Received: 6 November 2025 | Revised: 28 January 2026 | Accepted: 30 March 2026 | Published online: 20 April 2026

RESEARCH ARTICLE

Journal of Optics and Photonics Research
2026, Vol. 00(00) 1–14
DOI: 10.47852/bonviewJOPR62028116

BON VIEW PUBLISHING

Machine Learning–Enhanced UWOC System for Robust Underwater Communication with FEC Integration

Shreyas Jain[1], Harshit Kumar Sharma[2] and Appala Venkata Ramana Murthy[1,*]

[1] *Department of Applied Physics, Defence Institute of Advanced Technology, India*

[2] *School of AI and Emerging Technology, Lovely Professional University, India*

Abstract: Underwater wireless optical communication (UWOC) often suffers from channel-induced bit errors that degrade adaptive threshold detection, particularly due to dynamic turbidity conditions. Standard approaches of applying mitigation techniques may improve the link performance to an extent. In this study, we have incorporated an integrated approach based on machine learning (ML) into the UWOC system, along with forward error correction (FEC), and tested it on various turbidity levels. This has significantly reduced bit errors and improved link performance. To train the ML models, experimental data were collected using an underwater testbed, with various turbidity levels up to seven nephelometric turbidity units (NTU) to simulate different underwater conditions. The results revealed that across different water models, fluctuations due to angle of arrival (AoA) changes cause a one-order change in magnitude for clear water. In contrast, for harbor water, changes occur only at short ranges. Similarly, the signal-to-noise ratio changes sharply in harbor water compared to clear water as the communication range increases. When we studied the effect of ML models and FEC, we found that an additional 5 m of error-free range can be achieved in underwater environments for 100 Mbps UWOC links using a 2-watt green laser. In different scenarios, ML models alone were able to provide an error reduction from 10^{-1} to 10^{-2} across all water conditions.

Keywords: underwater wireless optical communication, machine learning, seawater, NTU, bit error rate (BER)

1. Introduction

One of the most promising technologies to date is underwater wireless optical communication (UWOC), which could play a key role in future ultra-wideband (UWB) communication networks. The main advantages of UWB radio frequency (RF) technology, which has become a communication solution in wireless sensor technology, are low power consumption, low transmit power, and immunity to multi-path fading for underwater networks [1]. We are motivated to create UWOC/UWB hybrid links since underwater UWOC/RF hybrid links are essential and are receiving a lot of attention. UWOC offers significant advantages over traditional acoustic and RF communication systems in underwater environments. These advantages include low latency, high bandwidth, and reduced signal absorption [2]. While acoustic communication is known for long-range capabilities, it suffers from high latency, low data rates, and interference caused by surface tides and ambient noise [3]. In contrast, UWOC operates in the blue-green region of the electromagnetic (EM) spectrum—identified through the EM absorption spectrum as having minimal absorption compared to other optical wavelengths [4]. Although optical signals experience higher attenuation than acoustic waves over long distances [5, 6], because they are much less vulnerable to interference from outside elements such as surface turbulence, UWOC is a great option for high-speed, short-range underwater communication. Reconfigurable intelligent surface (RIS)-enabled UWOC/RF links for IoUT, mixed RF/UWOC links without a relay node [7], cooperative relaying-based UWOC/RF links [8], and also free space optical/UWOC hybrid links were facilitated by space division multiplexing [9].

In recent years, various researchers have demonstrated the efficacy and potential of UWOC systems for underwater applications [10]. However, challenges still persist—particularly due to the inherent properties of the underwater channel, such as scattering, turbulence, and absorption variability [11]. Ongoing research continues to focus on mitigating these limitations and improving the reliability and adaptability of UWOC systems across different underwater conditions [8, 12–14].

To address the challenges inherent in UWOC systems, researchers have been actively studying underwater channel characteristics and developing receiver-side post-processing techniques. For instance, Nootz et al. [15] quantified the effects of optical turbulence, while Jamali et al. [16] performed statistical analyses of turbulence-induced fading. Building on these insights, subsequent studies aimed to improve system performance under

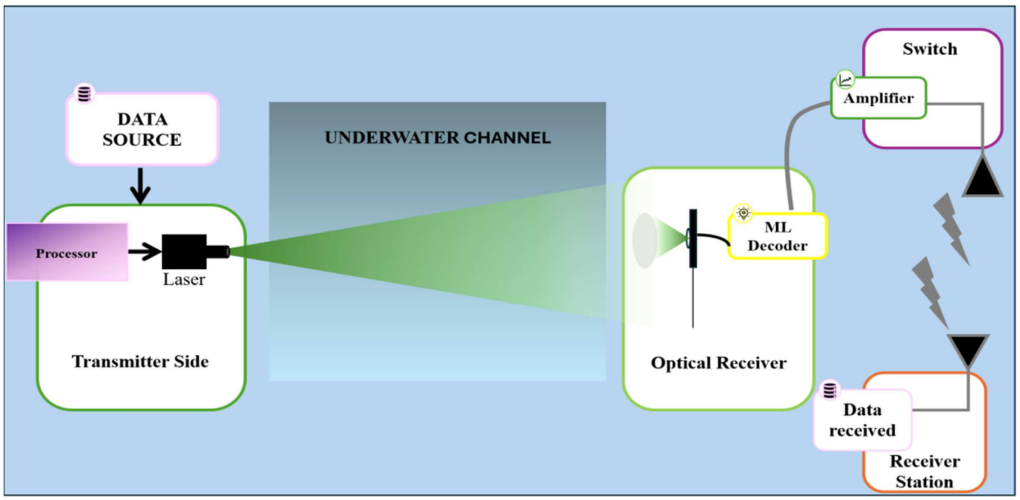
***Corresponding author:** Appala Venkata Ramana Murthy, Department of Applied Physics, Defence Institute of Advanced Technology, India. Email: avrmurthy@diat.ac.in

turbulent conditions. Gökçe et al. [17] demonstrated how aperture averaging could enhance bit error rate (BER) performance in such environments. In another notable study, Zhang et al. [18] proposed a diffuse-based receiver to mitigate the effects of scintillation. Using an experimental testbed, Devappa et al. [19] assessed the potential application of baseband modulation in conjunction with customized forward error correction (FEC) for error reduction. Priyalakshmi and Mahalakshmi [20] demonstrated how to employ HVD Turbo in UWOC when data are sent over a vertical link utilizing the M-ary OAM-PPM modulation approach. For UWOC links, Miroshnikova and Sattarova [21] have demonstrated the effective usage of low-density parity-check (LDPC), polar code with orthogonal chip division multiplexing (OCDM) modulation, and product and convolutional Turbo codes. By simulating a channel using a composite fading channel model, which integrates the combined effects of the generalized gamma distribution, oceanic turbulence, zero-boresight pointing error, and Elamassie underwater path, Wang et al. [22] attempted to analyze system performance using BER. This aids in precise channel modeling. Further contributions include Qasem et al. [23], who introduced a pilot-assisted UWOC system aimed at reducing the peak-to-average power ratio—a critical parameter in optical link performance. Wang et al. [24] demonstrated a real-time duplex UWOC system in a typical 50-m swimming pool based on several blue/green light sources using multi-pixel photon counters with 4-QAM orthogonal frequency division multiplexing and convolutional coding. For an orthogonal frequency division multiplexing (OFDM)-based UWOC system, Li et al. [25] suggested and demonstrated a straightforward sampling frequency offset estimation and compensation scheme based on two phase-conjugated pilots, which reduced the system's error vector magnitude. A predictive equalization method for underwater wireless optical camera communication, assuming color shift keying, was proposed and implemented by Shigenawa et al. [26] based on the varying wavelength-related attenuation of light intensity.

In recent years, the integration of machine learning (ML) techniques into UWOC systems has gained attention. For example, Zhao et al. [27] explored ML-based link adaptation, and Kamatchi et al. [28] investigated ML-aided signal detection for enhanced underwater optical communication. To lower BER in UWOC systems, Lu et al. [29] employed deep learning-based channel estimation and prediction. Salam and Aly [30] employed UWOC in underwater wireless sensor networks for object monitoring using ML models based on long short-term memory, spatial attention, recurrent neural networks, Transformer, and gated recurrent units. K-nearest neighbors (kNN) based demodulation method for QAM-16 and QAM-32 modulation developed by Nennouche et al. [31]. Kong et al. [32] showed how to use YOLOv5s to minimize pointing error and ensure appropriate acquisition during link formation. Analogously, ML is integrated into other aspects of the UWOC system, including beam optimization for multiple-input multiple-output (MIMO) systems, signal identification, RIS switching, neural network-based convolutional coding for error coding, adaptive optics, and many more.

In a conventional OWC system, it is the fixed voltage threshold that decides the bit stream and classifies '1's and '0's, known as hard decision decoding (HDD), which compares matched filter output against a threshold, assuming perfect transmitter–receiver synchronization. On the other hand, adaptive threshold decoding dynamically adjusts these levels and can be implemented with low complexity using ML models. These methods enhance current UWOC system architectures, such as MIMO, adaptive power transfer, and FEC, by improving robustness and reliability. This paper proposes a quantitative ML approach with experimental data. We have introduced ML classifiers along with FEC optimization, enabling dynamic adaptation to changing underwater conditions. Primarily, turbidity variations and angle-of-arrival (AoA) fluctuations. Classifiers trained on diverse turbidity datasets enable adaptive threshold decoding, improving communication reliability across scenarios. Among various combinations of ML and FECs, results suggest that the kNN and the Turbo code combination performed better in all turbidity conditions and are presented in this paper. Figure 1 illustrates the hybrid UWOC/UWB network that forms the foundation of this study. The remainder of the paper is organized as follows: Section 2 describes the underwater channel modeling approach for numerical replication of diverse underwater environments. Section 3 details the experimental setup for data collection and validation. Section 4 presents a comprehensive BER performance analysis across ML models and FEC schemes in various underwater conditions. Section 5 summarizes key findings, conclusions, and future research directions for intelligent adaptive UWOC systems.

