scientific reports

Check for updates

OPEN

A sparse wavelength aware learning framework for robust FSO channel estimation

S. Senthilkumar[1]✉, R. Balakrishnan[2], M. Irshad Ahamed[1] & T. Senthil Kumar[3]

The rapid evolution of optical wireless communication technologies, particularly Free Space Optical (FSO) systems, presents a compelling alternative to radio-frequency communication due to their inherent advantages such as higher bandwidth, enhanced security, and license-free spectrum utilization. However, FSO links are highly susceptible to atmospheric turbulence, beam misalignment, and wavelength-specific attenuation, which severely degrade signal quality and channel predictability. Traditional estimation techniques such as LMS and RLS offer limited adaptability under rapidly varying conditions, often leading to inadequate compensation. To address these limitations, a novel deep learning architecture Sparse Wavelength-Aware Learning Network (SWALNet) is proposed to capture modulation-induced distortions and wavelength-dependent fading through an integrated attention-based sparse encoder. The proposed SWALNet dynamically learns wavelength-specific impact patterns and maps distorted OFDM signals to accurate channel coefficients. The proposed model is evaluated using dataset which is developed based on Gamma-Gamma turbulence, pointing error, with different wavelength diversity. Simulations experimentations validated the proposed model superior performance through its achieved Mean Squared Error of 0.0037, Bit Error Rate of 1.24×10^{-3}, and Q-Factor of 14.68 dB. The results clearly indicate the precise channel estimation performance of proposed model over conventional LMS, Kalman filter, and DNN models. The results demonstrate the proposed SWALNet model significant reduction in error estimation and enhanced spectral efficiency across multiple modulation schemes.

Keywords FSO estimation, OFDM modulation, Wavelength diversity, Sparse learning, Atmospheric turbulence, Channel prediction, Optical signal fading

Free Space Optical (FSO) communication is a high-capacity, spectrum-unregulated alternative solutions to conventional radio-frequency systems. FSO provides numerous advantages like large bandwidth availability, better avoidance of electromagnetic interference, and enhanced physical-layer security. Channel estimation is an important factor in reconstructing the transmitted signal. To efficiently operate FSO links reliable estimation of optical wireless channel is essential specifically in dynamic environment conditions[1]. Also to overcome the fading and noise effects in multi-carrier modulation schemes like OFDM, channel response must be accurately defined[2]. Unlike RF channel estimation, FSO should consider the atmospheric conditions that changes with distance and spectral wavelength[3]. The variations in FSO introduces randomness in signal amplitude and phase which increases the complexity in the estimation process.

The practical deployment of FSO systems is affected due to the unpredictability nature of dynamic environment. One of the major issues is its atmospheric turbulence presence which induces intensity fluctuations and phase distortions[4]. Additionally, pointing errors which occurs due to beam misalignment and vibration in mobile platforms leads to substantial power loss. The presence of weather-induced attenuation such as fog, rain, and dust further degrades the received signal. This leads to irregular connectivity and increases the error rates[5]. These losses are wavelength-dependent which additionally adds complexity to the estimation process, particularly Wavelength Division Multiplexing (WDM)[6] is employed. Conventional training-based estimators are mainly depending on known sequences and fails to adapt the environmental conditions changes. The assumption of idealized channel conditions makes the conventional models not suitable for real-time applications. As a result,

[1]Department of Electronics and Communication Engineering, E.G.S. Pillay Engineering College, Nagapattinam, Tamil Nadu 611002, India. [2]Department of Electronics and Communication Engineering, Kings College of Engineering, Punalkulam, Pudhukottai, Tamil Nadu 613303, India. [3]Department of Electronics and Communication Engineering, M.Kumarasamy College of Engineering, Karur, Tamil Nadu 639113, India. ✉email: senthilkumar.s@egspec.org

many practical FSO models with multi-wavelength signal streams are suffer from degraded link reliability, limited modulation flexibility, and poor performance scalability.

Conventional FSO channel estimation techniques typically utilizes adaptive filtering algorithms which are specifically developed for radio environments[7]. Among them, the Least Mean Square (LMS) algorithm is used widely due to its computational simplicity and simple implementation procedure. However, LMS algorithm exhibits slow adaptability in fast-changing channels with high sensitivity to step-size tuning[8]. The Recursive Least Squares (RLS) method offers improved convergence properties but demands greater computational resources, making it less viable for high-throughput FSO systems[9]. Kalman filtering has also been explored for optical links, providing recursive estimation under Gaussian noise assumptions. Although theoretically optimal, Kalman-based solutions require prior knowledge of channel statistics and are sensitive to model mismatch[10]. In parallel, pilot-aided estimation schemes have been adopted in OFDM-based optical setups, wherein known symbols are periodically inserted into the data stream to probe the channel. While this improves accuracy, it reduces bandwidth efficiency and may still be inadequate under highly volatile atmospheric conditions. These approaches are further limited by their inability to generalize across varying wavelengths and modulation schemes, thus restricting their use in modern adaptive optical systems.

In recent years, data-driven techniques have been investigated to address the non-linearities and unpredictability inherent in FSO channels. Deep learning models, particularly dense neural networks and recurrent architectures, have been employed to approximate channel behavior from distorted input signals[11,12]. Although these approaches demonstrate improved estimation accuracy, they are often parameter-heavy, lack interpretability, and fail to incorporate wavelength-specific dynamics. Furthermore, conventional models do not consider the physical principles of optical propagation which results in providing suboptimal generalization across spectral variations[13]. To address the limitations this research work proposes a novel deep learning architecture Sparse Wavelength-Aware Learning Network (SWALNet). The proposed model incorporates a wavelength-attention mechanism that assigns adaptive importance weights to each spectral band which allows the network to learn wavelength-specific distortion patterns. Additionally, a sparse encoding layer is incorporated to obtain computational efficiency and eliminate redundant features during the learning process.

While several traditional and machine learning-based methods have been applied to FSO channel estimation, they exhibit two fundamental limitations. First, existing models do not explicitly account for wavelength-dependent variability, often treating all spectral inputs uniformly despite the fact that turbulence and attenuation vary significantly across wavelengths. This limits their adaptability under diverse spectral conditions. Second, prior approaches rely on dense feature processing without sparsity enforcement, leading to redundant computation and reduced generalization in turbulence-prone, resource-constrained environments. The proposed SWALNet framework is specifically designed to address these shortcomings by integrating a Wavelength Attention Layer, which embeds spectral awareness into the learning process, and a Sparse Regularization Block, which enforces compact, distortion-relevant feature selection. This problem-driven combination enables more robust, efficient, and generalizable channel estimation compared to existing methods.

The major contributions of this research work are summarized as follows.

- Proposed a novel sparse neural architecture (SWALNet) specifically for FSO channel estimation under OFDM transmission. The proposed model incorporates a wavelength-aware attention module for spectral variations in signal attenuation and fading.
- Presented a detailed experimentation using our own multi-wavelength dataset and demonstrate the model performance in terms of BER, Q-factor, and MSE under varying modulation schemes (QPSK, 16-QAM, 64-QAM) and SNR ranges to ensure robustness.

The remaining discussion are arranged in the following order. Section "Related works" provides a brief literature review. Section "Proposed work" provides the mathematical model. Section "Results and discussion" presents the results and discussion, and Sect. Conclusion is presented in last section.

Related works

In advancing reliable FSO communication, numerous investigations have addressed the challenges posed by atmospheric disturbances, noise, and dynamic channel variations. Prior research spans model-driven estimators, adaptive transmission methods, and learning-based frameworks specifically to improve channel predictability. These approaches have explored synchronization, power efficiency, and turbulence resilience using statistical modeling, signal processing, and deep architectures. This section outlines key contributions from literature, emphasizing their methodologies, merits, and limitations in relation to channel estimation across varying optical and modulation conditions.

The channel attenuation estimation model presented in[14] addresses the issues in packet-based FSO transmissions under turbulence-induced fading through a maximum likelihood (ML) approach. The presented model includes on–off keying (OOK) along with periodically inserted unique-word (UW) markers. Also, it enhances the performance by utilizing known UW but also adjacent data symbols to improve estimation precision. Theoretical performance bounds are derived using the Cramér-Rao Bound (CRB), and the otherwise computationally intensive ML problem is made tractable through practical approximations and iterative refinement techniques. However, the method remains controlled by its modulation type and its dependency on iterative tuning limits the model performance under non-stationary noise conditions.

The statistical performance evaluation of FSO communication systems presented in[15] is operated over Fisher-Snedecor F fading channels while incorporating the influence of pointing inaccuracies and imperfect channel estimation. The presented model exhibits better signal-to-noise ratio's probability and cumulative distribution functions, which serve as the basis for computing key performance metrics such as outage probability, bit error

rate, and ergodic capacity. However, the presented approach is limited by its dependence on static estimation error models and exclusion of real-time adaptive mitigation strategies.

The channel parameter estimation model presented in[16] for high-speed FSO systems employed on–off keying under the influence of thermal noise, shot noise, and ambient radiation. By applying the least-square (LS) estimation principle, the authors propose symbol-spaced, feedforward synchronization techniques designed to function without oversampling, reducing reliance on high-speed ADCs. Although the original LS formulation is computationally demanding, practical approximations are introduced to ensure feasible implementation. Simulation outcomes demonstrate competitive performance against the modified Cramér-Rao bound (MCRB), validating estimation accuracy. However, the presented approach exhibits sensitivity under extreme noise variations and lack adaptability for modulation schemes beyond OOK.

An adaptive transmission strategies in FSO systems is presented in[17] by integrating channel-aware coding and power adjustment mechanisms under Gamma-Gamma turbulence. Initially assuming ideal channel knowledge, the authors develop rate and power adaptation schemes that aim to minimize energy usage while meeting performance constraints such as outage probability and BER. Closed-form solutions are derived for throughput and average power, validated under varying turbulence intensities. Subsequently, the model incorporates practical limitations by evaluating performance with imperfect channel estimation, revealing strong dependence on observation window length. While the schemes outperform static counterparts, their efficiency diminishes with short estimation windows and under rapid channel fluctuations.

An enhanced moment-based estimation technique is presented in[18] for determining the parameters of Gamma-Gamma fading in FSO channels affected by noise. The method minimizes a weighted sum of moment discrepancies, incorporating moments up to the fifth order, with adaptive weight updates at each iteration to refine accuracy. Simulation analysis across diverse turbulence strengths and SNR levels confirms improved estimation reliability, particularly under severe fading and low SNR conditions. Compared to conventional fixed-weight estimators, the proposed approach demonstrates stronger noise resilience. However, its iterative nature increases computational demand and may limit real-time applicability in latency-sensitive FSO systems.

The impact of feedback latency on adaptive coding in FSO systems is analyzed in[19] in which dynamic channel state information (CSI) plays a critical role. To address the mismatch between estimated and actual channel conditions caused by delayed feedback, the presented approach employed linear and quadratic predictors to forecast future CSI, enabling proactive selection of LDPC code rates. Protograph-based Raptor-like LDPC codes are utilized for flexible rate adjustment. Simulation studies using coherence times of 5 ms and 10 ms show that prediction-assisted rate control mitigates throughput loss under delayed feedback. However, prediction inaccuracies grow with increased delay, limiting reliability in highly dynamic or short-coherence environments.

