

ML-Driven QoT Estimation for Adaptive P2MP Optical Network-as-a-Service in 6G X-Haul

Original

ML-Driven QoT Estimation for Adaptive P2MP Optical Network-as-a-Service in 6G X-Haul / Zeb, S., Dipto, I.C., Masood, M.U., Ali, A., Pedro, J., Napoli, A., Curri, V.. - (2026). (2026 IEEE International Conference on Machine Learning for Communication and Networking (ICMLCN) Abu Dhabi (UAE) 8 - 11 June 2026).

Availability:

This version is available at: 11583/3015548 since: 2026-09-15T21:52:07Z

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)

ML-Driven QoT Estimation for Adaptive P2MP Optical Network-as-a-Service in 6G X-Haul

Sanwal Zeb
Dept. of Electronics and Telecom.
Politecnico di Torino, Italy
sanwal.zeb@polito.it

Imran Chowdhury Dipto
Dept. of Electronics and Telecom.
Politecnico di Torino, Italy
imran.dipto@polito.it

Muhammad Umar Masood
Dept. of Electronics and Telecom.
Politecnico di Torino, Italy
muhammad.masood@polito.it

Ahtisham Ali
Consorzio Top-IX
Politecnico di Torino, Italy
ahtisham.ali@top-ix.org

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

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

Vittorio Curri
Dept. of Electronics and Telecom.
Politecnico di Torino, Italy
vittorio.curri@polito.it

Abstract—Point-to-Multipoint (P2MP) coherent transmission enables efficient converged metro-access networks for future B5G/6G services, but dynamic routing, Distributed Units (DUs) placement, and spectrum reconfiguration introduce physical-layer impairments that make conventional analytical Quality-of-Transport (QoT) estimation inaccurate and computationally expensive. This work proposes a scalable and topology-agnostic QoT prediction framework that formulates BER estimation as a regression problem using physical-layer and network related features. Models are trained on a small topology and directly applied to larger unseen networks to evaluate cross-topology generalization under limited training data and noisy telemetry, reflecting realistic operational conditions. Results show that lightweight boosting-based regression, particularly XGBoost, achieves near perfect BER prediction with sub-millisecond inference, enabling real-time feasibility checks without executing full physical-layer simulations. These findings demonstrate that ML-driven QoT regression can replace analytical evaluation and serve as a practical enabler for intelligent and adaptive Optical Network-as-a-Service (ONaaS) for future 6G transport.

Index Terms—Point-to-multipoint, ML-regression, Digital Subcarrier Multiplexing, 6G

I. INTRODUCTION

The rapid growth of AI-driven applications and data-intensive computing is dramatically increasing bandwidth demand while tightening end-to-end latency requirements across modern communication infrastructures. Beyond-5G (B5G) and 6G impose strict Service-Level-Agreement (SLA) constraints on x-haul segments, including fronthaul, midhaul, and backhaul [1]. These requirements place significant pressure on the transport network, which must evolve from static point-to-point provisioning toward flexible, resource-efficient, and dynamically reconfigurable architectures [2].

Converged metro-access networks are emerging as a key enabler to meet these demands by providing high-capacity optical transport that aggregates heterogeneous traffic from disaggregated Radio Access Network (RAN) deployments. In this context, coherent transmission technologies such as Digital Subcarrier Multiplexing (DSCM)-enabled Point-to-Multipoint (P2MP) allow capacity sharing across multiple endpoints, improving spectral efficiency and reducing transceiver counts [3].

However, this flexibility also increases operational complexity. Dynamic routing, variable traffic loads, Distributed Unit (DU) placement, and frequent resource reconfiguration introduce route-dependent physical-layer impairments that lead to significant Quality-of-Transport (QoT) variability across the network.

Frequent routing updates, traffic fluctuations, dynamic DU placement, and spectrum reallocation introduce path-dependent and nonlinear physical-layer impairments that are difficult to capture with conventional analytical QoT models. Such models rely on fixed assumptions and simplified propagation rules, which do not accurately represent the behavior of coherent systems under reconfiguration. As a result, QoT estimates may deviate from actual performance, leading to inefficient resource provisioning and SLA violations. Fast and reliable QoT prediction that adapts to changing network conditions is therefore essential for practical Optical Network-as-a-Service (Optical Network-as-a-Service (ONaaS)) operation.

