 4 ECEF to ENU coordinates transformation

Inputs: $r_{\mathrm{ECEF}}, \phi, \nu$ ▷ ϕ, ν: lat and long of the GS
Outputs: r_{ENU}

$$R = \begin{bmatrix} -\sin\nu & \cos\nu & 0 \\ -\cos\nu\sin\phi & -\sin\nu\sin\phi & \cos\phi \\ \cos\nu\cos\phi & \sin\nu\cos\phi & \sin\phi \end{bmatrix}$$

$r_{\mathrm{ENU}} = R r_{\mathrm{ECEF}}$

As the last step, from the ENU coordinates, it is possible to compute the Azimuth (A), Elevation (E) and Range (ρ) of the satellite w.r.t. the antenna. For our case, only the elevation angle will be used in the visibility analysis. In Algorithm 5 are detailed the mathematical steps to compute these three parameters.

Algorithm 5 ENU to azimuth, elevation, range parameters

Inputs: r_{ENU}
Outputs: A, E, ρ

norm = $||r_{\text{ENU}}||$
$\rho = r_{\text{ENU}}/\text{norm}$
$E = \arcsin \rho_z$ [rad]
$A = \arctan(\rho_x, \rho_y)$ [rad]

3.2.2 | Satellite to satellite visibility

Due to the short field of view of the LEO satellites, in order to exchange information between two sites far away from each other, a constellation of satellites is needed. This implies the creation of a communication link between two satellites of the same constellation in order to reach the remote site efficiently. The concept of visibility applies also in this case. To simplify the analysis, only the *geometric visibility* is considered. Algorithm 6 details the procedure for the geometric visibility check.

Algorithm 6 Geometric visibility check between satellite A and B

Inputs: r_A, r_B
Parameters: $R_{\text{Earth}} = 6378136.3$ [m]
Outputs: isSatVis [bool]

Initialise output
isSatVis = False
norm = $||r_A||$

if $r_A == r_B$ **then** ▷ Same point in space
    isSatVis = True
    return isSatVis
else
    $r_C = r_A - r_B$ ▷ Relative position vector
    min dist = Minimum distance between r_C and the centre of the Earth
    if min dist $\geq R_{\text{Earth}}$ **then**
        isSatVis = True
    end if
    return isSatVis
end if

3.3 | MDP formulation

The system dynamics described above has been translated to a MDP in order to exploit the RL framework.

The state space is

$$\mathcal{S} = < t >, t = 0, ..., T, \tag{13}$$

where t is the generic time step and T is the final step within the transmission window period.

The action space is

$$\mathcal{A} = < \mathrm{OGS_T}, \mathrm{SAT_1}, \mathrm{SAT_2}, \mathrm{OGS_R} >, \tag{14}$$

where $\mathrm{OGS_T}$ is the index of the transmitting OGS, $\mathrm{SAT_1}$ is the index of the first satellite receiving data from the transmitter, $\mathrm{SAT_2}$ is the index of the second satellite receiving data from the first one and $\mathrm{OGS_R}$ is the index of the receiver.

Eventually, the reward function models the success rate of the end-to-end handover and is defined in the following way:

$$R = \begin{cases} +1, \text{ if transmission is successful} \\ -1, \text{ otherwise.} \end{cases} \tag{15}$$

It is important to highlight the fact that, as the reward formulation suggests, the main goal of this work is to obtain the best link availability, without considering in the reward shaping other KPIs which may be relevant in general in space communication networks. One of them, as an example, would be the switching rate; that is, the amount of times it is decided to change path over the total number of instants in which the communication between the two end points is instantiated.

4 | SIMULATIONS AND RESULTS

In order to simulate and validate our control approach, two geographical areas from two different continents have been considered, namely,

1. the east coast of United States and Canada, with $N_{\mathrm{tr}} = 10$ transmitting OGSs located in the main cities, as in Figure 3;
2. the territory of Israel, with $N_{\mathrm{re}} = 6$ receiving OGSs, shown in Figure 4.

To perform the weather forecast for all the OGS zones, the LSTM deep neural network was trained on a publicly available weather dataset[57] covering approximately 5 years of weather data (from 1 October 2012 to 30 November 2017), with temporal resolution $T_R = 1\mathrm{h}$. Having this in mind, being T_R quite high (i.e. the acquisition frequency of time-series data about weather conditions is low), it is possible to neglect time lags due to data processing. This is also known as the time required to perform inference over the pre-trained neural network model. The available features for training are the following:

