

Article

Performance Monitoring of Hybrid All-Optical Fiber/FSO Communication Systems

Maged Abdullah Esmail 

Communications and Networks Engineering Department and Smart Systems Engineering Laboratory,
Prince Sultan University, Riyadh 11586, Saudi Arabia; mesmail@psu.edu.sa

Abstract: The demand for network capacity has increased due to the introduction of new digital applications and services, which rely heavily on optical communication networks. While fiber networks serve as the optical networks' backbone, deploying fiber in certain scenarios is not feasible, making it necessary to use other technologies conjointly. A hybrid all-optical fiber/free space optic (FSO) link is proposed to avoid such a challenge. The all-optical system avoids using electronics that have limited bandwidth. Hence, it supports high-capacity communication. However, the all-optical system comes with challenges arising from fiber and FSO channel impairments. To monitor the amount and type of distortion in the optical channel, machine learning (ML) techniques are exploited. In this work, Gaussian process regression (GPR) is utilized as an ML technique to predict three main channel impairments that arise in the hybrid all-optical fiber/FSO channels, which are turbulence, optical signal-to-noise ratio (OSNR), and chromatic dispersion (CD). The model's performance is evaluated using boxplot graphs, root mean square error (RMSE) metric, and R-squared metric. The results indicate that the model can predict the various impairments with high accuracy, except under strong amplified spontaneous emission (ASE) noise, where the model demonstrated lower accuracy in predicting light turbulence parameters. The proposed approach provides a self-aware and self-adaptive communication system and can optimize network resources in the future.

Keywords: all-optical network; hybrid fiber/FSO; machine learning; performance monitoring



Citation: Esmail, M.A. Performance Monitoring of Hybrid All-Optical Fiber/FSO Communication Systems. *Appl. Sci.* **2023**, *13*, 8477. <https://doi.org/10.3390/app13148477>

Academic Editor: Nuno Silva

Received: 29 June 2023

Revised: 15 July 2023

Accepted: 20 July 2023

Published: 22 July 2023



Copyright: © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Over the past few years, the need for network capacity has increased significantly due to the introduction of new digital applications and services such as online gaming, high-quality video streaming, and cloud computing [1]. Therefore, modern communication networks rely heavily on optical communication networks, which serve as their backbone because of the wide bandwidth the fiber provides. Fiber-based networks are the drivers for current 5G and future wireless communication networks, smart cities, the Internet of Things (IoTs), and industry 4 [2]. They have the capability to provide huge data rates with low latency. However, there are some scenarios where deploying fiber is not feasible or difficult such as over hills, rivers, and lakes. In addition, a fast communication link is necessary to establish during disasters, which is impossible with fiber. Therefore, there is a need to use other technologies such as microwave and millimeter wave links as a bridge to overcome such challenges. However, these techniques have limited bandwidth, which does not match with the fiber bandwidth. The feasible solution is using the free space optic (FSO) technique to avoid fiber discontinuity when facing obstacles [3]. FSO is similar to the fiber in that it has wide bandwidth that supports current and future high-demand bandwidth applications. Signal transmission over the FSO link can reach several kilometers. However, instead of transmitting the optical signal over fiber, FSO sends the optical signal over free space. This feature makes FSO the best solution when fiber installation is infeasible. Moreover, FSO is a good quick solution for service recovery when the fiber cable is damaged for certain reasons such as disasters and construction work. Also, FSO can be installed

as a backup for fiber in availability-sensitive networks [4,5]. FSO also offers several other advantages, including cost-effectiveness, rapid deployment, enhanced security compared to RF links, and the use of unlicensed spectrum. However, there are some challenges associated with FSO links. For example, link blockage by obstacles such as construction work may require re-planning of the FSO link. The possibility of unauthorized access to the communication link is another concern, although it is harder than the RF link to gain access because FSO uses a line-of-sight link. FSO link alignment is another challenge, as it requires precise alignment due to its narrow beam. The safety of living creatures such as birds is also a concern, which can be addressed by respecting the recommended maximum transmitted power.

In a traditional intensity modulation/direct detection (IM/DD) FSO system, when the fiber is connected to the FSO system on the transmitter side, the optical signal carried by the fiber should be converted to an electrical signal first to drive the laser. This is achieved using a photodetector (PD), which is an optoelectronic device. The drive circuit modulates the laser light to produce an output-modulated optical signal. Such conversion operation in the transmitter of the traditional hybrid fiber/FSO links from an optical signal to an electrical signal and then to an optical signal increases the system cost and reduces the available bandwidth due to using electronics that have limited bandwidth [6]. To overcome such a limitation, a hybrid all-optical communication fiber/FSO system is proposed [6,7]. In such a system, the optical signal delivered by the fiber in the transmitter side is coupled to the free space using a light collimating lens. Hence, there is no need for optical/electrical/optical conversion. In the FSO receiver side, the light is collimated to the fiber using another collimating lens. Hence, the signal over the entire fiber/FSO link remains optical.

The advantage of a hybrid all-optical fiber/FSO link comes with challenges arising from the channel impairments. In the fiber channel, the optical signal is prone to different types of impairments such as chromatic dispersion (CD) and amplified spontaneous emission (ASE) noise due to signal amplification. In the FSO channel, the signal is prone to turbulence and atmospheric attenuation. Such limitations degrade the optical communication system's performance. Therefore, it is important to have a feature in the system to monitor the amount and type of distortion in the optical channel. Such a feature could be used in improving the system performance using techniques such as adaptive modulation on the transmitter side or designing appropriate mitigation algorithms on the receiver side based on the type and amount of the distortion. This leads to having a self-aware and self-adaptive communication system.