Machine learning (ML) has recently emerged as a promising alternative for data-driven QoT inference. Regression models can learn complex impairment accumulation directly from monitoring data, enabling accurate prediction of Bit-Error-Rate (BER) and other performance indicators even under noisy or incomplete telemetry [4]–[6]. However, existing ML-based approaches remain topology-specific, require re-training for each network configuration, or assume ideal monitoring conditions, which limits their practical applicability in dynamic metro-access networks.

In this work, we address these limitations by formulating QoT estimation as a regression problem and developing a scalable, topology-agnostic prediction framework for DSCM-enabled P2MP converged metro-access networks. Using synthetic datasets generated from realistic physical-layer simulations, models are trained on a constrained small-scale topology and subsequently evaluated on larger and more complex configurations to assess cross-topology generalization. Furthermore, we explicitly investigate robustness under limited

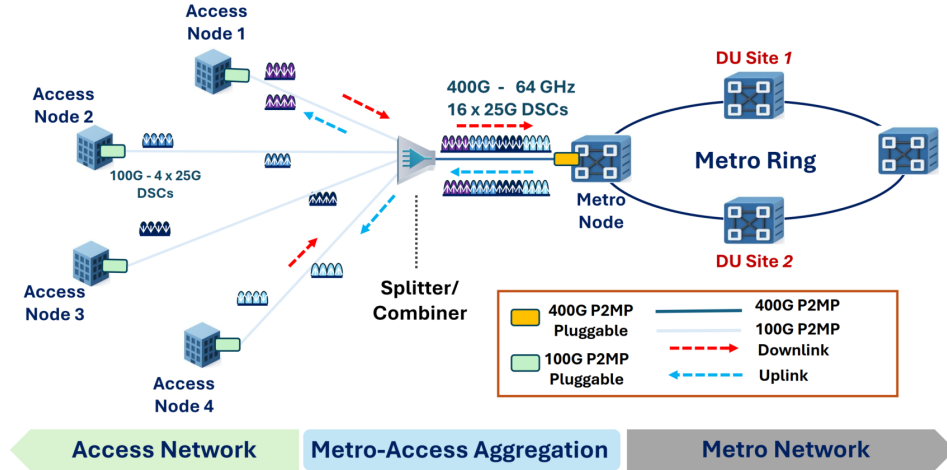


Fig. 1: P2MP converged metro-access network using DSCM pluggable transceiver

training data and noisy features to emulate practical monitoring constraints encountered in operational networks. Overall, this work demonstrates that ML-driven regression can replace analytical and resource exhaustive evaluations and serve as a practical enabler for intelligent, adaptive, and scalable metro-access convergence within future ONaaS architectures.

II. P2MP CONVERGED METRO-ACCESS NETWORK

The evolution toward B5G and 6G has driven a fundamental transformation of the RAN through disaggregation, virtualization, and open interfaces. In this architecture, processing is split into Central Unit (CU), DU, and Radio Unit (RU), interconnected through high-capacity fronthaul (RU-DU) and midhaul (DU-CU) segments with stringent latency, jitter, and capacity requirements [7]–[9]. As a result, the transport network must be built on a transparent fiber infrastructure that provides the high capacity, low attenuation, and low delay necessary to meet strict fronthaul and midhaul performance constraints.

Metro networks play a central role in enabling disaggregated RAN architectures, as they provide the fiber-based backbone interconnecting distributed edge computing nodes with centralized cloud resources. Their ability to support high-bandwidth optical transport, precise timing distribution, and logical isolation through slicing and SDN control is essential for carrying heterogeneous fronthaul, midhaul, and backhaul flows [10]. For B5G/6G, the adoption of functional splits causes fronthaul bandwidth to scale with radio bandwidth, MIMO order, and sector configuration, while midhaul load depends on DU placement in metro network and traffic aggregation strategies [11]. Under these conditions, conventional point-to-point fiber provisioning becomes economically and operationally impractical, since each DU or access node would require a dedicated wavelength. This motivates the transition toward shared, dynamically managed optical transport, capable of efficiently aggregating heterogeneous and time-varying RAN traffic across the metro-access networks [12].