- humidity
- atmospheric pressure

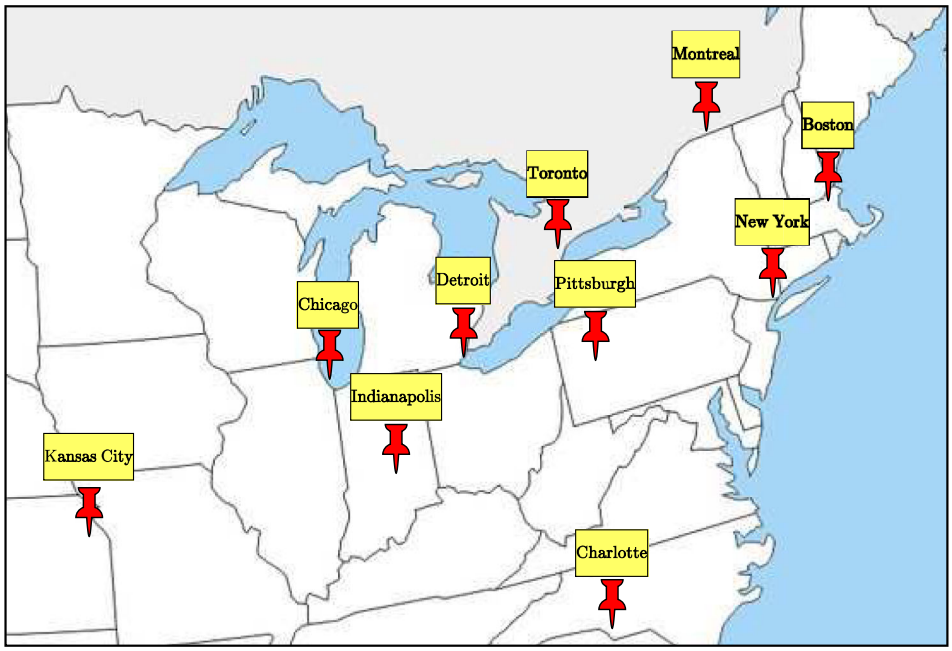

FIGURE 3 Map of the transmitting optical ground stations (OGSs) in the east coast of North America.

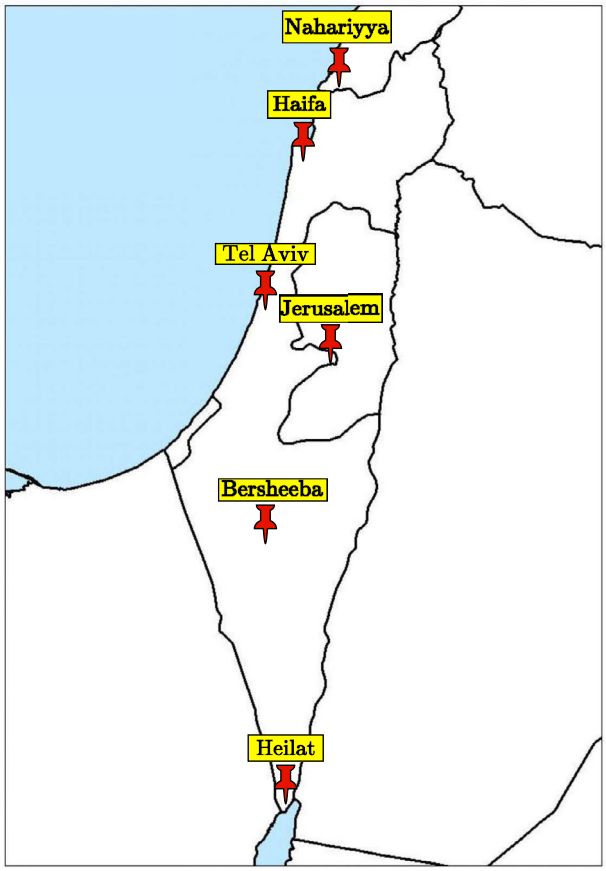

FIGURE 4 Map of the receiving optical ground stations (OGSs) in Israel.

- wind direction
- temperature
- wind speed
- month of the year
- weather conditions within the time window prediction $\mathcal{T}_p$.

The missing data for each feature were filled in by taking up the numerical value of the feature of the previous entry: This approach makes it possible not to break the hourly time sequence of the meteorological data.

As per the weather, the dataset contains a very detailed description of the weather conditions. The latter have been mapped into binary labels for training the model, thus leading to a binary classification problem. In particular, note that label 0 has been assigned to clear sky, few/scattered clouds and haze, which correspond to mild weather conditions allowing satellite-OGS communication, whereas label 1 indicates inclement weather (rain, thick clouds, snow and others) which does not allow a successful data transmission.