Various conventional techniques were proposed in the literature for monitoring different channel impairments such as ASE, CD, and polarization mode dispersion (PMD) [8]. However, such techniques have technical limitations and/or are costly. Recently, machine learning (ML) techniques are exploited in many applications [8–11]. One application is to monitor the optical network performance. The use of ML techniques holds great potential for enhancing the intelligence of optical network nodes. By utilizing ML, network nodes can gather insights from network conditions, which can then be applied to optimize network resources in the future.

Various ML techniques have been suggested in the scientific literature for monitoring optical performance in both fiber-based optical links and FSO-based links that exploit IM/DD. Performance monitoring of IM/DD fiber-based optical communication links using an artificial neural network (ANN) model was investigated in [12–15]. In [12], the ANN model was proposed for monitoring three parameters, which are optical signal-to-noise ratio (OSNR), CD, and differential group delay (DGD). The model was trained using features extracted from the eye diagrams of the received signal such as the root mean square and the jitter. However, this technique requires precise timing, which increases the system's complexity. In [13], the same parameters are monitored using an ANN model but trained with an asynchronous amplitude histogram (AAH). Using ANN was also proposed in [14,15] for monitoring the OSNR, CD, and DGD parameters, using empirical moments and the asynchronous delay-tap sampling (ADTS) technique, respectively. A

conventional neural network (CNN) model was proposed in [16] for monitoring OSNR and CD parameters. In this work, the input data for the CNN model are 2-D images generated using the ADTS technique. However, using images increases the complexity of the system. In [17], a deep neural network (DNN) model trained on features extracted from the signal eye diagram was proposed for monitoring the OSNR, CD, and DGD parameters.

Monitoring FSO channel parameters was reported in [18–21] using a CNN model. In [18,19], turbulence impairment monitoring is investigated using orbital angular momentum (OAM) modes as a training feature for the model. Similarly, the OAM modes were used in [20] to monitor three impairment parameters, which include OSNR, turbulence, and pointing errors. In [21], the OAM modes are exploited for monitoring the visibility range under dusty weather conditions. The work in [22] used the support vector machine (SVM) technique to monitor the turbulence and OSNR parameters in addition to the visibility range under fog weather conditions. Table 1 summarizes the literature work on monitoring the performance of IM/DD optical fiber and FSO systems.

Table 1. Literature review.

Ref.	Channel Type	ML Model	Training Feature	Monitored Parameter
[12]	Fiber	ANN	Egy diagrams	OSNR CD DGD
[13]	Fiber	ANN	AAH	OSNR CD DGD
[14]	Fiber	ANN	empirical moments	OSNR CD DGD
[15]	Fiber	ANN	ADTS	OSNR CD DGD
[16]	Fiber	CNN	ADTS	OSNR CD
[17]	Fiber	DNN	Eye diagrams	OSNR CD DGD
[18]	FSO	CNN	OAM	C_n^2
[19]	FSO	CNN	OAM	C_n^2
[20]	FSO	CNN	OAM	C_n^2 OSNR
[21]	FSO	CNN	OAM	Pointing errors Visibility range Visibility range
[22]	FSO	SVM	AAH	OSNR Pointing errors
This work	Hybrid Fiber/FSO	GPR	AAH	OSNR CD C_n^2

In this work, we consider monitoring three main channel parameters that arise in the hybrid all-optical fiber/FSO channels; these are turbulence, OSNR, and CD. To the best of the authors' knowledge, this is the first study that addresses monitoring these impairments in a hybrid optical system of this kind. To achieve this goal, Gaussian process regression (GPR) is utilized as an ML model to predict the three parameters. The model's performance is evaluated using the root mean square error (RMSE) metric and the R-squared metric. Additionally, boxplot graphs are used as a visualization tool to evaluate the model's performance. The results indicate that the model can predict the various impairments with high accuracy, except under strong ASE noise, where the model demonstrated lower accuracy in predicting light turbulence parameters.

The paper is structured as follows. Section 2 provides an overview of the system and the ML technique employed. The results' analysis and discussion are presented in Section 3. Finally, the paper concludes in Section 4.

2. Simulation Setup Description

2.1. Simulation Setup

The simulation setup of the all-optical system that includes the data acquisition system is shown in Figure 1. The simulations are carried out using Transmission Maker 11.1 simulator, which is a powerful tool for modeling and analyzing optical systems and components. The simulator enables us to generate a dataset that includes different channel conditions, which is a difficult task and requires a long measurement campaign to collect such a dataset in real scenarios. Moreover, it allows repeating the measurements under the same channel conditions. The software employs a combination of different simulation techniques, including Finite-Difference Time-Domain, Beam Propagation Method, Finite Element Method, and Eigenmode Expansion Method, to provide a comprehensive solution for the design and simulation of photonic components and systems.

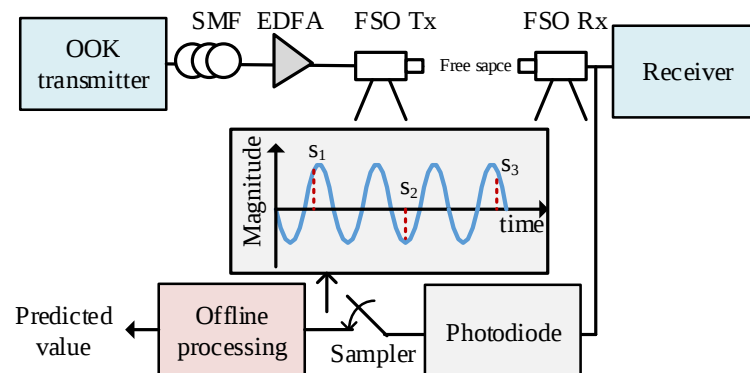


Figure 1. Block diagram of the hybrid all-optical fiber/FSO simulation setup including the acquisition system.

