electronics

MDPI

Article

Neural Network SNR Prediction for Improved Spectral Efficiency in Land Mobile Satellite Networks

Ivan Vajs [1,*], Srđan Brkić [2], Predrag Ivaniš [3] and Dejan Drajic [1,3]

[1] Innovation Center of the School of Electrical Engineering in Belgrade, Bulevar Kralja Aleksandra 73, 11120 Belgrade, Serbia; dejan.drajic@etf.bg.ac.rs
[2] Tannera Technologies LLC, Veljka Dugosevica 54, 11000 Belgrade, Serbia; srdjan@tannera.io
[3] School of Electrical Engineering, University of Belgrade, Bulevar Kralja Aleksandra 73, 11120 Belgrade, Serbia; predrag.ivanis@etf.bg.ac.rs
* Correspondence: ivan.vajs@ic.etf.bg.ac.rs; Tel.: +381-113-218-455

Abstract: The use of satellites to cover remote areas is a promising approach for increasing communication availability and reliability. The satellite resources, however, can be quite costly, and developing ways to optimize their usage is of great interest. Optimizing spectral efficiency while keeping the transmission error rate above a certain threshold represents one of the crucial aspects of resource optimization. This paper provides a novel strategy for adaptive coding and modulation (ACM) employment in land mobile satellite networks. The proposed solution incorporates machine learning techniques to predict channel state information and subsequently increase the overall spectral efficiency of the network. The Digital Video Broadcasting Satellite Second Generation (DVB-S2X) satellite protocol is considered as the use case, and by using the developed channel simulator, this paper performs an evaluation of the proposed machine learning solutions for channels with various characteristics, with a total of 90 different observed channels. The results show that a convolutional neural network with a modified loss function consistently achieves an improvement (over 100% in some scenarios) of spectral efficiency compared to the state-of-the-art ACM implementation while keeping the transmission error rate under 0.01 for single channel evaluation. When observing two channels, an improvement of more than 300% compared to the outdated information spectral efficiency was obtained in multiple scenarios, showing the effectiveness of the proposed approach and allowing optimization of the handover strategy in satellite networks that allow user-centric handover executions.

Keywords: land mobile satellite communications; DVB-S2X protocol; channel state prediction; spectral efficiency; SNR; machine learning; neural networks

Citation: Vajs, I.; Brkić, S.; Ivaniš, P.; Drajic, D. Neural Network SNR Prediction for Improved Spectral Efficiency in Land Mobile Satellite Networks. *Electronics* **2024**, *13*, 3659. https://doi.org/10.3390/electronics13183659

Academic Editor: Reza K. Amineh

Received: 11 August 2024
Revised: 4 September 2024
Accepted: 10 September 2024
Published: 14 September 2024

1. Introduction

The rapid development of digital technologies in different application areas demands an expansion of infrastructure to support the communication needs of the immense number of newly introduced devices. Whether it is broadband internet services [1], mobile communication [2], the concept of smart cities or smart homes [3], maritime and aviation communication [4,5], etc., stable, reliable and fast communication is essential. For certain applications and scenarios, terrestrial 5G networks offer more than suitable conditions, but on the other hand, based on specific needs, communication through satellite networks can be used to overcome the existing shortcomings of other available communication systems [6,7].

There are many types of satellites that can be used for establishing communications, and they can be generally divided into four categories based on their type of orbit: Geostationary Earth Orbit (GEO), Low Earth Orbit (LEO), Medium Earth Orbit (MEO), and Highly Elliptical Orbit (HEO) satellites [8]. Because of their consistently low distance

from the Earth's surface, LEO satellites offer the smallest amount of latency and higher throughput, which can be crucial for data transmission in many different applications.

Although LEO satellites offer low latency and potentially high throughput, they present certain challenges. They have a high speed relative to the Earth's surface, and although they can cover remote areas where there is no other infrastructure, their speed can cause a Doppler shift in frequency that needs to be accounted for. In addition, since they move quickly, it is necessary to have appropriate handover procedures since the LEO satellites circle the Earth in less than two hours, so from a single point, they are only visible for a few minutes [9]. Signal quality improvement in LEO satellite systems can also be performed from many different viewpoints, whether it is through antenna design, beamforming algorithms, channel state or signal-to-noise ratio (SNR) prediction, etc.; there are many opportunities for optimization in LEO satellite communication.

When considering the practical applications of satellite communication, it is important not only to consider the SNR or the channel state information (CSI) [10] but also consider the protocol that is used for communication, and in the proposed research, the Digital Video Broadcasting Satellite Second Generation extended (DVB-S2X) protocol [11] is adopted. The initial version of the protocol was developed for GEO satellite communication and is mostly used for video broadcasting, but the DVB-S2X also supports data transmission, which is more sensitive to propagation delay, and it can be applied to LEO satellite systems. The DVB-S2X offers a range of spectral efficiencies (0.2–5.6 b/s/Hz), and its flexibility in terms of adaptive coding and modulation allows for dynamic optimization of bandwidth usage, opening up many possibilities for optimization.

Considering that the optimization of communication processes can be quite complex and that ample amounts of data are available through both real communication setups and simulation implementations, machine learning (ML) presents itself as a promising approach that can improve communication performance in different scenarios [12]. ML algorithms rely on available data and are often used to perform direct input–output mappings. The process of ML parameter optimization that results in the mapping between the algorithms' inputs and outputs is called training, and this process is performed based on available data by trying to minimize the optimization (loss) function of the system. The complexity of possible mappings depends on the selection of the algorithm and available data, and a subset of ML that can provide not only simple mappings but mappings that consider additional limitations/constrictions and goals is called neural networks [13]. Neural networks take inspiration from the structure of the human brain and can have both simple and extremely complex architectures, depending on the intended application. The resources that neural networks require both in terms of available data for training and computation power can be extensive, but the results obtained using neural networks present state-of-the-art results in many fields, ranging from image [14,15] and signal [16,17] processing all the way to large language models [18] and data synthesis [19]. In line with everything previously stated, neural networks can often be found in the communication optimization literature alongside other conventional optimization techniques.

1.1. Related Work

In [20], the authors propose a joint user scheduling and beamforming algorithm for LEO satellites following the DVB-S2X standard [11]. The authors consider the influence of Doppler shift and phase disturbance on the CSI estimation, but no prediction of the CSI is performed. The optimization method is based on hierarchical clustering, the semidefinite programming algorithm and the concave–convex process (CCCP). The focus is placed on the optimization of massive multiple input multiple output (MIMO) communication systems, and the authors report that the observed metric, energy efficiency, is higher for the proposed solution when compared to the traditional decoupling design algorithms. The optimization of scheduling in satellite communications has also been analyzed with different optimization scenarios in [21,22] without focusing on the influence various factors can have on the CSI. The authors in [21] aimed to maximize the sum rate under the per-

beam power constraint and minimum SINR requirement of scheduled users, proposing a convex–concave procedure-based algorithm. The study described in [22] rather focuses on the separate analysis of intra-beam and inter-beam scheduling and takes the interference among beams into consideration during the scheduling. The authors in [23] observed the ergodic capacity of satellite-terrestrial links in the presence of co-channel interference and outdated CSI and presented an analytical form of the achievable ergodic capacity. The analysis was carried out on channels with both uncorrelated and correlated fading and the numerical and simulation results validate the analysis proposed by the authors. The estimation of a massive MIMO orthogonal frequency division multiplexing channel for LEO satellite communication is presented in [24]. The authors show that an asymptotic minimum mean square error (MMSE) of the estimation can be minimized under certain array response vector conditions. In addition, the authors proposed a two-stage channel estimation, i.e., per-subcarrier space domain processing, followed by per-user frequency domain processing. The simulation results show that the proposed solution can achieve a performance near the MMSE with a much lower complexity.

In terms of CSI prediction, the authors in [25] propose a scheme based on a deep neural network to assist LEO satellite massive MIMO systems. The implemented neural networks contain a convolutional and long short-term memory module, which is used to analyze the relationship of uplink–downlink channels between LEO satellites, the mapping between CSI and beamformers, and directly predicting the future downlink CSI of LEO satellites based on the observed uplink CSI. The implemented architectures could also generate the appropriate beamformers based on the predicted downlink CSI. The implemented systems are validated through numerical methods and show a promising application of neural networks in complex communication systems. A comparison between different machine learning methods and a modified autoregressive integrated moving average (ARIMA) model for CSI prediction is presented in [26]. The authors compare the performance of the proposed improved ARIMA model to neural networks (convolutional and long short-term memory (LSTM)), a Markov chain model and a basic ARIMA algorithm. The simulation shows that the performance of the improved ARIMA model has a significant increase in spectral efficiency compared to the other implemented prediction models. CSI prediction was analyzed through the influence of atmospheric attenuation in [27]. The authors implemented several types of neural networks, including a convolutional neural network, LSTM and a multilayer perceptron. The performed simulations show that the prediction of the interference period and types of atmospheric attenuation using the proposed architecture can be effectively obtained. In [28], the authors analyzed the massive MIMO channel model combined with LEO satellite characteristics and proposed an approach based on the LSTM neural network for a prediction scheme to mitigate the effects of outdated CSI. The proposed architecture achieved higher prediction accuracy compared to the adopted outdated CSI.

A review of the methods used for communication optimization in the work presented in the literature is given in Table 1.

Table 1. A review of related work.

Reference	Communication Method	Optimization Criteria	Optimization Method
[20] Liu et al.	LEO satellites and DVB S2X	energy efficiency	hierarchical clustering, the semidefinite programming algorithm and the CCCP
[21] Bandi et al.	satellite communications and DVB-S2X	maximum sum rate under the per-beam power constraint and minimum SINR requirement of scheduled users	convex–concave procedure-based algorithm

Table 1. *Cont.*

Reference	Communication Method	Optimization Criteria	Optimization Method
[22] Zhang et al.	high throughput GEO satellite	Fairness among users and spectral efficiency	fixed-size user grouping algorithm, calculation of equivalent CSIs and frame scheduling
[23] Bankey et al.	multiuser hybrid satellite-terrestrial amplify and forward relay network	the ergodic capacity with co-channel interference and outdated CSI	analytical form of the achievable ergodic capacity
[24] Li et al.	MIMO LEO satellite	CSI MMSE	Asymptotic MMSE and a two-stage channel estimation
[25] Zhang et al.	MIMO LEO satellite system	CSI MSE	A system based on neural networks
[26] Guo et al.	MIMO adaptive modulation and coding LEO satellite system	CSI absolute error and spectral efficiency	Multiple ML models and modified ARIMA
[27] Zhang et al.	LEO satellite	Atmospheric attenuation in CSI	Different types of neural networks
[28] Zhang et al.	Massive MIMO LEO satellite system	CSI NMSE	LSTM
The proposed paper	LEO satellites and DVB S2X	SNR MSE, spectral efficiency, error rate	CNN and adapted margin

1.2. Paper Contributions and Structure

The work presented in this paper focuses on the application of ML (including neural networks) for SNR predictions in LEO satellite communications. The approach is based on an implemented simulator that creates channels with the following:

- Varying expected SNR values;
- Levels of Doppler shift;
- Shadowing levels.

The simulator is used to create data from 90 different channels (six different expected SNR values, five different levels of Doppler shift, and three different levels of shadowing) in order to study the relationship between different channel conditions and the improvements that can be provided by various ML algorithms for single-channel communication in comparison to a baseline method. The proposed methods are evaluated not only in terms of SNR prediction but also in terms of spectral efficiency and transmission error rate using the DVB-S2X protocol, placing a focus on practical applications. Finally, the evaluation of algorithms is also performed on two separate satellite channels, observing each pair of channels within the same level of shadowing.

We propose a systematic approach, observing the relationship between different channel conditions and improvement in LEO satellites' spectral efficiency following the DVB-S2X protocol, that is, to the best of the authors' knowledge, not present in the literature. In addition to various scenarios, a proposed neural network algorithm with a modified loss function also provides an improvement when compared to the traditional, outdated information approach both when observing a single satellite and two separate satellite channels. This new approach to obtaining a reliably low transmission error rate while still consistently providing spectral efficiency improvement opens up possibilities for future work based on the modification of neural network loss functions in order to meet the required system optimization criteria. In terms of applications, the improved prediction and spectral efficiency also allow for user-centric handover procedures, and the proposed work also evaluates the possibilities for lowering outage probabilities when using two satellites instead of a single one.