Figure 1
Underwater wireless optical communication as part of ultra-wideband networks



2. Theoretical Framework

2.1. Machine learning

We intend to develop ML models for signal classification and decoding to improve receiver performance. Because they strike a compromise between accuracy and ease of implementation, five small, hardware-friendly ML algorithms were selected.

The five ML models are:

1) **Logistic regression** is a supervised ML algorithm used for binary classification. It models the probability of a data point belonging to a specific class using a logistic (sigmoid) function, providing outputs between 0 and 1 for decision-making [33]. The mathematical representation of the probability of an input belonging to class 1 is given by

$$P(Y=1) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n)}} \quad (1)$$

where $P(Y = 1)$ is the probability of the positive class, β are the coefficients, and X are the input features.

2) **Support vector machine** (SVM) is a learning algorithm used for classification. It finds the optimal hyperplane that maximizes the margin between different classes, ensuring robust separation and effective performance even with high-dimensional or nonlinear data [34].

The objective function of SVMs aims to maximize the margin while minimizing classification error, expressed as

$$\min_{w,b} \frac{1}{2} ||w||^2 + C \sum_{i=1}^{N} \xi_i \quad (2)$$

Subject to:

$$y_i (w \cdot x_i + b) \geq 1 - \xi_i$$
$$\xi_i \geq 0 \quad (3)$$

where x is the input feature vector, w is the weight vector, and b is the bias term. The margin, inversely proportional to the norm of the weight vector $||w||$, represents the distance between the decision boundary and the nearest data point.

3) **Decision tree** (DT) is an algorithm used for classification and regression. It splits data into branches based on feature values, forming a tree-like structure that makes decisions by following paths from the root node to leaf nodes [34]. Mathematically, a DT can be represented as follows:

$$f(x) = \sum_{m=1}^{M} c_m \cdot I(x \in R_m) \quad (4)$$

f(x) represents the prediction made by a DT for input x, where M is the total number of regions (leaf nodes), c_m is the predicted value for region R_m, and I ($x \in R_m$) is an indicator function that equals 1 if x belongs to R_m and 0 otherwise. This equation defines the DT's function, assigning predictions based on the input's region.

4) **K-nearest neighbors** (kNN) is a supervised algorithm that assigns a class to a sample by examining the k closest points in the dataset. The decision depends on the majority label among these neighbors, using distance-based similarity [35].

5) **Naïve Bayes**: Naïve Bayes (NB), based on Bayes' theorem, assumes feature independence [36]. Despite its simplicity, it performs well, especially in tasks such as text classification and spam filtering. It calculates the posterior probability $P(C_k|x)$ for each class C_k given the input features x, and assigns the class with the highest probability.

$$\hat{y} = \text{argmax} P(C_k|x) \quad (5)$$

Despite assuming feature independence, it often delivers competitive performance, particularly in text classification tasks.

2.2. Modulation formats and forward error correction codes

In optical wireless communication systems, modulation formats play a crucial role in determining transmission efficiency, power utilization, and noise resilience. They define how information is encoded onto the optical carrier using variations in light intensity, pulse position, or timing. Different schemes such as Hamming codes add parity bits at specific positions in the data to detect and correct single-bit errors. They provide efficient, low-overhead error correction and are widely used in memory and digital communication systems [37]. The other is Repeat codes that work by transmitting each bit multiple times to ensure redundancy. The receiver decides the original bit using majority voting, providing simple but low-efficiency protection against random transmission errors [38]. BCH (Bose–Chaudhuri–Hocquenghem) codes are cyclic error-correcting codes capable of detecting and correcting multiple random bit errors. They use polynomial algebra for encoding and decoding, offering strong reliability in noisy communication channels [39]. Reed–Solomon (RS) codes operate on blocks of data symbols rather than individual bits. They correct burst errors by using polynomial encoding over finite fields, making them ideal for storage, optical, and wireless systems [40]. Turbo codes combine two or more convolutional encoders with interleaving and iterative decoding. They achieve near–Shannon limit performance, offering exceptional error correction efficiency in modern communication systems such as deep-space and 4G/5G networks [41]. The modulation schemes are important. Some baseband modulation techniques include on–off keying (OOK), a simple intensity modulation method in which binary data are represented by the presence (on) or absence (off) of light. It offers easy implementation but is sensitive to noise and power fluctuations [42]. Return-to-zero is a line coding scheme where the signal returns to zero between each bit interval. A '1' is represented by a pulse shorter than the bit period, improving synchronization but increasing bandwidth requirements [43]. Pulse position modulation (PPM) conveys information through the position of a light pulse within a fixed time slot. It offers high power efficiency but requires precise synchronization and is sensitive to timing jitter [44]. Pulse interval modulation encodes data by varying the time interval between successive light pulses. It provides improved bandwidth efficiency and lower average power consumption, making it suitable for optical wireless and underwater communication systems [45].

2.3. Underwater channel

We modeled the underwater optical channel to study the effects of different impairments in the UWOC links. Underwater environments vary significantly depending on geographical and environmental conditions. However, based on previous studies,

Table 1
Optical properties of three different oceanic waters

Water type	Absorption (a m^{-1})	Scattering (b m^{-1})	Attenuation (c m^{-1})	Single-scattering albedo (ω_0)
Clear ocean	0.114	0.037	0.151	0.25
Coastal ocean	0.179	0.219	0.398	0.55
Turbid harbor	0.366	1.824	2.190	0.83

water types can be broadly categorized, as shown in Table 1, and it can be said that UWOC links typically suffer from three main impairments: (1) inherent properties such as absorption and scattering, (2) fading due to turbulence, and (3) AoA fluctuations as the optical signal propagates through water. The beam attenuation coefficient c serves as a key parameter in modeling the underwater channel, as it encapsulates the inherent optical properties contributing to signal degradation such as attenuation due to absorption and scattering. In a homogeneous medium, the beam attenuation over a finite path length can be mathematically defined as:

$$c = -\frac{\ln(\Delta L)}{d} \tag{6}$$

1) ΔL is the relative change in radiance along the path
2) d is the path length of propagation

Another important parameter is single-scattering albedo, denoted by ω_0, which indicates the scattering strength in a certain water type. According to Table 1 [46], it is evident that clear ocean is absorption dominant ($\omega_0 < 0.5$), and coastal waters have comparable absorption and scattering, while harbor water is scattering dominant ($\omega_0 > 0.5$). When scattering becomes more dominant in water, this leads to a greater divergence in the propagating optical beam. Also, there is a change in the scattering phase function of three different waters. Combining all these effects leads to a change in the scattering angle. The laser divergence is an apparent optical property of the UWOC system that is due to inherent properties; as a result, AoA at the receiving plane changes and can directly affect the received power and the AoA. The scattering angle and received power depend on the turbulence. In this study, we have used the Fisher–Snedecor F turbulence model, which fits better on experimental data than the Lognormal and Gamma-Gamma models [47].

Coming to the model, assuming that the optical wave propagates through the F turbulence channel along with attenuation and AoA fluctuations, the received signal can be written as:

$$y = \eta h\, x + n \tag{7}$$

where η is the photodiode quantum efficiency, x is the transmitted signal, and n is the additive white Gaussian noise (AWGN) noise, while h is the combined channel response that can be written as:

$$h = h_{al} h_{at} h_{af} \tag{8}$$

where h_{al}, h_{at}, and h_{af} represent the channel effects due to attenuation, turbulence, and AoA fluctuations, respectively. The instantaneous signal-to-noise ratio (SNR) for this system model, considering the IM/DD, can be written as [48]:

$$\mathrm{SNR} = \frac{(\eta h)^2}{N_0} \tag{9}$$

2.3.1. Attenuation

As mentioned before, all the water types cause the strong attenuation of the optical beam due to absorption and scattering, and this attenuation can be characterized using beam attenuation coefficient c, and hence, h_{al} can be further written as [49]:

$$h_{al} = e^{-c \cdot d} \tag{10}$$

2.3.2. Turbulence

Channel effect due to turbulence is taken as the F turbulence model because of its simpler mathematical shape in terms of fundamental elementary functions; the F distribution is more enticing. The probability density function of the UWOC channel, assuming that it is driven by the F distribution, can be expressed as [47]:

$$f_{h_{at}}(h_{at}) = \frac{a^a (b-1)^b h_{at}^{a-1}}{B(a,b)(a h_{at} + b - 1)^{a+b}}, \quad h_{at} > 0 \tag{11}$$

where the Beta function is represented by $B(\cdot\ ,\cdot)$ and $a = \frac{1}{\exp(\sigma^2_{\ln X}) - 1}$ and $b = 1 \exp\left(\sigma^2_{\ln Y}\right) - 1$ Here, the log variances showing the impact of large and small turbulent scales are denoted by $\sigma^2_{\ln X}$ and $\sigma^2_{\ln Y}$.