For the training phase, the data from 1 October 2012 up to 20 December 2016 have been selected. The model accuracy has been evaluated by splitting the remaining part of the dataset with respect to the four seasons, in accordance with the 2016 and 2017 astronomical tables.[58]

The chosen LSTM model architecture is depicted in Figure 5. It is worth noting that, like any other Machine Learning model, LSTM may be affected by the well-known drawback known as overfitting, consisting into overadapting to the training dataset, being not able to generalise to unseen data, hence showing poor performance when classifying samples within the test set. To prevent this, as it is possible to see in Figure 5, the dropout regularisation technique has been implemented.[59] During training, the dropout layers coming after each LSTM layer randomly select with a certain probability ζ neurons to be discarded (i.e. set to zero). This process helps prevent the network from relying too heavily on any individual neuron or feature, forcing it to learn more robust representations.

The selected hyperparameters of the LSTM neural networks are the following:

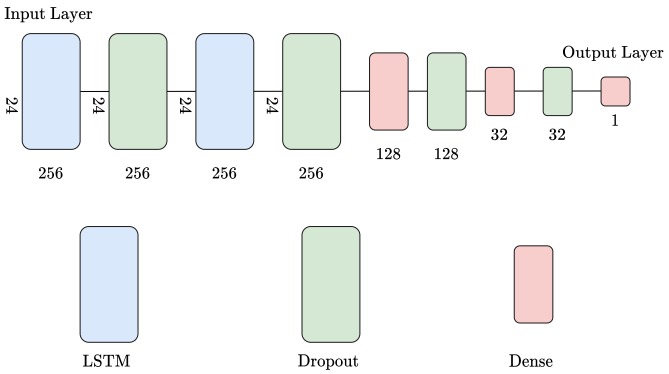

FIGURE 5 Long short-term memory (LSTM) neural network architecture. The network is made by two LSTM layers and two dense layers, each one followed by a dropout layer, with the final output layer having one neuron representing bad weather probability.

TABLE 1 LSTM accuracies per season.

City	Winter	Spring	Summer	Autumn
New York	0.984	0.986	0.982	0.989
Montreal	0.988	0.981	0.963	0.974
Boston	0.993	0.989	0.987	0.989
Chicago	0.973	0.978	0.952	0.971
Charlotte	0.976	0.976	0.972	0.975
Pittsburgh	0.991	0.984	0.979	0.988
Detroit	0.997	1.000	0.996	0.998
Kansas City	0.992	0.990	0.983	0.990
Toronto	0.999	0.999	0.985	0.988
Indianapolis	0.992	0.986	0.981	0.992
Beersheba	0.944	0.956	0.966	0.998
Tel Aviv District	0.977	0.955	0.966	0.991
Eilat	0.932	0.923	0.968	0.999
Haifa	0.944	0.982	0.987	0.999
Nahariyya	0.989	0.985	0.945	0.997
Jerusalem	0.991	0.981	0.959	0.998

Abbreviation: LSTM, long short-term memory.

- number of epochs $E = 5$
- Adam optimiser with constant learning rate $\eta = 0.001$
- time window length $T_W = 24$
- dropout rate $\zeta = 0.2$.

The LSTM model performance on unseen data (from 21 December 2016 to 30 November 2017) against all seasons and per each city, in terms of test accuracy, have been reported in Table 1. It is evident that the neural network model is able to predict correctly almost all the weather condition within the test set, thus representing a powerful tool to estimate in advance the precipitation or thick clouds probability over the zone in which the OGS is located.

The RL-based controller hyperparameters for the training phase have been selected as follows:

- $\gamma = 0.9$
- $\varepsilon_0 = 1$ with episodic decay law with respect to the generic episode η

$$\varepsilon(\eta) = \mathrm{e}^{-\frac{\eta}{\beta N_{\mathrm{ep}}}},$$

with $\beta = 0.2$ being the decay rate and N_{ep} the number of episodes.

- $\alpha_0 = 1$ with episodic decay law with respect to the generic episode η

$$\alpha(\eta) = \mathrm{e}^{-\frac{\eta}{1000}}.$$

As for the evaluation phase, the controller performance has been figured out over a transmission period of $T = 2$ days.

The AI-based control law has been evaluated in terms of link availability, defined as follows:

$$L_A = \frac{N_S}{N_T}, \tag{16}$$

where N_S is the number of times a successful data transmission is achieved and N_T is the total number of transmissions attempted.

Results with respect to the above-defined KPI have been compared with other benchmark routing approaches in the FSO domain, listed as follows:

- B1. Both transmitting and receiving OGSs and both LEO satellites are chosen randomly.
- B2. Transmitting and receiving OGSs are chosen with a reactive approach based on the current weather condition, and the satellites are chosen with the min range technique, following the reasoning and modelling provided in Henniger and Wilfert.[60]
- B3. Transmitting and receiving OGSs are chosen with a reactive approach based on the current weather condition, and the satellites are chosen as those with the maximum elevation.
- B4. Transmitting and receiving OGSs are chosen with the LSTM-based weather forecasts, and the satellites are chosen with the min range technique.
- B5. Transmitting and receiving OGSs are chosen with the LSTM-based weather forecasts, and the satellites are chosen as those with the maximum elevation.