On the transmitter side, a 10 Gbps on-off keying (OOK) signal is generated and transmitted over a single mode fiber (SMF) that introduced 0.2 dB/km attenuation and has 1.47 group refractive index. A laser diode (LD) emitting 16 dBm average power at 1550 nm wavelength is used at the transmitter to send the data optically over the fiber. The propagated optical signal over the fiber undergoes CD. The CD values vary between 100 ps/nm and 500 ps/nm with 100 ps/nm step size. The CD value is set by changing the value of the CD parameter in the fiber module in the software. To compensate for the signal attenuation in the system, signal amplification using an erbium-doped fiber amplifier (EDFA) is used, which introduces ASE noise. This noise is defined by the OSNR parameter. As the gain of the EDFA changes, the OSNR value changes too. The OSNR values range between 12 dB and 20 dB with a 2 dB step size. In the software, the value of the OSNR is set by changing the OSNR parameter value in the EDFA module. The signal is then transmitted over 1 km FSO channel, which introduces turbulence into the signal, modeled by the Gamma-Gamma model. The amount of turbulence impairment is defined by the index of refraction structure parameter, C_n^2 , with the following values: $10^{-17} \text{ m}^{-2/3}$ and $10^{-16} \text{ m}^{-2/3}$ for weak turbulence, $10^{-15} \text{ m}^{-2/3}$ for moderate turbulence, and $10^{-14} \text{ m}^{-2/3}$, and $10^{-13} \text{ m}^{-2/3}$ for strong turbulence. To emulate the turbulence in the software, the C_n^2 model, which is defined in [22], is used.

On the receiver side, a data acquisition system is used to build the dataset. It includes a PD, a sampler, and offline processing. The PD is a PIN photodiode that converts the optical signal into an electrical signal. Then, the sampler samples the signal and generates a realization of 8192 samples. Note that, if the received signal power is attenuated due to factors such as geometric loss, fiber attenuation, and channel impairments, and the

resulting power level is lower than the receiver sensitivity, the receiver will not be able to detect the signal.

After data acquisition, the collected dataset is processed offline for training and testing the ML model. The dataset includes 200 realizations for each value of the three parameters. This means the data size includes (200 realizations/value $\times$ 5 values/parameter $\times$ 3 parameters) 3000 realizations. This dataset is divided into 70% for training the model and 30% for testing.

2.2. Machine Learning Model

In this work, the three parameters under study introduce amplitude variation. Therefore, in order to have a power ML model that can predict each parameter accurately, the model should be trained using features that preserve the amplitude information of the received signals. The amplitude histogram is used in this work as a feature for training the proposed model since it preserves the optical signal's amplitude information. To reduce the acquisition system cost, we utilize an AAH. The AAH samples the signal asynchronously as depicted in the inset in Figure 1, which reduces the sampling speed and hence allows using a low-speed sampler. Each realization is sampled at 500 Msample/s. The AAH feature not only helps build a cost-effective ML model but also eliminates the need for timing recovery at the receiver, thereby avoiding the need for additional hardware.

Figure 2 shows examples of the AAH features extracted from different acquired signals using 100 bins. The signals are subjected to 500 ps/nm CD impairment and different OSNR impairment values. It is clear from the figures that the AAH signals have discrepancies, which can be exploited as features for the different types of impairments. The degree of correlation among distinct AAH features is a determining factor for the precision of the ML model in predicting the various parameters.

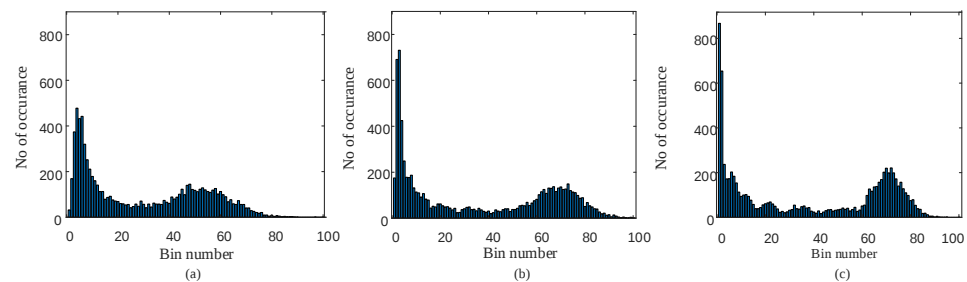


Figure 2. AAH features for an acquired signal under 500 ps/nm CD and (a) OSNR = 12 dB, (b) OSNR = 16 dB, and (c) OSNR = 20 dB.

