POLITECNICO DI TORINO
Repository ISTITUZIONALE

ML-Based Physics-Guided Method for Low-Margin Nonlinear Interference Estimation in Optical Links

Original
ML-Based Physics-Guided Method for Low-Margin Nonlinear Interference Estimation in Optical Links / Ietro, R., Napoli, A., Dipto, I.C., Zeb, S., Bhyri, S.K., Virgillito, E., Galimberti, G.M., Costa, N., Pedro, J., Parisi, G., Curri, V.. - (2026). (26th International Conference on Transparent Optical Networks (ICTON 2026) Prague (Cze) 12-16 July 2026).

Availability:
This version is available at: 11583/3015750 since: 2026-10-20T10:08:51Z

Publisher:
IEEE

Published
DOI:

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
IEEE postprint/Author's Accepted Manuscript

©2026 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

18 November 2026

ML-Based Physics-Guided Method for Low-Margin Nonlinear Interference Estimation in Optical Links

Imran Chowdhury Dipto
Politecnico di Torino, Italy
imran.dipto@polito.it

Rosario Ietro
Politecnico di Torino, Italy
rosario.ietro@polito.it

Emanuele Virgillito
Politecnico di Torino, Italy
emanuele.virgillito@polito.it

Sanwal Zeb
Politecnico di Torino, Italy
sanwal.zeb@polito.it

Sai Kishore Bhyri
Nokia, Bengaluru, India
sai.bhyri@nokia.com

Gabriele Maria Galimberti
Nokia, Vimercate, Italy
gabriele.galimberti@nokia.com

Nelson Costa
Nokia, Lisbon, Portugal
nelson.costa@nokia.com

Giuseppe Parisi
Nokia, Munich, Germany
giuseppe.parisi@nokia.com

João Pedro
Nokia, Carnaxide, Portugal
Instituto de Telecomunicações, IST, Lisboa, Portugal
joao.pedro@nokia.com

Antonio Napoli
Nokia, Munich, Germany
antonio.napoli@nokia.com

Vittorio Curri
Politecnico di Torino, Italy
vittorio.curri@polito.it

***Abstract*—Modern-day Optical Transport Networks run over a heterogeneous fiber infrastructure, with increasingly tight design margins and dynamic channel loading conditions. As a result, the estimation of Nonlinear Interference (NLI) is a central requirement for performance prediction and automation. In dispersion-managed (DM) links, NLI is typically described as the sum of Self-Channel Interference (SCI) and Cross-Channel Interference (XCI), but traditional Gaussian Noise (GN) predictions become inaccurate. This research focuses on the semi-analytical NLI model recently introduced for SCI and XCI in DM systems, where the GN contribution is corrected through two key parameters: the spatial correlation factor C_{ij}, accounting for nonlinear interactions between spans i and j, and the GN conservativity correction ρ_i, correcting the intrinsic overestimation of the GN model. As these terms lack closed-form expressions, we tested three regression Machine Learning models and proposed the PhysicsXGB model for predicting C_{ij} and ρ_i separately for SCI and XCI. The analysis investigated the impact of physics-derived engineered features and span-level heterogeneity on regression accuracy, evaluating all models under realistic DM link configurations. The results showed that the proposed model enabled real-time and spectrally disaggregated NLI estimation with significant accuracy, providing a viable path towards a fast and reliable Digital Twin engine for optical networks.**

***Index Terms*—Machine Learning, GN Modelling, NLI, Decision Tree, Random Forest, XGBoost**

I. Introduction

In optical fibre-based transport networks, coherent-optimized transmission infrastructure, hereafter referred to as Uncompensated Transmission (UT), is gradually replacing older infrastructure, designated as Dispersion Managed (DM), to satisfy the increasing amount of data traffic requests. Despite that, many DM links are still present in metro and subsea networks, whose substitution would require considerable investments. As a result, the deployment, even if sub-optimal, of phase-modulated coherent-detection modulation formats over DM optical links will continue to occur in the short and medium term. The problem that arises here is to find a suitable model to predict system performance. In a real network scenario, it is essential to have a software representation of network infrastructure (Digital Twin) to estimate system impairments and maximize throughput, where the noise introduced by each channel can be predicted independently of the other channels occupying the remaining portion of the spectrum (which, in a network context, is often unknown). This requires a spectrally disaggregated and computationally instantaneous approach.