The proposed control approach has been tested over three different case studies, in which the communication between the two OGS networks is realised with different LEO constellations:

- Case study 1. $N_{\mathrm{sat}} = 15$ satellites from the Iridium constellation
- Case study 2. $N_{\mathrm{sat}} = 15$ satellites from the Starlink constellation
- Case study 3. $N_{\mathrm{sat}} = 30$ satellites given by the combination of satellites from case study 1 and case study 2.

The satellite orbital parameters and generic data have been gathered via two-line elements (TLEs) files from CelesTrak.[61] A TLE file is a data format encoding a list of orbital elements of an Earth-orbiting object for a given point in time.

The propagation of the satellites motion over time is performed by using the fourth-order Runge–Kutta algorithm as integrator with fixed time step $dt = 1$ min.

4.1 | Iridium constellation

In this case study, the number of episodes for training the RL controller has been set as $N_{\mathrm{ep}} = 100$. The season-related cumulative reward trend over the training episodes is shown in Figure 6. In all the four cases, the reward converges to a steady-state value in terms of data transmission success rate, which is higher in the autumn season due to the presence of a higher number of hours with favourable weather conditions both at transmitting and receiving zones.

The comparison of the performance of the proposed approach with respect to the benchmark solutions is shown in Figure 7. It is worth noting that the RL controller together with an LSTM-based weather prediction achieves higher link availability with respect to the other standard techniques for FSO communication.

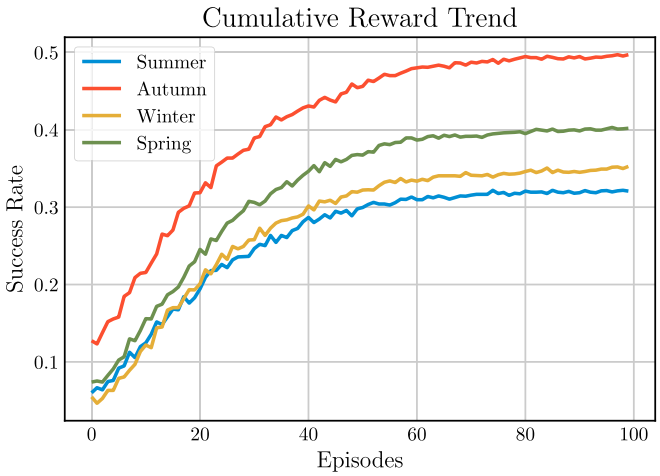

FIGURE 6 Season-related reward trend of the reinforcement learning (RL) controller for the Iridium case study.

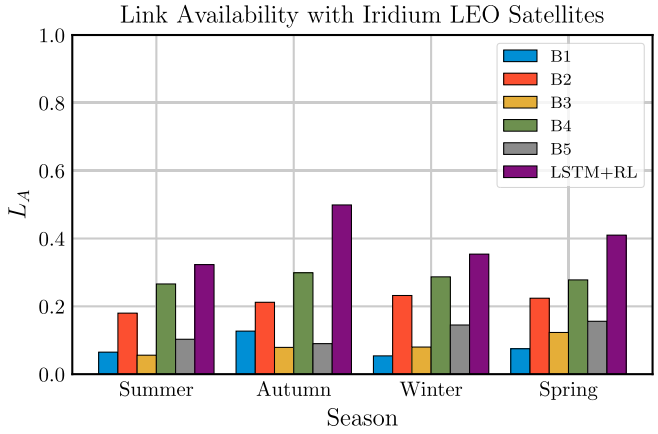

FIGURE 7 Season-related link availability comparison for the Iridium case study.

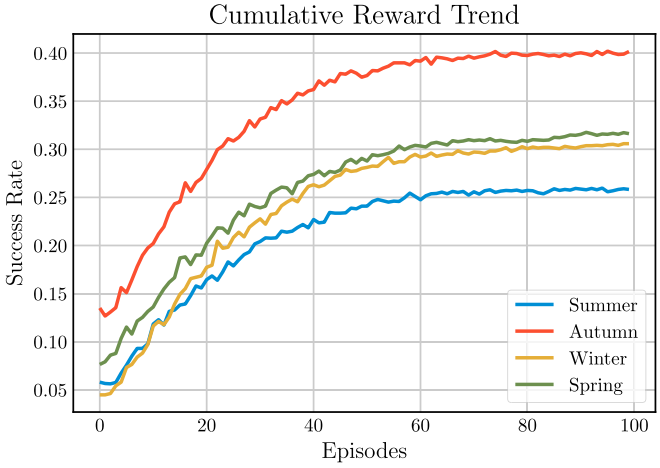

FIGURE 8 Season-related reward trend of the reinforcement learning (RL) controller for the Starlink case study.