In this work, we consider the GPR model for parameter prediction. GPR is a robust method with the ability to effectively handle complex problems, including nonlinearity, high dimensions, and small sample sizes. It has strong generalization capabilities and is more easily implemented compared to neural network approaches due to its fewer parameters. As a result, GPR is commonly utilized in a wide range of fields [23]. For a training dataset of size N , each training data pair (x_i, y_i) is assumed to have unknown distribution. A linear regression model in vector form is given by

$$\mathbf{y} = \mathbf{x}^T \boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (1)$$

where $\mathbf{x}$ is the input vector to the model and $\mathbf{y}$ is the predicted vector. $\boldsymbol{\varepsilon}$ is an error (noise) vector that has a normal distribution with zero mean and σ^2 variance, and $\boldsymbol{\beta}$ is a vector of weights. The $\boldsymbol{\beta}$ and $\boldsymbol{\varepsilon}$ vectors are estimated from the training data. The GPR regression model performs regression operation by introducing latent variables $f(x_i)$ from explicit basis function h and a Gaussian process (GP). The smoothness of the response is reflected in a covariance function, $k(\mathbf{x}, \mathbf{x}')$ of the latent variables, while the basis functions h maps the input vector $\mathbf{x}$ into a feature space vector $h(\mathbf{x})$. The covariance function is parameterized by a set of kernel parameters. The GPR converts (1) as follows [24]

$$\mathbf{y} = h(\mathbf{x}^T)\boldsymbol{\beta} + f(\mathbf{x}), \quad (2)$$

where $h(\mathbf{x})$ is a GP with zero mean and k covariance function. In this work, the GPR model uses a squared exponential covariance (kernel) function and a constant basis function. The initial values of β and σ are -17.19 and 0.55 , respectively. It is worth mentioning that although the GPR model is a powerful regression technique, it has limitations, including scalability and sensitivity to kernel function selection. The model can become computationally expensive when dealing with large datasets, making it impractical. Moreover, selecting the right kernel function can be challenging as the model's performance depends heavily on this choice. However, by using machine learning techniques such as quantum and graph neural networks, it may be possible to improve the performance of the GPR method in real-time scenarios. Quantum neural networks, for example, can exploit the properties of quantum mechanics to perform certain computations more efficiently than classical computers. Similarly, graph neural networks can exploit the structure of graph data to improve the accuracy of predictions. In this work, the dataset size is small, and the kernel function was carefully chosen, thus avoiding the aforementioned limitations.

3. Results and Discussions

In this section, we utilize two metrics to evaluate the performance of the proposed model in predicting the various values of the channel parameters. The metrics are the Coefficient of Determination (R-squared) metric and the RMSE metric. The R-squared metric is given mathematically by

$$\rho = 1 - \frac{\sum_{n=1}^N (m_n - \hat{m}_n)^2}{\sum_{n=1}^N (m_n - \bar{m})^2}. \quad (3)$$

The R-squared metric utilizes the actual and estimated data, sample mean ($\bar{m}$), and the total number of test samples (N), with m_n and $\hat{m}_n$ representing the actual and estimated data, respectively. The metric's range is between 0 and 1, with higher values indicating improved prediction accuracy as ρ approaches 1. Conversely, the prediction accuracy deteriorates as ρ approaches 0. On the other hand, the RMSE metric is defined mathematically by

$$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^N (m_n - \hat{m}_n)^2}. \quad (4)$$

In addition, we use a visualization method, the boxplot method, to examine the proposed model's performance. It is a graphical tool used to display the distribution of a dataset. It is constructed from five key statistics: the minimum and maximum values, the lower and upper quartiles, and the median. The box in the plot represents the interquartile range (IQR), which is the range between the first (25th percentile) and third (75th percentile) quartiles. The median is represented by a line inside the box. The whiskers extend from the box to the minimum and maximum values that lie within a certain distance from the box, typically 1.5 times the IQR. Any points outside this range are considered outliers and are represented as individual points. Boxplots are used to visualize the spread and skewness of data and to compare multiple datasets side-by-side. The boxplots in Figure 3 illustrate the prediction accuracy of the model under different channel types. High-accurate predictions are represented by narrow boxes.

Subsequently, we present the results of monitoring each parameter, taking into account the prediction of each parameter under different conditions: when there are no other impairments (single impairment: ASE, CD, turbulence), when there is a joint impairment (double impairment: ASE and CD, ASE and turbulence, CD and turbulence), and when all the impairments are present (triple impairment: ASE, CD, turbulence). It is worth noting that we considered the severe impairment conditions in evaluating the model performance

defined by 12 dB OSNR, 500 ps/nm CD, and $10^{-15} \text{ m}^{-2/3}$ turbulence. Moderate and light conditions will result in much better prediction accuracy.