In our previous research, we developed a tool similar to the semi-analytical model introduced in [1] for Self-Channel Interference (SCI), which was subsequently validated for Cross-Channel Interference (XCI) in [2]. We then introduced the elements of machine learning (ML) in [3], achieving a definitive form in [4]. The objective of these papers was the estimation of Nonlinear Interference (NLI), which is a crucial impairment in optical transmission.

In general, system performance in optical line systems is evaluated through the Generalized Signal-to-Noise Ratio (GSNR), given by

$$\frac{1}{GSNR} = \frac{1}{SNR_{NLI}} + \frac{1}{OSNR}, \tag{1}$$

where $OSNR$ is analytically known, and SNR_{NLI} is our target estimation, reducible to estimating NLI noise since the signal power is known. In a spectrally disaggregated context, NLI is described as the sum of SCI (the channel under test interfering with itself) and XCI_k (channel k interfering with the channel under test).

979-8-3195-4420-9/26/$31.00 © 2026 European Union

The central expression of our model predicts NLI for channel k at span i using

$$\Delta\eta_{XCI,k,i} = \Delta\eta_{XCI,pure,k,i} + 2\sum_{j=1}^{i-1} C_{k,ij}\sqrt{\Delta\eta_{XCI,pure,k,i}\,\Delta\eta_{XCI,pure,k,j}} \tag{2}$$

with

$$\Delta\eta_{XCI,pure,k,i} = \rho_{k,i}\Delta\eta_{GN,k,i}. \tag{3}$$

This formulation also encompasses the SCI case as $\text{XCI}_{k=0}$. Here, the $\Delta\eta_{XCI,pure,k,i}$ terms are computed through the Gaussian Noise (GN) model [5] and the correction factor $\rho_{k,i}$, which compensates for GN model conservativity, while $C_{k,ij}$ captures span-to-span correlation effects specific to DM links. Since $C_{k,ij}$ and $\rho_{k,i}$ do not have closed-form expressions, we estimate them using Machine Learning models for both SCI and XCI, which are physically distinct phenomena.

The rest of this paper is organized as follows. Section II describes the ML regression models considered in this work. In Section III, the descriptions of NLI estimation and the data collection process are provided. We then present a comprehensive analysis of the ML models for NLI estimation in Section IV and conclude the paper and outline future research directions in Section V.

II. Machine Learning Models

ML is the process of teaching computers the functional relationships in data. In this research, we are tackling a regression problem, which is supervised learning. In this type of ML, regression models are used to learn mappings from input features to continuous-valued outputs [6], [7]. For NLI prediction, regression is chosen because analytical models like the Gaussian Noise (GN) model capture the dominant physical behavior but do not consider residual components that rely on complex interactions among fiber parameters, span heterogeneity, and launch-power dynamics. ML-based regressors can capture these network patterns and enhance performance prediction accuracy [8]. Previous studies show that tree-based regression models have strong generalization capabilities and the ability to understand diverse signal-processing and optical communication tasks. They can efficiently model nonlinear interactions, avoid focusing on feature linearity, and naturally support physics-guided hybrid learning when used with analytical baseline features [9], [10]. In this section, we describe the different regression models considered for this research.

A. Types of Regression Models

Regression is a technique that aims to estimate an unknown function $F(X)$ mapping input feature vectors x to a continuous output y. ML models like Linear Regression and Kernel Regression are well established in statistical learning [7]; however, Decision Tree-based methods excel at understanding interactions between system-level parameters, which is central to our use case.