4.2 | Starlink constellation

In this case study, the number of episodes for training the RL controller has been set as $N_{ep} = 100$. The cumulative reward trend is shown in Figure 8. For all seasons, the reward converges to a steady-state value in terms of data transmission success rate, as in the previous case study with the Iridium constellation.

Figure 9 depicts the comparison of the performance of the RL controller with respect to the benchmark solutions. The proposed control strategy achieves the best performances in terms of link availability. However, it shall be noticed that the overall performances are worse with respect to those achieved by means of the Iridium constellation. As an example, in the autumn season, the RL controller guarantees $L_A = 0.401$ using Starlink satellites and $L_A = 0.499$ with iridium satellites: Similar results hold for the other seasons. This is due to the fact that the Starlink constellation orbits have been designed to cover mainly the North America continent, thus guaranteeing poor coverage within the Israel territory.

4.3 | Mixed constellation

In the last case study, the number of episodes for training the RL controller has been increased to $N_{ep} = 300$, in order to allow a broader exploration due to the availability of double the amount of LEO satellites with respect to the previous case studies.

Figures 10 and 11 show training and evaluation performances against all seasons. As expected, the increased number of satellites guarantees higher cumulative reward trend and, hence, link availability for the FSO transmission. This is due to the fact that increasing the number of satellites leads to a wider coverage over the Earth surface: This allows to establish a successful communication with guaranteed intersatellite visibility for longer periods. Also in this case, the performances of the proposed control algorithm are better than the benchmark ones.

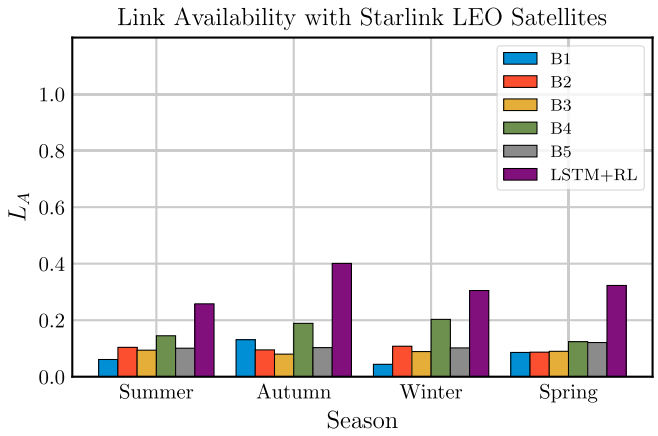

FIGURE 9 Season-related link availability comparison for the Starlink case study.

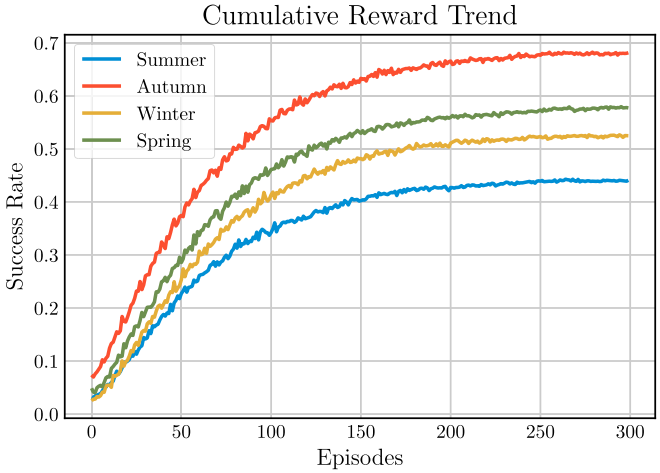

FIGURE 10 Season-related reward trend of the reinforcement learning (RL) controller for the mixed case study.

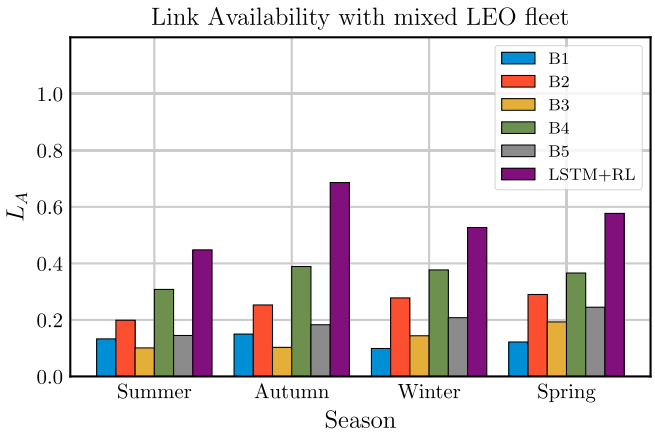

FIGURE 11 Season-related link availability comparison for the mixed case study.