The remainder of the paper is structured as follows. Section 2 contains the description of the system model, with Section 2.1 describing the satellite network and Section 2.2 describing the channel modeling and simulation. Section 3 presents the proposed channel state prediction and modulation and coding (MODCOD) selection strategy, with Section 3.1 describing the algorithms used for SNR prediction and Section 3.2 describing spectral efficiency evaluation protocol and MODCOD selection for both single channel evaluation and the scenario where two satellite channels are considered. Section 4 presents the obtained results and discussion, with Section 4.1 describing the SNR prediction results. Section 4.2 is the obtained spectral efficiency improvement for single channel evaluation, Section 4.3 is the evaluation of two satellite channels, and Section 4.4 is the discussion. Finally, the conclusion is given in Section 5, alongside the directions for future work.

2. System Model

2.1. Satellite Network Description

A typical radio access network (RAN) supports a set of MODCODs spectral efficiencies $M = \{M_1, M_2, \ldots, M_K\}$, each associated with monotonously increasing function $T : M \to \mathbb{R}$, where $T(M_i)$ represents minimal SNR value needed to operate in the MODCOD M_i, also called the threshold. If we denote the instantaneous SNR in the channel by γ then the *optimal MODCOD* is the greatest element in a set $S_\gamma \subseteq M$, $S_\gamma = \{M_i | T(M_i) \leq \gamma\}$, i.e., supremum of S_γ denoted as $\sup(S_\gamma)$. In a special case, when $S_\gamma = \varnothing$, the lowest spectral efficiency M_1 will be used for transmission.

Let us enumerate all of the total N RANs that the end-user can use to establish communication by a set I = {1,2,…,N}. At a time interval (t_{j-1},t_j], the end-user measures instantaneous SNRs for all the currently available RANs, i.e., RANs from a set $I_j \in I$. The principle of forming the I_j set is discussed later. The measured values form a set, $\Gamma_j = \{\hat{\gamma}_{j,i} | i \in I_j\}$, where $\gamma_{j,i}$ represents measured SNR in [dB] of the i-th RAN during time interval (t_{j-1},t_j]. Then, a prediction function $f : \mathbb{R}^3 \to M$ is executed, which for all the available RANs produces the MODCODs that will potentially be used until the next SNR measurements are completed, i.e., in the time interval (t_j,t_{j+1}]. The goal of the function $f()$ is to make predictions regarding the optimal MODCOD based on the current and previous SNR measurements. Let $S \subseteq M$ be a set of MODCOD efficiencies formed in the following way

$$P = \left\{ M_m \middle| \, g\left(\hat{\gamma}_{j,i}, \hat{\gamma}_{j-1,i}, \hat{\gamma}_{j-2,i}\right) \geq T(M_m) \right\}, \tag{1}$$

where $\gamma_{j,i} \in \Gamma_j$, while $g : \mathbb{R}^3 \to \mathbb{R}^+$ produces a prediction of channel state information (i.e., SNR value). Thus we have

$$f\left(\hat{\gamma}_{j,i}, \hat{\gamma}_{j-1,i}, \hat{\gamma}_{j-2,i}\right) = \sup(P). \tag{2}$$

The end-user simply chooses the RAN that will produce the MODCOD with the highest spectral efficiency, i.e.,

$$f\left(\hat{\gamma}_{j,n_j}, \hat{\gamma}_{j-1,n_j}, \hat{\gamma}_{j-2,n_j}\right) \geq f\left(\hat{\gamma}_{j,i}, \hat{\gamma}_{j-1,i}, \hat{\gamma}_{j-2,i}\right), \ \forall i \in I_j. \tag{3}$$

In the case where there are multiple RANs that achieve maximal efficiency, and the previously used RAN is among them, the end-user does not perform a handover. Otherwise, the least-used RAN in the current time window is selected in order to reduce the possibility that the selected RAN leaves a set of available RANs. If $n_j \neq n_{j-1}$, the handover procedure is triggered, and the information of the preferred RAN is sent to the network core via reverse link.

In our previous work [29], we optimized the threshold margin; however, our further findings revealed that the prediction of SNR could be directly made by state-of-the-art machine learning models, as discussed in the subsequent section.

2.2. Channel Modeling and Simulation

In this work, we assume that instantaneous SNR, denoted by $\gamma(t)$, is defined as the following [30]:

$$\gamma(t) = \frac{P_T}{\sigma^2 d^\beta}|h(t)|^2, \tag{4}$$

where $h(t)$ denotes the time-varying complex channel gain, P_T denotes transmitted power, σ^2 denotes the variance of the additive Gaussian noise, d is the distance between transmitter and receiver, and β is the corresponding path-loss factor.

Various channel models were developed to describe the propagation conditions in a narrowband land mobile satellite channel. Typically, it is assumed that the fluctuations are the result of weak scattered components (multipath fading) and random variations in the power of multipath components (shadowing).

It is typically assumed that the channel gain between the RAN access point and the end-user is composed of two time-varying components [31–33]:

$$h(t) = a(t)e^{j\alpha(t)} + z(t)e^{j\alpha_0}. \tag{5}$$

The first term in the above expression corresponds to the scattering component, and the second term corresponds to the line-of-sight (LOS) component. In particular, $a(t)$ represents the amplitude of the scattering component, $\alpha(t)$ is the random phase, $z(t)$ is the amplitude of the LOS component, and α_0 denotes the corresponding phase.

In Loo's model, the amplitude of the LOS component is modeled by using the random variable with log-normal distribution [31]. Although this model corresponds to the measurement results available in the literature, the corresponding mathematical analysis is usually too complex. The shadowed Rice model [32,33] is a simpler but accurate channel model for narrowband land mobile satellite channels. In this model, it is assumed that $a(t)$ is Rayleigh distributed, $\alpha(t)$ is uniformly distributed, $z(t)$ is Nakagami-m distributed, and α_0 is a deterministic value. The corresponding probability density function of the instantaneous power gain in the satellite-terrestrial channel, denoted by $\lambda(t) = |h(t)|^2$, was derived in the paper [32] and given in the following form:

$$f_\Lambda(\lambda) = \left(\frac{2b_0 m}{2b_0 m + \Omega}\right)^m \frac{1}{2b_0} e^{-\frac{\lambda}{2b_0}} {}_1F_1\left(m; 1; \frac{\Omega\lambda}{2b_0(2b_0 m + \Omega)}\right), \ \lambda \geq 0, \tag{6}$$

where ${}_1F_1()$ is the confluent hypergeometric function, m represents the parameter of Nakagami-m distribution, $2b_0$ is the average power of the scattering component (the first term in Equation (5)), and Ω denotes the average power of the LOS component (the second term in Equation (5)). Three degrees of freedom in this channel model, i.e., parameters m, b_0 and Ω, can be used to accurately describe different propagation conditions in the satellite-terrestrial channel.

An accurate simulator of shadowed Rice fading that generates the corresponding fading samples with arbitrary temporal properties was proposed in [34]. A slightly improved version of this simulator, described in [35], is used in this paper.

In real-world scenarios, observing channels with various characteristics in a systematic way can be difficult as these characteristics are not necessarily predictable or stable. It is, however, necessary to have a sufficient amount of data for all possible scenarios when developing algorithms that improve telecommunication performance, which can easily be achieved through the usage of simulations. To obtain the most benefit from the availability of the data through simulations, channels with varied characteristics are considered in this paper. Namely, the three characteristics that are varied for the different channels observed in this paper are as follows:

- **The expected SNR value of the channels;**
- **The level of Doppler frequency shift;**
- **The level of shadowing.**

The expected SNR value of a channel does not necessarily imply changes in the morphology of the signal, but it does heavily influence the possibility and speed of data transmission. Hence, it is important to observe this parameter in order to analyze its interactions and influence on spectral efficiency and the transmission error rate, in combination with the other two considered characteristics that change the morphology of the SNR signal for a given channel. In general, a higher expected SNR value of a channel allows for higher spectral efficiency to be obtained and is considered a favorable condition. In this paper, a set of expected SNR values is considered, {0 dB, 3 dB, 6 dB, 9 dB, 12 dB, 15 dB}, to cover both extremes, the very high SNR, the very low SNR and several steps in between.

The second characteristic that is observed in this paper is **the level of Doppler frequency shift**. This influence of Doppler frequency shift is modeled through the f_{Dm} parameter of the simulation and essentially influences the SNR characteristic of the channel in terms of speed of change. In terms of the SNR dynamic, the higher the f_{Dm} value, the quicker the changes in the SNR occur over time and the weaker the correlation of future SNR values to the previous ones. This makes the SNR value more complex to predict and can significantly impact communication performance. A set of values for the f_{Dm} parameter was considered, {5 Hz, 25 Hz, 50 Hz, 75 Hz, 100 Hz}, and these cover a range of movement influences that can be considered residual movement after considering the movement corrections based on the GPS signal.

The third characteristic of a channel that is considered in this paper is **the level of shadowing**. Based on the channel simulation described in the previous section, three scenarios were considered through the parameters shown in Table 2.

Table 2. Channel parameters for different shadowing scenarios.

Shadowing Scenario	b_0	Ω	m
light shadowing	0.158	1.29	19.4
average shadowing	0.126	0.835	10.1
heavy shadowing	0.063	0.000897	0.739

The level of shadowing modulates how often and to what extent the SNR values drop to very low values. This can add to the difficulty of predicting future SNR values, but more prominently, it creates intervals in which no transmission of data could occur. The lower level of shadowing represents more favorable conditions, but it is of interest to see to which extent channels with higher levels of shadowing can be used and how reliable they can be.

In combination with various levels of expected SNR and Doppler frequency shift, the observed levels of shadowing create a space of channel characteristics that can cover a wide range of scenarios. The total number of different generated channels amounts to 90 (6 expected SNRs $\times$ 5 different f_{Dm}s $\times$ 3 different shadowing levels). For each channel, at least 100.000 consecutive samples were generated for the purposes of training and evaluation. Each generated sample has a true SNR value present in the channel and an estimated SNR value available as input to the decision-making algorithms.

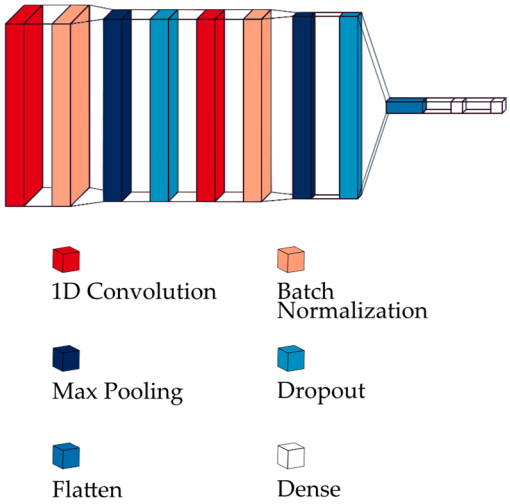
3. A Novel Channel State Prediction and MODCOD Selection Strategy

3.1. SNR Prediction Algorithms

For the purposes of SNR prediction, the outdated information (OI) approach was used as a baseline, and two different approaches using neural networks were implemented alongside two other machine learning (ML) models, support vector machine (SVM) and linear regression (LR). The OI approach simply considers that the future SNR value will be equal to the last estimated one. This approach conceptually works well when the channels have low levels of Doppler frequency shift and shadowing but falls short when channel estimation problems become more complex. On the other hand, the machine learning approaches use 10 consecutive estimated SNR values to predict the following true SNR value of the channel. The future SNR value is predicted based on a fixed input sequence

with a length of 10 since this length was considered to be sufficient for neural networks to recognize the characteristics of the channel without having a very high number of input features nor requiring long buffering of data.

The two simpler ML models were used with default parameters: ridge regularization for linear regression with the regularization parameter set to 1 and the radial basis function kernel with the regularization parameter set to 1 for the SVM. The goal of the paper was to see how simple ML algorithms would perform in different scenarios, not to explore multiple complex architectures. The neural networks used for the SNR prediction were convolutional neural networks with identical architecture but different loss functions. The architecture of the neural networks is given in Figure 1.



Type of layer	Parameters
Input layer	size = 10
1D Convolution	filters = 32, kernel size = 5, strides = 1, padding = same, activation = ReLU
Batch Normalization	
1D Max Pooling	pool size = 2, strides = 2
Dropout	rate = 0.5
1D Convolution	filters = 32, kernel size = 5, strides = 1, padding = same, activation = ReLU
Batch Normalization	
1D Max Pooling	pool size = 2, strides = 2
Dropout	rate = 0.5
Dense	10 neurons, activation = ReLU
Dense	1 neuron, no activation

Figure 1. The architecture of the implemented neural networks.

The proposed architecture is a rather simple convolutional neural network with a fully connected layer at the end. For the neural networks, smaller architectures were implemented to allow for potential practical applications, as larger neural networks have a higher inference time and might not be suitable for tasks such as real-time SNR predictions. The neural networks also have the possibility of modifying the loss function, which can greatly influence the performance of the network regarding certain areas of interest. The goal was to analyze this, comparing the two implemented neural network algorithms.