B. Decision Tree

Decision Tree (DT) is a type of supervised learning model that splits feature space by building decision trees based on a mathematical criterion such as the Gini index or mean-squared error [11]. Their hierarchical structure enables capturing nonlinear dependencies between features and NLI behavior.

C. Random Forest

Random Forest (RF) is an improved version of the Decision Tree model. It improves Decision Tree performance by training a collection of trees with random feature selection at each split [12]. This ensures a reduction in variance and enhances robustness to noise and parameter fluctuations, making it beneficial for GN modelling deviations from span-dependent nonlinear accumulation, power evolution, and similar effects.

D. XGBoost

XGBoost (Extreme Gradient Boosting) is an upgraded version of the gradient-boosted decision tree [13]. It iteratively builds trees that correct the residual errors of previous ones and adds second-order gradient information, regularization, and shrinkage. XGBoost delivers strong regression performance across different domains.

E. Proposed Model (PhysicsXGBoost)

We propose PhyXGB, which is a physics-guided extension of the XGBoost model developed for nonlinear interference NLI estimation in arbitrary optical system designs. The core idea is to enrich this model with physically meaningful interaction features that capture the dominant mechanisms governing nonlinear signal propagation. Unlike the baseline XGBoost, which is purely data-driven, our model incorporates optical-fiber physics directly into the training process with engineered features computed from dispersion, attenuation, spectral separation, and nonlinear parameters.

1) Model Structure: Given an input vector of physical parameters:

$$\mathbf{x} = \{L,\, L_{\text{eff}},\, \alpha,\, \gamma,\, D_{\text{acc}},\, \Delta f,\, \dots\}, \tag{4}$$

the model constructs a set of additional physics-based interaction terms, denoted:

$$\mathbf{x}_{\text{phy}} = f_{\text{physics}}(\mathbf{x}), \tag{5}$$

including, for example,

$$\gamma L_{\text{eff}}, \qquad \frac{D_{\text{acc}}}{L}, \qquad |D_{\text{acc}}^{(i)} - D_{\text{acc}}^{(j)}|, \qquad |D_{\text{acc}}^{(i)} - D_{\text{acc}}^{(j)}|(\Delta f)^2, \tag{6}$$

and symmetric span-pair relations such as:

$$\frac{x_i + x_j}{2}, \qquad |x_i - x_j|, \qquad x_i x_j. \tag{7}$$

The final feature input to XGBoost is, therefore,

$$\mathbf{x}_{\text{total}} = \left[\mathbf{x}\,;\, \mathbf{x}_{\text{phy}}\right]. \tag{8}$$

This addition of features enables our model to learn nonlinear dependencies and cross-span interactions that are otherwise difficult to capture using the raw parameters alone.

2) Purpose of Physics-Guided Features: NLI accumulation depends on joint interactions among dispersion, nonlinear coefficient, effective length, spectral separation, and span heterogeneity. These interactions typically appear through multiplicative or ratio-based relations, such as $\gamma^2 P^2 L_{\text{eff}}$, $D_{\text{acc}}\,\Delta f$, and αL, rather than through the raw features themselves from the datasets. Thus, adding these composite quantities into the feature set guides our model with representations that closely reflect the physical laws governing nonlinear propagation. The schematic design of the model is shown in Fig. 1.

F. Hyperparameter Tuning

For this experiment, we found that the baseline XGBoost model performed better than the Decision Tree and Random Forest regressors. Therefore, further improvements to this model can be achieved by means of hyperparameter tuning along with the integration of further physics-based features for the proposed model. Since XGBoost is a gradient-boosted tree-type model, it contains many hyperparameters that impact its regression capability for nonlinear interactions, its ability to avoid overfitting the training data, and its ability to learn heterogeneous optical-link conditions. Such parameters control the number of trees, the maximum depths of those trees, and so on [7]. As we have incorporated additional features into our proposed model, careful selection of these hyperparameters greatly improved its performance compared to the baseline XGBoost model.