5 | CONCLUSIONS

In this work, a mixed AI and RL approach for FSO communication has been proposed. This technique exploits weather prediction algorithms to improve the quality of the communication link, as well as a dynamical data-driven optimisation for maximising the link availability in a data transmission scenario between two terrestrial OGS networks communicating through LEO satellites. The proposed decision and control approach have been compared with several benchmark solutions, achieving better performances in all seasons over the three analysed case studies in which different LEO constellations have been exploited. This methodology is also scalable to larger networks comprising more satellites and/or more OGSs, since (i) the policy is gathered through offline learning and (ii) the decision about the path is made by direct inference, which can be realised with real-time computations.

However, some limitations hold. First of all, a 1-h temporal resolution may lead to inaccurate predictions, since in some cases clouds can move more quickly than 1 h from one zone to the other.

Secondly, despite the superiority of the proposed combined LSTM-RL approach in terms of link availability, the formulation of the reward as in (15) leads to a control algorithm which may select different actions at every time step, hence showing poor performance in terms of switching rate.

Future works may focus on the limitations defined above, proposing the exploitation of weather datasets with higher meteorological acquisition frequency and with different on-ground site diversity configurations. A possible goal would be synthesising reward policies capable of finding a trade-off optimisation between the link availability and the switching rate or possibly other KPIs relevant in the context of optical satellite communications. Moreover, it would be interesting to make the satellite network robust against cyberattacks, by implementing redundant checks and additional security measures on the input data, for example, introducing a federated and distributed approach for routing the data from the source to destination.

ORCID

Andrea Wrona https://orcid.org/0000-0002-4210-2641
Andrea Tantucci https://orcid.org/0000-0003-4322-9086

REFERENCES

1. Yaqoob A, Bi T, Muntean G-M. A survey on adaptive 360 video streaming: solutions, challenges and opportunities. *IEEE Commun Surv Tutor*. 2020; 22(4):2801-2838.
2. Sadiku MNO, Musa SM, Momoh OD. Cloud computing: opportunities and challenges. *IEEE Potentials*. 2014;33(1):34-36.
3. Li S, Xu LD, Zhao S. The internet of things: a survey. *Inform Syst Front*. 2015;17:243-259.
4. Kaushal H, Jain VK, Kar S. *Free Space Optical Communication*: Springer; 2017.
5. Ball DW. The electromagnetic spectrum: a history. *Spectroscopy*. 2007;22(3):14.
6. Killinger D. Free space optics for laser communication through the air. *Opt Photon News*. 2002;13(10):36-42.
7. Burchardt H, Serafimovski N, Tsonev D, Videv S, Haas H. VLC: beyond point-to-point communication. *IEEE Commun Mag*. 2014;52(7):98-105.
8. Biswas A, Srinivasan M, Piazzolla S, Hoppe D. Deep space optical communications. In: Free-space Laser Communication and Atmospheric Propagation XXX, Vol. 10524 SPIE; 2018:242-252.
9. Hauschildt H, Elia C, Moeller HL, Schmitt D. ScyLight ESA's secure and laser communication technology framework for SatCom. In: 2017 IEEE International Conference on Space Optical Systems and Applications (ICSOS) IEEE; 2017:250-254.
10. Andrews LC, Phillips RL, Hopen CY, Al-Habash MA. Theory of optical scintillation. *JOSA A*. 1999;16(6):1417-1429.