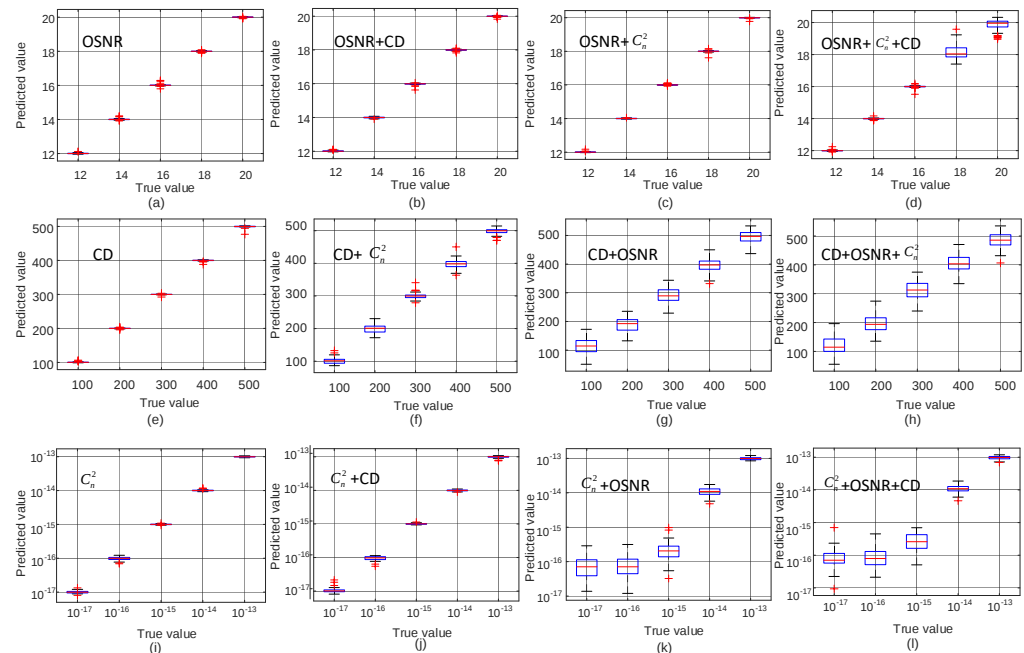


Figure 3. Prediction accuracy of the GPR model represented visually using boxplot. (a–d) Prediction accuracy of the OSNR parameter, (e–h) prediction accuracy of the CD parameter, and (i–l) prediction accuracy of the C_n^2 parameter. Severe impairments values are jointly included in predicting the impairments defined by 12 dB OSNR, 500 ps/nm CD, and $10^{-13} \text{ m}^{-2/3} C_n^2$. The channel conditions are depicted on each subfigure.

3.1. OSNR Parameter Prediction

Figure 3a illustrates the prediction accuracy of the OSNR parameter for the five different OSNR values. The narrow width of the boxplot indicates high prediction accuracy. The obtained R-squared and RMSE values were 1 and 0.09 dB, respectively, further confirming the high accuracy of the predictions. Figure 3b presents the application of the model to predict the OSNR parameter under the influence of 500 ps/nm CD. The results obtained are similar to those in Figure 3a, indicating that the model can still accurately predict the OSNR parameter with high precision. The achieved R-squared and RMSE values were 1 and 0.13 dB, respectively. Figure 3c depicts the performance of the model in predicting the OSNR parameter under the influence of OSNR and severe turbulence defined by $10^{-13} \text{ m}^{-2/3}$. The model continues to accurately predict the OSNR parameter, as evidenced by the achieved R-squared and RMSE values of 1 and 0.16 dB, respectively. Finally, the combined effect of OSNR with 500 ps/nm CD and $10^{-13} \text{ m}^{-2/3}$ turbulence is investigated, with the results presented in Figure 3d. The model is still able to accurately predict the low values of the OSNR parameter, although the accuracy slightly degrades for high values. The achieved R-squared and RMSE values were 0.96 and 0.54 dB, respectively, indicating high prediction accuracy. Overall, the model demonstrated high accuracy in predicting the various values of the OSNR parameter.

3.2. Chromatic Dispersion Parameter Prediction

The proposed model's prediction of the CD parameter is visually depicted in Figure 3e–h. Figure 3e shows the model's ability to predict the CD parameter alone without any other impairments, with very narrow boxes indicating high accuracy. The obtained R-squared and RMSE values were 1 and 0 ps/nm, respectively. When the signal was corrupted with severe turbulence, defined by $10^{-13} \text{ m}^{-2/3}$ turbulence, the R-squared value slightly dropped to 0.98, and the RMSE increased slightly to 0.2 ps/nm. Nonetheless, the model was still capable of

predicting the CD parameter with high accuracy, as shown in Figure 3f. Similarly, when the system was corrupted with ASE noise defined by 12 dB OSNR, the R-squared value dropped slightly to 0.96, while the RMSE increased slightly to 0.27 ps/nm. With this slight degradation in the model performance, the model still accurately predicted the CD parameter, as depicted in Figure 3g. Finally, Figure 3 shows the model's performance in predicting the CD parameter when the system was corrupted with both ASE noise defined by 12 dB OSNR and turbulence defined by $10^{-13} \text{ m}^{-2/3}$. The joint impairments slightly degraded the model's accuracy, with the R-squared value dropping to 0.93, and the RMSE value increasing to 0.37 ps/nm. In general, the model was able to predict the different CD parameter values with high accuracy.

3.3. Turbulence Parameter Prediction

This subsection discusses the proposed model's performance in predicting the C_n^2 parameter, with the results depicted in Figure 3i–l. Firstly, we examine the model's performance in predicting the C_n^2 parameter when no other impairments are present in the system. Figure 3i illustrates the boxplot of the prediction, indicating high accuracy. The obtained R-squared and RMSE metrics were 1 and $0.1 \text{ m}^{-2/3}$, respectively. Next, we evaluated the model's performance in predicting the C_n^2 parameter when the system was subject to 500 ps/nm CD impairment. The results in Figure 3j indicate that the prediction performance was similar to that in Figure 3i when there was no CD impairment. The achieved R-squared and RMSE metrics were 0.03 and $0.1 \text{ m}^{-2/3}$, respectively. Figure 3k reports the C_n^2 parameter prediction when the channel is subject to ASE noise defined by 12 dB OSNR. In this scenario, we observed that the model demonstrated high accuracy in predicting the harsh C_n^2 parameter, defined by high C_n^2 values. However, when the C_n^2 parameter values were low, the impact of ASE noise on the signal was high, leading to difficulty in predicting the OSNR parameter. The obtained R-squared and RMSE metrics were 0.84 and $0.57 \text{ m}^{-2/3}$, respectively. Finally, we evaluated the proposed model's capability in predicting the C_n^2 parameter when the system was subject to harsh conditions defined by 12 dB OSNR and 500 ps/nm CD. We obtained similar results to those in Figure 3k because the dominance of ASE noise complicates the model's capability of predicting the C_n^2 parameter. In general, the model demonstrated high accuracy in predicting the C_n^2 parameter when no ASE noise was present or under severely turbulent channels. However, when strong ASE noise was added to the system, the model's capability decreased significantly in predicting low C_n^2 values. Table 2 below summarizes the performance of the proposed model in predicting the various parameters in terms of R-squared and RMSE metrics.