We performed a hyperparameter search through the Optuna framework, which is an efficient optimization toolkit designed for automated ML. This uses a Bayesian optimisation method based on the Parzen Estimator, which models the likelihood of high-performing portions in the search space and adaptively selects the ideal parameters [14]. Thus, this method yields significantly higher sample efficiency compared to grid or random search methods [15]. Here, we used it to set an objective to maximize the validation R^2 score over a fixed 80/20 split of each of the training datasets. This resulted in the final model generalizing better when trained on these hyperparameters and tested on the unseen test sets.

We designed the search space to reflect the physics-based GN extensions. The number of estimators was chosen over the range of 300–1500, and the maximum tree depth was varied between 4 and 12. The learning-rate range was 0.01 to 0.3. Subsampling and feature-sampling ratios were set between 0.6 and 1.0 to control the variance and prevent overfitting. Finally, the L1 and L2 regularization parameters (`reg_alpha` and `reg_lambda`) were chosen to control the boosted ensemble, which mitigates complex tree formation. The parameter search setup was carefully designed within these ranges so that the model could capture the large feature diversities introduced by the additional physics features, such as multivariate interactions among dispersion, attenuation, nonlinear coefficient, and spectral separation, into the proposed PhysicsXGB model. For the four datasets (Cij-SCI, Cij-XCI, p-SCI, p-XCI), Optuna found the combination of hyperparameters that maximized the validation performance in terms of R^2; then, the best combinations of parameters were used to train the final PhysicsXGB model tested on the test datasets.

G. ML Evaluation Metrics

We have used four evaluation metrics to assess the regression capabilities of the base models and our proposed model, commonly used in ML literature, including Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2) [6]–[8]. To check model efficiency, we also considered the training time of the models.

1) Mean Squared Error (MSE): MSE is the average squared difference between the predicted NLI $\hat{y}_i$ and the true target values y_i. It is defined as:

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2, \tag{9}$$

where lower values indicate better predictive accuracy. For this research, MSE assesses the average modelling error between predicted and reference NLI contributions.

2) Root Mean Squared Error (RMSE): The RMSE is the square root of the MSE and is defined by:

$$\text{RMSE} = \sqrt{\text{MSE}}. \tag{10}$$

RMSE is a more interpretable measure of the typical prediction error. This metric is chosen for evaluating the magnitude of the difference between predicted and true NLI values.

3) R^2 Score: The coefficient of determination, or R^2 score, measures the amount of variance in the target NLI captured by the ML models. It is given by:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2}, \tag{11}$$

where $\bar{y}$ is the mean of the true target NLI values. Values close to 1 indicate excellent ML predictive performance, while negative values show that the model performs worse than simply predicting the mean. The R^2 metric provides a strong measure of the overall regression quality of the ML models.

III. NLI Estimation

As mentioned in the previous section, the formula in which we are interested to estimate NLI is (2), where the terms we predict through ML are $C_{k,ij}$ and $\rho_{k,i}$. From now on, we will refer to these quantities as C_{ij} and ρ_i, since the k subscript stands for channel index in the model formulation, and we focus on a generic frequency spacing Δf on which we train the models. The quantities of interest are modeled as functions of many physical parameters:

$$\begin{aligned} C_{ij} = f(&D_i, D_j, D_{acc,(1,i-1)}, D_{acc,(1,j)}, D_{acc,(j,i-1)}, \\ &\alpha_i, \alpha_j, L_{s,i}, L_{s,j}, L_{eff,i}, L_{eff,j}, R_s, \Delta f), \end{aligned} \tag{12}$$

$$\rho_i = f(D_i, D_{acc,(1,i-1)}, \alpha_i, L_{s,i}, L_{eff,i}, R_s, \Delta f). \tag{13}$$