2.3.3. AoA fluctuations

The Henyey–Greenstein phase function was used to simulate various photon trajectories under random scattering disturbances in order to characterize the AoA fluctuation. With a mean of μ_{AOA} and a standard deviation of σ_{AOA}, the distribution of AoA was determined to be roughly Gaussian. Table 2 summarizes all the simulation parameters used in this study, which were employed to generate. Figure 2 shows the characterization of different water types.

3. Experimental Setup

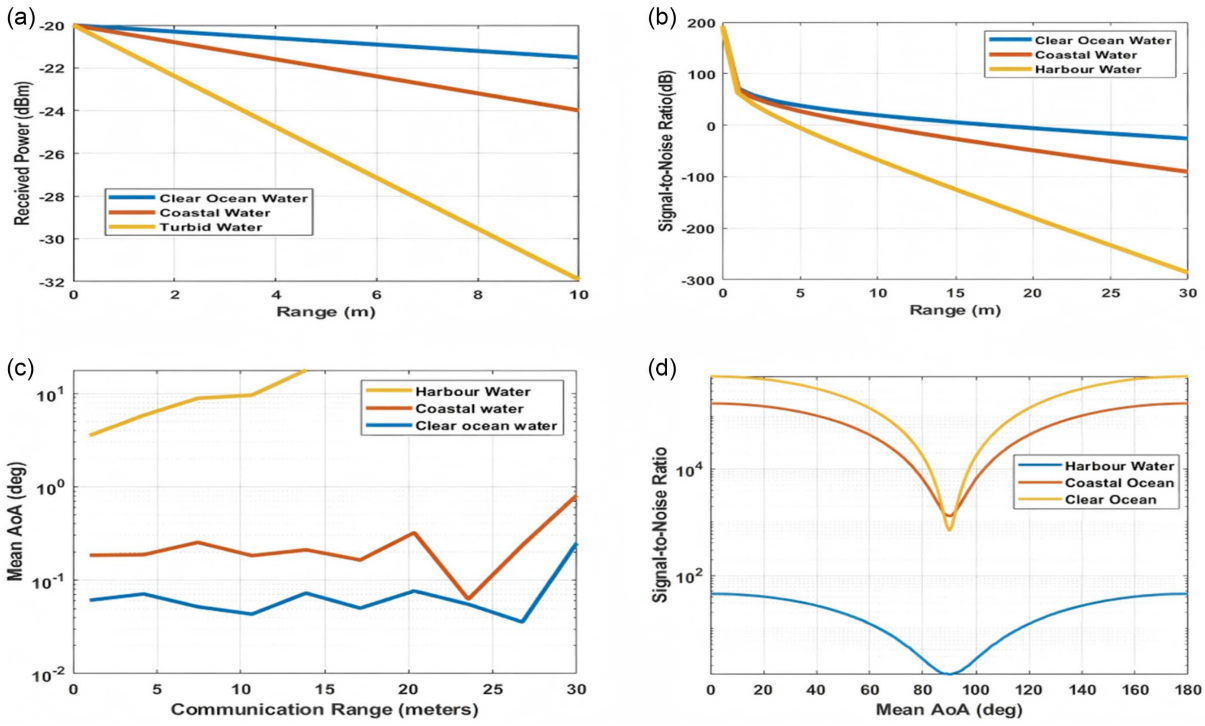
The experimental setup was specifically designed for data collection to support ML model training and validation. The goal of this system was to generate a diverse dataset that captures the effects of varying water turbidity levels (quantified in nephelometric turbidity unit (NTU)) on the performance of an UWOC ink. By systematically varying the NTU values, the attenuation coefficient and scattering effects within the water medium were altered. An increase in NTU corresponds to a higher concentration of suspended particles, which in turn enhances light scattering and absorption, thereby degrading the received optical signal. This controlled variation allowed us to replicate different underwater visibility conditions and establish a realistic test environment for ML-based analysis.

To complement the simulation studies, we constructed a laboratory-scale underwater optical link, as illustrated in Figure 3. In this setup, a green laser diode was used to transmit a binary data sequence of 10,000 bits through a 0.5-m-long water tank filled

Table 2
The detailed system parameters of simulations

Parameter	Symbol	Value/range	Unit	Description
Transmitted optical power	Pt	2	W	Optical power emitted by the transmitter
Solar luminance	LSol	119	–	Background solar luminance in underwater environment
Field of view	FOV	$\pi/180$	rad	Receiver field of view
Integration time	Δt	12×10^{-9}	s	Photodetector integration period
Receiver aperture diameter	Dr	0.1	m	Effective receiver aperture size
Electron charge	q	1.6×10^{-19}	C	Elementary charge of an electron
Photodiode responsivity	s	0.58	A/W	Responsivity of the photodiode
Excess noise factor	F	1	–	Noise amplification factor of photodiode
Bandwidth	Bw	1.0×10^{8}	Hz	System electrical bandwidth
Equivalent noise bandwidth	Bwen	$\pi \times$ B_w / 2	Hz	Noise-equivalent bandwidth
Photodetector efficiency	η	0.8	–	Quantum efficiency of photodetector
Receiver lens parameter	d	0.025	m	Focal length/effective optical parameter of receiver
Mean angle of arrival	$\theta 0$	0	rad	Mean AoA (nominal optical alignment)
Std. deviation of AoA jitter	$\sigma\theta$	2°, 5°, 10°	deg	Standard deviation of AoA fluctuation to simulate misalignment
Random AoA fluctuation	θfluc	N ($\theta 0$, $\sigma\theta 2$)	rad	Randomly generated AoA samples for each range value

Figure 2
Comparative analysis of underwater optical communication in varying water environments: (a) power vs range, (b) SNR vs range, (c) mean AoA vs range, and (d) SNR vs mean AoA



with water samples of varying NTU levels. The transmitted optical signal was detected at the receiver end using a photoreceiver.

Due to the influence of ambient light, background noise, and fluctuations in the scattering medium, the received signal intensity exhibited variations and distortions. The analog signal from the photoreceiver was continuously sampled using a processing unit at a rate in the kilobits per second (kbps) range. The processing unit is specifically designed to digitize the analog values and transfer the data to a laptop via serial communication for storage and further analysis.

This procedure was repeated for different NTU levels, and for each condition, 10,000 transmitted bits and their corresponding analog voltage samples were collected. The resulting dataset forms the foundation for training and validating ML classification

Figure 3
(a) Experimental setup for underwater optical communication, where turbidity levels were varied using Maalox to simulate different scattering conditions, and (b) variation of the attenuation coefficient with respect to NTU, showing the increase in optical loss with higher turbidity

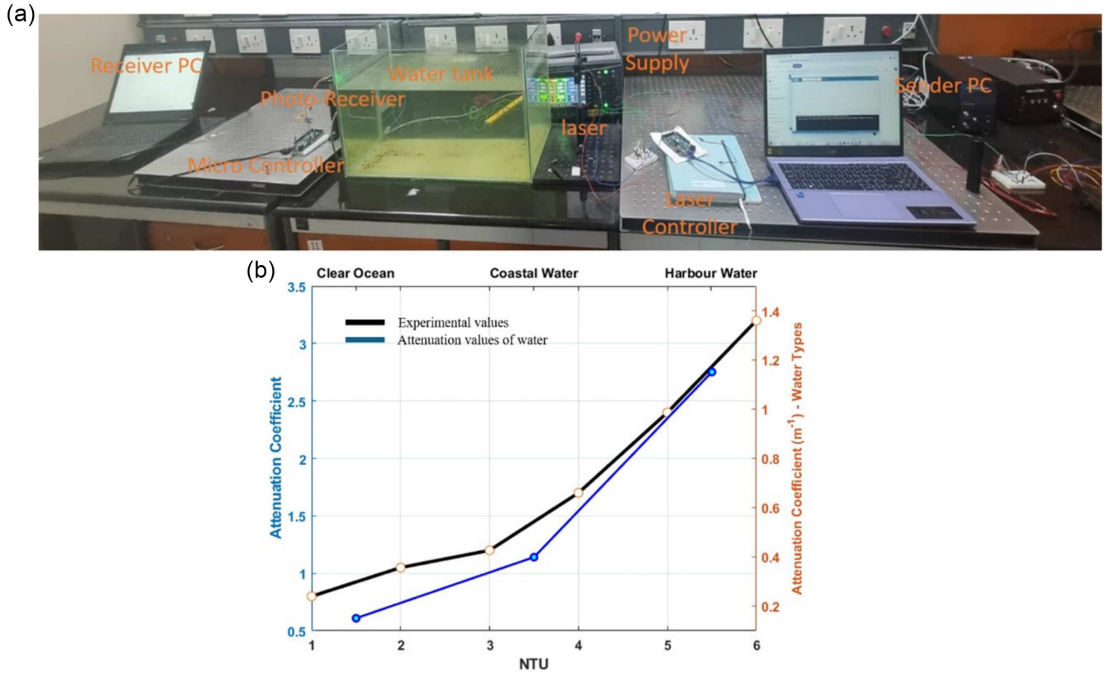


Table 3
Optimized hyperparameters of the trained models

Models	SVM	DT	kNN	NB
Hyperparameters	Kernel: Radial Basis Function, $C = 0.525$, $\gamma = 0.256$	Trees: 106, Depth: 700	Distance: cosine; Number of neighbors: $k = 5$	Gaussian Naïve Bayes Using variance smoothing: 10–8

models, enabling the algorithm to learn the statistical relationships between transmitted bits and their noisy received analog counterparts under varying channel conditions. After applying the classification algorithms to the received data, the BER was computed by comparing the reconstructed bitstream with the original transmitted sequence, thus evaluating the ML model's performance across different turbidity conditions.