P2MP coherent transmission based on DSCM provides the foundation for the converged metro-access architecture

shown in Fig. 1. In DSCM-enabled P2MP architecture, a single high capacity hub transceiver digitally synthesizes a set of tightly packed Nyquist Digital Subcarriers (DSCs), each modulated independently, and broadcasts DSCs towards downstream through a simple power splitter [3]. Leaf nodes employ narrowband filtering to extract only their assigned DSCs. Upstream transmission reuses the same SCs in the reverse direction, where each leaf remodulates its allocated subcarriers and sends them back toward the hub [13], [14].

As illustrated in Fig. 1, the converged metro-access architecture is structured around a 400G DSCM-enabled P2MP coherent hub transceiver deployed at the metro node, which digitally generates 16×25 Gb/s Nyquist DSCs, each occupying a 4GHz spectral width modulated with 16QAM (dual polarization). A single 400G P2MP coherent Transceiver (TRX) broadcasts the generated DSCs towards the leaf or access nodes, where each 100G P2MP pluggable terminates only its assigned DSCs. In the uplink, the same DSCs are modulated towards the hub or metro node for selective subcarrier coherent detection. The aggregated traffic is then forwarded onto the metro ring, interconnecting DU sites [15]. By replacing multiple Point-to-Point (P2P) wavelengths with a single shared P2MP, it enhances spectral efficiency, reduces coherent transceiver counts, removes the need for intermediate regeneration, and provides a scalable and cost-effective optical transport infrastructure for x-haul in dense B5G/6G deployments.

III. METHODOLOGY

This section describes the methodology adopted to evaluate the proposed QoT estimation framework. First, realistic converged metro-access topologies are simulated to generate datasets. Then, the resulting QoT metrics are used to train and assess ML-regression models for BER estimation under varying network scales and operational conditions.

A. Network Simulation and Dataset Generation

This study considers two converged metro-access network topologies, denoted as N_1 and N_2 , to emulate realistic B5G/6G optical transport deployments with different scales and routing

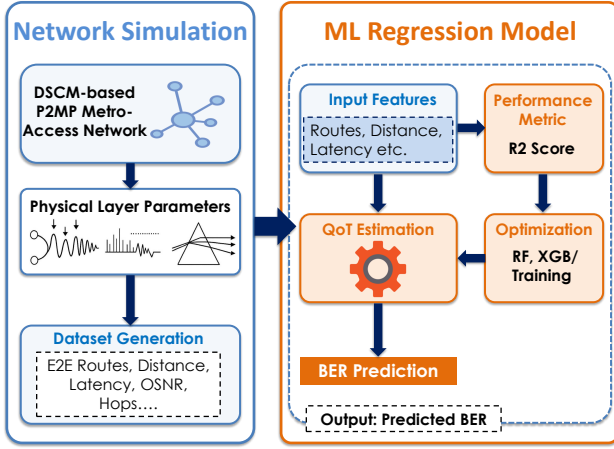


Fig. 2: Simulation framework

diversity. N_1 represents a moderate-size network with fewer nodes and limited path diversity, while topology N_2 models a larger and denser metro-access infrastructure with increased connectivity and a substantially higher number of feasible end-to-end routes. The main structural characteristics of both scenarios are summarized in Table I.

Fig. 2 shows simulation based framework to emulate DSCM-enabled coherent P2MP transmission. For each feasible route between an access node and a DU site, the simulator accumulates fiber attenuation, ROADM and passive-device insertion losses, amplifier noise, and transceiver impairments. Based on these calculations, several QoT metrics are computed for each end-to-end routes.

TABLE I: Characteristics of P2MP Converged Metro-Access Topologies

Network attributes	Topology N_1	Topology N_2
Total nodes	37	63
Metro nodes (M)	12	18
Access nodes (A)	25	45
Max DUs	5	5
Avg. metro degree	3.42	4.06
Avg. link length (km)	6.9	7.31
Max link length (km)	10	12
Access design	4A/splitter	4A/splitter
Reliability	Moderate	High
Scalability	Moderate	High
Dominant risk	fault tol.	latency/QoT
Routes (1DU / 5DU)	309 / 1544	1089 / 5471

** Bold values indicate higher capacity and scalability in N_2 .

Each route therefore corresponds to one labeled datapoint, where the measured physical-layer and topology-related metrics form the input features and the resulting BER actually shows transmission quality. The QoT parameters extracted for each route are summarized in Table II.