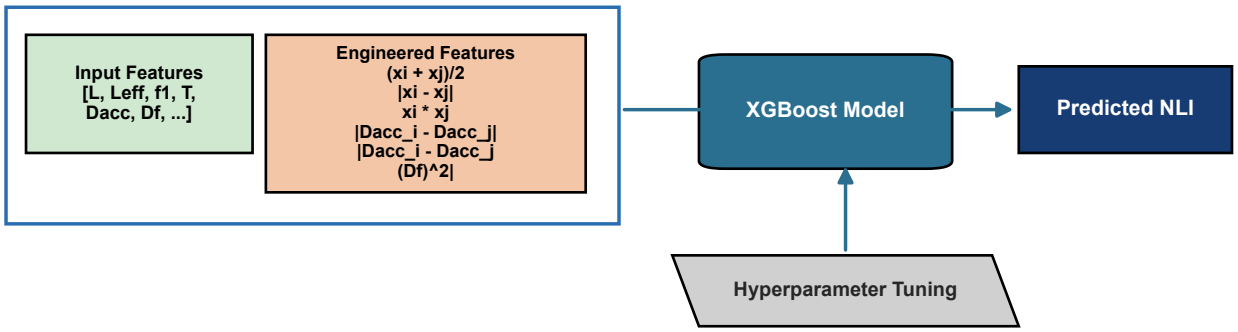


Fig. 1: PhysicsXGB Model Design

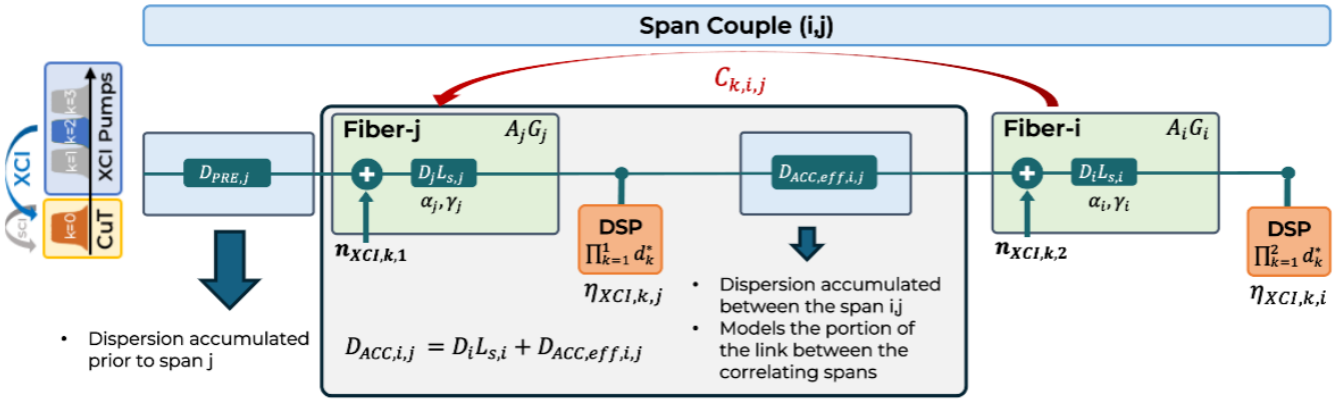


Fig. 2: Training Abstraction Process

These parameters describe, from a physical point of view, each fiber span and the transmitted signal:

- α_i [dB/km] is the distributed fiber loss per km of span i.
- D_i [ps/nm/km] is the distributed dispersion per km at span i.
- $D_{acc,(j,i)}$ [ps/nm] is the total accumulated dispersion between spans j and i.
- $L_{s,i}$ [km] is the length of span i.
- $L_{eff,i}$ [km] is the length after which NLI generated at span i can be considered negligible.
- R_s is the symbol rate of the signal.
- Δf [Hz] is the frequency spacing between the two channels used; $\Delta f = 0$ in the SCI case since we have just one channel.

We use features we know to be relevant from a physical modeling point of view, leading to Physics-Informed Machine Learning models, which can also be exploited to retrieve useful information about distributions of quantities of interest with respect to specific parameters.