The difference between the regular neural network implementation (NN) and the modified one (NN2) was that the NN had a standard mean square error loss function, while the NN2 had a mean absolute error loss function with an added factor of $0.5 \times (\hat{y} - y)$, where $\hat{y}$ represents the prediction of the network, and y represents the true value. The motivation behind the implementation of the modified loss function of NN2 is that predicting higher SNR values than the true ones often results in unsuccessful data transmission, while a lower SNR prediction is suboptimal, but communication exists. In terms of simple SNR prediction, NN2 is expected to perform worse than a regular NN, but when performing further evaluations using various MODCODs, its intentionally lower predictions could be beneficial for both a lower error rate and a higher spectral efficiency. Both neural networks were trained using an Adam optimizer, batch size of 256, validation split of 0.2 and patience of 20 epochs.

The implementation of the ML models, neural networks, signal processing, evaluation and visualization was performed in the Python programming language [36], using the numpy [37], keras [38], scikit-learn [39] and matplotlib [40] libraries. Any parameters of the implemented neural networks and ML models that are not stated are left as default in the keras version 2.13.1 and scikit version 1.0.2 libraries. The initial evaluation of the proposed algorithms (OI, LR, SVM, NN, NN2) was performed using the mean square error (MSE) between the predictions $\hat{y}$ and labels y (with n samples):

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2. \quad (7)$$

The training of the algorithms was performed on the first 75% of the signal and the evaluation was performed on the last 25% of the generated SNR signal. The inputs to the algorithm were estimated SNR values, while the labels were the true SNR values, as generated by the simulation. On each set (both training and test), input–output pairs were created by sliding a window of 10 samples for creating inputs and taking the next consecutive sample as the label that should be predicted.

3.2. MODCOD Selection Evaluation

After the initial evaluation that shows the ability of the algorithms to perform SNR predictions, the next step was to evaluate the algorithms in terms of spectral efficiency. For spectral efficiency evaluation, all operation points of the DVB-S2X protocol [11] (short-frame communication) were used. For the considered MODCODs, the spectral efficiencies (M_i) and the SNR thresholds ($T(M_i)$), minimal SNR value needed to operate with the efficiency, are listed in Table 3. Based on the algorithms' predicted SNR, the system takes the highest possible MODCOD that could be successfully used with that predicted SNR.

Table 3. MODCODs, the spectral efficiencies (M_i) and the SNR thresholds ($T(M_i)$).

MODCOD	M_i [b/s/Hz]	$T(M_i)$ [dB]
BPSK-S 1/5	0.1	−9.9
BPSK-S 11/45	0.12	−8.3
BPSK 1/5	0.2	−6.1
BPSK 4/15	0.27	−4.9
BPSK 1/3	0.33	−3.72
QPSK 11/45	0.49	−2.5
QPSK 4/15	0.53	−2.24
QPSK 14/45	0.62	−1.46
QPSK 7/15	0.93	0.6
QPSK 8/15	1.07	1.45
QPSK 32/45	1.42	3.66
8PSK 8/15	1.60	4.71
8PSK 26/45	1.73	5.52
16APSK 7/15	1.87	5.99
16APSK 8/15	2.13	6.93
16APSK 26/45	2.31	7.66
16APSK 3/5	2.40	8.1
16APSK 32/45	2.84	9.81
32APSK 2/3	3.33	11.41
32APSK 32/45	3.56	12.18

The first level of evaluation was performed for each channel separately. So, the evaluation for each ML algorithm was performed for 90 different channels. To ensure the evaluations relate to practical use cases, margins were determined on the training set so that the transmission error on the training set is lower than 0.01 where possible or lower than 0.001+unavoidable error (samples where the SNR is lower than the lowest operation threshold). This margin was extracted for each algorithm and scenario separately (simple search with a resolution of 0.5 dB) and was then applied to the predictions on the appropriate test set. In this way, all algorithms have a comparatively fixed error rate, and the performances in terms of spectral efficiency could be adequately performed. The determined error rate threshold of 0.01 was selected as the highest possible error rate that might be considered usable for transmission, and further evaluation was performed on two channels as it is considered that combining two channels would create a scenario where, at least in theory, a lower error rate could be obtained.

The final step of the evaluation was performed using two channels, where each combination of channels within the same level of shadowing is considered. This was performed for OI and NN2 to provide an easier comparison since NN2 has been shown to have better performance in terms of spectral efficiency than all other algorithms. For this evaluation, a greedy selection was performed by the algorithms, i.e., for each point in time, the channel with the higher predicted SNR was selected. This creates a single array of SNR prediction values for the final algorithm predictions. The true labels were selected based on the labels for the channel that was selected for that corresponding point in time. The same principle for margin estimation was performed, extracting the margin on the training set and then applying it to the test set for OI and NN2 separately. The resolution for the search was also 0.5 dB, but the goal transmission error was 0.001 since two channels would often allow for a lower transmission error. Of course, if the transmission error of 0.001 was unattainable, the margin would be determined to reach the unavoidable error +0.0001 on the training set. The comparison between the NN2 and OI is performed by showing a relative improvement in spectral efficiency from the NN2 algorithm, calculated as $\left(M_i^{NN2} - M_i^{OI}\right)/M_i^{OI}$.

4. Results and Discussion

4.1. SNR Prediction

The initial testing results present the MSE between the predictions and labels on the test set for various channels and for the five observed algorithms. The OI is considered a baseline algorithm as it is simple to implement and is often used in the literature, while the ML models were expected to offer improvement. In terms of MSE performance (as well as spectral efficiency), the implemented LR and SVM performed the same or worse than the regular NN, so for an easier comparison, only the results of the NN are presented in this and further sections as it will be a good representative of the best performing simple ML algorithms. The MSE values for all observed scenarios are shown in Table 4, while Figures 2–4 show the same metric visually for light shadowing, average shadowing and heavy shadowing, respectively.

Table 4. The MSE [dB^2] achieved on the test set for various scenarios, OI/NN/NN2.

	MSE [dB^2]				
SNR [dB]	f_{Dm} **= 5 Hz**	f_{Dm} **= 25 Hz**	f_{Dm} **= 50 Hz**	f_{Dm} **= 75 Hz**	f_{Dm} **= 100 Hz**
Light shadowing					
0	1.0/0.5/0.6	2.2/1.4/2.1	4.5/2.9/3.9	7.4/4.8/6.4	11.7/7.2/9.1
3	0.7/0.3/0.5	1.8/1.3/1.7	3.6/2.4/3.5	6.8/4.4/5.7	10.5/6.8/8.8
6	0.5/0.3/0.4	1.5/1.0/1.4	3.6/2.2/2.9	7.3/4.6/5.7	10.2/6.8/8.5
9	0.4/0.3/0.3	1.2/0.9/1.2	3.2/1.9/2.5	6.3/3.9/5.4	9.8/6.5/8.3
12	0.3/0.3/0.3	1.3/0.9/1.1	3.3/1.9/2.8	7.0/4.2/5.2	9.5/6.3/8.2
15	0.2/0.2/0.2	1.0/0.7/0.9	3.5/2.0/2.7	6.7/4.1/5.2	10.2/6.8/8.8
Average shadowing					
0	1.5/0.8/1.1	7.4/3.8/4.7	17.2/7.3/9.5	31.4/11.9/14.5	38.2/13.8/17.0
3	1.5/0.8/0.9	6.8/3.3/4.1	17.0/7.2/9.2	27.8/10.8/13.7	35.5/13.2/16.2
6	0.7/0.6/0.6	7.1/3.0/4.1	17.7/7.2/9.1	30.1/11.5/13.9	35.9/13.6/16.3
9	0.7/0.5/0.6	6.1/2.6/3.4	17.2/7.1/9.1	27.5/11.2/13.5	33.5/13.2/16.3
12	0.4/0.4/0.4	5.4/2.3/3.3	16.9/6.8/8.3	26.7/10.8/13.2	35.4/13.6/16.6
15	0.6/0.5/0.6	6.2/2.6/3.2	16.8/6.8/8.3	29.4/11.1/13.9	34.3/13.3/16.8
Heavy shadowing					
0	3.4/1.2/1.8	7.9/3.9/5.4	15.7/7.8/10.3	26.2/13.5/17.1	34.0/18.5/23.2
3	2.3/1.1/1.6	6.2/2.9/4.1	14.9/7.2/9.4	24.7/12.5/16.1	34.2/18.0/23.3
6	1.1/0.6/1.3	5.7/2.7/3.5	13.9/6.4/8.0	22.8/11.4/14.8	32.5/17.6/23.1
9	0.8/0.5/0.6	5.0/2.3/2.9	12.7/5.7/7.6	22.2/11.1/15.0	31.8/17.3/22.1
12	0.8/0.5/0.8	5.2/2.0/3.3	12.6/5.7/7.5	22.8/11.0/14.0	30.5/17.6/21.3
15	0.6/0.4/0.5	4.6/1.8/2.5	12.7/5.4/6.8	22.9/11.5/15.0	30.4/17.3/22.1

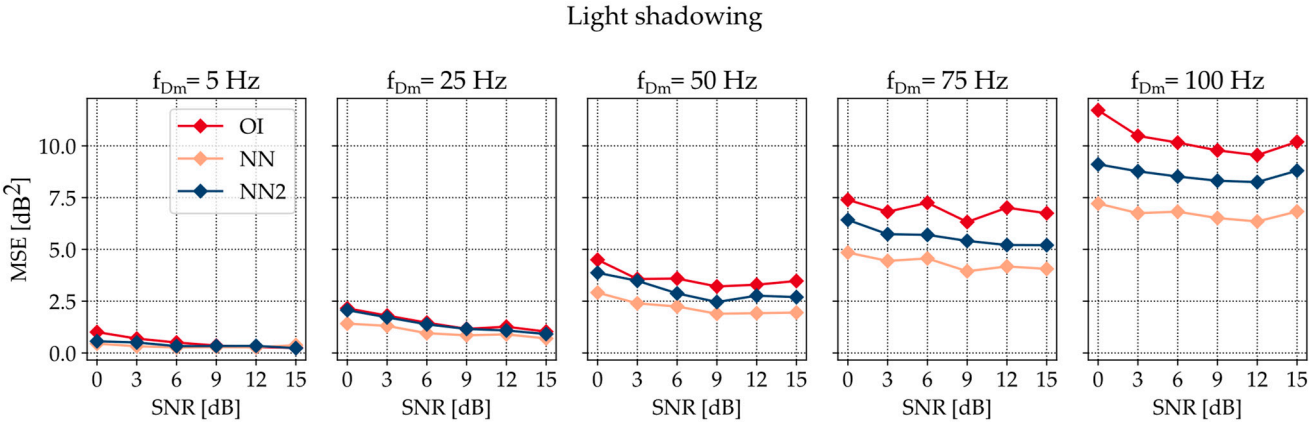


Figure 2. The achieved MSE for SNR prediction on the test set for various scenarios under light shadowing conditions.

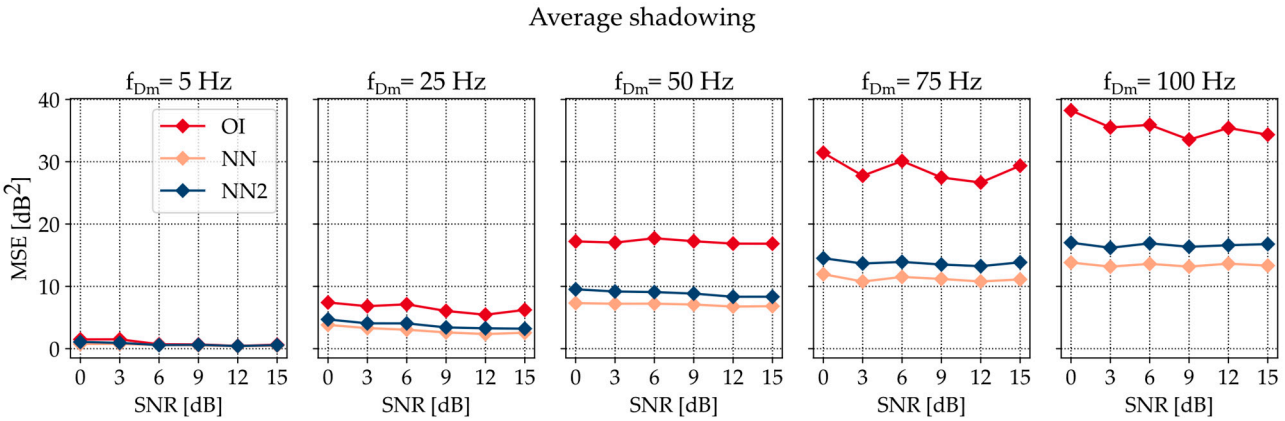


Figure 3. The achieved MSE for SNR prediction on the test set for various scenarios under average shadowing conditions.