Collecting all feasible routes across both topologies generate datasets containing thousands of samples depending on network size.

TABLE II: QoT metrics

Parameter	Description
boundaries	source–destination pair
path	route between nodes
distance	total length (km)
gsnr	generalized SNR (dB)
rop	received power (dBm)
snr	signal-to-noise ratio (dB)
ber	bit-error-rate
latency	propagation delay (ms)
path_id	route identifier
last_link_km	final hop length (km)
boundary_id	boundary identifier

B. ML Regression Framework for QoT Estimation

QoT estimation is formulated as a regression problem. For each route, a feature vector describing topology and end-to-end route conditions is constructed as

$$\mathbf{x} = [d, gsnr, rop, snr, latency, hops, \dots] \quad (1)$$

where d denotes the total path distance and the remaining parameters correspond to the QoT metrics listed in Table II. The objective is to learn a nonlinear mapping

$$f(\mathbf{x}) \rightarrow \hat{y}, \quad (2)$$

where $\hat{y}$ represents the predicted BER.

Once trained, the learned model enables fast QoT inference for unseen routes without executing computationally exhaustive physical-layer simulations. This capability supports real-time feasibility checks and dynamic provisioning decisions within an Optical Network-as-a-Service (ONaaS) control plane. BER is selected as the primary prediction target since it directly determines transmission feasibility, while metrics such as latency are treated as input features.

To evaluate scalability and topology transferability, models are trained on topology N_1 under the 1-DU configuration and tested on topology N_2 under the 5-DU configuration. This asymmetric train–test setup reflects practical scenarios where models trained offline must generalize to expanded or reconfigured networks. To emulate operational monitoring constraints, we vary the fraction of available training data and inject Gaussian noise into the testing features, assessing robustness to limited telemetry and measurement uncertainty.

Four regression algorithms are considered: Decision Tree (DT), Random Forest (RF), Gradient Boosting (GB), and XGBoost. These models are selected for their ability to capture nonlinear impairment relationships while maintaining low computational complexity suitable for real-time inference. Hyperparameter tuning is performed across multiple trials to ensure fair comparison and stable generalization.

Performance is evaluated using the coefficient of determination (R^2), which quantifies how well the predicted BER matches the measured values. Higher R^2 indicates more accurate QoT prediction.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (3)$$

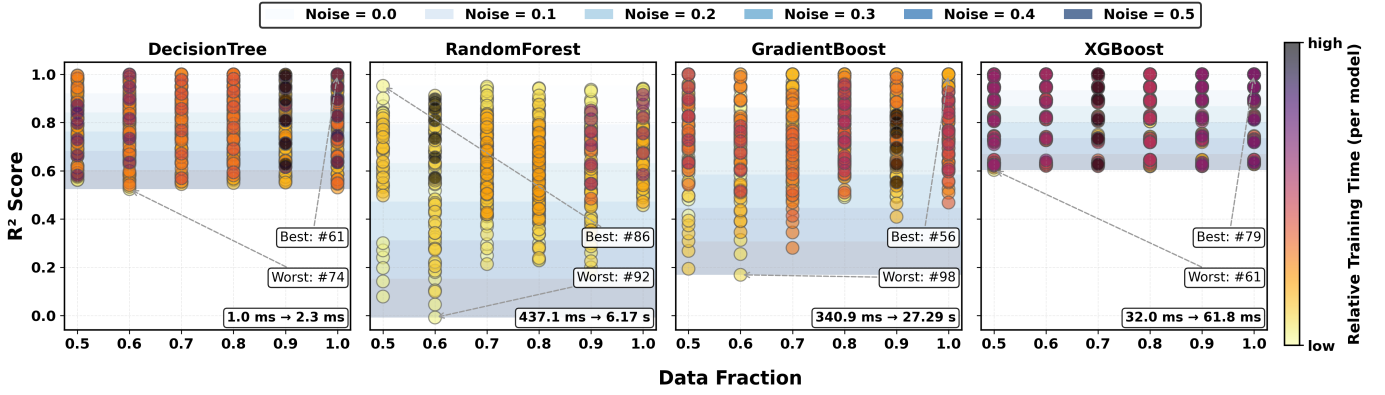


Fig. 3: R^2 score across varying data fractions and noise levels under hyperparameter tuning.