The data obtained from this setup form the foundation for training and validating ML classification models. We used K-fold cross-validation to optimize parameters before training (Table 3). Training enabled the algorithm to learn the statistical relationships between transmitted bits and their noisy received analogue counterparts under varying channel conditions.

4. Result and Discussions

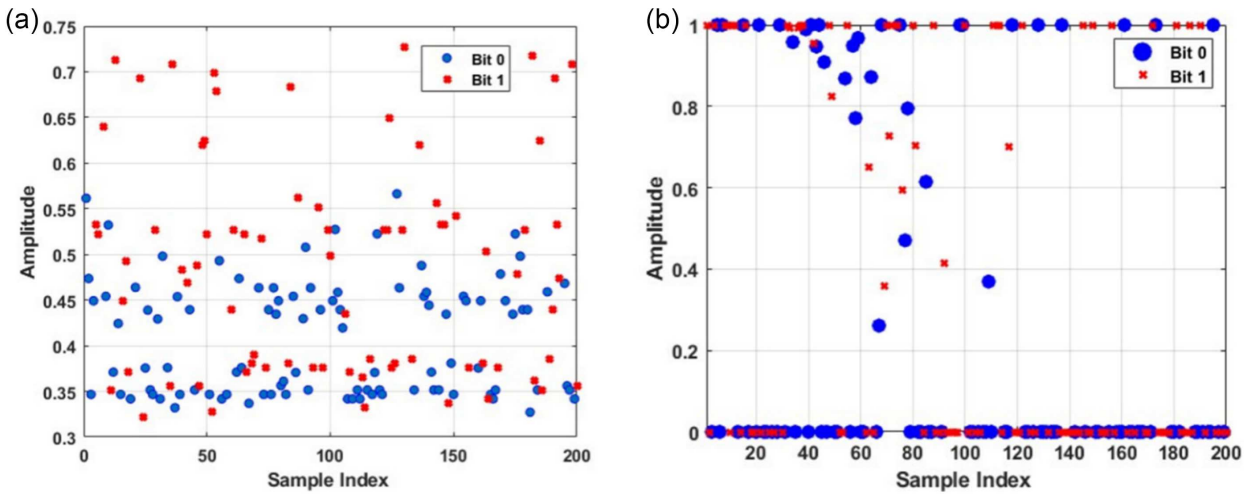
This section investigates the correlation between training and testing datasets, the evaluation of the ML model's data decoding capability using a validation dataset, how key channel and system parameters affect the UWOC system performance, and the collaboration of FEC codes at the receiver node to minimize BER, thereby enhancing system reliability.

4.1. Relationship between training and testing data

To establish the reliability of the ML-assisted adaptive threshold detection (ATD) model, both training and testing datasets have been analyzed using a single-dimensional feature: the amplitude of received symbols/bits. We obtained the training dataset from the experimental setup (Section 3). Figure 4 shows the feature (amplitude) distributions of the training and testing datasets. The feature values in both cases span similar ranges but exhibit different variances and distributions. This demonstrates the dataset's statistical compatibility despite using kbps/0.5 m experimental link versus our modeled UWOC link. The datasets are mutually independent from temporal tendencies because HDD is used after the matched filter; therefore, the matched filter outputs produce consistent feature ranges regardless of data rate or distance scaling. The experimental 0.5 m tank attenuation (via controlled NTU variation) matches the UWOC link model in different types of waters through Beer–Lambert scaling laws, with ergodic channel assumptions [50, 51] ensuring fast-fading averages over 100 Mbps symbol periods.

The testing data were generated through MATLAB-based simulations of clear, coastal, and harbor waters. The experimental data exhibited higher randomness and noise. The similarity

Figure 4
The amplitude value distribution of bits '1' and '0' of (a) test data and (b) training data



in statistical trends between the two datasets ensured that the ML model could generalize well across different channel conditions, forming a strong foundation for adaptive detection and node switching. The statistical measurements reported in Table 4 (variance, inter-class distance, Fisher's discriminant ratio, and silhouette score) quantitatively validate that the training data are more ambiguous. Figure 3(b) shows the class distribution overlap in the training data, which is explained by low inter-class distance and low total variance in Table 4.

Table 4
Summary table of ambiguity and separability analysis for training vs test data

Metric	Training data	Testing data
Variance	0.0084	0.5500
Intra-class variance	0.0077	0.5501
Inter-class distance	0.0802	0.0097
Fisher's discriminant ratio	0.8323	0.0002
Silhouette coefficient	0.2995	–0.0009

4.2. Validation of training

A subset of the experimental data that was not used for training has been used to validate the training model (around 20%). After several repetitions, the validation accuracy stabilized, suggesting convergence and successful learning of turbidity-induced intensity changes. Figure 5 shows a confusion matrix of various ML models. It is suggested that DT performs best with this validation set, with an accuracy of 77%, while NB performs the least with an accuracy of 74.27%, and all other models were able to achieve an accuracy of more than 75% on this validation set.

We tested the performance of ML models on each batch of data with a different NTU. Figure 6 shows plot precision and recall, which indicates that the model becomes inefficient for higher NTUs because of more ambiguous data. As NTU exceeds three NTU, the data amplitude values of 1 and 0 overlap more and more, which lowers system performance.

4.3. Link performance in clear waters

Figure 7 shows the BER performance of the UWOC system with integration of different FEC codes and ML models in clear water conditions. Till around 20 m, the link remains nearly error-free across all combinations. Beyond this range, there is an increase in BER as signal attenuation and scattering effects become dominant. Figure 7 shows that without FEC, ML-based models, particularly DT and kNN, show significant improvement over the non-ML baseline (W/O ML), maintaining lower BER values across longer ranges. The Hamming codes are used, as we can see that all ML models exhibit enhanced performance, with SVM, DT, and kNN achieving stable low BER up to approximately 25 m, emphasizing the benefit of coupling simple coding with ML-assisted decoding. The Repeat 3 code (Figure 7(c)) further improves reliability up to around 27 m; however, its effectiveness decreases beyond this distance. Still, kNN and DT maintain consistent decoding accuracy. Incorporating BCH coding (Figure 7(d)) results in a more robust system, sustaining BER below 10^{-3} up to nearly 28 m, indicating strong error resilience when coupled with ML-based approaches, particularly kNN and SVM. RS coding (Figure 7(e)) enhances link stability over a wider range, where DT and kNN deliver the lowest BER values even under intensified signal degradation. The Turbo code demonstrates the best overall performance, with all ML-assisted models achieving extremely low BER even beyond 30 m. Among these, kNN and NB stand out for their robustness and adaptability to noise and scattering variations. Overall, these results confirm that integrating ML classifiers with FEC codes significantly improves UWOC link reliability and provides a more error-free range. Among all combinations, the combination of Turbo or RS coding with kNN or DT models provides the most effective balance between computational efficiency and error resilience. The superior performance of DTs arises from their ability to form piecewise nonlinear decision regions, which better accommodate the locally varying signal statistics induced by AoA jitter.

4.4. Link performance in coastal waters

As seen in Figure 2(b), the coastal water model has higher levels of attenuation, scattering, and turbulence, all of which

Figure 5
Confusion matrix of different classifiers: (a) logistic, (b) SVM, (c) DT, (d) kNN, (e) NB, and (f) BER comparison of classifiers using the validation dataset