11. Berman GP, Chumak AA, Gorshkov VN. Beam wandering in the atmosphere: the effect of partial coherence. *Phys Rev E*. 2007;76(5):056606.
12. Wang Z, Zhang J, Gao H. Impacts of laser beam divergence on lidar multiple scattering polarization returns from water clouds. *J Quantit Spectrosc Radiat Transfer*. 2021;268:107618.
13. Chitre P, Yegenoglu F. Next-generation satellite networks: architectures and implementations. *IEEE Commun Mag*. 1999;37(3):30-36.
14. Mao X, Arnold D, Girardin V, Villiger A, Jäggi A. Dynamic GPS-based leo orbit determination with 1 cm precision using the bernese GNSS software. *Adv Space Res*. 2021;67(2):788-805.
15. Lim CB, Montmerle-Bonnefois A, Petit C, Sauvage J-F, Meimon S, Perrault P, Mendez F, Fleury B, Montri J, Conan J-M. Single-mode fiber coupling with adaptive optics for free-space optical communication under strong scintillation. In: 2019 IEEE International Conference on Space Optical Systems and Applications (ICSOS) IEEE; 2019:1-6.
16. Zhang C, Jin J, Kuang L, Yan J. LEO constellation design methodology for observing multi-targets. *Astrodynamics*. 2018;2:121-131.
17. Tantucci A, Wrona A, Pietrabissa A. Precise orbit determination on leo satellite using pseudorange and pseudorange-rate measurements. In: 2023 31st Mediterranean Conference on Control and Automation (MED) IEEE; 2023:341-347.
18. Su Y, Liu Y, Zhou Y, Yuan J, Cao H, Shi J. Broadband leo satellite communications: architectures and key technologies. *IEEE Wirel Commun*. 2019;26(2):55-61.
19. Yue P, An J, Zhang J, Ye J, Pan G, Wang S, Xiao P, Hanzo L. Low earth orbit satellite security and reliability: issues, solutions, and the road ahead. *IEEE Commun Surv Tutor*. 2023.
20. Chowdhury PK, Atiquzzaman M, Ivancic W. Handover schemes in satellite networks: state-of-the-art and future research directions. *IEEE Commun Surv Tutor*. 2006;8(4):2-14.
21. He S, Wang T, Wang S. Load-aware satellite handover strategy based on multi-agent reinforcement learning. In: Globecom 2020-2020 IEEE Global Communications Conference IEEE; 2020:1-6.
22. Kasper M, Fedrigo E, Looze DP, Bonnet H, Ivanescu L, Oberti S. Fast calibration of high-order adaptive optics systems. *JOSA A*. 2004;21(6):1004-1008.
23. Tyson RK, Frazier BW. *Principles of Adaptive Optics*: CRC press; 2022.
24. Vosselman G, Maas H. Adjustment and filtering of raw laser altimetry data. In: Proceedings of OEEPE Workshop on Airborne Laserscanning and Interferometric Sar for Detailed Digital Terrain Models, Stockholm, Sweden OEEPE; 2001.
25. Jolly SW, Gobert O, Quéré F. Spatio-temporal characterization of ultrashort laser beams: a tutorial. *J Opt*. 2020;22(10):103501.
26. Wrona A, De Santis E, Priscoli FD, Lavacca FG. An intelligent ground station selection algorithm in satellite optical communications via deep learning. In: 2023 31st Mediterranean Conference on Control and Automation (MED) IEEE; 2023:493-499.
27. Rödiger B, Ginthör D, Labrador JP, Ramirez J, Schmidt C, Fuchs C. Demonstration of an FSO/RF hybrid-communication system on aeronautical and space applications. In: Laser Communication and Propagation Through the Atmosphere and Oceans IX, Vol. 11506 SPIE; 2020:1150603.
28. Sinka C, Bitó J. Site diversity against rain fading in LMDS systems. *IEEE Microw Wirel Components Lett*. 2003;13(8):317-319.
29. Nakatani T, Maekawa Y, Shibagaki Y, Hatsuda K. Relationship between rain front motion and site diversity in ku-band satellite links. In: 25th AIAA International Communications Satellite systems Conference (Organized by APSCC) APSCC; 2007:3173.
30. Petropoulou P, Michailidis ET, Panagopoulos AD, Kanatas AG. Radio propagation channel measurements for multi-antenna satellite communication systems: a survey. *IEEE Anten Propag Mag*. 2014;56(6):102-122.
31. Baptista JPVP, Davies PG. Reference book on attenuation measurement and prediction. In: 2nd Workshop Opex OPEX; 1994.
32. Goldhirsh J, Musiani BH, Dissanayake AW, Lin K-T. Three-site space-diversity experiment at 20 GHz using acts in the Eastern United States. *Proc IEEE*. 1997;85(6):970-980.
33. Lin SH, Bergmann HJ, Pursley MV. Rain attenuation on earth-satellite paths-summary of 10-year experiments and studies. *Bell Syst Techn J*. 1980;59(2):183-228.
34. Luglio M, Mancini R, Riva C, Paraboni A, Barbaliscia F. Large-scale site diversity for satellite communication networks. *Int J Satellite Commun*. 2002;20(4):251-260.
35. Bruni C, Priscoli FD, Koch G, Pietrabissa A, Pimpinella L. Network decomposition and multi-path routing optimal control. *Trans Emerg Telecommun Technol*. 2012;24(2):154-165.
36. Pietrabissa A, Celsi LR, Cimorelli F, Suraci V, Priscoli FD, Giorgio AD, Giuseppi A, Monaco S. Lyapunov-based design of a distributed wardrop load-balancing algorithm with application to software-defined networking. *IEEE Trans Control Syst Technol*. 2019;27(5):1924-1936.
37. Chiariotti F, Vikhrova O, Soret B, Popovski P. Information freshness of updates sent over leo satellite multi-hop networks. arXiv preprint arXiv:200705449; 2020.
38. Fang Z, Wang J, Ren Y, Han Z, Poor HV, Hanzo L. Age of information in energy harvesting aided massive multiple access networks. *IEEE J Sel Areas Commun*. 2022;40(5):1441-1456.
39. Poulenard S, Crosnier M, Rissons A. Ground segment design for broadband geostationary satellite with optical feeder link. *J Opt Commun Netw*. 2015;7(4):325-336.
40. Rossi T, De Sanctis M, Maggio F, Ruggieri M, Hibberd C, Togni C. Smart gateway diversity optimization for EHF satellite networks. *IEEE Trans Aerosp Electron Syst*. 2019;56(1):130-141.
41. Efrem CN, Panagopoulos AD. Globally optimal selection of ground stations in satellite systems with site diversity. *IEEE Wirel Commun Lett*. 2020;9(7):1101-1104.
42. Fuchs C, Moll F. Ground station network optimization for space-to-ground optical communication links. *J Opt Commun Netw*. 2015;7(12):1148-1159.
43. Lyras NK, Efrem CN, Kourogiorgas CI, Panagopoulos AD. Optimum monthly based selection of ground stations for optical satellite networks. *IEEE Commun Lett*. 2018;22(6):1192-1195.
44. Erdogan E, Altunbas I, Kurt GK, Bellemare M, Lamontagne G, Yanikomeroglu H. Site diversity in downlink optical satellite networks through ground station selection. *IEEE Access*. 2021;9:31179-31190.
45. Zhu R, Li G, Zhang Y, Fang Z, Wang J. Load-balanced virtual network embedding based on deep reinforcement learning for 6g regional satellite networks. *IEEE Trans Veh Technol*. 2023.
46. Hu X, Zhang Y, Liao X, Liu Z, Wang W, Ghannouchi FM. Dynamic beam hopping method based on multi-objective deep reinforcement learning for next generation satellite broadband systems. *IEEE Trans Broadcast*. 2020;66(3):630-646.