In Eq. (3), y_i denotes the actual measured BER for the i th data-point, $\hat{y}_i$ is the predicted BER, and $\bar{y}$ is the mean of all observed BER values. The numerator represents the model’s residual error, while the denominator captures the total variance in the dataset.

IV. PERFORMANCE ANALYSIS

Fig. 3 compares the performance of four regression models (DT, RF, GB, and XGBoost) under varying levels of data availability and input noise. The x-axis denotes the data fraction used for training, spanning 0.5 to 1.0 (corresponding to 50% to 100% of the available training dataset). For each model and for every combination of data fraction and noise level, 100 trials were executed, each using a distinct hyperparameter configuration sampled from the model-specific search space. This extensive sampling enables a rigorous assessment of both model robustness and hyperparameter sensitivity.

Noise conditions are visualized through horizontal shaded bands, where lighter tones represent lower noise and progressively darker shades correspond to higher noise levels. The bottom annotation in each plot reports the absolute training time range for that model, from the fastest to the slowest run (in seconds or milliseconds) capturing the computational variability introduced by different hyperparameter tuning. Each circle in the scatter plot corresponds to a single trial, and its color encodes the relative training time (normalized within each model), providing a clear view of computational cost distribution while keeping inter-model comparisons interpretable.

For each model, the trial with the highest R^2 score and the trial with the lowest R^2 score across all 100 hyperparameter samples are highlighted and connected with dashed arrows. These annotations illustrate the spread between the best and worst achievable outcomes, reflecting how strongly model performance can vary with respect to hyperparameter selection and noise conditions. Overall, the figure illustrates the interplay between predictive accuracy, noise robustness, data availability, and computational efficiency. GB and XGBoost maintain strong and consistent performance across noise levels, while RF exhibits significant variability in training cost. DT is computationally lightweight but display reduced resilience under higher noise levels.

Fig. 4 provides a detailed hyperparameter sensitivity analysis for each model under clean ($Noise = 0$) and noisy ($Noise = 0.5$) conditions, revealing how each model’s R^2 score evolves as its functional complexity is adjusted. Each subplot isolates a single hyperparameter such as `max_depth`, `min_samples_leaf`, `n_estimators`, `learning_rate`, or `subsampling_ratio`, and reports the resulting performance across the full search space that was explored through 100 hyperparameter trials. The blue and orange curves highlight a consistent pattern across all models: increasing functional complexity through deeper trees, smaller leaves, larger ensembles, or more aggressive learning rates improves performance when noise is absent, but the same settings cause performance degradation once noise is introduced because variance increases and regularization becomes insufficient.

DT displays the sharpest fluctuations, where small changes in depth or leaf-size constraints produce abrupt shifts in R^2 , particularly at ($Noise = 0.5$). This indicates high sensitivity to hyperparameters and limited robustness against noisy inputs. RF reduces some of this instability due to ensembling, although its curves still deteriorate noticeably when trees become too deep or leaf sizes become too small, since correlated tree structures can still capture noise when randomness in the feature space is inadequate. GB exhibits smoother curvature across hyperparameter ranges, and its performance is strongly determined by the interaction between learning rate and number of estimators. Small learning rates require sufficiently large ensembles to avoid underfitting, while large learning rates improve convergence speed but increase the risk of fitting noise, especially in the high-noise setting. XGBoost provides the most stable behavior, maintaining consistently high R^2 under ($Noise = 0$) across a wide hyperparameter range because of its second-order optimization, shrinkage, column sampling, and explicit tree-level regularization. When noise is introduced, XGBoost’s performance declines more gradually compared to the other models, and even extreme hyperparameter choices avoid the abrupt performance drops seen in single-tree or bagging-based methods.