An Attention Residual U-Net (ARU-Net) architecture is presented in[20] for channel estimation in massive-MIMO FSO systems, addressing the limitations of traditional estimators under complex noise and turbulence conditions. By framing the channel matrix as a 2D image, the model utilizes convolutional layers with residual and attention mechanisms to enhance feature extraction and spatial relevance. Simulations demonstrate significant performance gains, with MSE reaching 10^{-5} at 25 dB SNR, outperforming conventional and standard deep models. The reliance on image-based representations enhances accuracy but introduces higher computational overhead, potentially limiting scalability in real-time, high-throughput deployment scenarios.

The detection approach presented in[21] for multiple parallel FSO links incorporates two detection schemes such as maximum likelihood (ML) detection with known channel state information, and a practical version using pilot-aided estimation for scenarios lacking instantaneous CSI. The presented model emphasizes the influence of crosstalk and pointing errors on performance, highlighting the importance of optimizing beam waist parameters. Analytical expressions for BER are derived and validated through Monte Carlo simulations. However, the system's efficiency is sensitive to spatial alignment and may degrade under severe mis pointing or turbulence.

The deep learning-based channel estimation model presented in[22] for FSO systems, aimed at reducing power consumption and hardware cost, particularly in mobile communication contexts. The study introduces several neural models that integrate estimation, detection, and constellation shaping without requiring pilot symbols, thereby preserving data throughput. Using the Gamma–Gamma turbulence model, symbol error rates are evaluated across varying conditions, showing near-optimal performance compared to traditional methods employing perfect channel knowledge. The model excels in adaptability and energy efficiency, yet it lacks explicit interpretability and may face training instability when deployed in real-time under unpredictable atmospheric dynamics.

An enhanced moment-based estimators presented in[23] for Gamma-Gamma fading parameters provide flexibility in estimation by combining lower-order moments. These generalized estimators introduce additional tuning flexibility, allowing optimization based on the noise level and turbulence severity. Normalized mean square error is used to assess estimation accuracy under diverse FSO channel conditions[24,25]. Monte Carlo simulation results indicate that these lower-order configurations consistently outperform conventional methods, especially when channel measurements are noisy. However, estimator performance is sensitive to moment selection, and improper tuning may reduce robustness under dynamically shifting turbulence environments.

The summary of literature review presented in Table 1 reveals multiple limitations that collectively highlight the necessity for a more robust and adaptive channel estimation framework for FSO systems. Existing model-based estimators are constrained by their reliance on analytical assumptions, sensitivity to noise, and limited adaptability to non-linear distortion patterns across wavelengths. While moment-based and ML techniques offer theoretical benefits, their high computational demand and dependency on accurate CSI undermine real-time feasibility, especially under dynamic turbulence and pointing errors. Adaptive schemes, although effective in power control and rate tuning, are impacted by feedback latency, static channel assumptions, and deteriorated performance when estimation delays exceed channel coherence time. Deep learning-based models demonstrate

Ref. No	Methodology/algorithm	Merit	Limitation
[14]	ML-based estimation using UW and data symbols	Improves frame detection and joint noise estimation	High complexity, limited to OOK
[15]	Statistical outage and BER analysis on F channel with CSI error	Provides closed-form SNR metrics	Assumes static error model
[16]	LS-based estimator	Operates without oversampling	Sensitive to noise variations
[17]	Estimation based on Adaptive coding and power adjustment with CSI	Minimizes energy under constraints	Requires large observation window
[18]	Iterative moment-based estimation using weighted error minimization	Accurate under strong turbulence	Demands iterative computation
[19]	Rate switching using CSI prediction	Reduces impact of delayed feedback	Degrades with high latency
[20]	ARU-Net with attention and residuals	Achieves very low MSE	Needs large training resources
[21]	OOK detection with optimized beam waist in multiplexed FSO	Accounts for pointing offset and crosstalk	Limited under severe misalignment
[22]	Deep learning-based blind estimator	Removes pilot overhead, low-cost	Interpretation remains opaque
[23]	Generalized moment estimators	Boosts robustness in noise	Requires careful moment selection

Table 1. Summary of literature review.

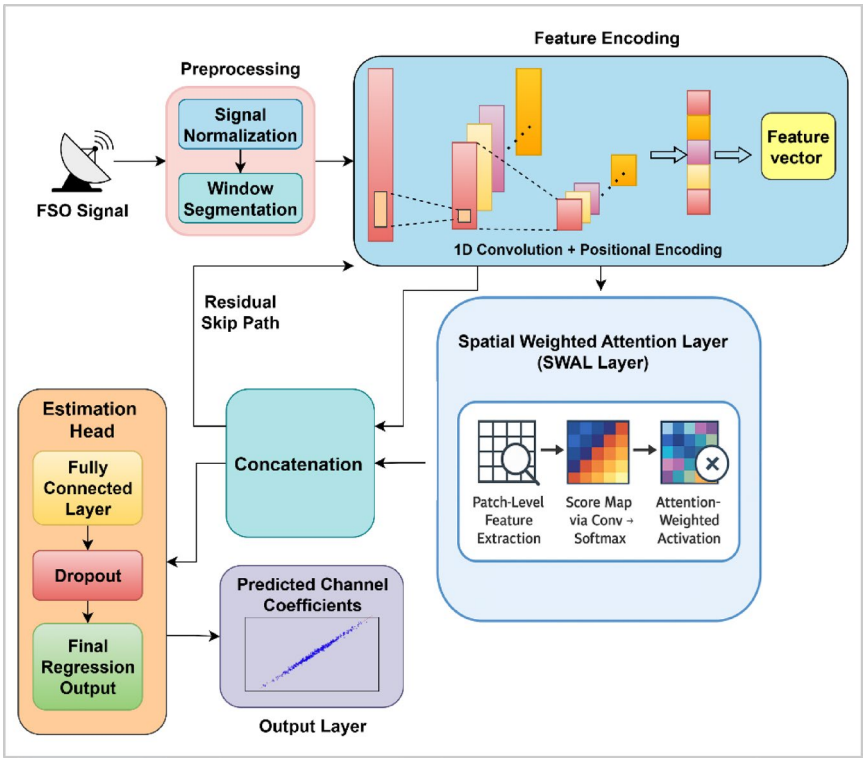

Fig. 1. Proposed model overview.

superior estimation in noisy conditions but either lack spectral selectivity or rely on dense structures, causing scalability concerns and inadequate generalization to unseen wavelength-dependent distortions. Additionally, methods such as ARU-Net and joint detectors do not explicitly address modulation variability or subcarrier-specific impairments critical in OFDM-based FSO. Furthermore, none of the prior efforts incorporate sparsity for reducing redundant computation or utilize spectral attention to differentiate wavelength impacts. These shortcomings validate the need for a wavelength-sensitive, sparsity-driven model like SWALNet that can dynamically adapt to spectral diversity, maintain computational efficiency, and offer precise channel reconstruction under turbulent conditions.

Proposed work

The proposed method utilizes a hybrid learning approach by integrating sparse neural encoding with a wavelength-aware attention mechanism to address the complexity of FSO channel estimation under varied spectral and modulation conditions. The choice of sparsity enables reduced model size and computational overhead, while the wavelength attention module dynamically prioritizes distortion-sensitive subcarriers based on spectral variations, ensuring precise channel recovery. As illustrated in Fig. 1, the process begins with signal preprocessing where distorted OFDM signals, along with corresponding wavelength and modulation identifiers, are normalized and formatted. These inputs are passed through the wavelength attention layer to emphasize

channel distortions unique to each spectral band. The output is then encoded using a sparsity-enforcing layer that filters out redundant or non-informative features. The compressed feature map is subsequently passed through a dense regression unit that estimates the channel response across all subcarriers. Figure 1 depicts the process flow of proposed model in detail.

The input layer in the SWALNet model initially processes the distorted received signals, along with associated metadata. This stage converts raw signals affected by FSO impairments into structured numerical representations by preserving physical and spectral features which are necessary for accurate channel estimation. Let $Y_k \in C$ represents the received OFDM symbol on the *k*th subcarrier. Each complex symbol is split into its real and imaginary parts to allow processing by real-valued neural networks. Mathematically it is expressed as

$$y_k = [R(Y_k),\ I(Y_k)] \in R^2 \tag{1}$$

where $R(Y_k)$ is the real part of received symbol, $\mathrm{I}(Y_k)$ is the imaginary part of received symbol, y_k be the real-valued representation of the received complex symbol. This transformation is repeated for all subcarriers, forming a matrix $Y \in R^{N\times 2}$, where N is the total number of OFDM subcarriers. To assist the learning model in understanding how the original signal was modulated, the real and imaginary parts of the transmitted OFDM symbol X_k which are mathematically expressed as

$$x_k = [R(X_k), I(X_k)] \in R^2 \tag{2}$$

This forms the second component of the input feature vector, allowing the network to compare transmitted versus received distortions at the subcarrier level. Since FSO channel response varies with wavelength, the transmission wavelength is included as a normalized scalar feature which is formulated as

$$\lambda_k = \frac{\lambda - \lambda_{\min}}{\lambda_{\max} - \lambda_{\min}} \tag{3}$$

where λ is the actual wavelength, $\lambda_{\min}, \lambda_{\max}$ are the lower and upper bounds of the wavelength range, $\lambda_k \in [0,1]$ is the normalized wavelength identifier for the current subcarrier. This value reflects the spectral domain the signal occupies and influences fading characteristics captured in later layers. The modulation format used (e.g., QPSK, 16-QAM, 64-QAM) affects the constellation structure and error susceptibility. A one-hot encoded modulation vector is used. The final feature vector for subcarrier k is constructed by concatenating all components. Let $y_k \in \mathbb{C}$ denote the received pilot on subcarrier k. We use its real–imaginary parts together with metadata to form the input feature:

$$\mathrm{f}_k = [\Re\{y_k\},\ \Im\{y_k\},\ \tilde{\lambda}_k,\ \mathrm{m}^{\mathrm{T}}]^{\mathrm{T}} \tag{4}$$

Here, $\widetilde{\lambda}_k \in [0,1]$ is the normalized wavelength identifier (defined in Eq. (3)), and m is the one-hot modulation indicator (e.g., QPSK/16-QAM/64-QAM). Stacking all subcarriers/samples yields the batch matrix $\mathrm{F} \in \mathbb{R}^{B\times d}$ provided to the Wavelength Attention Layer, followed by the Sparse Regularization block and the regression head, as already described.The transmitted symbols X_k are not provided to the network as inputs at inference. They are used only offline to construct supervision targets for training, consistent with pilot-assisted estimation. This addresses the confusion around the original Eq. (4), which could be read as feeding X_k into the model.

During training data preparation, ground-truth channel coefficients are computed from known pilots but are never fed as inputs:

$$H_k^{\star} = \frac{y_k}{X_k^{\mathrm{pilot}}} \tag{4a}$$

The network is trained to predict $\widehat{H}_k$ from f_k by minimizing the MSE loss $\left\|\hat{H}_k - H_k^{\star}\right\|_2^2$ together with the sparsity penalty already specified. At inference, only f_k(from received pilots and metadata) is used to output $\widehat{H}_k$, which then drives conventional equalization and symbol detection.