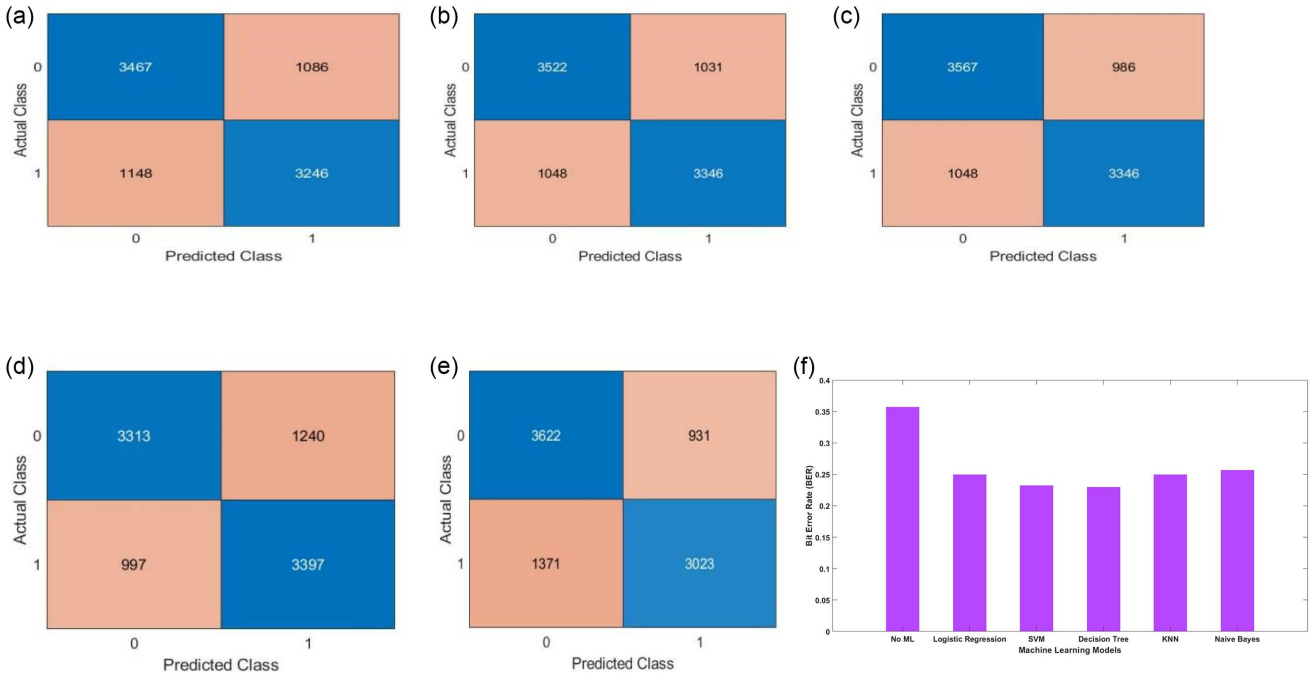
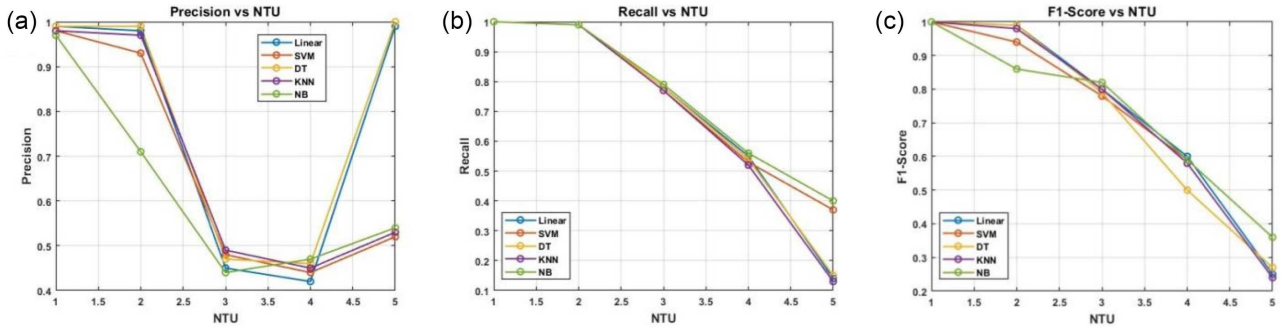
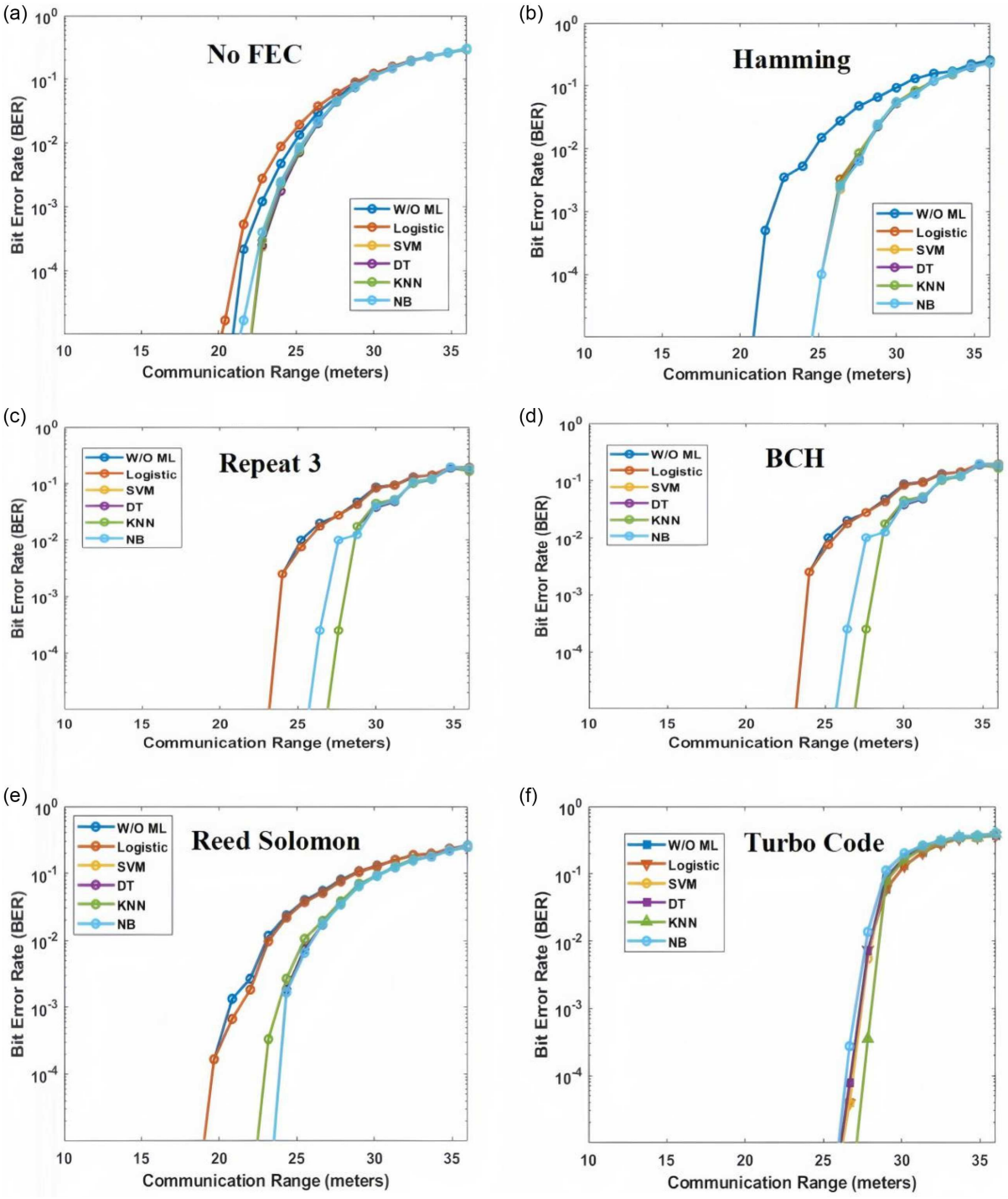


Figure 6
The figure shows the variation of three key classification metrics—(a) precision, (b) recall, and (c) F1-score—with respect to each NTU dataset



contribute to a lower SNR, making coastal waters a much more challenging environment for communication links. In Figure 8(a), transmission remains error-free at lower BERs up to approximately 16 m, where SVM, DT, and kNN outperform the traditional approach, particularly maintaining BER around 1×10^{-3}~10^{-4}, at 20 m. Beyond this range, BER escalates sharply across all decoders, with NB and Logistic regression providing only marginal improvements over conventional methods, demonstrating the critical need for error correction in extended-range underwater optical communication. In Figure 8(b), the integration of Hamming coding decreases the BER, with SVM, DT, and kNN achieving approximately 1×10^{-4}. at greater ranges compared to the no-FEC scenario. However, NB exhibits degraded performance under Hamming coding, suggesting lower compatibility with this error correction scheme, while logistic regression remains marginally stable but inferior to tree-based and distance-metric ML bit classifiers. In Figure 8(c), repeat coding delivers near-error-free transmission up to 20 m for DT and kNN, which demonstrates exceptional robustness in the near range; however, its benefits diminish beyond 23 m, where BER increases steeply, and the conventional approach and logistic regression notably lag behind memory-based and tree-based models. In Figure 8(d), BCH coding provides better system reliability, maintaining error-free up to approximately 20 m for all classifiers except conventional, with DT and kNN consistently delivering the lowest BER across extended ranges. The kNN benefits from the local clustering of received signal features caused by scattering-induced angular spread, enabling robust classification despite global distribution distortion. Even conventional and logistic regression methods achieve error-free operation up to 17 m, underscoring BCH's superior error correction capability compared to previous schemes. Figure 8(e) – Reed Solomon: Reed Solomon coding promotes link stability with all classifiers sustaining error-free transmission close to 14 m, though its effectiveness is less

Figure 7
BER vs communication range for OOK transmission in clear ocean water using different machine learning models with various FEC techniques: (a) no FEC, (b) Hamming, (c) Repetition codes, (d) BCH, (e) Reed–Solomon, and (f) Turbo codes