In order to train ML models with a scalable and modular methodology able to generalize any couple of spans i and j, we created a DM architecture (observable in Fig. 2) over which we simulate one or two channels (respectively the SCI and XCI cases) through an SSFM (Split-Step Fourier Method) simulator and extract ground-truth C_{ij} and ρ_i to populate our four different training datasets.

With respect to the parameter-value combinations to launch in each simulation, we first select a feasible range of values for each parameter and then apply the Latin Hypercube Sampling (LHS) algorithm [16] to extract a combined distribution spreading these data samples along the parameter hyperspace such that hyperspace coverage with a limited number of samples is optimized.

IV. Results and Discussion

The performance of all the machine learning models across the four datasets is summarized in Fig. 3. The PhysicsXGB model consistently achieves higher R^2 values and lower MSE and RMSE scores compared to the baseline models, particularly for the Cij-SCI and p-SCI datasets. In these cases, the physics-based features effectively capture self-channel interference effects, allowing the model to learn nonlinear physical behaviour that is not directly visible to standard tree-based models. The additional nonlinear, dispersion, and effective-length interaction terms significantly improve its ability to model the dominant physical phenomena. However, the p-XCI dataset remains the most challenging for all models, and

PhysicsXGB shows only marginal improvement over baseline XGBoost.

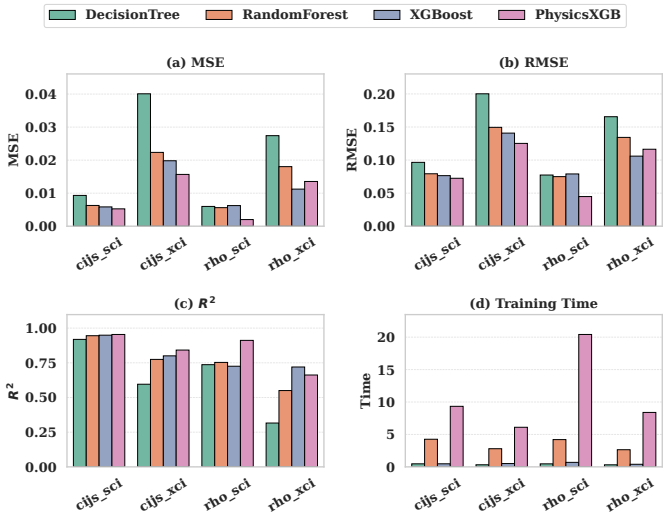


Fig. 3: Model performance across all datasets.

Model efficiency is further analyzed in Fig. 4, which plots training time relative to the R^2 and RMSE scores obtained on each dataset. Lightweight models such as DecisionTree yield low predictive performance but minimal computational cost, whereas RandomForest and XGBoost strike a balance between accuracy and training time. PhysicsXGB requires the longest training time, yet consistently achieves the highest predictive accuracy, demonstrating a clear accuracy–efficiency tradeoff.

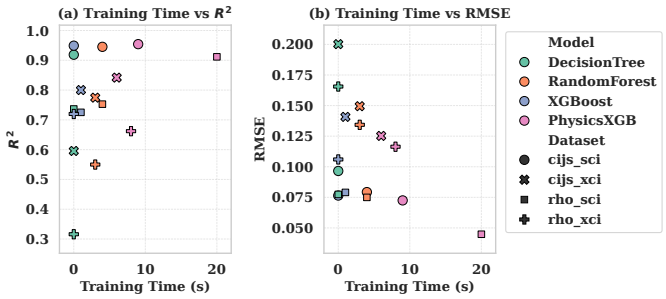


Fig. 4: Training time versus achieved R^2 and RMSE across all datasets.

A. *Performance Analysis of PhysicsXGB*

A deeper assessment of PhysicsXGB is provided in Fig. 5, which compares predicted versus true values and visualizes residual distributions for the four datasets. The SCI datasets exhibit tightly concentrated predictions along the ideal $y = x$ line with symmetric residuals, demonstrating strong generalization capability supported by the engineered physics-informed features. The XCI datasets show greater dispersion, especially at lower true values, and exhibit heavier-tailed residuals. Even for the challenging p-XCI dataset, the model maintains a clear linear trend and reasonable predictive fidelity.