Table 2. Summary of the proposed model performance.

Predicted Parameter	Channel Impairments	R-Squared	RMSE
OSNR	ASE	1.00	0.09 dB
	ASE+CD	1.00	0.13 dB
	ASE+Turbulence	1.00	0.16 dB
	ASE+CD+ Turbulence	0.96	0.54 dB
CD	CD	1.00	0.00 ps/nm
	CD+ Turbulence	0.98	0.20 ps/nm
	CD+ASE	0.96	0.27 ps/nm
	CD+ASE+ Turbulence	0.93	0.37 ps/nm
C_n^2	Turbulence	1.00	$0.10 \text{ m}^{-2/3}$
	Turbulence +CD	1.00	$0.03 \text{ m}^{-2/3}$
	Turbulence +ASE	0.84	$0.57 \text{ m}^{-2/3}$
	Turbulence +CD+ASE	0.84	$0.56 \text{ m}^{-2/3}$

3.4. Performance Comparison with Other ML Models

The present section compares the performance of the GPR model with that of two other commonly used ML models, namely SVM and random forest (RF) models. Figure 4 illustrates the performance comparison using R-squared and RMSE metrics. For the

turbulence parameter monitoring, it was observed that the GPR model outperformed the other models in terms of both RMSE and R-squared metrics. Specifically, the achieved R-squared was 1 under turbulence conditions alone and also under turbulence and 500 ps/nm CD impairment, as indicated in Figure 4a. Similarly, for these two-channel conditions, the RMSE value was less than 1. However, when the signal was corrupted with turbulence and ASE noise (12 dB OSNR), the R-squared value dropped to 0.84 for the GPR model, which was still slightly better than that of the RF and SVM models. Similarly, the RMSE degraded to 0.55, but still remained better than that of the RF and SVM models as depicted in Figure 4d. The performance of predicting the turbulence parameter under all harsh conditions, i.e., 12 dB OSNR and 500 ps/nm, remained relatively unchanged, as shown in Figure 4a,d.

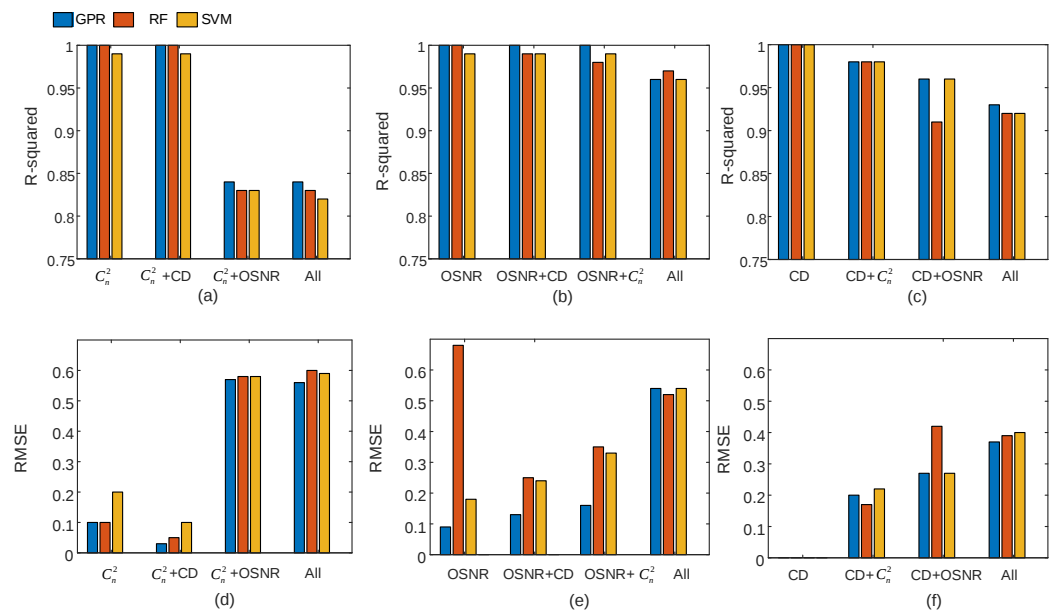


Figure 4. Performance comparison of GPR, RF, and SVM models in terms of R-squared and RMSE metrics for predicting (a,d) C_n^2 parameter, (b,e) OSNR parameter, and (c,f) CD parameter. Severe channel conditions are considered as follows: 12 dB OSNR, 500 ps/nm CD, and $10^{-13} \text{ m}^{-2/3} C_n^2$ turbulence.

In the following analysis, we compare the performance of the three models in predicting the OSNR parameter under different channel conditions, as shown in Figure 4b,e. The results indicate an R-squared value of 1 using the GPR model under ASE noise alone, as well as when the channel included 500 ps/nm CD and $10^{-13} \text{ m}^{-2/3}$ turbulence. In comparison, the two other models exhibited slightly lower R-squared values. Furthermore, the GPR model demonstrated significantly better performance in terms of RMSE compared to the two other models. However, when the channel was subjected to both 500 ps/nm CD and $10^{-13} \text{ m}^{-2/3}$ turbulence, the R-squared value declined to 0.96, while the RMSE degraded to 0.55, which was similar to the obtained performance of the two other models.