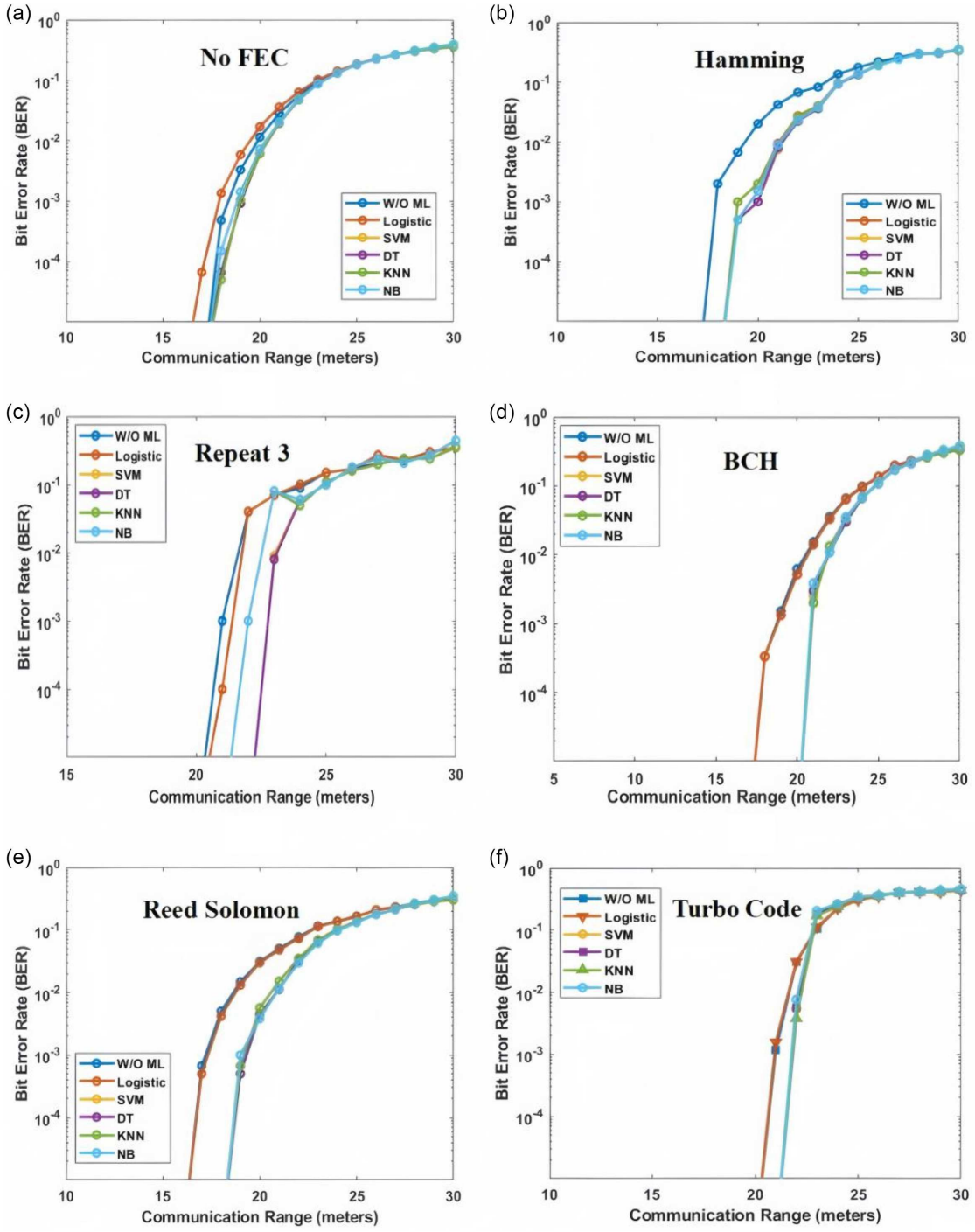


pronounced than BCH, with DT and kNN enabling marginally extended ranges while ML models excluding logistic regression maintain sustained low BER at moderate distances. Figure 8(f) – Turbo Code: Turbo code produces the flattest and most resilient BER curve, achieving error-free operation up to 17 m for all classifiers and extending robust performance at even larger ranges, where both logistic regression and the conventional approach show marked improvement. The superior stability of DT and kNN persists across all FEC schemes, demonstrating their exceptional reliability and adaptability, making the combination of BCH or Turbo coding with DT or kNN the optimal choice for energy-efficient, low-power underwater optical communication in coastal environments.

4.5. Link performance in harbor waters

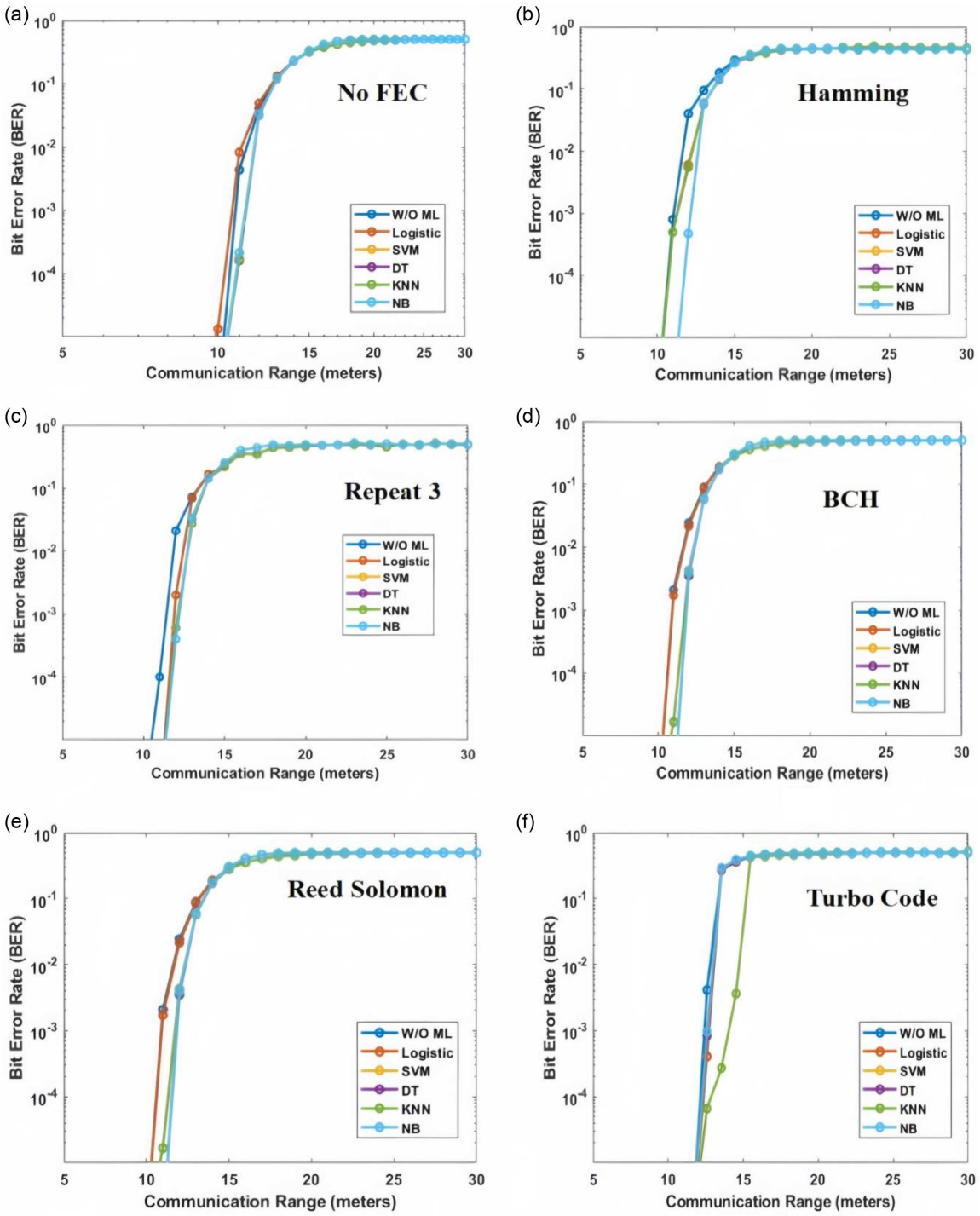
Figure 9 presents the BER variation of the UWOC system under turbid water conditions for different FEC codes integrated

Figure 8
BER vs communication range for OOK transmission in coastal waters using different machine learning models with various FEC techniques: (a) no FEC, (b) Hamming, (c) Repetition codes, (d) BCH, (e) Reed–Solomon, and (f) Turbo codes



with ML models. Due to stronger scattering and absorption, BER rises more rapidly compared to clear water, with reliable communication achieved up to around 10–12 m in most cases. Without FEC (Figure 9(a)), DT and kNN again show notable improvements over W/O ML, achieving lower BER and extending the communication range slightly. Applying Hamming coding (Figure 9(b)) stabilizes the link, where SVM, DT, and kNN maintain BER values below 10^{-2} up to nearly 13 m, showing effective correction even under increased turbidity. The Repeat 3 code (Figure 9(c)) offers better consistency up to 14 m, though its benefits diminish beyond this range. kNN continues to perform reliably, minimizing BER growth as range increases. The BCH scheme (Figure 9(d)) enhances robustness further, keeping BER below 10^{-3} up to 15 m for DT and kNN, reflecting stronger correction capability. With RS coding (Figure 9(e)), the ML-assisted system maintains improved stability, especially for kNN and SVM, which outperform other classifiers across the full range. The Turbo code (Figure 9(f)) demonstrates

Figure 9
BER vs communication range for OOK transmission in harbor waters using different machine learning models with various FEC techniques: (a) no FEC, (b) Hamming, (c) Repetition codes, (d) BCH, (e) Reed–Solomon, and (f) Turbo codes

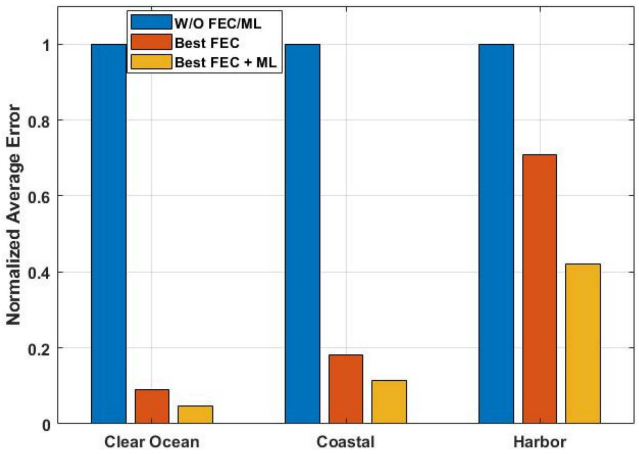


superior error suppression; all ML models maintain nearly error-free performance up to around 16 m, with kNN and DT achieving the lowest BER values overall.