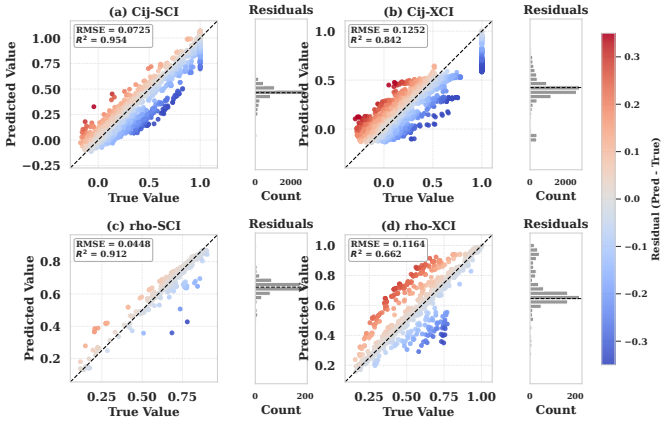


Fig. 5: True vs predicted values and residual histograms for PhysicsXGB.

Physics-based feature engineering has immensely helped the performance gain of the proposed regressor. The analysis of the effects of the top four engineered features is analyzed in Fig. 6 along with model calibration across the datasets. The top row shows the four most impactful engineered features ranked by the XGBoost gain method, which calculates the average reduction in the loss of the model when a feature is used to split the decision trees. Features with higher gain values strongly affect the model's predictive capability, exhibiting their physical relevance in explaining the nonlinear interference. In all the datasets, dispersion-based features, namely "Dispersion product", "Span-dispersion product", and "Dispersion accumulation", consistently showed the highest gain values, followed by terms like "Walk-off" and "Walk-off $\times$ Df^2," and attenuation-based terms like "Loss $\times$ length" and "Dispersion/length," which are key physical mechanisms aiding the model predictions across the SCI and XCI conditions.

The bottom row of the figure shows calibration curves that compare predicted values with average true values. Points falling on the dashed line indicate perfect calibration. On all the datasets, the proposed PhysicsXGB model shows strong calibration performance. This further confirms that the introduced engineered features not only improve predictive performance but also make it reliable in its probabilistic behavior. These results show that the proposed physics-informed features help the model capture the dominant physical properties that control SCI and SCI generation while allowing stable and well-calibrated predictions in various datasets.

V. Conclusion and Future Work

In this paper, we propose PhysicsXGB: a physics-guided ML method for the estimation of Nonlinear Interference in heterogeneous and low-margin optical transmission. We combined analytical knowledge from the Gaussian Noise model and added engineered physics-related interaction features. Our model is able to map dominant nonlinear propagation effects along with span-dependent deviations not accounted for by traditional analytical models. The experimental results across

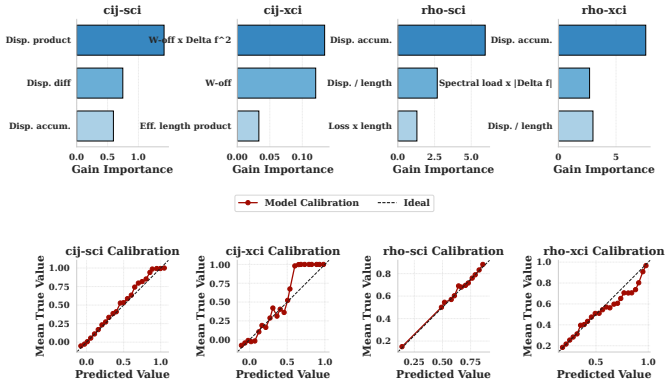


Fig. 6: Physics-based Feature Importance and Calibration Curves of PhysicsXGB

four heterogeneous datasets showed that our model achieved higher R^2 scores and lower MSE and RMSE values compared to the baseline Decision Tree, Random Forest, and XGBoost models. Our model was particularly effective in SCI-dominated conditions, where nonlinear accumulation mechanisms are strongly related to single-channel properties. The proposed model showed a clear linear relationship between predicted and actual NLI values, which shows the model's robustness and generalization capabilities.