Finally, we compare the performance of the GPR model with the two other models in predicting the CD parameter, as illustrated in Figure 4c,f. The results exhibit similar performance in terms of R-squared and RMSE for all three models. Specifically, the achieved R-squared was 1 and the RMSE was 0. However, when a severe turbulence condition, defined by $10^{-13} \text{ m}^{-2/3}$ turbulence value, was added to the channel, the R-squared metric value degraded to 0.98 for all models, and the RMSE increased to 0.2, which was slightly better than that of the SVM model but worse than that of the RF model. When ASE noise was introduced to the channel, defined by 12 dB OSNR, the GPR model's performance slightly degraded in terms of R-squared and RMSE. It exhibited the same performance as the SVM model and was better than that of the RF model. Finally, when both CD and OSNR

severe conditions were considered, the GPR model achieved slightly better performance than the two other models.

It is worth mentioning that the GPR technique is often preferred over supervised learning techniques such as SVM and RF in situations where the dataset is small or noisy, and high prediction accuracy is required. This is because GPR is a non-parametric model that can handle complex input–output relationships without imposing strong assumptions about the underlying data distribution. Moreover, GPR can provide probabilistic estimates of the output variable, making it particularly useful in situations where uncertainty estimates are required. Additionally, GPR can handle high-dimensional input spaces and limited training data more effectively than SVM and RF models.

4. Conclusions

This work proposes the use of the GPR ML technique for monitoring the performance of optical wireless communication systems. The model demonstrated high accuracy in predicting the system's impairment parameters, except when high ASE noise was introduced, where it struggled to predict low values of the C_n^2 parameter. Comparisons with RF and SVM models showed that the proposed GPR model achieved equal or better performance in terms of R-squared and RMSE metrics in all channel conditions. The amount of improvement depends on the type of the involved impairments in the system and their severity. The proposed model can be utilized by telecommunication network operators who use FSO links in their networks to accurately predict the level of impairment in their optical signals. Based on the model output, operators can determine the best course of action to mitigate channel impairments. This may involve implementing adaptive modulation formats, adjusting the baud rate, or applying digital signal processing techniques to the transmitter, receiver, or both. Future work could extend the model to predict other system impairments, such as DGD in fiber channels and pointing errors in outdoor channels. In addition, techniques such as quantum and graph neural networks could be investigated to improve the performance of predicting the channel impairments.

Funding: This research was funded by Prince Sultan University and the APC was funded by Prince Sultan University.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to acknowledge the support of Prince Sultan University for paying the Article Processing Charges (APCs) for this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AAH	asynchronous amplitude histogram
ADTS	asynchronous delay-tap sampling
ANN	artificial neural network
ASE	amplified spontaneous emission
CD	chromatic dispersion
CNN	conventional neural network
DD	direct detection
DGD	differential group delay
DNN	deep neural network
EDFA	erbium-doped fiber amplifier
FSO	free space optic
GP	Gaussian process

GPR	Gaussian process regression
IM	intensity modulation
IoTs	Internet of Things
GPR	Gaussian process regression
IM	intensity modulation
IoTs	Internet of Things
IQR	interquartile range
LD	laser diode
ML	machine learning
OAM	orbital angular momentum
OOK	on-off keying
OSNR	optical signal-to-noise ratio
PD	photodetector
PMD	polarization mode dispersion
RMSE	root mean square error
SMF	single mode fiber
SVM	support vector machine
SVM	random forest