These results highlight that even in turbid environments, combining ML models with suitable FEC significantly improves underwater link performance. Among all configurations, Turbo and RS codes paired with kNN or DT offer the best trade-off between complexity and reliability, ensuring stable communication despite strong scattering and attenuation effects.

The normalized average error analysis reveals a substantial improvement in system performance when FEC and ML techniques are applied under varying water conditions. Without any FEC or ML, the baseline errors were highest in harbor water and lowest in clear ocean water, consistent with the increasing turbidity and optical scattering effects. Incorporating the best FEC schemes notably reduced errors across all conditions—by approximately 91%, 82%, and 29% for clear, coastal, and harbor waters, respectively, as shown in Figure 10. Among the tested

Figure 10
Normalized error across different water conditions for the best-performing FEC and ML schemes: clear ocean waters using kNN with Turbo codes, coastal waters using DT with BCH codes, and harbor waters using kNN with Turbo codes



FEC methods, the Turbo code demonstrated the best performance in clear and harbor waters, while the Repeat code provided superior results in coastal waters. Further integration of ML algorithms with FEC led to additional improvements, lowering the normalized error to about 4.6%, 11.3%, and 42% of the baseline for clear, coastal, and harbor water, respectively. The best-performing ML models varied with the water type: DT and kNN were most effective for clear water, NB performed best for coastal water, and DT again yielded the lowest error in harbor conditions.

5. Conclusions

This study explored the effectiveness of various ML–FEC combinations for ATD-based switching in UWOC/UWB hybrid networks, emphasizing reliability across different water conditions. The findings demonstrate that integrating ML with advanced FEC schemes significantly enhances link stability and error resilience. Among the evaluated techniques, Turbo codes consistently outperformed other FEC methods, exhibiting strong synergy with ML classifiers in both clear and turbid environments. In particular, DT models paired with Turbo codes achieved superior performance in harbor waters, while kNN with Turbo codes excelled in clear ocean conditions. Furthermore, NB, SVM, and DT classifiers also showed similar results, just differing by 5~10% from kNN across varying turbidity levels, highlighting the adaptability of ML-driven decoding. Overall, the combination of optimized FEC and ML classifiers not only extends the reliable communication range but also achieves substantial BER reduction, confirming their potential as robust candidates for future adaptive and intelligent UWOC systems.

Acknowledgment

The authors would like to acknowledge IIT Guwahati Technology Innovation and Development Foundation for this project funding and Defence Institute of Advanced Technology for financial and infrastructural support.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

Data are available from the corresponding author upon reasonable request.

Author Contribution Statement

Shreyas Jain: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Visualization. **Harshit Kumar Sharma:** Software, Validation, Investigation. **Appala Venkata Ramana Murthy:** Conceptualization, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration.

References

[1] Xiang, J., Yang, L., Guo, K., Zlatanov, N., & Wu, Y. (2024). Performance evaluation of UAV-aided radio frequency-UAC relaying systems. *IEEE Internet of Things Journal*, *11*(13), 24183–24195. https://doi.org/10.1109/JIOT.2024.3388581

[2] Chang, C., Han, X., Li, G., Li, P., Nie, W., Liao, P., . . . , & Xie, X. (2024). Extending UWOC system applications through photon transmission dynamics study in harbor waters. *Applied Sciences*, *14*(6), 2493. https://doi.org/10.3390/APP14062493

[3] Aissaoui, A., & Hacini, L. (2024). Enhancing UWOC link performance using a hybrid OFDM/SAC-OCDMA system. *Optical and Quantum Electronics*, *56*(1), 28. https://doi.org/10.1007/s11082-023-05599-7

[4] Mirza, M., Sommers, J., Barford, P., & Zhu, X. (2010). A machine learning approach to TCP throughput prediction. *IEEE/ACM Transactions on Networking*, *18*(4), 1026–1039. https://doi.org/10.1109/TNET.2009.2037812

[5] Sun, X., Kang, C. H., Kong, M., Alkhazragi, O., Guo, Y., Ouhssain, M., . . . , & Ooi, B. S. (2020). A review on practical considerations and solutions in underwater wireless optical communication. *Journal of Lightwave Technology*, *38*(2), 421–431. https://doi.org/10.1109/JLT.2019.2960131

[6] Huang, L., Wang, Y., Zhang, Q., Han, J., Tan, W., & Tian, Z. (2022). Machine learning for underwater acoustic communications. *IEEE Wireless Communications*, *29*(3), 102–108. https://doi.org/10.1109/MWC.2020.2000284

[7] Govinda Waduge, T., Yang, Y., & Seet, B.-C. (2025). A review of reconfigurable intelligent surfaces in underwater wireless communication: Challenges and future directions. *Journal of Sensor and Actuator Networks*, *14*(5), 97. https://doi.org/10.3390/jsan14050097

[8] Choudhary, N., Joshi, S., & Chaubey, V. K. (2023). Outage analysis of IRS-assisted dual-hop mixed RF-UWOC network. In *OCEANS 2023 - Limerick* (pp. 1–5). https://doi.org/10.1109/OCEANSLimerick52467.2023.10244287

[9] Kumar, L. B., & Krishnan, P. (2020). Multi-hop convergent FSO-UWOC system to establish a reliable communication

link between the islands. *Optics Communications*, *474*, 126107. https://doi.org/10.1016/j.optcom.2020.126107

[10] Pei, Q., Zhang, Y., Sun, S., Lian, W., & Xuan, Z. (2024). END-to-end performance of mixed RF/UWOC system with AF protocol based on Mellin transform. *Proceedings of the Romanian Academy, Series A*, *25*(3), 253–260. https://doi.org/10.59277/PRA-SER.A.25.3.10

[11] Solonenko, M. G., & Mobley, C. D. (2015). Inherent optical properties of Jerlov water types. *Applied Optics*, *54*(17), 5392–5401. https://doi.org/10.1364/ao.54.005392

[12] Song, J., Liu, X., Luo, W., & Zhang, J. (2024). Performance analysis of STAR-RIS assisted UWOC-RF systems with NOMA. In *2024 9th International Conference on Intelligent Computing and Signal Processing*, 1316–1320. https://doi.org/10.1109/ICSP62122.2024.10743385

[13] Zhang, Q., Yue, D.-W., Jin, S.-N., & Xu, X.-Y. (2025). Outage probability analysis of RSMA-based multihop UWOC systems with relay selection over thermocline turbulence. *IEEE Internet of Things Journal*, *12*(23), 50771–50784. https://doi.org/10.1109/JIOT.2025.3610535

[14] Pi, X., Jia, S., Zhang, D., Du, P., Xu, Y., Lou, Y., & Li, X. (2025). Outage probability of UAV-assisted dual-hop RF-UWOC systems leveraging NOMA. *IEEE Transactions on Vehicular Technology*. https://doi.org/10.1109/TVT.2025.3598083

[15] Nootz, G., Jarosz, E., Dalgleish, F. R., & Hou, W. (2016). Quantification of optical turbulence in the ocean and its effects on beam propagation. *Applied Optics*, *55*(31), 8813. https://doi.org/10.1364/AO.55.008813

[16] Jamali, M. V., Mirani, A., Parsay, A., Abolhassani, B., Nabavi, P., Chizari, A., . . . , & Salehi, J. A. (2018). Statistical studies of fading in underwater wireless optical channels in the presence of air bubble, temperature, and salinity random variations. *IEEE Transactions on Communications*, *66*(10), 4706–4723. https://doi.org/10.1109/TCOMM.2018.2842212

[17] Gökçe, M. C., Baykal, Y., Kamacioglu, C., & Uysal, M. (2015). Aperture averaging in multiple-input single-output free-space optical systems. *Optical Engineering*, *54*(6), 066103. https://doi.org/10.1117/1.OE.54.6.066103

[18] Zhang, Y., He, Y., Ma, X., Yuan, S., & Xu, J. (2025). A diffuser-based underwater wireless optical communication receiver for scintillation mitigation. *IEEE Photonics Technology Letters*, *37*(11), 617–620. https://doi.org/10.1109/LPT.2025.3543929

[19] Devappa, B. C. D., Pawar, K., Jain, S., Narayanan, S. L. S., Mahale, K. P., & Murthy, A. V. R. (2025). Practical implementation and performance evaluation of different modulation and forward error correction techniques in underwater optical communication testbed. *Journal of Optics*. https://doi.org/10.1007/S12596-025-02847-9