Although the performance of the proposed model for NLI estimation was ideal across all four datasets, it was seen that it had a lower R^2 score compared to the baseline XGBoost model for the estimation of NLI for the p-XCI dataset. This is mainly due to a lack of available testing data for that particular dataset, so it would be feasible to train and test the PhysicsXGB model on a larger dataset for better p-XCI estimation and finally test the model on a live cable for real-time integration into network-management applications.

ACKNOWLEDGMENTS

Funded by the European Union's Horizon Europe research and innovation programme under project NESTOR G.A. No. 101119983, in collaboration with Nokia Optical Networks (Carnaxide, Portugal; Munich, Germany) & Nokia Bengaluru, India.

REFERENCES

[1] E. Virgillito, A. Castoldi, A. D'Amico, S. Straullu, A. Bovio, R. Pastorelli, and V. Curri, "Spatially disaggregated model for self-channel interference in mixed fiber optical network segments," *Journal of Optical Communications and Networking*, vol. 15, no. 8, pp. C271–C287, Aug. 2023.

[2] E. Virgillito, R. Ietro, S. C. Ramesh, S. Varughese, A. Napoli, S. K. Bhyri, G. Galimberti, W. Wakim, and V. Curri, "Spatially disaggregated approach to cross-channel interference in dispersion-managed optical links," in *2025 Optical Fiber Communications Conference and Exhibition (OFC)*, 2025, pp. 1–3.

[3] R. Ietro, E. Virgillito, A. Napoli, S. K. Bhyri, G. Galimberti, W. Wakim, and V. Curri, "Ml-assisted gaussian noise modeling of nli accumulation in dispersion managed optical links," in *Advanced Photonics Congress (IPR, Networks, NOMA, SOLITH, SPPCom)*. Optica Publishing Group, 2025, p. NeTu2C.2. [Online]. Available: https://opg.optica.org/abstract.cfm?URI=Networks-2025-NeTu2C.2

[4] E. Virgillito, R. Ietro, A. Napoli, S. K. Bhyri, G. Galimberti, W. Wakim, and V. Curri, "Combining machine learning and the gn model for fast nli prediction in dispersion-managed links," in *51st European Conference and Exhibition on Optical Communication (ECOC 2025)*, Sep. 2025, pp. 1–3.

[5] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "A detailed analytical derivation of the gn model of non-linear interference in coherent optical transmission systems," 09 2012.

[6] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.

[7] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed. Springer, 2009.

[8] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

[9] Y. Pointurier *et al.*, "Machine-learning techniques for quality of transmission estimation in optical networks," *J. Opt. Commun. Netw.*, vol. 13, no. 4, pp. B60–B69, 2021.

[10] F. Musumeci, C. Rottondi, A. Nag, I. Macaluso, D. Zibar, M. Ruffini, and M. Tornatore, "An overview on application of machine learning techniques in optical networks," *IEEE Communications Surveys Tutorials*, vol. 21, no. 4, pp. 1383–1408, 2019.

[11] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and Regression Trees*. Wadsworth, 1984.

[12] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[13] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794.

[14] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A next-generation hyperparameter optimization framework," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2019, pp. 2623–2631.

[15] J. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl, "Algorithms for hyper-parameter optimization," in *Advances in Neural Information Processing Systems*, 2011, pp. 2546–2554.

[16] I. Iordanis, C. Koukouvinos, and I. Silou, "On the efficacy of conditioned and progressive latin hypercube sampling in supervised machine learning," *Applied Numerical Mathematics*, vol. 208, 01 2024.