Reliable QoT Estimation for Imbalanced Data from Deployed Optical Networks Using MDN-ResKAN

Shaowen Xiang
State Key Laboratory of Photonics and Communications, School of Information Science and Electronic Engineering, Shanghai Jiao Tong University
Shanghai 200240, China
Starryspace@sjtu.edu.cn

Qizhi Qiu
State Key Laboratory of Photonics and Communications, School of Information Science and Electronic Engineering, Shanghai Jiao Tong University
Shanghai 200240, China
qizhi.qiu@sjtu.edu.cn

Yihao Zhang
State Key Laboratory of Photonics and Communications, School of Information Science and Electronic Engineering, Shanghai Jiao Tong University
Shanghai 200240, China
yihao.zhang@sjtu.edu.cn

Siyuan Wu
State Key Laboratory of Photonics and Communications, School of Information Science and Electronic Engineering, Shanghai Jiao Tong University
Shanghai 200240, China
wsy23211992521114@sjtu.edu.cn

Ruidong Liu
State Key Laboratory of Photonics and Communications, School of Information Science and Electronic Engineering, Shanghai Jiao Tong University
Shanghai 200240, China
David123@sjtu.edu.cn

Weisheng Hu
State Key Laboratory of Photonics and Communications, School of Information Science and Electronic Engineering, Shanghai Jiao Tong University
Shanghai 200240, China
wshu@sjtu.edu.cn

Qunbi Zhuge*
State Key Laboratory of Photonics and Communications, School of Information Science and Electronic Engineering, Shanghai Jiao Tong University
Shanghai 200240, China
qunbi.zhuge@sjtu.edu.cn

***Abstract*—This paper proposes MDN-ResKAN to model Q-factor distributions for reliable QoT estimation under imbalanced deployed-network data. It reduces nominal RMSE by 32.5% and achieves >97% anomaly recall on a field-deployed WDM testbed.**

***Keywords*—Optical networks, quality of transmission, digital twin, imbalanced data modeling**

I. INTRODUCTION

Digital twins (DTs) are emerging as a key enabler for intelligent optical-network operation, supporting service provisioning, resource optimization, failure management, and lifecycle management [1], [2]. Within a DT, quality-of-transmission (QoT) estimation is one of the core physical-layer modeling functions, since it estimates whether a candidate network configuration can satisfy transmission requirements before being deployed. Physics-based models, such as the Gaussian-noise (GN) model, provide analytical QoT estimates [3], but their accuracy in deployed systems is often limited by uncertain or time-varying parameters, especially span and connector insertion losses [4]. Data-driven machine-learning (ML) models have therefore been widely studied to learn the mapping from network configurations, including erbium-doped fiber amplifier (EDFA) gains and tilts, to QoT metrics from measured data [5], [6].

Despite these advances, reliable QoT estimation remains challenging in field-deployed optical networks. Operational data are usually highly imbalanced: most records correspond to nominal transmission states, whereas low-Q samples are rare but critical for service assurance. Such samples can arise from anomalous QoT fluctuations, device drift, measurement noise, or transient impairments [7], and missing them may lead to overly optimistic provisioning or delayed failure response [8], [9]. A deterministic estimator trained with squared-error loss must represent nominal and abnormal responses using a single prediction function. As a result, rare low-Q samples can either be smoothed into the nominal trend or, when their deviations are large, distort the nominal QoT estimate.

Recent efforts have improved the reliability of QoT modeling from several directions, including active input refinement for uncertain physical parameters [4], lifelong QoT prediction under network evolution [10], and bias mitigation in ML-based QoT estimation [11]. However, these works mainly address parameter uncertainty, temporal adaptation, or performance imbalance from the viewpoint of prediction accuracy. The joint problem of preserving nominal-sample QoT estimation while explicitly detecting rare low-Q states remains less explored. This motivates a probabilistic QoT model that can separate the nominal and low-Q response modes rather than forcing them into one deterministic output.

In this paper, we propose a mixture-density-based QoT estimation framework for imbalanced data from deployed optical networks. For each measured wavelength-division multiplexing (WDM) channel, the model represents the conditional Q-factor distribution by a two-component Gaussian mixture, where the mixture expectation provides the QoT estimate and the lower-mean component weight gives an intrinsic anomaly score [12]. A lightweight ResKAN encoder is used to extract nonlinear features from EDFA gain and tilt settings [13]. We validate the proposed method on a 25-channel WDM field-deployed testbed. Compared with an NN point-prediction baseline, the proposed method reduces nominal-sample RMSE by 32.5%, achieves anomaly recall above 97%, and keeps the false-positive rate (FPR) below 0.5%.

II. PRINCIPLE OF THE PROPOSED METHOD

The proposed framework is designed to improve QoT estimation reliability under imbalanced data from deployed optical networks. As shown in Fig. 1(a), a ResKAN feature encoder maps the EDFA settings to a feature representation, and an MDN head models the conditional Q-factor distribution. The mixture mean gives the QoT estimate, while the lower-mean component weight provides an anomaly score. This design reduces the impact of rare low-Q samples on nominal QoT estimation while enabling anomaly detection.

A. ResKAN Feature Encoder

Before the linear MDN head, a feature encoder maps the input EDFA settings $x \in \mathbb{R}^{12}$ to a 64-dimensional representation. The ResKAN encoder uses two parallel branches: an MLP branch and a KAN branch. The MLP branch captures smooth dependence between EDFA settings and Q factors, while the KAN branch provides polynomial nonlinear features. Their outputs are concatenated to form the 64-dimensional representation. This design provides complementary features for the MDN head. For comparison, we also evaluate MLP, KAN, ResNet, and Transformer encoders with the same 64-dimensional output size.