[20] Priyalakshmi, B., & Mahalakshmi, K. (2020). Channel estimation and error correction for UWOC system with vertical non-line-of-sight channel. *Wireless Networks*, *26*(7), 4985–4997. https://doi.org/10.1007/S11276-020-02376-2

[21] Miroshnikova, N. E., & Sattarova, A. I. (2022). Analysis of error-correction codes properties for underwater optical communication systems with OCDM modulation. In *2022 Systems of Signal Synchronization, Generating and Processing in Telecommunications* (pp. 1–6). https://doi.org/10.1109/SYNCHROINFO55067.2022.9840984

[22] Wang, W., Li, Y., Hao, B., Lv, Y., & Ju, M. (2025). Error performance and power allocation for MIMO-UWOC systems in composite oceanic fading scenario. *Optics Communications*, *597*, 132301. https://doi.org/10.1016/j.optcom.2025.132301

[23] Qasem, Z. A. H., Ali, A., Deng, B., Li, Q., & Fu, H. Y. (2024). Spectral and energy efficient pilot-assisted PAPR reduction technique for underwater wireless optical communication systems. *Journal of Lightwave Technology*, *42*(2), 841–852. https://doi.org/10.1109/JLT.2023.3319080

[24] Wang, H., Wang, L., Wu, X., Peng, H., Liao, W., Cai, C., . . . , & Xu, J. (2025). A real-time MPPC-based duplex UWOC system. *Journal of Lightwave Technology*, *43*(18), 8737–8746. https://doi.org/10.1109/JLT.2025.3594481

[25] Li, H., Chen, T., Wang, Z., Cao, B., Li, Y., & Zhang, J. (2022). Low-complexity sampling frequency offset estimation and compensation scheme for OFDM-based UWOC system. *Photonics*, *9*(4), 216. https://doi.org/10.3390/photonics9040216

[26] Shigenawa, A., Onodera, Y., Takeshita, E., Hisano, D., Maruta, K., & Nakayama, Y. (2022). Predictive equalization for underwater optical camera communication. In *2022 IEEE 95th Vehicular Technology Conference*, 1–5. https://doi.org/10.1109/VTC2022-Spring54318.2022.9860954

[27] Zhao, X., Qi, Z., & Pompili, D. (2024). Link adaptation in underwater wireless optical communications based on deep learning. *Computer Networks*, *242*, 110233. https://doi.org/10.1016/j.comnet.2024.110233

[28] Kamatchi, K., V, A., Paramanandham, N., Yogarajan, G., & Krishnan, P. (2024). Machine learning aided signal detection in underwater wireless optical communication for IoUT applications. In *2024 International Conference on Smart Systems for Electrical, Electronics, Communication and Computer Engineering*, 774–778. https://doi.org/10.1109/ICSSEECC61126.2024.10649510

[29] Lu, H., Jiang, M., & Cheng, J. (2020). Deep learning aided robust joint channel classification, channel estimation, and signal detection for underwater optical communication. *IEEE transactions on communications*, *69*(4), 2290–2303. https://doi.org/10.1109/TCOMM.2020.3046659

[30] Salama, W. M., & Aly, M. H. (2025). Improved channel estimation for underwater wireless optical communication OFDM systems by combining deep learning and machine learning models. *Optical and Quantum Electronics*, *57*(3), 181. https://doi.org/10.1007/S11082-025-08090-7

[31] Nennouche, M., Khalighi, M. A., Dowhuszko, A., Merad, D., & Boï, J.-M. (2024). Application of machine learning to signal detection in underwater wireless optical communication links. In *2024 14th International Symposium on Communication Systems, Networks and Digital Signal Processing* (pp. 534–538). https://doi.org/10.1109/CSNDSP60683.2024.10636416

[32] Kong, Y., Li, Y., Zhang, X., & Zhou, Z. (2024). YY-YOLO: Improved YOLOv5 for object detection on traffic signs. In *Proceedings of the 2024 3rd International Symposium on Intelligent Unmanned Systems and Artificial Intelligence*, 221–226. https://doi.org/10.1145/3669721.3669724

[33] Cheng, Q., Varshney, P. K., & Arora, M. K. (2006). Logistic regression for feature selection and soft classification of remote sensing data. *IEEE Geoscience and Remote Sensing Letters*, *3*(4), 491–494. https://doi.org/10.1109/LGRS.2006.877949

[34] Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, *14*(3), 199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88

[35] Goodman, R. M., & Smyth, P. (1988). Decision tree design from a communication theory standpoint. *IEEE Transactions*

on Information Theory, *34*(5), 979–994. https://doi.org/10.1109/18.21221

[36] Aslam, M. W., Zhu, Z., & Nandi, A. K. (2012). Automatic modulation classification using combination of genetic programming and KNN. *IEEE Transactions on Wireless Communications*, *11*(8), 2742–2750. https://doi.org/10.1109/TWC.2012.060412.110460

[37] Hamming, R. W. (1950). Error detecting and error correcting codes. *Bell System Technical Journal*, *29*(2), 147–160. https://doi.org/10.1002/J.1538-7305.1950.TB00463.X

[38] Sathiya Narayanan, S. L., Dhanush Devappa, B. C., Pawar, K., Jain, S., & Venkata Ramana Murthy, A. (2024). Implementation of forward error correction for improved performance of free space optical communication channel in adverse atmospheric conditions. *Results in Optics*, *16*, 100689. https://doi.org/10.1016/J.RIO.2024.100689

[39] Bose, R. C., & Ray-Chaudhuri, D. K. (1960). On a class of error correcting binary group codes. *Information and Control*, *3*(1), 68–79. https://doi.org/10.1016/S0019-9958(60)90287-4

[40] Aspreas, A., & Yiannopoulos, K. (2024). Reed Solomon error correction in pre-amplified pulse position modulation receivers. *AEU - International Journal of Electronics and Communications*, *175*, 154859. https://doi.org/10.1016/j.aeue.2023.154859

[41] Pietrobon, S. S. (1998). Implementation and performance of a turbo/MAP decoder. *International Journal of Satellite Communications*, *16*(1), 23–46. https://doi.org/10.1002/(SICI)1099-1247(199801/02)16:1%3C23::AID-SAT590%3E3.0.CO;2-W

[42] Darwesh, L., & Kopeika, N. S. (2020). Deep learning for improving performance of OOK modulation over FSO turbulent channels. *IEEE Access*, *8*, 155275–155284. https://doi.org/10.1109/ACCESS.2020.3019113

[43] Dhanush Devappa, B. C., Banerjee, S., Pawar, K., & Murthy, A. V. R. (2025). Practical realization and performance analysis of Rivest-Shamir-Adleman encryption for secure underwater optical communication. *Next Research*, *2*(2), 100225. https://doi.org/10.1016/J.NEXRES.2025.100225

[44] Xu, F., Khalighi, M.-A., & Bourennane, S. (2009). Coded PPM and multipulse PPM and iterative detection for free-space optical links. *Journal of Optical Communications and Networking*, *1*(5), 404. https://doi.org/10.1364/JOCN.1.000404

[45] Aldibbiat, N. M., Ghassemlooy, Z. F., & McLaughlin, R. (2001). Performance of dual header-pulse interval modulation (DH-PIM) for optical wireless communication systems. In E. J. Korevaar (Ed.), *Proceedings of SPIE: Optical Wireless Communications III* (*4214:*, pp. 144–152). SPIE. https://doi.org/10.1117/12.417505

[46] Yuan, R., Zhang, T., Li, C., Gao, H., & Hu, L. (2025). Laser transmission characteristics of seawater for underwater wireless optical communication. *Sensors*, *25*(10), 3057. https://doi.org/10.3390/S25103057

[47] Zedini, E., Ata, Y., Al-Sallami, Mahdi., F, & Rajbhandari, S. (2025). Evaluation of underwater optical wireless channels over F turbulence for different detection types. *IEEE Transactions on Communications*, *73*(3), 1925–1937. https://doi.org/10.1109/TCOMM.2024.3454696

[48] Ali, M. F., Jayakody, D. N. K., & Li, Y. (2022). Recent trends in underwater visible light communication (UVLC) systems. *IEEE Access*, *10*, 22169–22225. https://doi.org/10.1109/ACCESS.2022.3150093

[49] Qiu, H., Huang, Z., Xu, J., Motani, M., & Ji, Y. (2025). Channel modeling, performance analysis, and probabilistic shaping for underwater wireless optical communications. *IEEE Journal on Selected Areas in Communications*, *43*(5), 1568–1581. https://doi.org/10.1109/JSAC.2025.3543508

[50] Jiang, H., Qiu, H., He, N., Popoola, W., Ahmad, Z., & Rajbhandari, S. (2022). Ergodic capacity and error performance of spatial diversity UWOC systems over generalized gamma turbulence channels. *Optics Communications*, *505*, 127476. https://doi.org/10.1016/j.optcom.2021.127476

[51] Vegni, A. M., & Ata, Y. (2024). Ergodic channel capacity analysis for underwater OWC under different detection schemes. In *2024 IEEE International Mediterranean Conference on Communications and Networking*, 131–136. https://doi.org/10.1109/MeditCom61057.2024.10621388

How to Cite: Jain, S., Sharma, H. K., & Murthy, A. V. R. (2026). Machine Learning–Enhanced UWOC System for Robust Underwater Communication with FEC Integration. *Journal of Optics and Photonics Research*. https://doi.org/10.47852/bonviewJOPR62028116