47. Hu X, Liu S, Wang Y, Xu L, Zhang Y, Wang C, Wang W. Deep reinforcement learning-based beam hopping algorithm in multibeam satellite systems. *IET Commun*. 2019;13(16):2485-2491.
48. Liu S, Hu X, Wang W. Deep reinforcement learning based dynamic channel allocation algorithm in multibeam satellite systems. *IEEE Access*. 2018;6: 15733-15742.
49. Hu X, Liu S, Chen R, Wang W, Wang C. A deep reinforcement learning-based framework for dynamic resource allocation in multibeam satellite systems. *IEEE Commun Lett*. 2018;22(8):1612-1615.
50. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput*. 1997;9(8):1735-1780.
51. Li M, Zhang T, Chen Y, Smola AJ. Efficient mini-batch training for stochastic optimization. In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining ACM; 2014:661-670.
52. Sutton RS, Barto AG. *Reinforcement Learning: An Introduction*: MIT press; 2018.
53. Watkins CJCH, Dayan P. Q-learning. *Mach Learn*. 1992;8:279-292.
54. Seeber G. *Satellite Geodesy*: Walter de gruyter; 2003.
55. Gill OME, Montenbruck O. *Satellite Orbits*: Springer; 2013.
56. Ashby N. The sagnac effect in the global positioning system. *Relativity in Rotating Frames: Relativistic Physics in Rotating Reference Frames*: Springer; 2004:11-28.
57. Beniaguev D. Historical hourly weather data 2012-2017. Accessed June 28, 2023; 2017.
58. GMT. Astronomical tables for equinoxes and solstices. Accessed July 6, 2023.
59. Do Santos CFG, Colombo D, Roder M, Papa JP. Maxdropout: deep neural network regularization based on maximum output values. In: 2020 25th International Conference on Pattern Recognition (ICPR) IEEE; 2021:2671-2676.
60. Henniger H, Wilfert O. An introduction to free-space optical communications. *Radioengineering*. 2010;19(2).
61. CelesTrak. Norad GP element sets current data. Accessed: 2023-09-09.

AUTHOR BIOGRAPHIES

Andrea Wrona received his Bachelor, Master's Degree and Doctoral Degree in Control Engineering from Sapienza University of Rome in 2018, 2020 and 2024, respectively. He is currently a Postdoctoral Researcher in Automatic Control within the same institution. His research interests include data-driven control methods applied to terrestrial and satellite-based telecommunication systems and the design of decision support strategies for the healthcare sector.

Andrea Tantucci received his Bachelor, Master's Degree and Doctoral Degree in Control Engineering from Sapienza University of Rome in 2018, 2020 and 2024, respectively. He is currently a Navigation System Engineer within Thales Alenia Space Italy. His research interests include AI-based approaches to space applications like satellite navigation and Earth Observation.

How to cite this article: Wrona A, Tantucci A. Artificial intelligence-based data path control in low Earth orbit satellites-driven optical communications. *Int J Satell Commun Network*. 2024;1-19. doi:10.1002/sat.1528