Overall, Fig. 4 shows that boosting model exhibit superior robustness to hyperparameter variation and noise, whereas tree-based models require careful tuning to avoid sharp drops

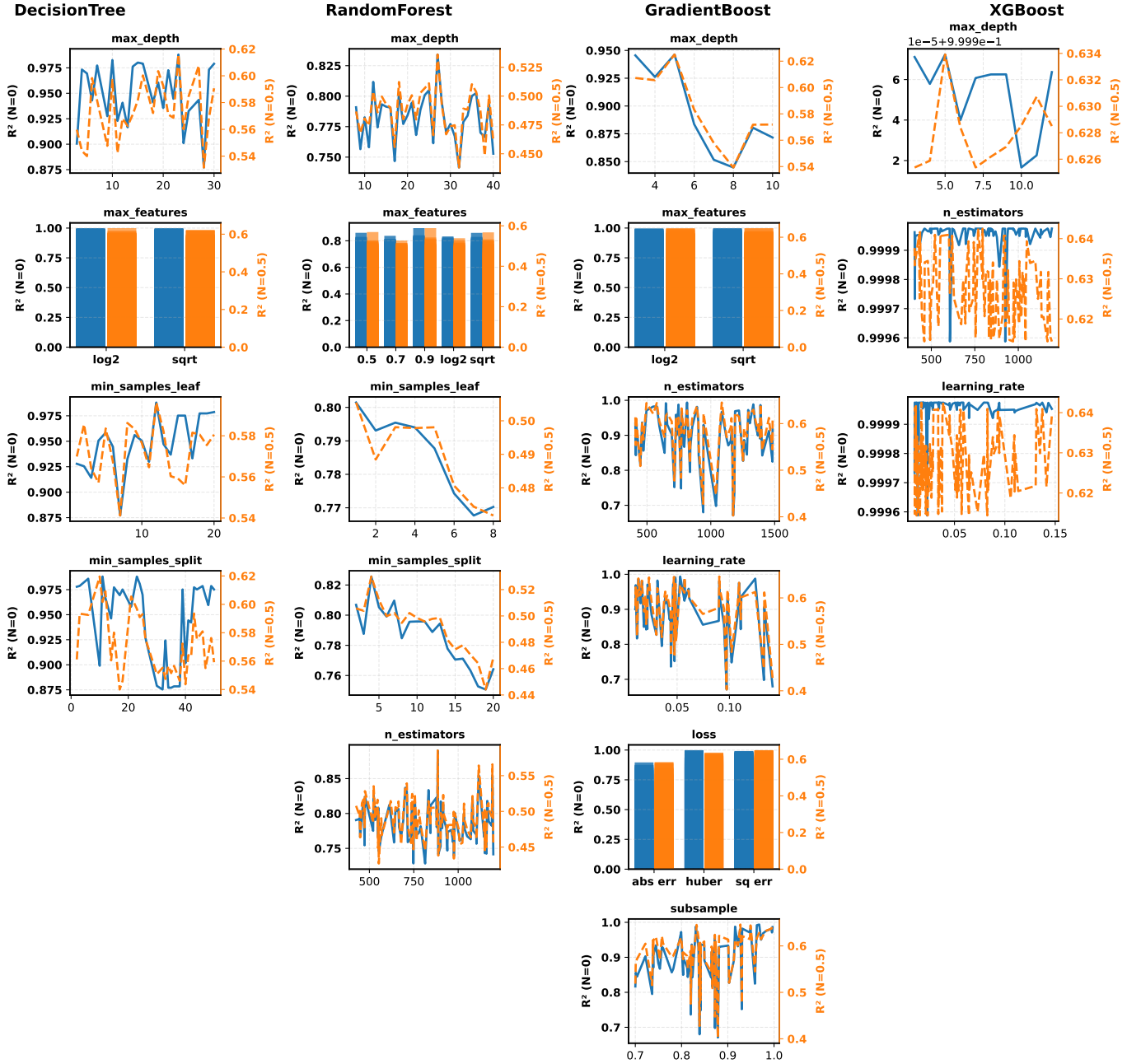


Fig. 4: Impact of hyperparameters sensitivity on R^2 score under (data fraction = 1) and noise levels $N = (0, 0.5)$.

in performance. The combined visualization of low and high noise conditions clearly exposes which hyperparameter range supports stable R^2 performance in noisy conditions.