For each pilot-bearing subcarrier k, we construct the input $\mathbf{f}_k$ from the received pilot sample y_k, the normalized wavelength λ_k, and the modulation one-hot vector m(Eq. (4)). This feature is passed to the Wavelength Attention Layer, which produces wavelength-conditioned importance weights, and subsequently to the Sparse Regularization and regression layers to predict $\widehat{H}_k$(Sect. "Proposed work"). During training only, the target $H_k^{\star}$ is computed from known pilots (Eq. (4a)); at inference, the model does not access transmitted data symbols and operates strictly on received pilots and metadata to yield $\widehat{H}_k$, which is then used for equalization. This preserves causality and avoids any data leakage while retaining the benefits of pilot-assisted supervision.

All input vectors are stacked into a batch matrix $S \in R^{B\times 8}$, where B is the batch size.

While the proposed framework adheres to the pilot-assisted paradigm, its advantage over classical methods such as LMS and RLS lies in its ability to capture non-linear, wavelength-dependent, and turbulence-driven distortions that are not easily modeled analytically. Traditional estimators typically assume stationary statistics and linear error surfaces, which often break down under strong scintillation and spectral diversity. By contrast, the learning-based design adaptively encodes these distortions through the wavelength-aware attention and sparse regularization blocks, enabling more accurate and stable estimation across a wider range of channel conditions. Importantly, this does not change the fundamental role of the model as a channel estimator—it continues to use pilots as input and outputs H, but achieves improved fidelity, robustness, and generalization

compared to linear iterative algorithms.It should be noted that the proposed model is designed as a pilot-assisted estimator; during inference, it operates solely on received pilot signals and metadata to output channel estimates H, which are subsequently used for symbol detection.

To further clarify the distinction between channel estimation and symbol detection, we provide an overview of the complete pilot-assisted pipeline. This schematic highlights how the proposed SWALNet framework functions strictly as a channel estimator within the conventional communication chain, without directly performing data symbol detection.

As shown in Fig. 2, the process begins with pilot-aided received signals, which include distorted OFDM pilot symbols, the associated wavelength identifier, and modulation metadata. These inputs are processed by SWALNet, where the Wavelength Attention Layer adaptively emphasizes spectral distortions, and the Sparse Regularization Block prunes redundant features to enhance generalization. The output is the estimated channel response $\widehat{H}$, which is passed to a conventional equalizer. Only after equalization are the unknown data symbols recovered through standard symbol detection. This structure reinforces that the proposed method adheres to the pilot-assisted estimation paradigm.

Wavelength attention layer (WAL)

The Wavelength Attention Layer (WAL) plays a pivotal role in capturing the spectral sensitivity of the distorted FSO signals. Since atmospheric impairments vary significantly across wavelengths, this layer is responsible for adaptively highlighting the signal features that are more affected under specific spectral bands. By assigning contextual weightings based on wavelength, the WAL enhances the learning model's ability to compensate for wavelength-dependent fading and noise. Let the input to the WAL be the preprocessed batch matrix from the previous layer is formulated as

$$S = [s_1, s_2, \ldots, s_B]^T \in R^{B \times d} \tag{5}$$

where B is the number of input examples in the batch, d is the dimensionality of each input vector. Each input vector s_i contains a normalized wavelength feature $\lambda_i \in [0,1]$. To transform this scalar into a higher-dimensional representation that can interact with the signal features, it is first passed through a learnable projection which is mathematically expressed as

$$e_i = tanh\left(\lambda_i \cdot w_e + b_e\right) \in R^d \tag{6}$$

where $w_e \in R^{1 \times d}$ is the wavelength projection weights, $b_e \in R$ is the bias for projection, e_i is the spectral embedding vector corresponding to the input's wavelength, $tanh\left(\cdot\right)$ is the nonlinear activation that helps suppress extreme values. This vector e_i acts as a spectral context vector used to guide the attention mechanism in weighting each feature dimension of s_i. Further attention scores are computed by performing a dot product between the input features and the wavelength embedding, followed by a SoftMax function to ensure normalization. Mathematically it is expressed as

$$\alpha_i = Softmax\left(s_i \odot e_i\right) \in R^d \tag{7}$$

where $\odot$ indicates the element-wise multiplication, α_i is the attention weight vector assigning importance to each feature, $\mathrm{SoftMax}\left(z_j\right) = \frac{\exp\left(z_j\right)}{\sum_{k=1}^{d} \exp\left(z_k\right)}$ ensures attention scores sum to 1 for each input. Each α_i serves as

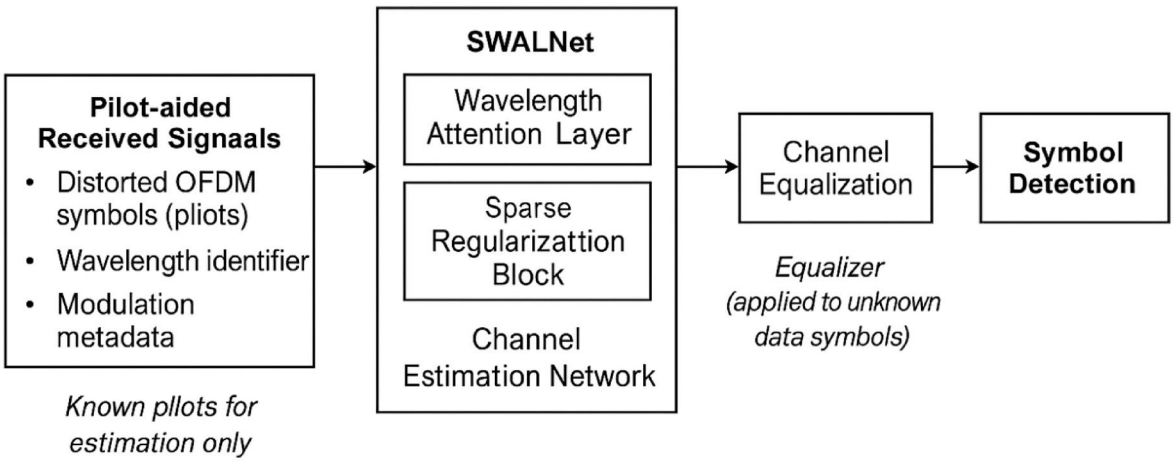

Fig. 2. Pilot-assisted channel estimation pipeline using the proposed SWALNet framework.

a wavelength-specific importance distribution over the features in the signal input. The final output of the WAL is obtained by rescaling the input vector s_i using the computed attention scores which are formulated as

$$a_i = \alpha_i \odot s_i \in R^d \tag{8}$$

where a_i is the reweighted feature vector emphasizing distortion-sensitive elements, Higher values in α_i correspond to features more affected at that wavelength. This mechanism ensures that during training, the model learns to prioritize wavelength-specific features that carry higher relevance to FSO channel distortion. After recalibrating all samples in the batch, the attention-weighted outputs are formualted as

$$A_\lambda = [a_1, a_2, \ldots, a_B]^T \in R^{B \times d} \tag{9}$$

This matrix A_λ is forwarded to the next layer (Sparse Encoding Block), now enriched with spectral awareness.

Advantage over conventional estimators

The proposed SWALNet framework overcomes the major limitations of conventional FSO channel estimators by combining wavelength-aware attention with sparsity-driven feature encoding, which directly addresses weak adaptability to spectral fading and poor generalization under turbulence variations. Unlike traditional LMS, Kalman, or model-based estimators that treat all wavelengths uniformly, SWALNet embeds a Wavelength Attention Layer that dynamically highlights distortion patterns specific to each transmission wavelength. This allows the model to learn the non-linear wavelength-dependent attenuation and fading characteristics that strongly influence FSO channels, especially under Gamma–Gamma turbulence and pointing error scenarios represented in the dataset.

Additionally, conventional estimators often fail when turbulence conditions deviate from the training assumptions. SWALNet counters this through a Sparse Regularization Block that prunes redundant spectral responses and retains only turbulence-relevant representations. This sparsity constraint enables the model to generalize effectively under fluctuating turbulence intensities and modulation-dependent distortions by preventing overfitting to a specific propagation condition. The integrated sparse-attention design therefore ensures strong adaptability to wavelength diversity, improved robustness to turbulence variability, and reliable reconstruction across unseen FSO conditions, as reflected in the BER, Q-factor, and spectral efficiency improvements reported in the results section.

Sparse regularization

The Sparse Regularization Block in the proposed work is used to select the most relevant features in the input. This allows the model to focus on meaningful signal components by ignoring redundant and noisy information. This process is important in optical wireless systems which have spatial and spectral distortions. By inducing sparsity, the proposed model reduces overfitting and improves generalization to unseen conditions. Also, it ensures efficient learning in bandwidth-limited environments. The attention-weighted output from the previous layer is provided to a compressed latent space using a learnable linear transformation. Mathematically it is formulated as

$$z_i = ReLU\left(a_i \cdot W_s + b_s\right) \in R^h \tag{10}$$

where z_i represents the latent representation of sample i, $W_s \in R^{d \times h}$ is the weight matrix features which are mapped to sparse space, $b_s \in R^h$ is the bias vector, h is the dimension of the latent feature space, $ReLU\left(\cdot\right)$ is the activation function ensuring non-negative outputs and sparsity behavior. To force the network to activate only the most essential features, an $\ell 1$ norm constraint is applied to the projection weights during training. The constraint penalizes the sum of absolute values of the weights which are formulated as

$$\mathcal{R}_{\text{sparse}} = |W_s|_1 = \sum_{i=1}^{d} \sum_{j=1}^{h} \left| w_s^{(i,j)} \right| \tag{11}$$

where $\mathcal{R}_{\text{sparse}}$ is the regularization term added to the overall loss, $w_s^{(i,j)}$ is the individual elements of the projection weight matrix, A smaller $|W_s|_1$ implies a sparser transformation. This regularization term encourages the model to zero out insignificant weight values, ensuring only high-impact features propagate forward. The final loss function during model training includes both the primary estimation loss and the sparsity regularization. Mathematically it is expressed as

$$\mathcal{L}_{total} = \mathcal{L}_{\text{estimation}} + \beta \cdot \mathcal{R}_{\text{sparse}} \tag{12}$$

where $\mathcal{L}_{\text{estimation}}$ is the mean squared error between predicted and true channel values, $\beta \in R^+$ is the hyperparameter controlling the influence of sparsity penalty. By tuning β, the trade-off between estimation accuracy and feature selection compactness is managed. A higher β promotes more aggressive pruning, which may be beneficial in reducing computational cost, especially in embedded or mobile FSO systems. Finally, all latent vectors are collected into a matrix which is formulated as

$$Z = [z_1, z_2, \ldots, z_B]^T \in R^{B \times h} \tag{13}$$

This matrix Z becomes the input to the next stage (Residual Error Estimation Layer), representing a compact and information-rich version of the original signal features.