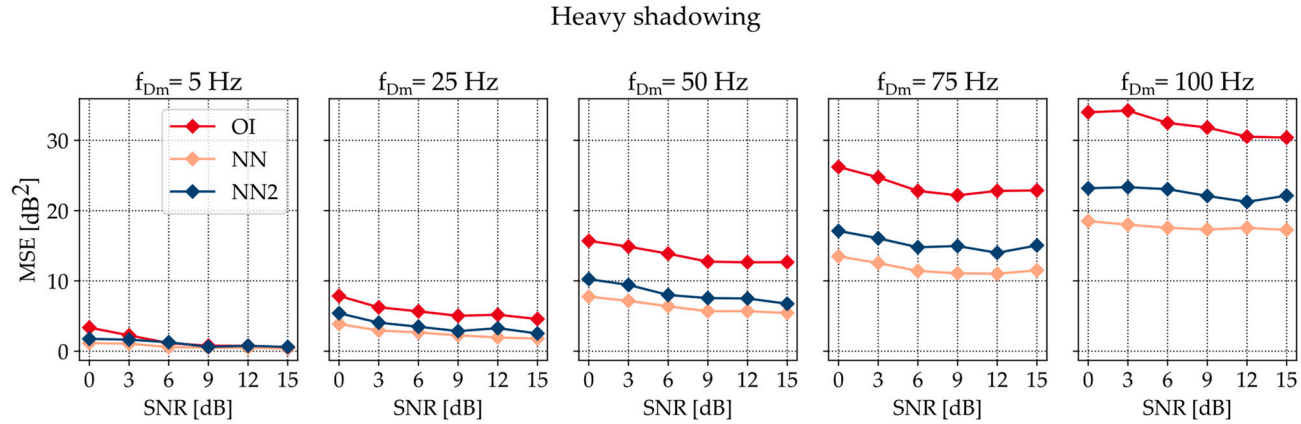


Figure 4. The achieved MSE for SNR prediction on the test set for various scenarios under heavy shadowing conditions.

When observing the values from Table 4, it can be concluded that based on different scenarios, the MSE can have quite a large range of values, from 0.2 dB^2 to 38.2 dB^2 for the outdated information. The vast range of these values shows that based on different channel characteristics, different expectations for SNR prediction quality should be present. The ranges of MSE for NN and NN2 have the same minimum value as the OI, and this value is so low that it can be concluded that the implementation of additional algorithms

outside of OI can be completely redundant in certain scenarios, e.g., f_{Dm} = 5 and higher expected SNR values. On the other hand, the maximum MSE for NN is quite lower than the maximum value for outdated information, amounting to 18.5 dB^2, which is less than half of the maximum value of the OI and almost half of the OI MSE value for that respective scenario. Notably, the NN2 has a consistently higher MSE than the NN but also provides an improvement when compared to the OI. This is to be expected as the loss function of the NN2 is not made to optimize for the MSE or exact prediction; rather, it is created to force the network to rarely overestimate the SNR value. For pure SNR prediction, this is impractical, but it will later be shown that this has benefits when observing practical applications and spectral efficiency.

The results shown in Figures 2–4 intuitively show the overall trends of the MSE in terms of various algorithms and channel scenarios. Notably, it can be observed that for all algorithms, the MSE is higher when the f_{Dm} parameter rises. This is expected as the higher values of f_{Dm} correspond to channels that have less predictable changes, and both OI and neural networks have difficulties performing when more rapid SNR changes are present. It is also evident that for a very low f_{Dm}, regardless of SNR, the algorithm performs quite similarly. As presented in Table 4, improvements exist in most scenarios, but they are so minor that the development of special algorithms for the prediction of SNR can be considered redundant. On the other hand, for high f_{Dm} values, the improvements that the neural networks provide become more evident. It is interesting to note that the discrepancy between the performance of NN and NN2 also increases with the increase in f_{Dm}, regardless of shadowing conditions. This is most likely because a higher f_{Dm} creates a more unpredictable channel, and since the NN2 is penalized for overestimating the SNR, it starts to consistently predict much lower values in order to avoid the overestimating. This results in a higher MSE and, therefore, worse performance when compared to the NN approach.

Another interesting characteristic worth noting is that the OI has a higher MSE for heavy shadowing as opposed to average shadowing for f_{Dm}s of 50 Hz and above (75 Hz and 100 Hz) for all SNR values. This is an interesting find, as heavy shadowing is considered a worse scenario than average shadowing. On the other hand, for unpredictable channels, such as those with high f_{Dm}s values, heavy shadowing actually creates a more predictable pattern; although the SNR values are lower, more frequent shadowing occurrences create more regularity in the pattern and create a correlation between previous samples and future ones. The neural networks, however, compensate for these characteristics, and the same aspect is not prominent in neural networks' performance.

4.2. Single Channel Spectral Efficiency

Considering the achieved results, it is clear that the neural networks can provide clear improvement in terms of SNR prediction when compared to the baseline OI approach, and the next step represents the evaluation of spectral efficiencies for all observed channels. The initial step is to analyze the outage probabilities if perfect SNR predictions would have been performed for the test channels, presented in Table 5. These outage probabilities present the unavoidable error on the test set and are relevant for the interpretation of further results.

The results show that for the set threshold of 0.01 for single channel evaluation, outage probability is too high in some scenarios, making the set threshold unobtainable. These outage probabilities, however, should not impair the improvements in spectral efficiency that the neural networks should offer. The outage probabilities are more prominent for higher levels of shadowing, which is expected as 1 percent of the test set corresponds to 250 samples, and more frequent, heavier shadowing can easily make more than 250 samples have values lower than the minimum operational threshold.

Table 6 shows the results in terms of spectral efficiency and achieved error transmission rate for all considered conditions and algorithms. Figures 5–7 show the relationship between the SNR and the achieved spectral efficiency for light, average and heavy shadowing, respectively, while Figures 8–10 show the results visually in terms of improvement percentage compared to the OI spectral efficiency for light, average and heavy shadowing, respectively.

Table 5. The outage probability on the test set for each observed scenario. The probabilities higher than or equal to 0.01 are presented in red.

Outage Probability					
SNR [dB]	**f_{Dm} = 5 Hz**	**f_{Dm} = 25 Hz**	**f_{Dm} = 50 Hz**	**f_{Dm} = 75 Hz**	**f_{Dm} = 100 Hz**
Light shadowing					
0	0.016	0.022	0.019	0.02	0.021
3	0.01	0.009	0.007	0.008	0.009
6	0.003	0.004	0.004	0.004	0.004
9	0.001	0.001	0.001	0.002	0.002
12	0	0.001	0.001	0.001	0.001
15	0	0	0	0	0
Average shadowing					
0	0.03	0.035	0.031	0.037	0.037
3	0.016	0.015	0.014	0.014	0.015
6	0.005	0.009	0.007	0.008	0.007
9	0.002	0.003	0.004	0.003	0.004
12	0.001	0.001	0.001	0.002	0.002
15	0	0.001	0.001	0.002	0.001
Heavy shadowing					
0	0.091	0.095	0.096	0.098	0.094
3	0.048	0.048	0.05	0.052	0.051
6	0.021	0.026	0.026	0.026	0.025
9	0.009	0.012	0.012	0.012	0.012
12	0.007	0.007	0.007	0.007	0.007
15	0.002	0.003	0.003	0.004	0.003

Table 6. The achieved spectral efficiency [b/s/Hz] achieved on the test set for a single channel for various scenarios, OI/NN/NN2. The spectral efficiencies in bold represent the ones for which the transmission error rate was lower than 0.01.

M_i[b/s/Hz]					
SNR [dB]	**f_{Dm} = 5 Hz**	**f_{Dm} = 25 Hz**	**f_{Dm} = 50 Hz**	**f_{Dm} = 75 Hz**	**f_{Dm} = 100 Hz**
Light shadowing					
0	0.33/0.4/0.41	0.27/0.29/0.28	0.19/0.19/0.2	0.14/0.14/0.14	0.11/0.11/0.12
3	0.73/0.77/0.79	**0.55/0.54/0.59**	**0.37/0.39/0.41**	**0.24/0.24/0.27**	**0.17**/0.18/**0.17**
6	**1.25/1.24/1.27**	**1.0/1.06/1.06**	**0.73/0.85/0.89**	**0.54/0.57/0.63**	**0.4/0.4/0.42**
9	**1.84/1.82/1.96**	**1.63/1.63/1.65**	**1.28/1.45/1.41**	**0.92/1.05/1.12**	**0.77/0.75/0.79**
12	**2.62/2.53/2.66**	**2.26/2.36/2.36**	**1.86/2.05/2.1**	**1.44/1.62/1.62**	**1.18/1.19/1.25**
15	**3.16/3.06/3.2**	**2.94/3.04/3.04**	**2.51/2.74/2.77**	**2.05/2.25/2.25**	**1.73/1.79/1.82**
Average shadowing					
0	0.33/0.35/0.36	0.15/0.2/0.19	0.1/0.12/0.13	0.1/0.1/0.1	0.1/0.1/0.1
3	0.6/0.66/0.67	0.24/0.35/0.37	0.11/0.15/0.16	0.1/0.12/0.12	0.1/0.1/0.1
6	**1.15/1.07/1.25**	**0.5/0.69/0.79**	**0.19**/0.33/**0.31**	**0.12/0.18/0.19**	**0.11/0.13/0.14**
9	**1.7/1.68/1.71**	**0.98/1.2/1.24**	**0.44/0.66/0.7**	**0.23/0.37/0.38**	**0.17/0.28/0.27**
12	**2.49/2.37/2.52**	**1.46/1.83/1.98**	**0.78/1.12/1.17**	**0.49/0.76/0.77**	**0.32/0.57/0.54**
15	**2.99/3.01/3.02**	**2.13/2.45/2.62**	**1.24/1.56/1.74**	**0.78/1.25/1.24**	**0.6/0.96/1.0**
Heavy shadowing					
0	0.34/0.42/0.39	0.22/0.26/0.27	0.14/0.18/0.19	0.11/0.13/0.14	0.1/0.11/0.11
3	0.6/0.67/0.71	0.36/0.42/0.5	0.2/0.26/0.3	0.14/0.19/0.19	0.11/0.14/0.14
6	0.94/1.0/1.04	0.52/0.7/0.83	0.25/0.49/0.51	0.17/0.26/0.27	0.13/0.16/0.17
9	**1.4/1.46/1.51**	0.74/1.03/1.2	0.36/0.62/0.71	0.2/0.36/0.41	0.14/0.2/0.23
12	**2.05/2.22/2.14**	**1.21/1.73/1.78**	**0.67/1.09/1.26**	**0.36/0.61/0.68**	**0.3/0.35/0.44**
15	**2.73/2.67/2.77**	**1.94/2.41/2.47**	**1.2/1.7/1.94**	**0.78/0.97/1.21**	**0.51/0.62/0.7**

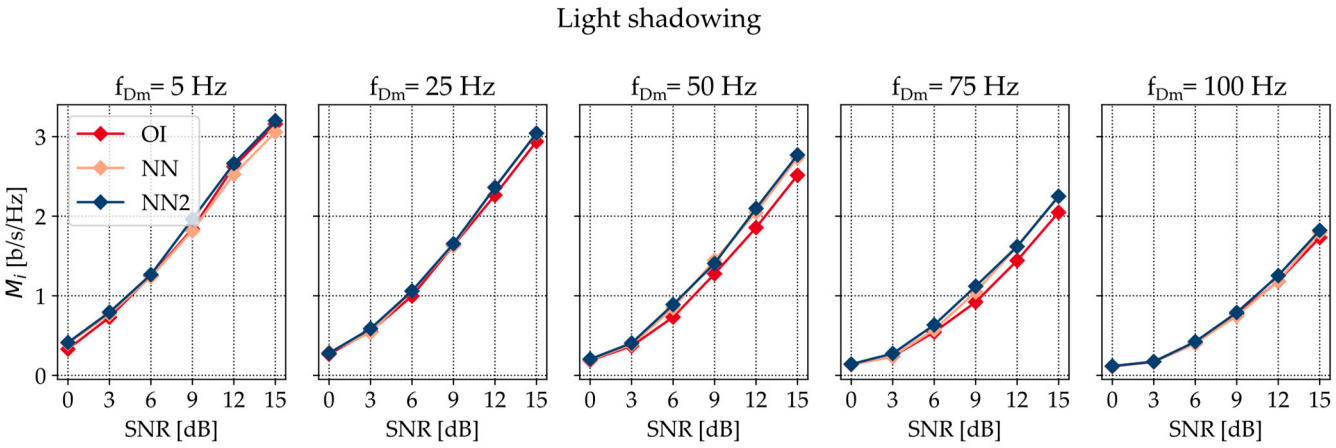


Figure 5. The achieved spectral efficiency on the test set for various scenarios under light shadowing conditions using a single channel.