Fig. 5a illustrates the variation in BER prediction accuracy across 100 hyperparameter trials for four models: DT, RF, GB, and XGBoost at noise level ($Noise = 0$). Each trend line reports prediction error, defined as the difference between the true BER and the predicted BER for each trial per model. The y-axis is plotted on a logarithmic scale in order to highlight differences spanning several orders of magnitude. It shows substantial instability in both DT and GB, where trial-wise errors fluctuate sharply. These oscillations indicate a high degree of sensitivity to hyperparameter choices, with many con-

figurations producing significant deviations from actual BER. RF exhibits consistently elevated errors on the order of 10^{-3} , suggesting that although its predictions are stable, they remain systematically biased across trials. In contrast, XGBoost demonstrates the most reliable behavior, with prediction errors consistently near 10^{-5} and showing minimal variation across all 100 trials. This stability reflects the combined influence of shrinkage, regularization, and second-order optimization, which collectively suppress prediction variance and prevent the extreme deviations observed in the other models.

Fig. 5b compares the training and inference times of the best performing hyperparameter configurations at ($Noise = 0$). For each model, the training time in (s) is shown on left y-axis,

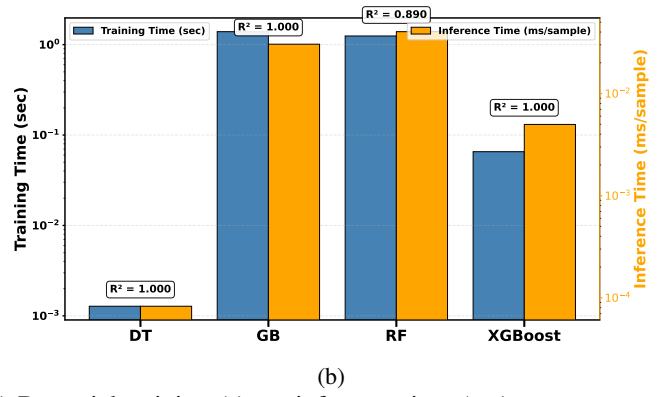
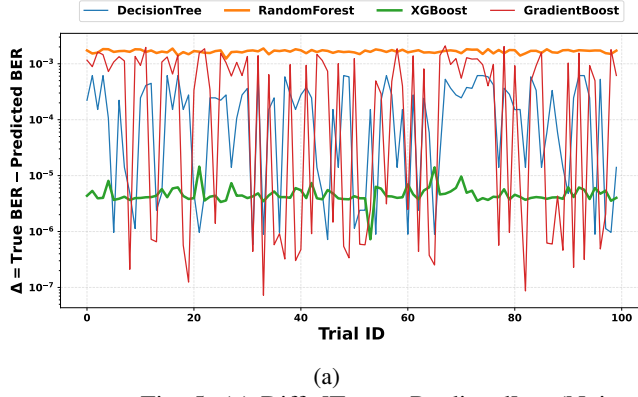


Fig. 5: (a) Diff. [True – Predicted] at (Noise = 0), (b) Best-trial training (s) vs. inference time (ms).

while the inference time per sample in (ms) is depicted on the right y-axis. Annotated labels specify the corresponding trial identifier and the achieved R^2 score for the selected configuration. DT achieves perfect prediction accuracy with extremely low computational demand, exhibiting training times near one millisecond and inference times on the order of 10^{-3} ms per sample. GB also attains $R^2 \approx 1.0$, although at substantially higher computational cost, with training times exceeding one second. RF provides a balanced but less favorable profile: despite moderately fast training, its best configuration reaches only $R^2 = 0.89$, and its inference time is the longest among all models due to the need to evaluate a large ensemble of trees. XGBoost attains perfect accuracy while maintaining relatively modest computational cost, with training times well below one second and inference times close to 10^{-4} ms per sample.

Overall, these results show that XGBoost provides the most desirable balance between accuracy and computational efficiency. DT is the fastest model but exhibits lower robustness across trials, GB is accurate but computationally heavy, and RF is both slower and less accurate. This comparison highlights why XGBoost is the preferred choice for QoT estimation demanding both higher precision and reliability.