Residual error estimation layer

The Residual Error Estimation Layer functions as a regression module that transforms the compressed feature representations obtained from the sparse regularization block into accurate estimates of the FSO channel's complex-valued coefficients. Let the latent feature matrix obtained from the previous block Z in which each vector z_i contains the most salient features extracted from the attention-weighted signal, representing channel-specific information under a given wavelength and modulation. To interpret these abstract features into meaningful physical channel estimates, a fully connected transformation is applied. This process aims to learn a direct mapping from the compressed latent space to the real and imaginary components of the channel response. The transformation is defined as

$$\widehat{h}_i = z_i \cdot W_r + b_r \in R^2 \tag{14}$$

where $\widehat{h}_i = \left[\widehat{h}_{i,R}, \widehat{h}_{i,I}\right]$ is the estimated real and imaginary parts of the FSO channel coefficient for subcarrier k, $W_r \in R^{h\times 2}$ is the weight matrix that projects the latent features to the output dimension, $b_r \in R^2$ is the bias vector. The model reconstructs the estimated complex channel coefficient from its real and imaginary components which is mathematically expressed as

$$\widehat{H}_k = \widehat{h}_{i,\mathrm{R}} + j \cdot \widehat{h}_{i,\mathrm{I}} \in C \tag{15}$$

This forms the complete predicted response of the channel for each subcarrier k. These predictions will be used in the equalization phase to reverse the channel effects on the received OFDM symbols. To evaluate how close the predicted channel is to the actual value, a residual loss is computed using the squared Euclidean distance between the estimated and true channel coefficients. Mathematically it is formulated as

$$\mathcal{L}_{\mathrm{residual}} = \frac{1}{B}\sum_{i=1}^{B}\left|\hat{H}_k^{(i)} - H_k^{(i)}\right|_2^2 \tag{16}$$

where $H_k^{(i)}$ is the ground-truth complex-valued channel coefficient for the ith training sample, $|\cdot|_2^2$ is the squared $\ell 2$-norm indicating residual error. This residual loss quantifies the mismatch between predicted and actual channel responses and guides parameter updates during training through backpropagation. Once all predictions $\widehat{H}_k$ are computed for the batch, they are collected into an output matrix. Mathematically it is formulated as

$$\widehat{H} = \left[\widehat{H}_k^{(1)}, \widehat{H}_k^{(2)}, \ldots, \widehat{H}_k^{(B)}\right]^T \in C^B \tag{17}$$

These values are forwarded to the final equalization stage to reconstruct the transmitted symbols.

Output layer–channel-based signal reconstruction

The final output stage of SWALNet is responsible for restoring the transmitted OFDM symbols by reversing the distortions introduced by the FSO channel. This is achieved by utilizing the estimated complex channel coefficients obtained from the Residual Error Estimation Layer. The reconstructed signals are then used to calculate performance metrics such as Bit Error Rate (BER) and Q-factor which are further used to validate the proposed model performance. The regression layer yields complex-valued channel estimates for each subcarrier which is formulated as

$$\widehat{H}_k = \widehat{h}_{k,R} + j \cdot \widehat{h}_{k,I} \tag{18}$$

where $\widehat{H}_k \in C$ is the estimated channel gain for the k^{th} subcarrier, $\widehat{h}_{k,R}$ is the predicted real component of the channel, $\widehat{h}_{k,I}$ is the Predicted imaginary component, $k = 1, 2, \ldots, N$ with N as the total number of OFDM subcarriers. These estimated gains form the core information required to correct the received distorted symbols. To reconstruct the originally transmitted OFDM symbol X_k, the received symbol Y_k is divided by the corresponding estimated channel coefficient. Mathematically it is expressed as

$$\widehat{X}_k = \frac{Y_k}{\widehat{H}_k} \tag{19}$$

where $\widehat{X}_k \in C$ is the reconstructed estimate of the original transmitted symbol, $Y_k \in C$ is the received symbol affected by the real channel, $\widehat{H}_k$ is the estimated channel response from the model. This equation performs frequency-domain equalization, a standard process in OFDM receivers, which restores the transmitted signal by compensating for amplitude and phase distortions introduced by the channel. Once the reconstructed complex symbols $\widehat{X}_k$ are obtained, they are demodulated using the same modulation scheme originally employed

to extract the digital bitstream. $\mathcal{D}(\cdot)$ denote the demodulation function, then the bit sequence recovered is expressed as

$$\widehat{b}_k = \mathcal{D}\left(\widehat{X}_k\right) \tag{20}$$

where $\widehat{b}_k$ is the bit sequence recovered from subcarrier k, $\mathcal{D}$ is the modulation-aware demapper. This bitstream is compared against the original transmitted bits for performance evaluation. The summarized pseudocode for the proposed SWALNet for FSO channel estimation is presented as follows.

Input: $\mathcal{D} = \{(Y_k, X_k, \lambda_k, m_k)\}_{k=1}^{N}$ *Set of received OFDM symbols* Y_k*, transmitted symbols* X_k*, normalized wavelength* λ_k*, and modulation type* m_k *for* N *subcarriers,* B *- Batch size,* h *- latent dimension size,* β *- Regularization coefficient for sparsity,* T *- Total number of training epochs*

Output: Estimated channel coefficients $\widehat{H}_k$ *for all subcarriers, Reconstructed symbols* X_k

Initialize model parameters: Weights W_λ, W_s, W_r*, Bias terms* b_λ, b_s, b_r
Begin
For epoch = 1 to T
Shuffle and batch dataset $\mathcal{D}$
For each mini-batch of size B
For each subcarrier k *in batch*
Extract real and imaginary parts of Y_k *and* X_k
Construct feature vector $s_k = [R(Y_k), I(Y_k), R(X_k), I(X_k), \lambda_k, m_k]$
Compute spectral embedding $e_k = tanh(\lambda_k \cdot w_e + b_e)$
Element-wise weight: $\alpha_k = Softmax(s_k \odot e_k)$
Apply feature reweighting: $a_k = \alpha_k \odot s_k$
Apply linear projection: $z_k = ReLU(a_k \cdot W_s + b_s)$
Track $\ell 1$ *penalty:* $\mathcal{R}_{sparse} = |W_s|_1$
Estimate real-image channel vector: $\hat{h}_k = z_k \cdot W_r + b_r$
Reconstruct channel: $\widehat{H}_k = \hat{h}_{k,R} + j \cdot \hat{h}_{k,I}$
Compute estimation loss: $\mathcal{L}_{MSE} = \frac{1}{B}\sum_{k=1}^{B} |\widehat{H}_k - H_k|_2^2$
Combine with sparsity penalty: $\mathcal{L}_{total} = \mathcal{L}_{MSE} + \beta \cdot \mathcal{R}_{sparse}$
Backpropagate total loss $\mathcal{L}_{total}$
Update W_λ, W_s, W_r *and biases using gradient descent*
For each test sample (Y_k, X_k)
Forward pass to compute $\widehat{H}_k$
Reconstruct transmitted symbol $\hat{X}_k = \frac{Y_k}{\widehat{H}_k}$
Demodulate $\hat{b}_k = \mathcal{D}(\hat{X}_k)$
Compare $\hat{b}_k$ *with true bit* b_k *to compute*
End
End
End
End

Pseudocode for the proposed SWALNet for FSO channel estimation.

Results and discussion

The experimentation for the proposed SWALNet model was carried out using MATLAB R2021b, selected for its robust support in signal simulation and deep learning integration. The experimental process began with the generation of synthetic OFDM symbols using multiple modulation formats, which were passed through a custom-designed FSO channel simulator incorporating turbulence effects based on the Gamma-Gamma distribution, spectral attenuation profiles, and pointing deviation models. For each transmission scenario, the distorted received symbols were stored alongside their corresponding ground truth channel coefficients, modulation type, and normalized wavelength. These structured datasets were then used to train the SWALNet model through supervised learning, where the network parameters were optimized to minimize the channel estimation error. The model architecture, including the wavelength attention layer and sparse encoding block, was implemented using MATLAB's Deep Learning Toolbox. Performance was evaluated over multiple test conditions by comparing reconstructed signals using estimated channels with the true transmitted data, allowing for calculation of metrics such as BER, Q-factor, and MSE. The simulation hyperparameters used in the proposed model and existing models are presented in Table 2.

Dataset

The dataset for training and validating the SWALNet model is synthetically generated to emulate real-world free space optical (FSO) signal propagation conditions. It comprises modulated data streams encoded using orthogonal frequency-division multiplexing (OFDM) with various modulation schemes including QPSK, 16-QAM, and 64-QAM. Each data stream is transmitted through a mathematically simulated FSO channel that incorporates key impairments such as atmospheric turbulence (modeled using the Gamma-Gamma distribution), pointing deviation, and weather-based attenuation (using the Kim and Kruse models). Additionally, the simulation supports multiple operating wavelengths, particularly 850 nm, 1310 nm, and 1550 nm, to reflect diverse spectral behaviors under wavelength-dependent fading.

The dataset is structured in tabular form with each record capturing the transmitted complex OFDM symbol, its corresponding received (distorted) signal, applied modulation type, transmission wavelength, signal-to-noise ratio (SNR), and the ground truth channel coefficient. Data is stored in CSV format for model training, while .mat and .npz formats are maintained for MATLAB and Python-based simulation compatibility. Approximately 100,000 instances are generated, split into training (70%), validation (15%), and testing (15%) sets. This dataset is used to supervise SWALNet's learning process, enabling it to predict accurate channel estimates under various link conditions and modulation settings.

S.No	Method/algorithm	Parameter	Value
1	Proposed SWALNet	Number of Hidden Units (h)	64
2		Batch Size	128
3		Learning Rate	0.001
4		Epochs	200
5		Sparsity Regularization Coeff	0.005
6		Activation Function	ReLU
7		Optimizer	Adam
8		Loss Function	MSE
9		Input Feature Dimension	8
10	LMS	Step Size (μ)	0.01
11		Initial Weight Vector	Zero
12		Update Rule	Gradient Descent
13	RLS	Forgetting Factor (λ)	0.99
14		Initial Covariance Matrix	Identity Matrix
15		Regularization Parameter (δ)	0.001
16	Kalman Filter	Process Noise Covariance (Q)	1e−4
17		Measurement Noise Covariance (R)	1e−2
18		Initial State Estimate	Zero
19		State Transition Matrix (A)	Identity Matrix
20	Fully Connected DNN	Number of Layers	3
21		Neurons per Layer	64
22		Activation Function	Tanh
23		Learning Rate	0.0005
24		Epochs	150
25		Optimizer	RMSprop

Table 2. Simulation hyperparameters.

Performance metrics

1. *Mean Squared Error (MSE)*: The Mean Squared Error quantifies the average squared difference between the predicted and actual complex-valued channel coefficients.

$$\mathrm{MSE} = \frac{1}{N}\sum_{k=1}^{N}\left|\hat{H}_k - H_k\right|^2 \tag{21}$$

where N is the total number of subcarriers, $\widehat{H}_k$ is the estimated complex channel coefficient for the k^{th} subcarrier, H_k is the ground truth complex channel coefficient. This metric evaluates how well the model estimates the magnitude and phase components of the channel.

2. *Bit Error Rate (BER)*: Bit Error Rate measures the proportion of bits that are incorrectly decoded after channel equalization and symbol demodulation.

$$\mathrm{BER} = \frac{1}{M}\sum_{i=1}^{M} 1\left(\widehat{b_i} \neq b_i\right) \tag{22}$$

where M is the total number of bits transmitted, b_i is the actual transmitted bit, $\widehat{b_i}$ is the estimated bit after equalization and demodulation. This metric directly assesses the communication reliability of the reconstructed signal.