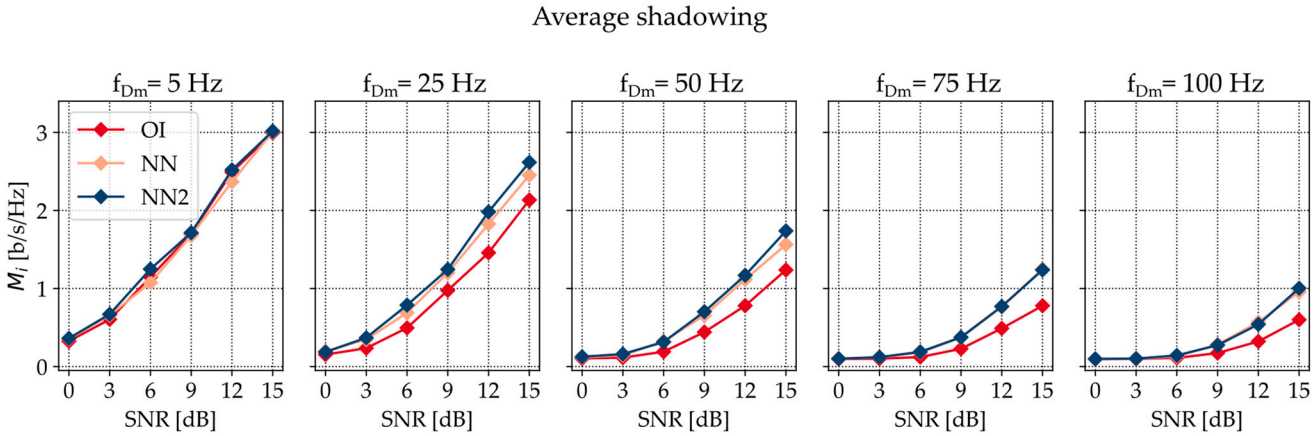


Figure 6. The achieved spectral efficiency on the test set for various scenarios under average shadowing conditions using a single channel.

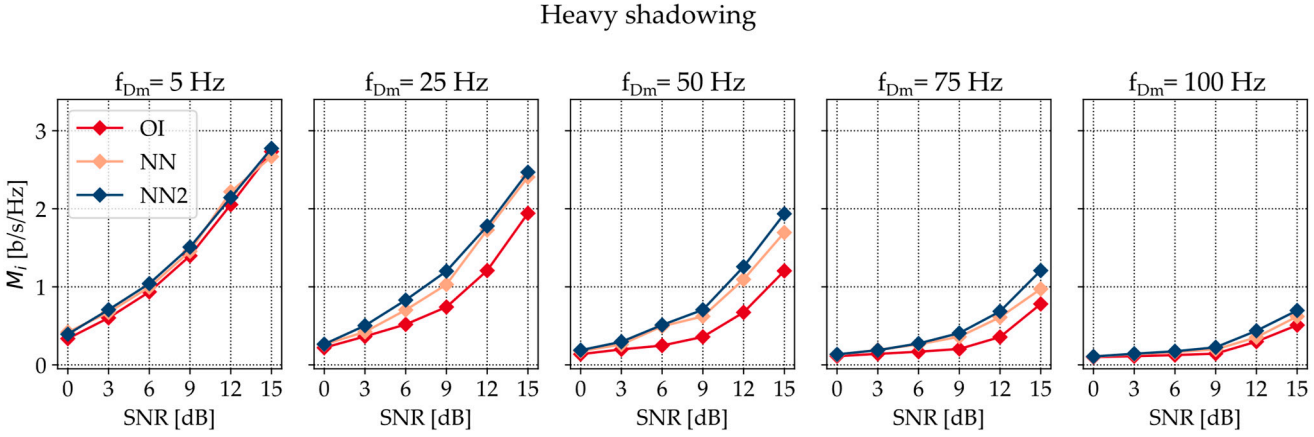


Figure 7. The achieved spectral efficiency on the test set for various scenarios under heavy shadowing conditions using a single channel.

The results in Table 6 show the wide range of performances that can be achieved for various scenarios and again point out that different channel characteristics can quite heavily influence the performance of the algorithms. Expectedly, as opposed to the MSE results, the expected SNR plays a significant role in achieving higher spectral efficiency. More specifically, the higher the expected SNR, the higher the achieved spectral efficiency.

This stands regardless of the implemented algorithm or f_{Dm} value. It is also important to note that there are many scenarios for which the transmission error rate is higher than 0.01, essentially making reliable communication impossible, regardless of the spectral efficiency that can be achieved. As can be seen in Table 5, all of the scenarios for which the desired error rate of 0.01 is not achieved have an unavoidable transmission error higher than 0.01. The implemented approach for margin determination does not make achieving an error of less than 0.01 on the test set certain, as the margin is determined on the training set and only then applied on the test set. However, the results indicate that this approach works quite well, as all the scenarios in which the error is larger than 0.01 correspond to the ones where the unavoidable outage probability is above 0.01.

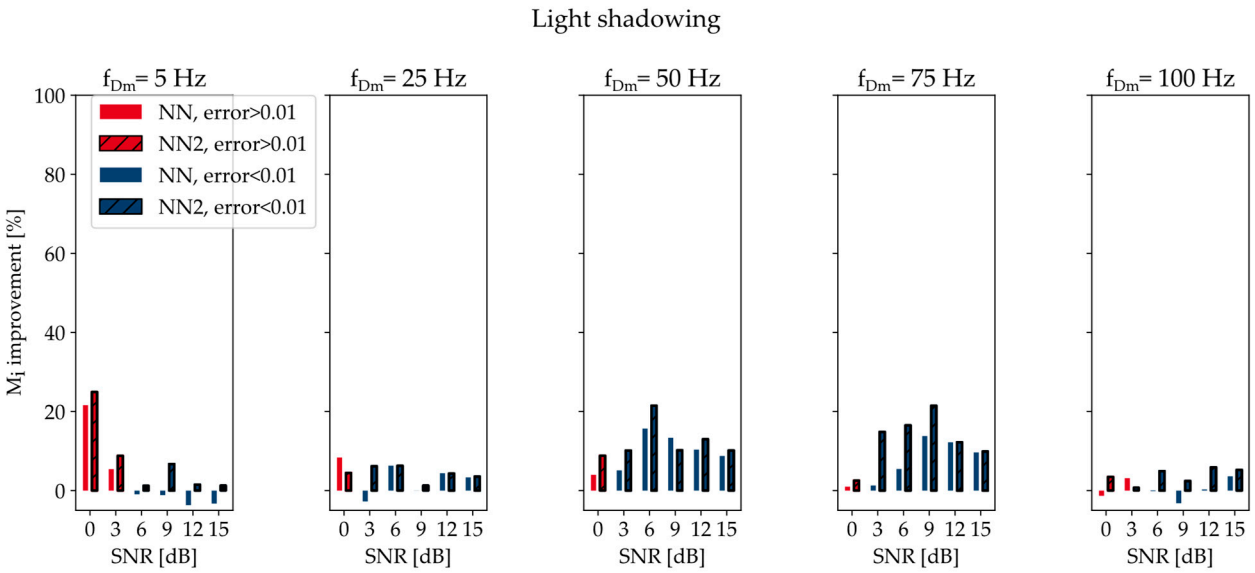


Figure 8. The achieved spectral efficiency improvement in % compared to the OI, on the test set for various scenarios under light shadowing conditions using a single channel. Scenarios where the transmission error is <0.01 are shown in blue, and scenarios where the transmission error is >0.01 are shown in red.

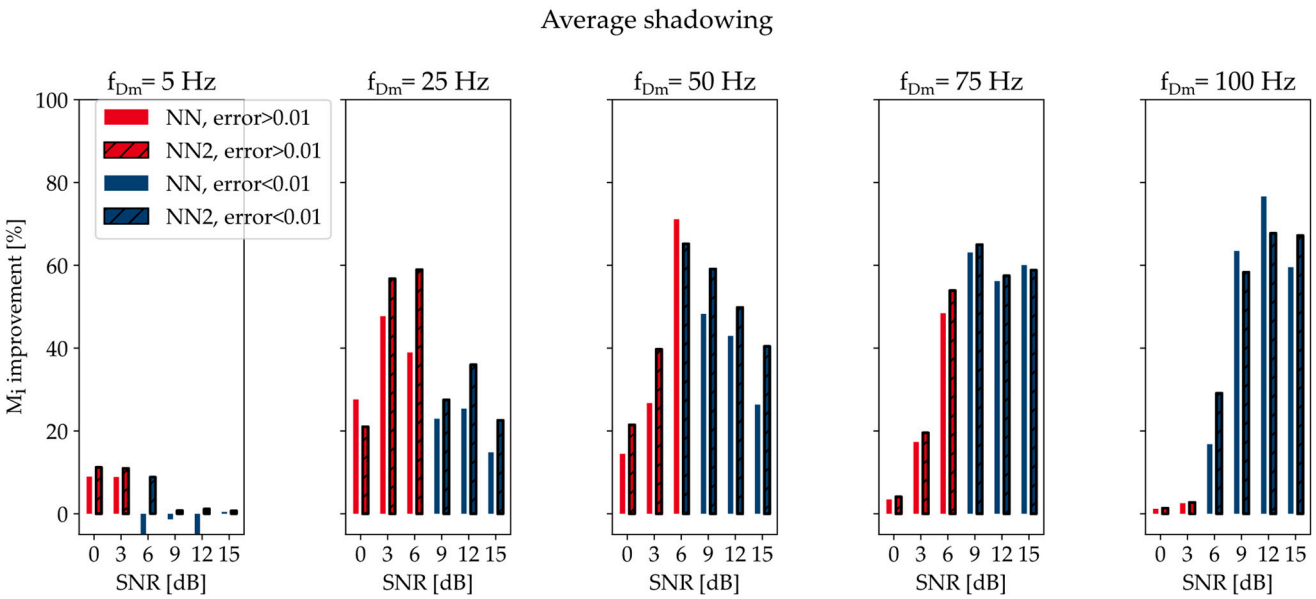


Figure 9. The achieved spectral efficiency improvement in % compared to the OI, on the test set for various scenarios under average shadowing conditions using a single channel. Scenarios where the transmission error is <0.01 are shown in blue, and scenarios where the transmission error is >0.01 are shown in red.

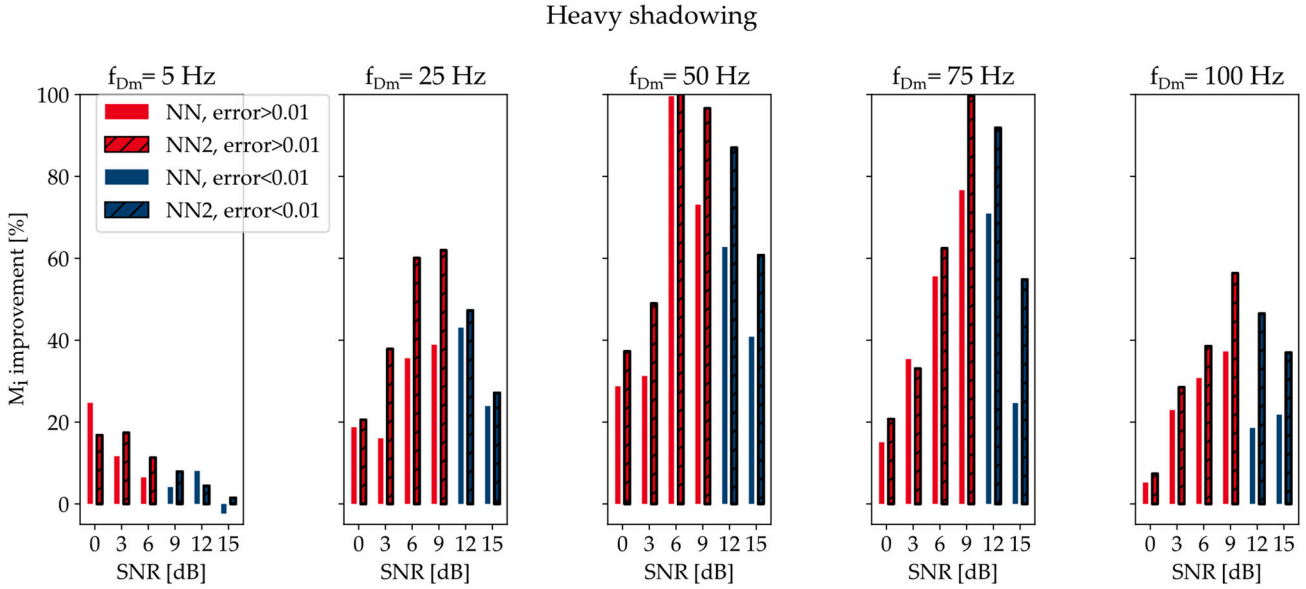


Figure 10. The achieved spectral efficiency improvement in % compared to the OI, on the test set for various scenarios under heavy shadowing conditions using a single channel. Scenarios where the transmission error is <0.01 are shown in blue, and scenarios where the transmission error is >0.01 are shown in red.