V. CONCLUSION AND FUTURE WORK

This paper presented a scalable and topology-agnostic QoT estimation framework for DSCM-enabled P2MP converged metro-access networks. By formulating QoT prediction as a regression task using physical layer and topological features, the proposed approach enables fast BER inference without executing computationally expensive analytical simulations. A key contribution is the demonstration of cross-topology generalization, where models trained on a small network reliably transfer to larger and more complex deployments, even under limited training data and noisy measurements. Experimental results show that boosting-based regression, particularly XGBoost, provides high prediction accuracy with sub-millisecond inference, enabling real-time QoT-aware routing and provisioning for ONaaS operation. Overall, the proposed method offers a practical and deployable alternative to conventional QoT evaluation, improving scalability and operational efficiency for metro-access convergence.

Future work will focus on integrating the framework with dynamic traffic-aware resource allocation and validating the approach using real network telemetry.

ACKNOWLEDGMENTS

Funded by the European Union’s Horizon Europe research and innovation programme under Marie Skłodowska-Curie project NESTOR (G.A. No. 101119983), EWOC (G.A. No. 101073265) and (ALLEGRO G.A. No. 101092766).

REFERENCES

- [1] C. Papapavlou *et al.*, “Forthcoming optical x-haul infrastructure supporting 6G mobile network requirements,” *Journal of Optical Communications and Networking*, 2025.
- [2] R. Winkelmann *et al.*, “The Future Role of Transport Networks in 6G,” Nokia White Paper, 2023, accessed: 2025-11-29.
- [3] D. Welch *et al.*, “Point-to-multipoint optical networks using coherent digital subcarriers,” *Journal of Lightwave Technology*, 2021.
- [4] Y. Pointurier, “Machine learning techniques for quality of transmission estimation in optical networks,” *Journal of Optical Communications and Networking*, 2021.
- [5] C. Natalino *et al.*, “AI/ML-as-a-Service for optical network automation: use cases and challenges,” *Journal of Optical Communications and Networking*, 2024.
- [6] S. Allogba *et al.*, “Machine-learning-based lightpath QoT estimation and forecasting,” *Journal of Lightwave Technology*, 2022.
- [7] M.-E. Kim *et al.*, “A Novel Radio Access Network Converged 6G Mobile Core Network Architecture,” *IEEE Communications Magazine*, 2025.
- [8] A. Ali *et al.*, “Benefits of Selective Hollow-Core Fiber Deployment for 6G X-Haul in Converged Metro-Access Networks,” in *2026 International Conference on Optical Network Design and Modeling (ONDM)*, 2026, pp. 1–6.
- [9] A. Ali *et al.*, “Sustainable 6G Optical Any-Haul with Coherent Pluggables in Converged Metro-Access,” in *2026 IEEE 11th International Conference on Network Softwarization (NetSoft)*, 2026, pp. 1–6.
- [10] E. Kosmatos *et al.*, “Real-time orchestration of QoS-aware end-to-end slices across a converged Metro and Access network exploiting burst-mode technology,” *Journal of Optical Communications and Networking*, 2022.
- [11] S. González *et al.*, “5G-Crosshaul: An SDN/NFV control and data plane architecture for the 5G integrated Fronthaul/Backhaul,” *Transactions on Emerging Telecommunications Technologies*, 2016.
- [12] C. Castro *et al.*, “Point-to-multipoint coherent transceivers for next-generation mobile transport,” in *2023 23rd International Conference on Transparent Optical Networks (ICTON)*. IEEE, 2023.
- [13] D. Welch *et al.*, “Digital subcarrier multiplexing: Enabling software-configurable optical networks,” *Journal of Lightwave Technology*, 2023.
- [14] J. Pedro *et al.*, “Extended network applications of coherent pluggable transceivers,” *Journal of Optical Communications and Networking*, 2025.
- [15] S. Zeb *et al.*, “AI-driven converged metro-access optical network-as-a-service with point-to-multipoint coherent optics for 6G X-Hauling,” *Computer Networks*, vol. 284, p. 112338, 2026.