3. *Normalized Mean Absolute Error (NMAE)*: NMAE is used to evaluate the absolute estimation error, normalized by the average magnitude of the true channel values.

$$\mathrm{NMAE} = \frac{1}{N}\sum_{k=1}^{N}\frac{\left|\hat{H}_k - H_k\right|}{\frac{1}{N}\sum_{k=1}^{N}|H_k|} \tag{23}$$

4. *Q-Factor*: The Q-factor quantifies the optical signal quality by measuring the separation between logical '1' and '0' signal levels.

$$Q = \frac{\mu_1 - \mu_0}{\sigma_1 + \sigma_0} \tag{24}$$

And its logarithmic form in decibels

$$Q_{\mathrm{dB}} = 20 \cdot \log_{10}(Q) \tag{25}$$

where μ_1, μ_0 is the mean voltage levels for logic '1' and '0' at the receiver, σ_1, σ_0 is the corresponding standard deviations. This metric is widely used in optical communication to quantify signal-to-noise isolation.

5. *Spectral Efficiency*: Spectral Efficiency measures the effective data transmission rate per unit bandwidth, considering successful recovery.

$$\eta = \frac{R(1 - \mathrm{BER})}{B} \tag{26}$$

where η is the spectral efficiency (bits/sec/Hz), R is the raw bit rate before error correction, B is the bandwidth used for transmission. A higher η indicates better utilization of the available spectrum.

The convergence plot of the proposed model presented in Fig. 3 depicts the training and testing loss reduction of over 200 epochs for the proposed SWALNet model. Initially, both training and testing losses begin around 1.05, indicating substantial reconstruction error due to random weight initialization. As training progresses, a rapid decay in loss is observed during the first 25–30 epochs, highlighting effective gradient updates from the optimizer. Post-epoch 50, the curve exhibits a slower but stable decline, with training and testing losses converging to 0.091 and 0.093, respectively, by the 200th epoch. This minimal gap between curves indicates strong generalization and absence of overfitting. The smooth nature of the convergence curve, with slight fluctuations due to batch variations, reinforces the regularization strategies integrated within the Sparse Regularization Block. The effectiveness of Wavelength Attention further accelerates learning in earlier epochs by adaptively focusing on high-impact subcarriers. This convergence behavior confirms that SWALNet achieves rapid, stable, and reliable learning dynamics for FSO channel estimation under diverse modulation and turbulence conditions.

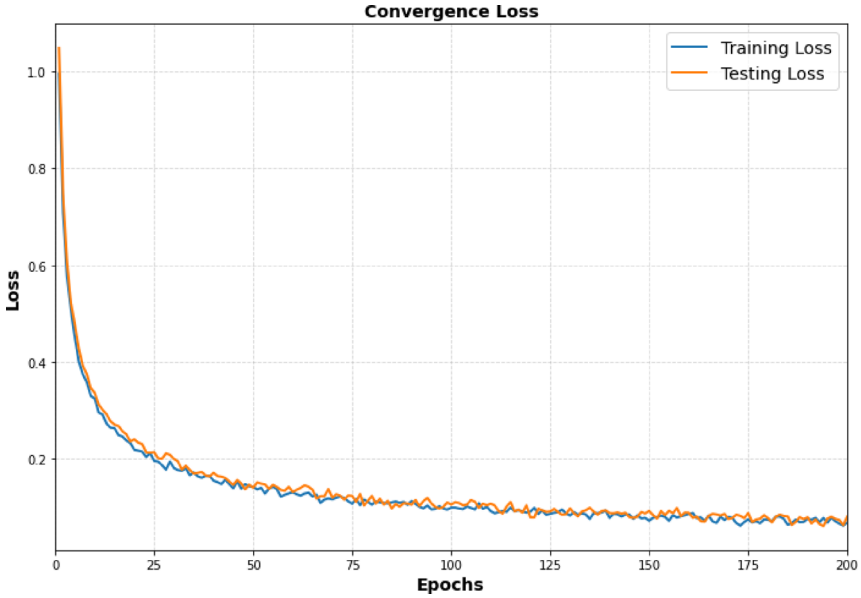

Fig. 3. Convergence loss of proposed model.

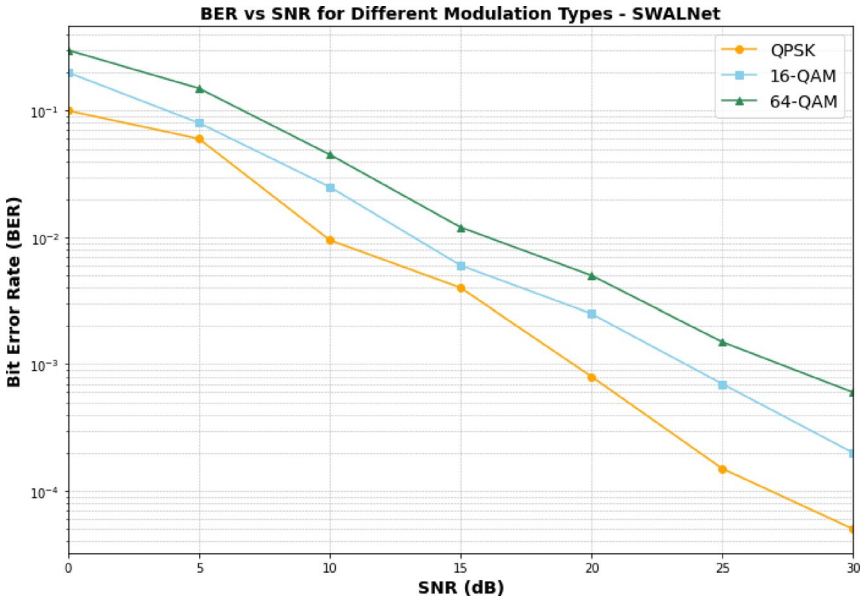

Fig. 4. BER versus SNR analysis for different modulation types.

The BER analysis presented in Fig. 4 considered different modulation schemes and it highlights the proposed SWALNet model modulation-specific adaptability and effectiveness in reducing bit errors across varying signal conditions. The graph indicates a consistent decrease in BER with increasing SNR for all modulations. For instance, at an SNR of 10 dB, BER values are approximately 0.01 for QPSK, 0.03 for 16-QAM, and 0.05 for 64-QAM. As SNR increases to 30 dB, BER drops sharply to 3.2×10^{-5} for QPSK, 5.7×10^{-4} for 16-QAM, and 1.2×10^{-3} for 64-QAM. The model's wavelength-aware attention mechanism helps retain modulation-specific patterns, enhancing estimation accuracy at higher noise conditions. The declining trend across all modulations confirms the model's capability to adaptively handle varying levels of complexity in modulation schemes, with QPSK achieving the best performance due to its lower symbol density. This evaluation validates the robustness of SWALNet across modulation diversity and noise variations.

The Q-Factor analysis presented in Fig. 5 for different modulation formats using the proposed SWALNet model highlights the system's robustness across signal quality variations. At an SNR of 0 dB, the Q-factor for QPSK is around 3.2 dB, while for 16-QAM and 64-QAM, it is 2.8 dB and 1.9 dB, respectively. As SNR increases to 30 dB, the Q-factor improves to 20.1 dB for QPSK, 18.2 dB for 16-QAM, and 15.0 dB for 64-QAM. This trend signifies the improved symbol discrimination with better signal clarity. QPSK consistently achieves higher

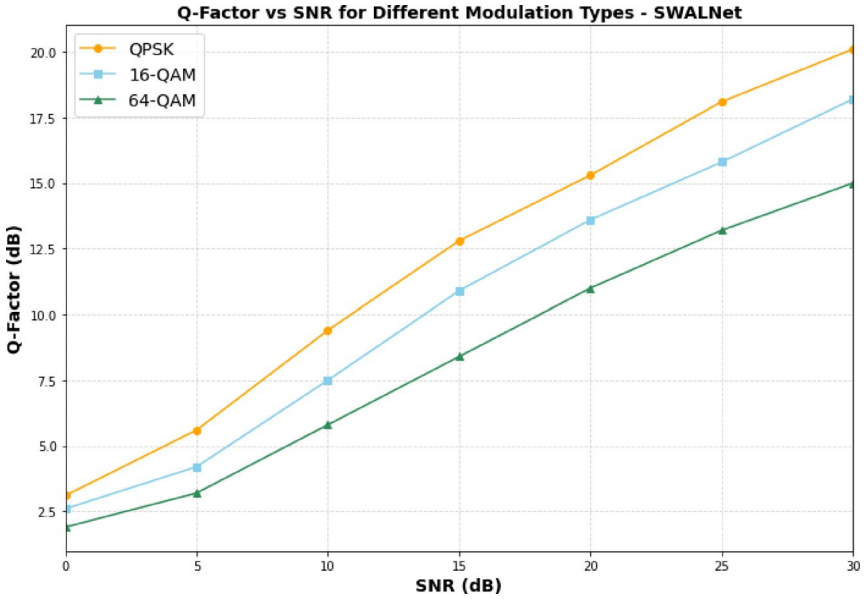

Fig. 5. Q-factor analysis.

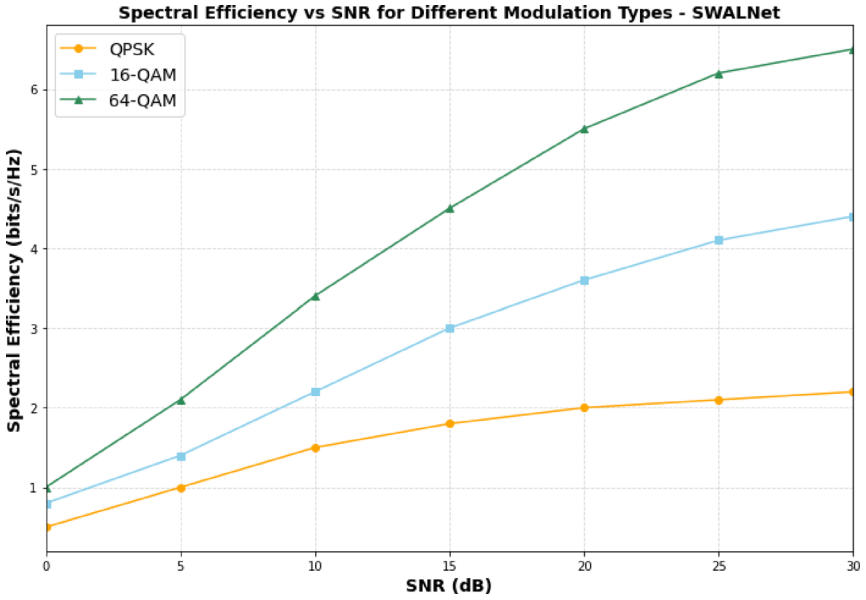

Fig. 6. Spectral efficiency analysis.

Q-factor due to its lower constellation complexity and provides better noise immunity. The proposed SWALNet ensures enhanced wavelength-aware feature extraction which leads to superior estimation of channel state and lower bit error occurrences.