The visual representations of the obtained spectral efficiencies are shown in Figures 5–7, indicating clearly how the trends of spectral efficiencies behave for different shadowing, f_{Dm} and SNR conditions. NN2 is consistently better than the OI, with NN also being better in most cases. For an SNR of 0 dB, the improvement seems negligible, but as the SNR increases, the difference between spectral efficiency becomes more prominent. The shapes of the presented curves also change based on the amount of shadowing. It can be seen in Figure 5 that the spectral efficiency rises mostly linearly with the SNR for most of the light shadowing conditions, regardless of the algorithm, but the trends become more curved as the f_{Dm} increases. For average shadowing, the trends seem more curved, and for heavy shadowing, almost no curve looks linear. This conclusion also stands for all algorithms, meaning that regardless of the achieved improvement, the relationship between SNR and spectral efficiency has a different trend based on the type of channel that is observed.

Observing the results from Figure 8, it can be seen more clearly that NN2 has a consistently better performance than OI and almost always a better performance than NN. The improvement in spectral efficiency that the NN2 provides is not drastic in comparison to OI or NN, as can be seen, but it is consistent. This clearly indicates the importance of considering different scenarios, as developing complex algorithms can provide limited improvement in certain ones. The consistency shows that the proposed method is conceptually good, but for practical applications, it is important to weigh the benefits of the spectral efficiency improvement against the complexity of integrating complex models into a system.

When comparing NN2 and NN, it is important to note that, as opposed to the simple SNR prediction, NN2 has better performance. This is due to the introduced margin and the need for a transmission error rate no greater than 0.01. Since there are many MODCODs considered for communication, their operations thresholds are not that far apart. Hence, if the neural networks predict an SNR value that is much higher than the operation point of the best possible MODCOD, no data will be transmitted, and a transmission error will occur. This is why the NN2 approach of underestimating values is useful because it is less likely to make such mistakes, so the determined margin ensuring a low error rate will not be as high and will not bring down the spectral efficiency improvements as much as it will for NN.

In terms of achieving a transmission error rate no greater than 0.01 for light shadowing specifically, Figure 8 recapitulates that for an expected SNR of 0 dB, it is not possible, and no approach achieves this, but it is also shown that for the SNR of 3 dB and an f_{Dm} of 5 Hz, none of the algorithms could obtain a transmission error rate lower than 0.01. This channel does not fall under the category of difficult or unpredictable, as there is light shadowing, and the f_{Dm} is quite small. However, since the SNR is not high, it is always possible that the SNR values happen to be distributed in such a way that the outage probability is higher than 0.01, which is exactly what happened in this case. This stands in line with the results obtained for average and heavy shadowing, as for both, there was an error rate higher than 0.01 for all scenarios where the expected SNR was equal to 3 dB. One more interesting occurrence is that for an f_{Dm} of 100 Hz and an expected SNR of 3 dB, the NN did not achieve an error rate lower than 0.01, while NN2 and OI did. This simply shows that the errors that NN makes can be such that the SNR prediction itself is better, but the overestimating of values that sometimes occur can have a negative impact on reaching certain goals, such as low rates of transmission error. In terms of obtainable improvement for light shadowing, for a frequency range of 40 MHz, if the best relative improvement scenario is considered for NN2 (f_{Dm} = 75 Hz, SNR = 9 dB), the contribution of NN2 would be $(1.12 - 0.92)$ b/s/Hz $\times$ 40 MHz = 8 Mb/s.

Observing average shadowing, similar patterns can be observed for light shadowing, with some changes. Firstly, for average shadowing, for an expected SNR of 3 dB, the transmission error rate was always higher than 0.01. This is because more frequent or heavier shadowing increases the intervals in which no communication can occur, thus raising the unavoidable error, which exceeds 0.01 in these scenarios. It can also be seen, in comparison to the low shadowing conditions, that for higher f_{Dm} values, the achieved spectral efficiencies are overall quite lower, while for the lower f_{Dm} values, this is not as prominent. This is to be expected as the combination of quicker changes in SNR in combination with more frequent shadowing makes predictions significantly more difficult, whereas if more shadowing but for slower changing SNR channels (lower f_{Dm}), the SNR pattern during shadowing can be more easily predicted and therefore not hinder the performance as severely. One more important observation is that although the absolute values are overall lower for higher f_{Dm} when compared to the light shadowing, the relative improvement between OI and NN2 is more pronounced. This would indicate that for less favorable scenarios, such as average shadowing and a high f_{Dm}, although the absolute spectral efficiency cannot be high, introducing more complex algorithms for SNR prediction could provide a significant benefit. Another occurrence that has happened for light shadowing was that in certain scenarios, NN has a transmission error rate higher than 0.01 while OI and NN2 do not. This has now happened for f_{Dm} of 50 Hz and an expected SNR of 6 dB for the same reason described in the light shadowing scenario. Once again, in the results obtained for heavy shadowing, it can be seen that the expected SNR of 6 dB does not allow for communication under any f_{Dm} conditions. For the average shadowing, in terms of obtainable absolute improvement (for a frequency range of 40 MHz), the best relative improvement scenario for NN2 (f_{Dm} = 100 Hz, SNR = 12 dB), the contribution of NN2 would be $(0.54 - 0.32)$ b/s/Hz $\times$ 40 MHz = 8.8 Mb/s.

The analysis of the results obtained for heavy shadowing is quite similar to the one for previous scenarios. The transmission error rate was higher than 0.01 (due to outage probability) for almost all f_{Dm}s and the expected SNR up to 9, with the only exception being the SNR of 9 dB and $f_{Dm}s$ of 5 Hz. This shows that, as in the examples above, sometimes channel SNR values can play out in such a way that they allow for communication to be established in a way that is not possible for similar scenarios. One more important observation that stands for all shadowing conditions but can best be seen for heavy shadowing is that even in the scenarios where the transmission error rate is higher than 0.01, the improvements of spectral efficiency exist between OI and NN2 and the absolute value of spectral efficiency rises with the rise of the expected SNR. This is extremely important because even if a desired error rate is unattainable, this approach

will still provide an improvement in spectral efficiency, which is crucial when considering multiple channels with different characteristics and the usability of the provided method. The relative improvement provided in certain scenarios for heavy shadowing is the highest among the observed scenarios and exceeds 100% in some cases. For the best relative improvement scenario for NN2 (f_{Dm} = 75 Hz, SNR = 12 dB), considering a frequency range of 40 MHz, the NN2 contribution amounts to (0.68 − 0.36) b/s/Hz × 40 MHz = 12.8 Mb/s.

4.3. Double Channel Spectral Efficiency

The final evaluation step is the one where the performance of the proposed method is evaluated for two communication channels. Here, only the NN2 and OI are compared for an easier overview of the results, especially considering that the NN2 approach has provided better results for the single channel spectral efficiency improvement. Figure 11 shows the results for light shadowing, Figure 12 for average shadowing and Figure 13 for heavy shadowing.

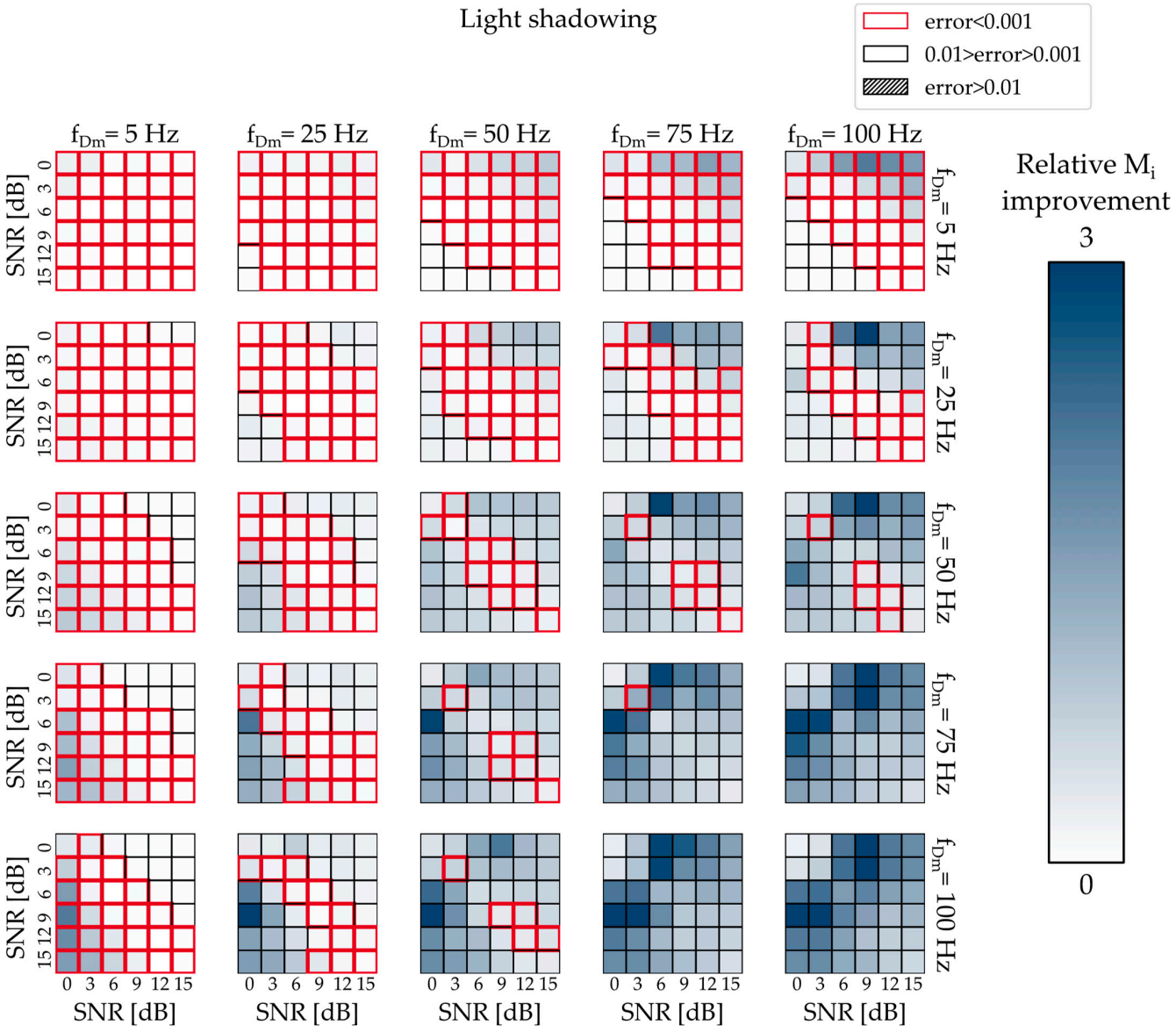


Figure 11. The achieved spectral efficiencies on the test set for various scenarios under light shadowing conditions using two channels, comparing OI and NN2. The color of each square represents a relative improvement of the achieved spectral efficiency calculated as $\left(M_i^{NN2} - M_i^{OI}\right)/M_i^{OI}$.

Figure 11 shows how the combination of two channels can influence the performance of the proposed system. Each larger square represents a scenario where the channels f_{Dm} s are fixed (e.g., second row, third column, $f_{Dm}1$ = 25 Hz, $f_{Dm}2$ = 50 Hz), while the smaller squares correspond to various combinations of expected SNR. The type of square, such as red outline, regular outline and hatched, corresponds to the range of transmission errors for that scenario, and the color scale corresponds to the relative improvement in spectral efficiency. The relative improvement is above 0 for all scenarios, i.e., there are no scenarios where the OI outperformed the NN2. Secondly, the red squares outline the scenarios in which the desired transmission error rate was achieved, i.e., it was under 0.001. It can be seen that for a low f_{Dm}, it is always achievable, but as the f_{Dm} rises, this becomes more difficult, and for the f_{Dm} = 100 Hz, regardless of SNR, the proposed method achieved a transmission error rate lower than 0.01 but not lower than 0.001. On the other hand, it can be seen that the relative improvement of spectral efficiency provided by the NN2 is much more prominent for the scenarios with a higher f_{Dm} (as seen in dark blue) as opposed to the ones for lower f_{Dm} (seen in white or light blue). This shows that depending on the scenario, different goals can be achieved and that the final goal has to be considered through the design of the algorithm since the most straightforward solution (such as NN) might not provide the best results. Overall, scenarios where lower errors are obtainable present ones where SNR is easier to predict; hence, OI initially had good performance, which is why the relative improvement offered by the NN2 is not as high as for some other scenarios.