References

1. Cai, W.; Shea, R.; Huang, C.Y.; Chen, K.T.; Liu, J.; Leung, V.C.M.; Hsu, C.H. A Survey on Cloud Gaming: Future of Computer Games. *IEEE Access* **2016**, *4*, 7605–7620. [\[CrossRef\]](#)
2. Chowdhury, M.Z.; Hossan, M.T.; Islam, A.; Jang, Y.M. A Comparative Survey of Optical Wireless Technologies: Architectures and Applications. *IEEE Access* **2018**, *6*, 9819–9840. [\[CrossRef\]](#)
3. Khalighi, M.A.; Uysal, M. Survey on Free Space Optical Communication: A Communication Theory Perspective. *IEEE Commun. Surv. Tutor.* **2014**, *16*, 2231–2258. [\[CrossRef\]](#)
4. Al-Gailani, S.A.; Mohd Salleh, M.F.; Salem, A.A.; Shaddad, R.Q.; Sheikh, U.U.; Algeelani, N.A.; Almohamad, T.A. A Survey of Free Space Optics (FSO) Communication Systems, Links, and Networks. *IEEE Access* **2021**, *9*, 7353–7373. [\[CrossRef\]](#)
5. Jahid, A.; Alsharif, M.H.; Hall, T.J. A contemporary survey on free space optical communication: Potentials, technical challenges, recent advances and research direction. *J. Netw. Comput. Appl.* **2022**, *200*, 103311. [\[CrossRef\]](#)
6. Esmail, M.A.; Ragheb, A.M.; Fathallah, H.A.; Altamimi, M.; Alshebeili, S.A. 5G-28 GHz Signal Transmission Over Hybrid All-Optical FSO/RF Link in Dusty Weather Conditions. *IEEE Access* **2019**, *7*, 24404–24410. [\[CrossRef\]](#)
7. Esmail, M.A.; Ragheb, A.; Fathallah, H.; Alouini, M.S. Investigation and Demonstration of High Speed Full-Optical Hybrid FSO/Fiber Communication System Under Light Sand Storm Condition. *IEEE Photonics J.* **2017**, *9*, 1–12. [\[CrossRef\]](#)
8. Saif, W.S.; Esmail, M.A.; Ragheb, A.M.; Alshawhi, T.A.; Alshebeili, S.A. Machine Learning Techniques for Optical Performance Monitoring and Modulation Format Identification: A Survey. *IEEE Commun. Surv. Tutor.* **2020**, *22*, 2839–2882. [\[CrossRef\]](#)
9. Akkad, K.; Mehboob, H.; Alyamani, R.; Tarlochan, F. A Machine-Learning-Based Approach for Predicting Mechanical Performance of Semi-Porous Hip Stems. *J. Funct. Biomater.* **2023**, *14*, 156. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Kaliappan, V.K.; Gnanamurthy, S.; Yahya, A.; Samikannu, R.; Babar, M.; Qureshi, B.; Koubaa, A. Machine Learning Based Healthcare Service Dissemination Using Social Internet of Things and Cloud Architecture in Smart Cities. *Sustainability* **2023**, *15*, 5457. [\[CrossRef\]](#)
11. Farooq, O.; Singh, P.; Hedabou, M.; Boulila, W.; Benjdira, B. Machine Learning Analytic-Based Two-Staged Data Management Framework for Internet of Things. *Sensors* **2023**, *23*, 2427. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Jargon, J.A.; Wu, X.; Willner, A.E. Optical Performance Monitoring Using Artificial Neural Networks Trained With Eye-Diagram Parameters. *IEEE Photonics Technol. Lett.* **2009**, *21*, 54–56. [\[CrossRef\]](#)
13. Shen, T.S.R.; Meng, K.; Lau, A.P.T.; Dong, Z.Y. Optical Performance Monitoring Using Artificial Neural Network Trained With Asynchronous Amplitude Histograms. *IEEE Photonics Technol. Lett.* **2010**, *22*, 1665–1667. [\[CrossRef\]](#)
14. Khan, F.N.; Shen, T.S.R.; Zhou, Y.; Lau, A.P.T.; Lu, C. Optical Performance Monitoring Using Artificial Neural Networks Trained With Empirical Moments of Asynchronously Sampled Signal Amplitudes. *IEEE Photonics Technol. Lett.* **2012**, *24*, 982–984. [\[CrossRef\]](#)
15. Jargon, J.A.; Wu, X.; Willner, A.E. Optical performance monitoring by use of artificial neural networks trained with parameters derived from delay-tap asynchronous sampling. In Proceedings of the 2009 Conference on Optical Fiber Communication, San Diego, CA, USA, 22–26 March 2009; pp. 1–3. [\[CrossRef\]](#)
16. Mrozek, T. Simultaneous Monitoring of Chromatic Dispersion and Optical Signal to Noise Ratio in Optical Network Using Asynchronous Delay Tap Sampling and Convolutional Neural Network (Deep Learning). In Proceedings of the 2018 20th International Conference on Transparent Optical Networks (ICTON), Bucharest, Romania, 1–5 July 2018; pp. 1–4. [\[CrossRef\]](#)
17. Huang, Y.; Chen, Y.; Yu, J. Optical Performance Monitoring of 56 Gbps Optical PAM4 Signal Using Artificial Neural Networks. In Proceedings of the Asia Communications and Photonics Conference, Guangzhou, China, 10–13 November 2017; Optica Publishing Group: Washinton, DC, USA, 2017. [\[CrossRef\]](#)

18. Li, J.; Zhang, M.; Wang, D.; Wu, S.; Zhan, Y. Joint atmospheric turbulence detection and adaptive demodulation technique using the CNN for the OAM-FSO communication. *Opt. Express* **2018**, *26*, 10494–10508. [[CrossRef](#)]
19. Lohani, S.; Glasser, R.T. Turbulence correction with artificial neural networks. *Opt. Lett.* **2018**, *43*, 2611–2614. [[CrossRef](#)]
20. Esmail, M.A.; Saif, W.S.; Ragheb, A.M.; Alshebeili, S.A. Free space optic channel monitoring using machine learning. *Opt. Express* **2021**, *29*, 10967–10981. [[CrossRef](#)]
21. Ragheb, A.; Saif, W.; Trichili, A.; Ashry, I.; Esmail, M.A.; Altamimi, M.; Almainan, A.; Altubaishi, E.; Ooi, B.S.; Alouini, M.S.; et al. Identifying structured light modes in a desert environment using machine learning algorithms. *Opt. Express* **2020**, *28*, 9753–9763. [[CrossRef](#)] [[PubMed](#)]
22. Esmail, M.A. Optical Wireless Performance Monitoring Using Asynchronous Amplitude Histograms. *IEEE Photonics J.* **2021**, *13*, 1–9. [[CrossRef](#)]
23. Zhang, X.; Gao, S.; Wang, T.; Li, Y.; Ren, P. Correcting Predictions from Simulating Wave Nearshore Model via Gaussian Process Regression. In Proceedings of the Global Oceans 2020: Singapore—U.S. Gulf Coast, Biloxi, MS, USA, 5–30 October 2020; pp. 1–4. [[CrossRef](#)]
24. Seeger, M. Gaussian processes for machine learning. *Int. J. Neural Syst.* **2004**, *14*, 69–106. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.