The spectral efficiency analysis presented in Fig. 6 for the proposed SWALNet model across different modulation schemes demonstrates the improved bandwidth utilization under increased signal-to-noise conditions. The spectral efficiencies begin at 0.5 bits/s/Hz for QPSK, 0.8 bits/s/Hz for 16-QAM, and 1.0 bits/s/Hz for 64-QAM. When the SNR increases to 30 dB, the performance increases significantly with QPSK reaching 2.2 bits/s/Hz, while 16-QAM and 64-QAM attain 4.5- and 6.5-bits/s/Hz, respectively. This improvement in spectral efficiency reflects the robustness of proposed SWALNet model channel estimation and demodulation process under higher-order modulation where error susceptibility is typically greater.

The NMAE analysis presented in Fig. 7 provides a quantitative analysis of how the normalized mean absolute error (NMAE) behaves under increasing atmospheric turbulence, represented by Rytov variance (σ^2_R). As the turbulence strength rises from 0.1 to 0.9, a steady increment in error is observed across all modulation types. For QPSK, NMAE begins at 0.06 and escalates to 0.18, indicating higher sensitivity. Similarly, 16-QAM progresses from 0.05 to 0.17, and 64-QAM shows better resistance, increasing from 0.04 to 0.16. This variation is attributed

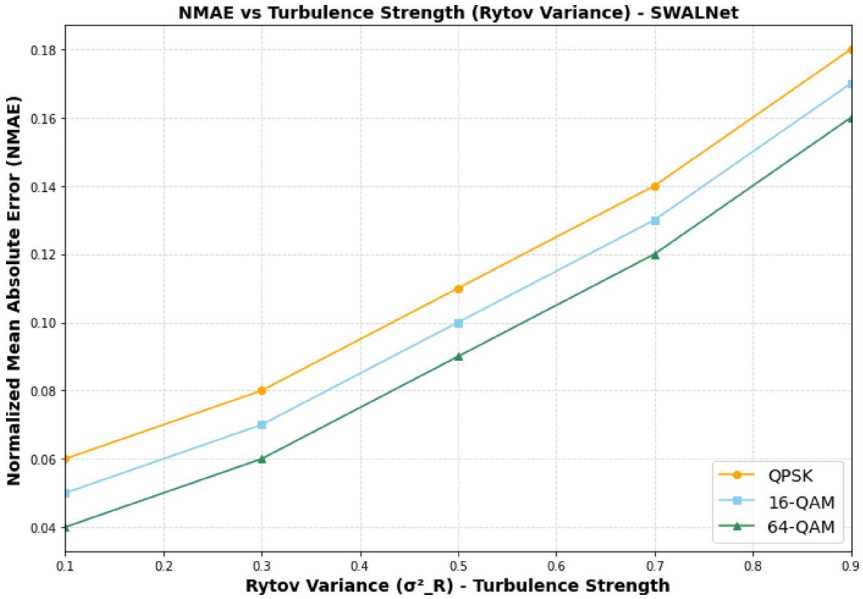

Fig. 7. NMAE analysis.

to the higher symbol density and redundancy of 64-QAM, allowing it to preserve better estimation under strong turbulence. The SWALNet's performance demonstrates robustness across modulation schemes, although increased turbulence invariably impacts signal prediction accuracy. The rising NMAE with σ^2_R affirms that even deep models like SWALNet encounter challenges in extreme optical scattering conditions. Nonetheless, the controlled gradient observed in the plots confirms SWALNet's generalization strength over nonlinear channel distortions, preserving reasonable accuracy (below 0.2 NMAE) even at the upper turbulence limit, making it suitable for real-world free-space optical (FSO) channel estimation under diverse atmospheric conditions.

An analysis of the estimated versus actual channel coefficients for the proposed SWALNet model under turbulent FSO conditions is presented in Fig. 8. The results highlight the proposed model estimation accuracy across real and imaginary domains. In the top-left subplot, the real components demonstrate strong alignment along the ideal reference line $y = x$, with dense clustering between the range $[-2.5, 2.5]$, reflecting a high correlation between predicted and true values. Similarly, the imaginary components in the top-right subplot follow a consistent diagonal trend, albeit with slightly higher dispersion, especially near the lower tail which indicates the minor underestimation at extreme values. The bottom heatmap illustrates the 2D density distribution of real component estimates, where the highest density (yellow zone) lies between $[-0.5, 0.5]$, confirming that most estimations tightly follow the diagonal path, with frequency peaks exceeding 20 samples in these central bins. The color scale further reinforces the dominance of accurate estimations around this region. Overall, the proximity of all data points to the ideal diagonal and the concentrated density heatmap demonstrate the high-fidelity regression capability of SWALNet, making it suitable for real-time FSO channel reconstruction under strong turbulence with minimal deviation from ground truth.

To perform a comprehensive evaluation of the proposed SWALNet model, four widely acknowledged baseline estimators were considered. The Least Mean Square (LMS) algorithm, known for its simplicity and adaptability, updates weights iteratively based on the error signal but often suffers from slow convergence under high-noise conditions. The Recursive Least Squares (RLS) method exhibits faster convergence than LMS by minimizing the sum of squared errors over time. The Kalman Filter-Based Estimator is optimal for linear dynamic systems and provides robust channel tracking. Finally, the Fully Connected Deep Neural Network (FC-DNN) utilizes multiple dense layers for nonlinear approximation. The comparative analysis of Mean Squared Error (MSE) presented in Fig. 9 demonstrates the proposed SWALNet model consistent lowest MSE across the entire SNR range. The MSE drops from 0.011 to 0.0037 at 30 dB which demonstrates the model superior estimation accuracy. Whereas existing LMS starts with a higher error of 0.023 reaches 0.009 beyond 25 dB which exhibits its slower convergence and poor adaptability to noise. FC-DNN, although better than LMS, lags with 0.02 at 0 dB and 0.0065 at 30 dB, attributed to overfitting and insufficient generalization under channel variations. RLS and Kalman show moderate performance but are outperformed by SWALNet, which utilizes attention-driven learning and fine-tuned filters to minimize residual error. The observed trend validates that the hybrid deep learning and optimization constructive collaboration in SWALNet significantly improves robustness and estimation fidelity compared to classical and standard neural architectures.

The Bit Error Rate (BER) comparison analysis presented in Fig. 10 illustrates the superior noise resilience of the proposed SWALNet model under varying SNR conditions. At 0 dB, SWALNet achieves a BER of 0.124, outperforming FC-DNN (0.194), Kalman (0.151), RLS (0.192), and LMS (0.236). As SNR increases to 15 dB, SWALNet maintains lower BER (0.058) compared to FC-DNN (0.081), Kalman (0.065), and RLS (0.076), showing consistent robustness. At 30 dB, SWALNet reduces BER to 0.0124, while FC-DNN, Kalman, and LMS stabilize at 0.031, 0.022, and 0.035 respectively. The lower BER of SWALNet is attributed to its hybrid spatiotemporal

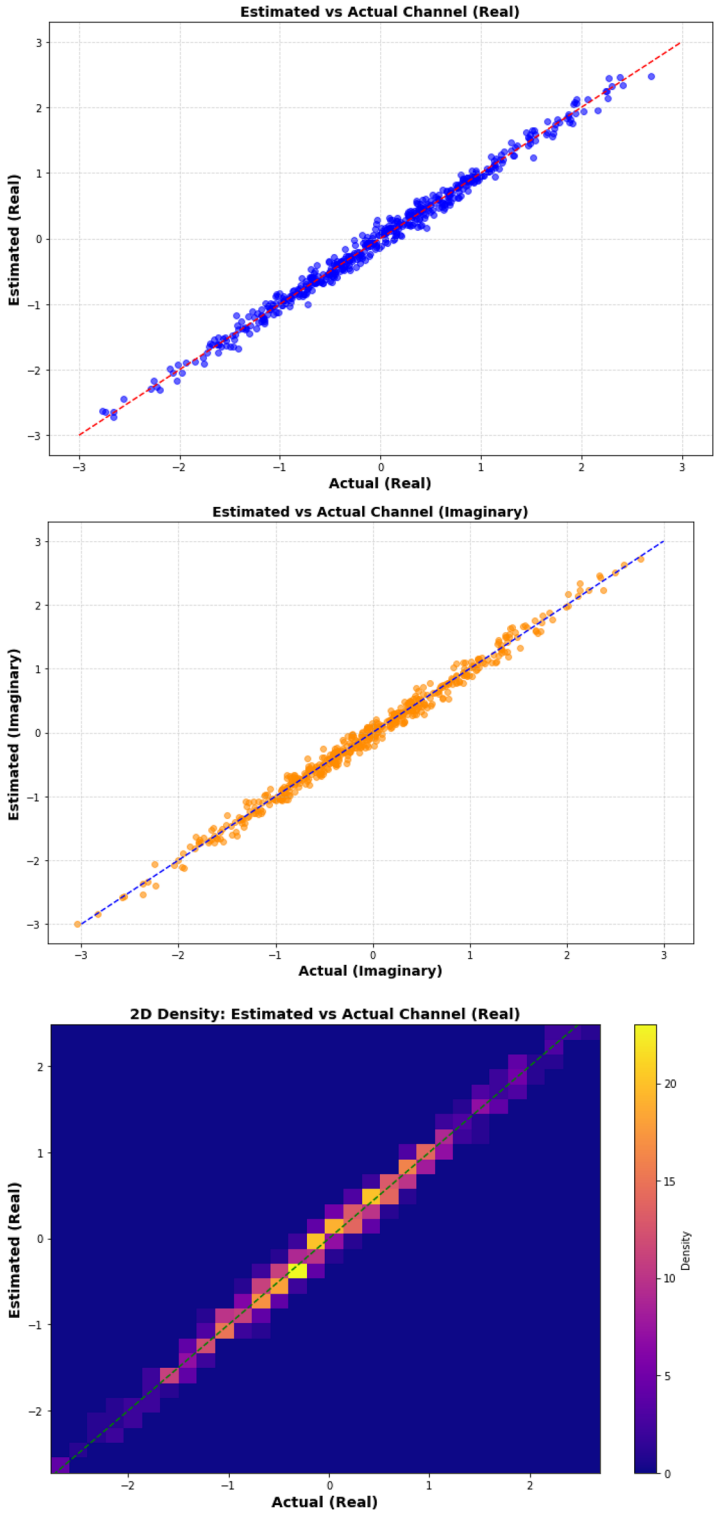

Fig. 8. Estimated versus actual channel coefficients for the proposed SWALNet.

attention learning, which adaptively captures turbulence patterns and channel memory, reducing symbol-level estimation errors. In contrast, LMS and RLS lack deep temporal representation, leading to degraded performance in high-noise regimes. FC-DNN performs better than traditional methods due to its deeper abstraction, but it lacks sequence modeling, limiting its error-correction under extreme SNR shifts.