Figure 12 shows how the increase in shadowing affects the performance of the system. When compared to the light shadowing conditions, many of the results are in darker blue, showing a greater relative improvement than the one achieved for light shadowing. Secondly, it can be seen that aside from a couple of scenarios of both channels having an f_{Dm} of 25 Hz, an error rate lower than 0.001 could not be achieved if one of the channels does not have an f_{Dm} of 5 Hz. Thirdly, it can be seen that for some scenarios of higher f_{Dm}s and lower expected SNRs, not even an error rate of 0.01 could be achieved. This is due to the unavoidable error rate, in the same manner as it was present for single channel evaluation.

The results shown in Figure 13 show several outcomes that could be considered expected and several ones that provide new information. Firstly, the scenarios where the unavoidable error is above 0.01 are more prominent, as can be seen for lower expected SNR scenarios where multiple fields are hatched. Secondly, there are more scenarios where the transmission error could be lower than 0.001 for higher f_{Dm}s when compared to the average shadowing. This might seem unexpected as more frequent and more heavy shadowing is not a favorable condition. However, it is possible that for higher f_{Dm}s, more frequent shadowing adds a level of order to the noisy signal, making the NN2 better at predicting what future SNR values will be. The "shadowed" parts of the signal provide very low SNR values, and if the NN2 can predict these values to be quite low, then the margin that is introduced might not need to be as high, and the overall performance could be better. The shadowed intervals have accurately low predicted SNRs, but regular parts of the signal could also have adequate predictions that will not be hindered by an extremely high margin.

When observing the error transmission ranges and the obtained results, it is important to note that the next order of magnitude, i.e., having a transmission error rate of 0.0001 or lower, was only unobtainable since the current test set has 25,000 samples, and such an error rate would imply no more than two samples could be allowed to be incorrectly transmitted. Furthermore, even with a larger test set, two channels and the considered shadowing conditions would probably not allow for such a low error to be theoretically obtainable anywhere where there is average or heavy shadowing. This could direct future work towards analyzing three or more channels or simply analyzing the performance obtainable when two channels of different shadowing levels are combined. These scenarios are understandably of interest but would simply be

out of scope for this paper as the goal was to perform a sort of grid analysis in terms of ML algorithm performance for various channels and to evaluate whether a neural network that purposefully underestimates values could be of interest considering fixed transmission error rates.

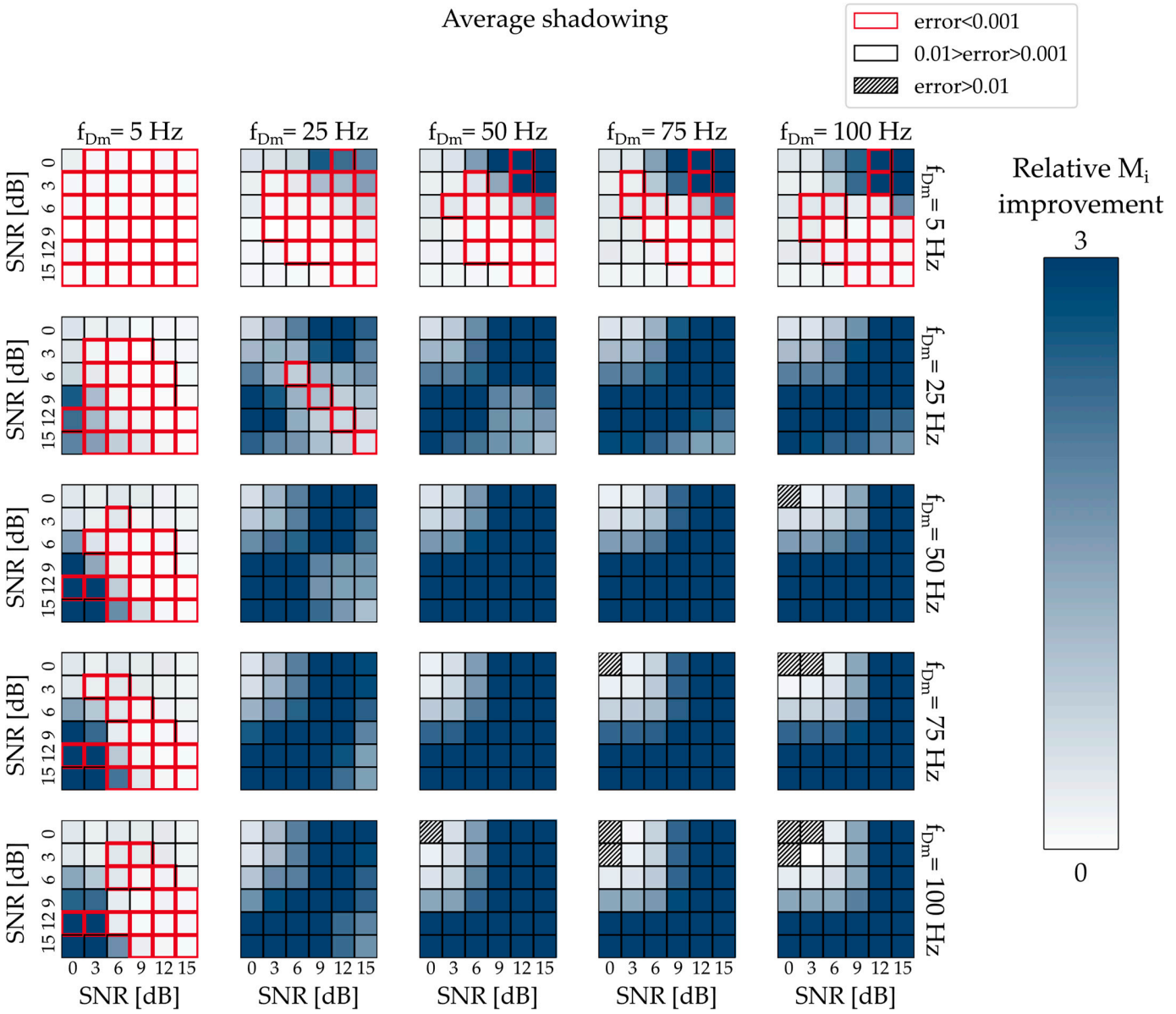


Figure 12. The achieved spectral efficiencies on the test set for various scenarios under average shadowing conditions using two channels, comparing OI and NN2. The color of each square represents a relative improvement of the achieved spectral efficiency calculated as $\left(M_i^{NN2} - M_i^{OI}\right)/M_i^{OI}$.

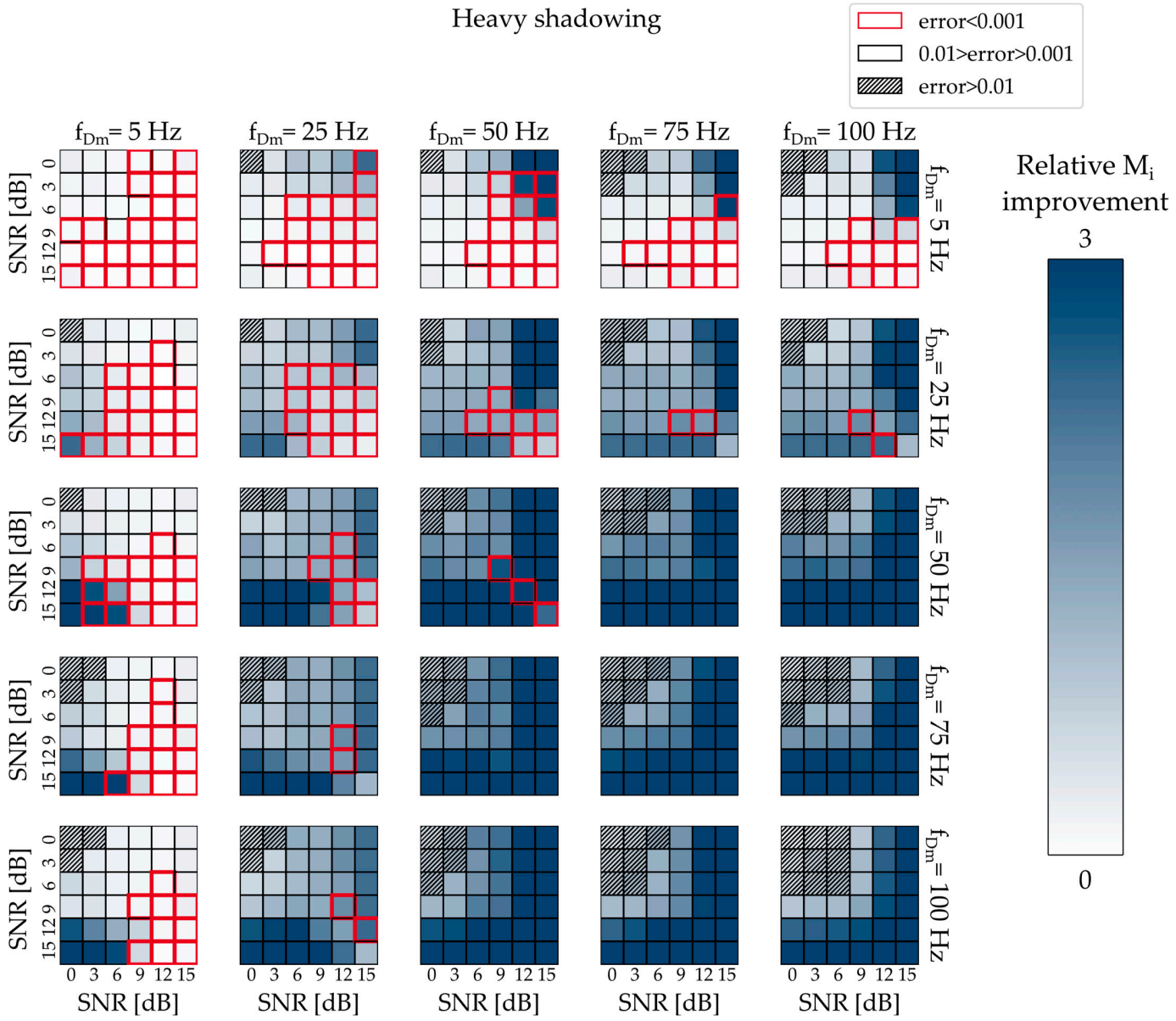


Figure 13. The achieved spectral efficiencies on the test set for various scenarios under heavy shadowing conditions using two channels, comparing OI and NN2. The color of each square represents a relative improvement of the achieved spectral efficiency calculated as $\left(M_i^{NN2} - M_i^{OI}\right)/M_i^{OI}$.

4.4. Discussion

When discussing the results obtained in this paper, it is not feasible to perform a direct comparison with the metrics obtained from other related work as the experimental setup and goals are so diverse. On the other hand, when comparing the methods and approaches that are present, it is possible to see how the performed research is compatible with other approaches and what could be some of the limitations of the approach considered in this paper.

The results obtained in this paper show that neural networks can successfully be used to predict the SNR values for channels with various characteristics. This stands in line with the results obtained from existing work where different neural network architectures were also shown to be successful in predicting CSI in LEO satellite systems [25,26,28]. The presented paper also provides a thorough evaluation of a multitude of different channels by using the implemented simulator and provides a novel approach in terms of improving

spectral efficiency while keeping the transmission error rate above a fixed threshold. The observed thresholds for the transmission error rate (0.01 and 0.001) are not suitable for all applications, but the paper presents a foundation that can be used for future improvements and provides a proof of concept that can be combined with other optimization strategies, such as weather influence [27] and energy efficiency [20].

When discussing the barriers to the practical implementation of the proposed method in terms of energy consumption and computation time, it is important to note that the approach uses a relatively small convolutional neural network that contains 6261 parameters in combination with a simple subtraction of the estimated margin. The time needed to infer a prediction on a single batch containing 64 inputs is 18.7 ms on an AMD Ryzen 7 7840HS CPU. Depending on the usage, it is understood that the current complexity might present an obstacle, but smaller architectures of neural networks could be tested, and the length of the prediction could also be changed to not be a single point but rather several points so the need for inference is not as frequent.