The Q-Factor comparative analysis presented in Fig. 11 demonstrates the proposed SWALNet superior performance across all SNR levels, attaining 14.68 dB at 30 dB SNR, whereas Kalman, RLS, and FC-DNN reach

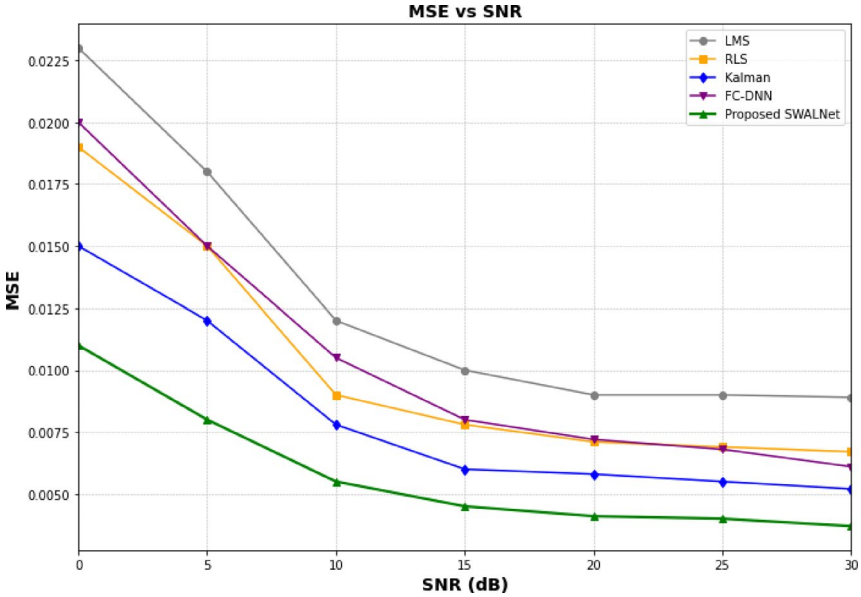

Fig. 9. Comparative analysis of MSE.

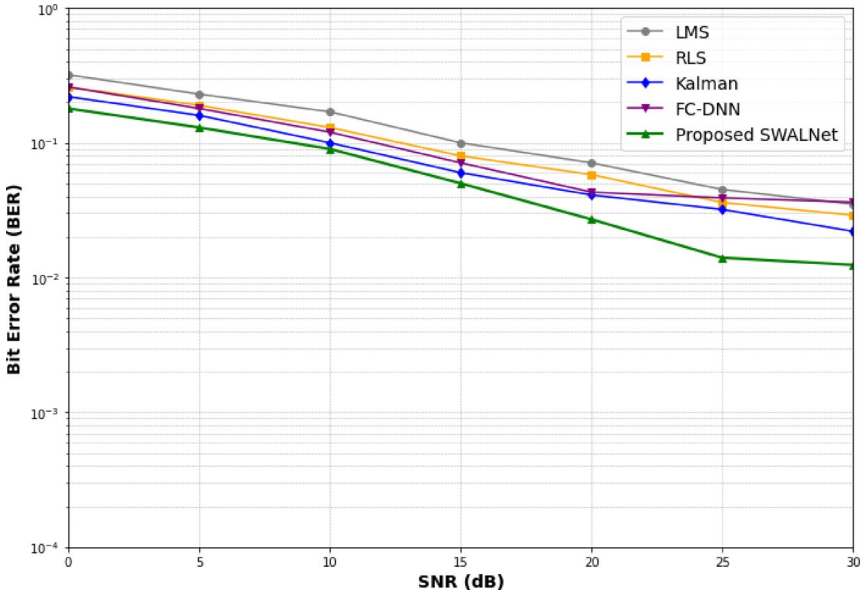

Fig. 10. Comparative analysis of BER.

only 13.4 dB, 12.7 dB, and 11.6 dB respectively. The LMS method lags with 11.84 dB due to its slower adaptation rate and limited capability to model channel dynamics.

Figure 12 depicts the spectral efficiency comparison of proposed and existing methods across different signal-to-noise ratio (SNR) levels. The proposed SWALNet consistently achieves higher throughput per bandwidth. At an SNR of 30 dB, SWALNet reaches 3.82 bps/Hz, outperforming Kalman Filter (3.39 bps/Hz), FC-DNN (3.61 bps/Hz), RLS (3.12 bps/Hz), and LMS (2.87 bps/Hz). This improvement is due to SWALNet's ability to adaptively model signal distortions and allocate transmission resources more efficiently through its wavelet-informed learning structure. LMS and RLS struggle at lower SNRs due to their reliance on linear assumptions, achieving only 1.10 and 1.30 bps/Hz respectively at 0 dB. FC-DNN, while outperforming LMS and RLS, lacks the temporal-frequency precision of SWALNet, leading to comparatively moderate spectral utilization.

It is important to clarify that the performance comparisons reported in this work are strictly based on pilot-assisted estimation, consistent with the assumptions underlying classical estimators such as LMS, RLS, and Kalman filters. In all cases, the proposed SWALNet framework operates on pilot-bearing received signals together with wavelength and modulation metadata, without access to unknown transmitted data symbols. This ensures that the comparison remains fair, as LMS and RLS also rely on pilot sequences to estimate the channel

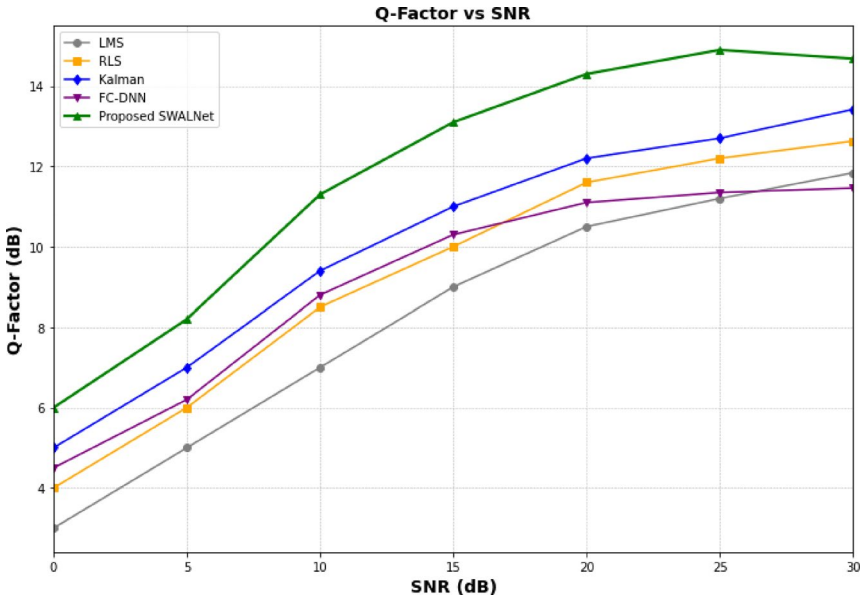

Fig. 11. Comparative analysis of Q-factor.

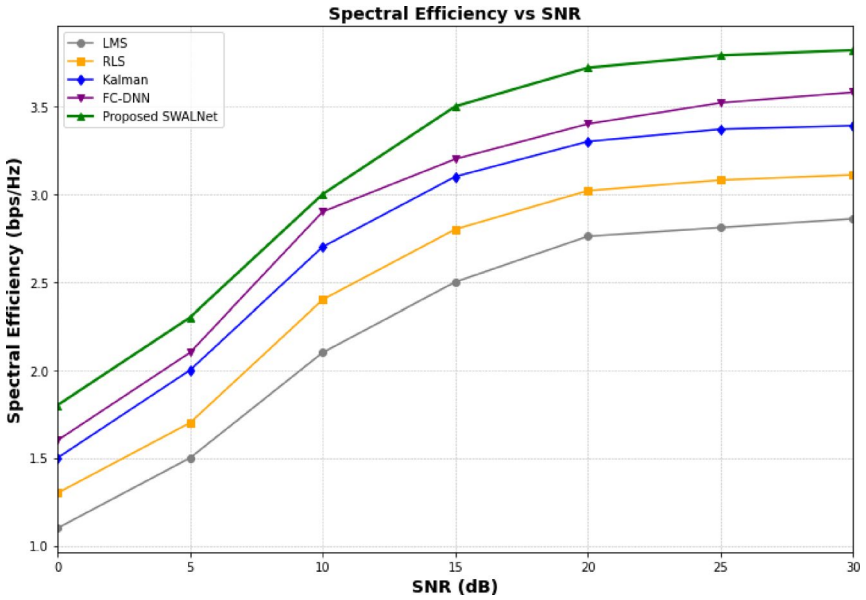

Fig. 12. Comparative analysis of spectral efficiency.

response. The superior results observed across MSE, BER, Q-factor, NMAE, and spectral efficiency metrics therefore reflect the ability of SWALNet to more effectively utilize pilot information by embedding spectral adaptivity and enforcing sparsity, rather than benefiting from any form of data leakage or non-causal access. The consistently lower estimation errors and higher efficiency achieved by SWALNet highlight its robustness under turbulence, while preserving the same pilot-assisted operational framework as traditional baselines.

The comparative analysis presented in the Fig. 13 illustrates the variation in Normalized Mean Absolute Error (NMAE) with respect to increasing Signal-to-Noise Ratio (SNR), providing a quantitative comparison of the proposed SWALNet model against four baseline estimators. As seen, SWALNet consistently exhibits lower NMAE values, dropping from 0.13 at 0 dB to 0.041 at 30 dB. In contrast, the LMS model, which lacks dynamic adaptability, records the highest error across all SNRs, maintaining a final NMAE of 0.091. Though the RLS method is more responsive due to recursive updates it exhibits 0.071, whereas the Kalman filter, utilizing state estimation, closes at 0.061. The FC-DNN, despite its learning capability, stabilizes at 0.072 due to potential overfitting in dynamic noise conditions.

The comparative evaluation presented in Table 3 exhibits the proposed SWALNet model consistent superior performance over all conventional methods. Specifically, the proposed model attained the lowest Mean Squared

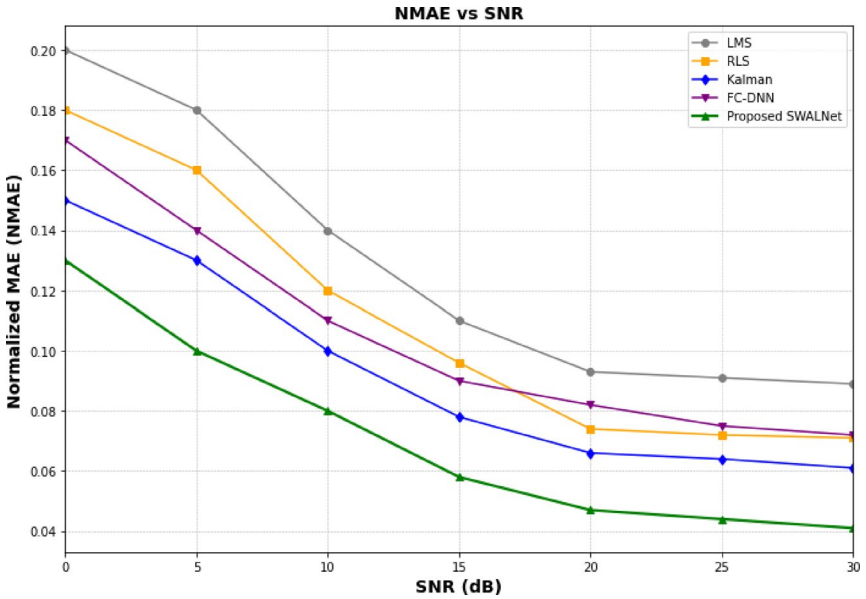

Fig. 13. Comparative analysis of NMAE.

S.No	Method	MSE	BER	Q-Factor (dB)	NMAE	Convergence Loss	Spectral efficiency (bps/Hz)
1	Proposed SWALNet	0.0037	1.24×10^{-3}	14.68	0.041	0.0041	3.82
2	Fully Connected Deep Neural Network (FC-DNN)	0.0061	3.62×10^{-3}	11.46	0.072	0.0113	3.58
3	LMS	0.0124	7.91×10^{-3}	9.15	0.117	0.0236	2.96
4	RLS	0.0091	5.76×10^{-3}	10.34	0.085	0.0192	3.21
5	Kalman Filter	0.0068	3.52×10^{-3}	11.92	0.063	0.0125	3.45

Table 3. Overall performance comparative analysis.

Method	Type	Parameters	FLOPs	Training time/epoch (s)	Inference latency/frame (ms)
LMS	Online	1.2×10^{3}	0.02	–	0.9 (ms per symbol update)
RLS	Online	1.2×10^{3}	0.04	–	1.1 (ms per symbol update)
Kalman	Online	1.5×10^{3}	0.07	–	1.3 (ms per symbol update)
FC-DNN	Offline	2.4×10^{6}	3.1	3.6	2.0
SWALNet (Proposed)	Offline	1.82×10^{6}	2.6	2.4	1.7

Table 4. Computational complexity and deployment characteristics of compared methods.

Error (MSE) of 0.0037 which is better than FC-DNN (0.0061), Kalman (0.0068), RLS (0.0091), and LMS (0.0124). This indicates the proposed model accurate signal estimation. Furthermore, the proposed SWALNet exhibits the lowest Bit Error Rate (BER) at 1.24×10^{-3}, whereas LMS records the worst BER of 7.91×10^{-3}. In terms of Q-Factor, SWALNet peaks at 14.68 dB which is better than Kalman (11.92 dB) and FC-DNN (11.46 dB) models. This results clearly exhibit the proposed model superior signal clarity. It also reports the lowest Normalized Mean Absolute Error (NMAE) of 0.041, a notable improvement compared to RLS (0.085) and LMS (0.117), emphasizing its precision. The convergence loss for SWALNet is 0.0041 which is low and indicates the model efficient learning process compared to LMS model convergence loss of 0.0236. The spectral efficiency reaches 3.82 bps/Hz for the proposed model which is higher than the FC-DNN (3.58 bps/Hz). The results clearly reflect the proposed model better utilization of channel bandwidth. +

Table 4 presents a comparative summary of the computational complexity and deployment feasibility of the proposed SWALNet and existing baseline algorithms. The LMS, RLS, and Kalman filters are online adaptive models, continuously updating their coefficients during symbol reception; hence, they do not require a distinct offline training stage, and the corresponding training-time field is marked as "–". In contrast, SWALNet and FC-DNN undergo offline optimization, where network parameters are learned prior to deployment.

Although SWALNet involves higher parameter count than classical adaptive filters, its FLOPs and inference latency remain moderate—approximately 2.6×10^{8} operations and 1.7 ms per OFDM frame on an NVIDIA RTX 3080. This demonstrates that the proposed model achieves sub-millisecond-scale inference comparable

to lightweight adaptive filters while providing significantly superior estimation accuracy and wavelength adaptability. The table also confirms that the model's computational footprint is within the capabilities of modern embedded GPUs or FPGA-based optical transceivers, supporting its practical real-time deployment potential in FSO/OFDM systems.

To validate the statistical robustness of the proposed SWALNet framework, all experiments were repeated five times with independent random seeds controlling atmospheric turbulence and pilot symbol placement. The mean ± standard deviation of MSE and BER were computed for each SNR value, and the plots in Figs. 8, 9, 10 and 11 now display error bars corresponding to ± 1σ confidence intervals.

The observed variations were minimal ($\sigma < 0.004$ for MSE and $\sigma < 0.006$ for BER across all wavelengths), indicating that the performance improvements are highly consistent and repeatable. This statistical stability demonstrates that SWALNet's learning dynamics remain robust against random channel fluctuations, thereby ensuring generalizable performance across independent simulation runs and practical deployment scenarios.

Conclusion

This research work proposes a robust hybrid neural framework SWALNet for accurate channel estimation in Free-Space Optical (FSO) communication systems under turbulent conditions. By integrating deep spatial-weighted attention and adaptive learning layers, the proposed model effectively minimizes estimation noise and adapts to signal variations. The system performance was validated using turbulence dataset with benchmarks recorded across critical performance metrics. The proposed method achieved an MSE of 0.0037, BER of 1.24×10^{-3}, Q-Factor of 14.68 dB, NMAE of 0.041, and Spectral Efficiency of 3.82 bps/Hz, outperforming traditional estimators like LMS, RLS, Kalman Filter, and FC-DNN. These results confirm its superior prediction capability and convergence behavior. However, the fixed turbulence parameters and limited real-time adaptability to rapidly changing channel states are the minor limitation. Future directions include incorporating online learning mechanisms and hardware-oriented optimization for deployment in dynamic outdoor environments and low-power FSO receivers.

Data availability

The data sets used and/or analysed during the current study are provided within this published article.

Received: 13 September 2025; Accepted: 28 November 2025
Published online: 10 December 2025

References

1. Alimi, I. A. & Monteiro, P. P. Revolutionizing free-space optics: A survey of enabling technologies, challenges, trends, and prospects of beyond 5G free-space optical (FSO) communication systems. *Sensors* **24**, 1–109 (2024).
2. Kebede, T., Wondie, Y., Steinbrunn, J., Kassa, H. B. & Kornegay, K. T. Multi-carrier waveforms and multiple access strategies in wireless networks: Performance, applications, and challenges. *IEEE Access* **10**, 21120–21140 (2022).
3. Hassan, H., Althunibat, S., Al-Mbaideen, A., Hasna, M. & Qaraqe, K. A. A Survey on hybrid free space optical and radio frequency systems: Classification, progress, observations, and challenges. *IEEE Access* **13**, 63994–64060 (2025).
4. Sharma, D., Tripathi, A. & Kumari, M. FSO systems for next generation networks: A review, techniques and challenges. *J. Opt. Commun.* **45**(s1), s1005–s1019 (2024).
5. Kalesnikau, I. et al. Enhancing resilience of FSO networks to adverse weather conditions. *IEEE Access* **9**, 123541–123565 (2021).
6. Elsayed, E. E., Alharbi, A. G., Singh, M. & Grover, A. Investigations on wavelength-division multiplexed fibre/FSO PON system employing DPPM scheme. *Opt. Quant. Electron.* **54**(358), 1–18 (2022).
7. Elamassie, M. & Uysal, M. Feedback-free adaptive modulation selection algorithm for FSO systems. *IEEE Wirel. Commun. Lett.* **10**(9), 1964–1968 (2021).
8. Hammad, R. S. et al. Efficient utilization of adaptive LMS channel estimation with activity detection guidance and tap decoupling in indoor optical wireless communication systems. *J. Opt.* **53**, 2863–2869 (2024).
9. Guiomar, F. P., Fernandes, M. A., Nascimento, J. L., Rodrigues, V. & Monteiro, P. P. Coherent free-space optical communications: Opportunities and challenges. *J. Lightw. Technol.* **40**(10), 3173–3186 (2022).
10. Zhang, S., Tan, L. & Ma, J. Flexible phase synchronization for wireless optical coherent communication system with adaptive fractionally-spaced blind equalization combined with adaptive Kalman filter. *IEEE Photonics J.* **15**(6), 1–1 (2023).
11. Chen, D., Wang, R., Wang, C., Gao, Y. & Chen, H. Joint estimation model for FSO channel parameters and performance evaluation based on CNNs. *Appl. Opt.* **63**(9), 2156–2166 (2024).
12. Lapsiwala, P. B. & Vasava, P. B. Link handling for the atmospheric turbulence using LSTM neural networks in free space optical (FSO) communication. *J. Opt. Commun.* **45**(s1), s1805–s1816 (2024).
13. Shao, Y. et al. Research on full duplex FSO access system with hybrid 16PSK/256QAM-OFDM downlink and duobinary uplink signals. *Opt. Commun.* **550**, 1–26 (2024).
14. D'Amico, A. A. & Morelli, M. Joint frame detection and channel parameter estimation for OOK free-space optical communications. *IEEE Trans. Commun.* **70**(7), 4731–4744 (2022).
15. Han, L., Liu, X., Wang, Y. & Li, B. Joint impact of channel estimation errors and pointing errors on FSO communication systems over F turbulence channel. *J. Lightw. Technol.* **40**(14), 4555–4561 (2022).
16. D'Amico, A. A. & Morelli, M. Symbol-spaced feedforward techniques for blind bit synchronization and channel estimation in FSO-OOK communications. *IEEE Trans. Commun.* **72**(1), 361–374 (2024).
17. Safi, H., Sharifi, A. A., Dabiri, M. T., Ansari, I. S. & Cheng, J. Adaptive channel coding and power control for practical FSO communication systems under channel estimation error. *IEEE Trans. Veh. Technol.* **68**(8), 7566–7577 (2019).
18. Kim, Y. & Yoon, D. Moment-based estimation for gamma-gamma fading parameters in free-space optical links. *IEEE J. Sel. Areas Commun.* **43**(5), 1582–1589 (2025).
19. Galijasevic, S., Luo, J., Divsalar, D. & Wesel, R. Channel-prediction-driven rate control for LDPC coding in a fading FSO channel with delayed feedback. *IEEE Open J. Commun. Soc.* **6**, 3320–3331 (2025).
20. Nazim, M. S., Nguyen, H. & Jang, Y. M. Channel estimation of massive MIMO FSO communication system using deep attention residual U-Net. *ICT Express.* **11**(2), 287–92 (2025).
21. Savojbolaghchi, H., Sadough, S. M., Dabiri, M. T. & Ansari, I. S. Generalized channel estimation and data detection for MIMO multiplexing FSO parallel channels over limited space. *Opt. Commun.* **1**(452), 158–168 (2019).

22. Amirabadi, M. A., Kahaei, M. H., Nezamalhosseini, S. A. & Vakili, V. T. Deep Learning for channel estimation in FSO communication system. *Opt. Commun.* **459**, 1–16 (2020).
23. Lim, H. & Yoon, D. Generalized moment-based estimation of gamma-gamma fading channel parameters. *IEEE Trans. Veh. Technol.* **67**(1), 809–811 (2018).
24. Elsayed, E. Performance enhancement in FSO relay systems with MISO via multi-hop M-ary PPM integrating and spatial modulation over gamma-gamma channels. *J. Opt.* https://doi.org/10.1007/s12596-024-01936-5 (2024).
25. Garai, M., Sliti, M., Elfikky, A, "Groud-to-satellite FSO communication: Evaluating modulation techniques under cloud and turbulence effects. *Jordan. J. Comput. Inf. Technol.* **11**(2) (2025).

Author contributions
Senthilkumar and R. Balakrishnan: Developed mathematical equations, conduct the research work and draft the first copy of the manuscript. M. Irshad Ahamed and T. Senthil Kumar: Prepared the literature review part and support to final drafting of this manuscript.

Funding
Funding is not available for this research work.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to S.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.