The presented approach has several limitations that should be mentioned. The presented work analyzes three different factors (shadowing, f_{Dm}, and expected SNR) but does not take into account other factors that could influence the channel, so this could be considered in future research. Furthermore, the paper observes a maximum of two satellites, and the scaling of the proposed method into larger systems, both in terms of implementation complexity and efficiency, is yet to be analyzed. The proposed solution is also focused on the DVB-S2X protocol, and its applications in different communication scenarios or with different satellite types were not in the scope of the paper. Finally, there is also an in-depth analysis of different neural network architectures, which was omitted in this paper as it is not the focus, but obtaining better results with smaller models always presents a broad area for future research.

5. Conclusions

The presented paper describes a novel strategy for the application of neural networks to optimize spectral efficiency in land mobile satellite communications. The simulated channels are described by the Rice-shadowed model, and the DVB-S2X satellite protocol is considered. For channel simulation, a range of expected SNR from 0 dB to 15 dB was observed, with five different values of the f_{Dm} alongside three levels of shadowing, resulting in a total of 90 different channels. The proposed machine learning algorithms have shown a consistent improvement in the SNR prediction compared to the baseline OI algorithm. For spectral efficiency improvement, the best results were obtained by using NN2, a neural network that is penalized when overestimating predicted values, in order to provide a sufficiently low error rate while still using the best available MODCOD. The improvements are obtained for all single-channel scenarios but are most prominent for f_{Dm}s of 50 and 75 Hz for single-channel evaluation, regardless of the shadowing level. For dual-channel evaluation, the improvements were most prominent for average shadowing but were also present for all scenarios. The obtained results show promise for future applications of neural networks with specialized loss functions for the optimization of LEO satellite communications, including optimization for user-centric handover procedures.

Author Contributions: Conceptualization, I.V. and S.B.; methodology, I.V., S.B., P.I. and D.D.; software, I.V., P.I. and S.B.; validation, I.V., S.B., P.I. and D.D.; formal analysis, I.V.; investigation, I.V., S.B, P.I., and D.D.; resources, I.V. and S.B.; data curation, I.V.; writing—original draft preparation, I.V. and S.B.; writing—review and editing, P.I. and D.D.; visualization, I.V.; supervision, D.D.; project administration, P.I. and D.D. All authors have read and agreed to the published version of the manuscript.

Funding: Authors acknowledge the support of the Science Fund of the Republic of Serbia, grant No 7750284 (Hybrid Integrated Satellite and Terrestrial Access Network—hi-STAR). This work was also supported by the Ministry of Science, Technological Development, and Innovation of the Republic of Serbia under contract numbers: 451-03-65/2024-03/200103 and 451-03-66/2024-03/200223.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: Author Srdjan Brkić was employed by the company Tannera Technologies LLC. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Mack, E.A.; Loveridge, S.; Keene, T.; Mann, J. A Review of the Literature About Broadband Internet Connections and Rural Development (1995–2022). *Int. Reg. Sci. Rev.* **2023**, *47*, 231–292. [CrossRef]
2. Rappaport, T.S. *Wireless Communications: Principles and Practice*; Cambridge University Press: Cambridge, UK, 2024; ISBN 1009489836.
3. Tripathi, A.; Sindhwani, N.; Anand, R.; Dahiya, A. Role of IoT in Smart Homes and Smart Cities: Challenges, Benefits, and Applications BT. In *IoT Based Smart Applications*; Sindhwani, N., Anand, R., Niranjanamurthy, M., Chander Verma, D., Valentina, E.B., Eds.; Springer International Publishing: Cham, Switzerland, 2023; pp. 199–217. ISBN 978-3-031-04524-0.
4. Wang, J.; Jiang, C.; Kuang, L.; Han, R. Satellite Multi-Beam Collaborative Scheduling in Satellite Aviation Communications. *IEEE Trans. Wirel. Commun.* **2024**, *23*, 2097–2111. [CrossRef]
5. Koo, H.; Chae, J.; Kim, W. Design and Experiment of Satellite-Terrestrial Integrated Gateway with Dynamic Traffic Steering Capabilities for Maritime Communication. *Sensors* **2023**, *23*, 1201. [CrossRef] [PubMed]
6. Sun, Y.; Peng, M.; Zhang, S.; Lin, G.; Zhang, P. Integrated Satellite-Terrestrial Networks: Architectures, Key Techniques, and Experimental Progress. *IEEE Netw.* **2022**, *36*, 191–198. [CrossRef]
7. Kokez, H.A.-D.F. On Terrestrial and Satellite communications for telecommunication future. In Proceedings of the 2020 2nd Annual International Conference on Information and Sciences (AiCIS), Fallujah, Iraq, 24–25 November 2020; pp. 58–67.
8. Gaber, A.; ElBahaay, M.A.; Maher Mohamed, A.; Zaki, M.M.; Samir Abdo, A.; AbdelBaki, N. 5G and Satellite Network Convergence: Survey for Opportunities, Challenges and Enabler Technologies. In Proceedings of the 2020 2nd Novel Intelligent and Leading Emerging Sciences Conference (NILES), Giza, Egypt, 24–26 October 2020; pp. 366–373.
9. Emery, W.; Camps, A. Orbital Mechanics, Image Navigation, and Cartographic Projections. *Introd. Satell. Remote Sens.* **2017**, 565–596. [CrossRef]
10. Tropea, M.; De Rango, F. A Comprehensive Review of Channel Modeling for Land Mobile Satellite Communications. *Electronics* **2022**, *11*, 820. [CrossRef]
11. *ETSI: EN 102 376-1*; Digital Video Broadcasting (DVB); Implementation Guidelines for the Second Generation System for Broadcasting, Interactive Services, News Gathering and Other Broadband Satellite Applications; Part 1: DVB-S2. ETSI: Sophia Antipolis, France, 2015.
12. Fontanesi, G.; Ortíz, F.; Lagunas, E.; Baeza, V.M.; Vázquez, M.Á.; Vásquez-Peralvo, J.A.; Minardi, M.; Vu, H.N.; Honnaiah, P.J.; Lacoste, C. Artificial intelligence for satellite communication and non-terrestrial networks: A survey. *arXiv* **2023**, arXiv:2304.13008.
13. Jain, A.K.; Mao, J.; Mohiuddin, K.M. Artificial neural networks: A tutorial. *Computer* **1996**, *29*, 31–44. [CrossRef]
14. Mall, P.K.; Singh, P.K.; Srivastav, S.; Narayan, V.; Paprzycki, M.; Jaworska, T.; Ganzha, M. A comprehensive review of deep neural networks for medical image processing: Recent developments and future opportunities. *Healthc. Anal.* **2023**, *4*, 100216. [CrossRef]
15. Dhillon, A.; Verma, G.K. Convolutional neural network: A review of models, methodologies and applications to object detection. *Prog. Artif. Intell.* **2020**, *9*, 85–112. [CrossRef]
16. Mehrish, A.; Majumder, N.; Bharadwaj, R.; Mihalcea, R.; Poria, S. A review of deep learning techniques for speech processing. *Inf. Fusion* **2023**, *99*, 101869. [CrossRef]
17. Ardeti, V.A.; Kolluru, V.R.; Varghese, G.T.; Patjoshi, R.K. An overview on state-of-the-art electrocardiogram signal processing methods: Traditional to AI-based approaches. *Expert Syst. Appl.* **2023**, *217*, 119561. [CrossRef]
18. Chang, Y.; Wang, X.; Wang, J.; Wu, Y.; Yang, L.; Zhu, K.; Chen, H.; Yi, X.; Wang, C.; Wang, Y.; et al. A Survey on Evaluation of Large Language Models. *ACM Trans. Intell. Syst. Technol.* **2024**, *15*, 3. [CrossRef]
19. Jabbar, A.; Li, X.; Omar, B. A Survey on Generative Adversarial Networks: Variants, Applications, and Training. *ACM Comput. Surv.* **2021**, *54*, 8. [CrossRef]
20. Liu, Y.; Li, C.; Li, J.; Feng, L. Joint User Scheduling and Hybrid Beamforming Design for Massive MIMO LEO Satellite Multigroup Multicast Communication Systems. *Sensors* **2022**, *22*, 6858. [CrossRef] [PubMed]
21. Bandi, A.; Chatzinotas, S.; Ottersten, B. Joint scheduling and precoding for frame-based multigroup multicasting in satellite communications. In Proceedings of the 2019 IEEE Global Communications Conference (GLOBECOM), Waikoloa, HI, USA, 9–13 December 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–6.
22. Zhang, S.; Jia, M.; Wei, Y.; Guo, Q. User scheduling for multicast transmission in high throughput satellite systems. *EURASIP J. Wirel. Commun. Netw.* **2020**, *2020*, 133. [CrossRef]
23. Bankey, V.; Upadhyay, P.K. Ergodic Capacity of Multiuser Hybrid Satellite-Terrestrial Fixed-Gain AF Relay Networks With CCI and Outdated CSI. *IEEE Trans. Veh. Technol.* **2018**, *67*, 4666–4671. [CrossRef]
24. Li, K.-X.; Gao, X.; Xia, X.-G. Channel Estimation for LEO Satellite Massive MIMO OFDM Communications. *IEEE Trans. Wirel. Commun.* **2023**, *22*, 7537–7550. [CrossRef]

25. Zhang, Y.; Liu, A.; Li, P.; Jiang, S. Deep Learning (DL)-Based Channel Prediction and Hybrid Beamforming for LEO Satellite Massive MIMO System. *IEEE Internet Things J.* **2022**, *9*, 23705–23715. [CrossRef]
26. Guo, R.; Wang, K.; Deng, Z.; Lin, W.; Song, R. A Prediction Model for Channel State Information in Satellite Communication System. In Proceedings of the 2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications, Virtual, 31 August–3 September 2020; pp. 1–6.
27. Zhang, H.; Song, W.; Liu, X.; Sheng, M.; Li, W.; Long, K.; Dobre, O.A. Intelligent Channel Prediction and Power Adaptation in LEO Constellation for 6G. *Netwrk. Mag. Glob. Internetwkg.* **2023**, *37*, 110–117. [CrossRef]
28. Zhang, Y.; Wu, Y.; Liu, A.; Xia, X.; Pan, T.; Liu, X. Deep Learning-Based Channel Prediction for LEO Satellite Massive MIMO Communication System. *IEEE Wirel. Commun. Lett.* **2021**, *10*, 1835–1839. [CrossRef]
29. Milojković, J.; Brkić, S.; Ivaniš, P.; Čiča, Z. Performance of Handover Execution in Satellite Networks with Shadowed-Rician Fading. In Proceedings of the 2023 16th International Conference on Advanced Technologies, Systems and Services in Telecommunications (TELSIKS), Niš, Serbia, 25–27 October 2023; pp. 155–158.
30. Pan, G.; Ye, J.; An, J.; Alouini, M. Latency Versus Reliability in LEO Mega-Constellations: Terrestrial, Aerial, or Space Relay? *IEEE Trans. Mob. Comput.* **2023**, *22*, 5330–5345. [CrossRef]
31. Loo, C. A statistical model for a land mobile satellite link. *IEEE Trans. Veh. Technol.* **1985**, *34*, 122–127. [CrossRef]
32. Abdi, A.; Lau, W.C.; Alouini, M.-S.; Kaveh, M. A new simple model for land mobile satellite channels: First- and second-order statistics. *IEEE Trans. Wirel. Commun.* **2003**, *2*, 519–528. [CrossRef]
33. Abdi, A.; Lau, W.C.; Alouini, M.-S.; Kaveh, M. On the second-order statistics of a new simple model for land mobile satellite channels. In Proceedings of the IEEE 54th Vehicular Technology Conference. VTC Fall 2001. Proceedings (Cat. No.01CH37211), Atlantic City, NJ, USA, 7–11 October 2001; Volume 1, pp. 301–304.
34. Ivaniš, P.; Blagojević, V.; Đorđević, G.T. The method of generating shadowed Ricean fading with desired statistical properties. In Proceedings of the 2023 22nd International Symposium INFOTEH-JAHORINA (INFOTEH), East Sarajevo, Bosnia and Herzegovina, 15–17 March 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 1–6.
35. Ivaniš, P.; Milojković, J.; Blagojević, V.; Brkić, S. Capacity Analysis of Hybrid Satellite–Terrestrial Systems with Selection Relaying. *Entropy* **2024**, *26*, 419. [CrossRef] [PubMed]
36. Van Rossum, G.; Drake, F.L. *Python 3 Reference Manual*; CreateSpace: Scotts Valley, CA, USA, 2009; ISBN 1441412697.
37. Harris, C.R.; Millman, K.J.; van der Walt, S.J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N.J.; et al. Array programming with NumPy. *Nature* **2020**, *585*, 357–362. [CrossRef]
38. Chollet, F. Keras. 2015. Available online: https://keras.io/getting_started/faq/#how-should-i-cite-keras (accessed on 9 September 2024).
39. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
40. Hunter, J.D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95. [CrossRef]