B. Mixture Density Network

To reduce the impact of low-Q samples on nominal QoT estimation, the MDN estimates conditional Q-factor distributions rather than a single-valued output. For each measured WDM channel c, the MDN parameterizes the conditional Q-factor distribution as a K-component Gaussian mixture:

$$p(Q_c|x) = \sum_{k=1}^{K} \pi_{c,k}(x)\mathcal{N}\left(Q_c, \mu_{c,k}(x), \sigma_{c,k}^2(x)\right) \quad (1)$$

where $\pi_{c,k}$, $\mu_{c,k}$, $\sigma_{c,k}^2$ denote the mixing weight, mean, and variance of the k-th component for channel c. We set $K = 2$ so that the mixture distribution can capture both nominal and low-Q abnormal Q-factor responses. The mixing weights are normalized, and the standard deviations are constrained to be positive.

The model is trained by minimizing the negative log-likelihood (NLL) across all channels:

$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{5} \log p(Q_c^{(i)} \,|\, x^{(i)}) \quad (2)$$

At inference, the mixture mean gives the Q-factor estimate:

$$\hat{Q}_c = \sum_{k} \pi_{c,k}\mu_{c,k} \quad (3)$$

The lower-mean component is interpreted as the abnormal component, and its mixing weight is used as an anomaly score. Specifically, for each channel, we identify the abnormal component as the one with the lowest mean:

$$k^* = \mathrm{argmin}_k \mu_{c,k}(x), \ \pi_{abn,c} = \pi_{c,k^*}(x) \quad (4)$$

The anomaly score is then defined as:

$$s(x) = \max_c \pi_{abn,c} \quad (5)$$

so that degradation in any measured WDM channel can trigger an alarm. An alarm is raised when $s(x) > \tau$, where τ is selected on the validation set.

III. DATASET AND EXPERIMENTAL SETUP

A. Field-Deployed Testbed and Dataset

The proposed model is evaluated on a QoT dataset collected from a field-deployed optical-network testbed, as shown in Fig. 1(b). The testbed contains a 440-km loop link with four G.652.D fiber spans and six EDFAs. A 25-channel WDM transmission system on a 75-GHz fixed grid is emulated. Five 63.9-Gbaud QPSK signals generated by commercial transponders are measured as real service channels, while the remaining channels are loaded with ASE-shaped dummy signals.

Each data sample consists of the gain and tilt settings of the six EDFAs, forming a 12-dimensional input vector. The output labels are the measured Q-factors of the five real service channels. After removing 83 complete-failure samples, where all measured Q-factors are zero, the dataset contains approximately 15,150 samples. A sample is regarded as abnormal if any of the five measured Q-factors is below the 6.0-dB FEC threshold. According to this criterion, abnormal samples account for 1.1% of the dataset, indicating a highly imbalanced operational data distribution.

B. Training and Evaluation Protocol

The dataset is divided into 7272 training samples, 4848 validation samples, and 3030 test samples. All learning-based

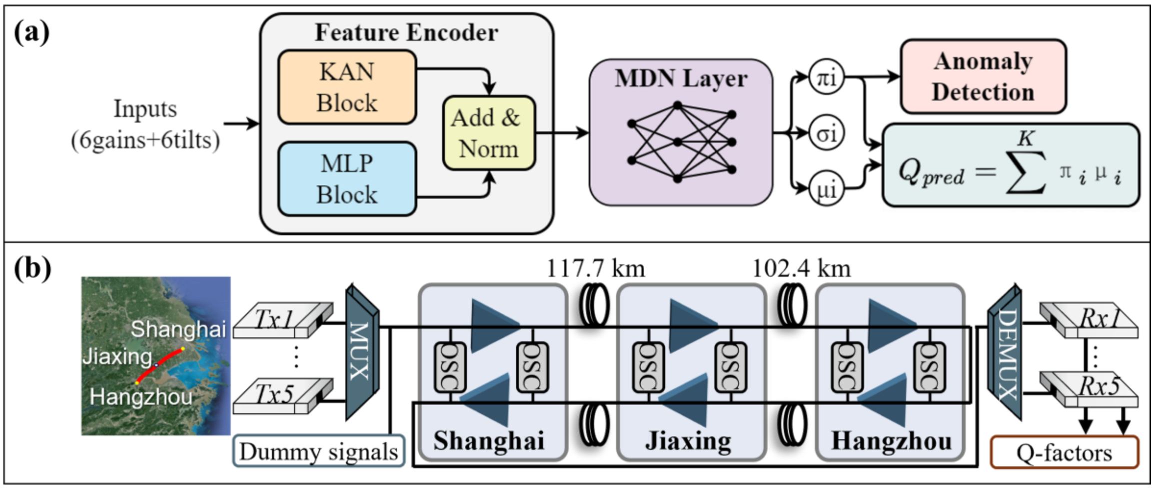

Fig. 1: (a) Proposed MDN-ResKAN framework for Q-factor estimation and anomaly detection. (b) Shanghai–Jiaxing–Hangzhou field-deployed testbed.

TAB. I. PERFORMANCE COMPARISON (3 SEEDS, MEAN ± STD). ESTIMATION METRICS ARE COMPUTED ON NOMINAL SAMPLES ONLY; DETECTION METRICS ARE COMPUTED ON THE FULL TEST SET. **BOLD & UNDERLINE**: BEST, **BOLD**: SECOND BEST.

Method	Estimation			Detection
	RMSE (dB) ↓	MAE (dB) ↓	R² ↑	Recall ↑
Deterministic baselines				
GN (Physical)	0.574 ± 0.004	0.458 ± 0.006	0.690 ± 0.010	0.183 ± 0.024
NN (MSE)	0.123 ± 0.017	0.067 ± 0.008	0.990 ± 0.009	0.929 ± 0.015
Uncertainty baselines				
Single Gaussian (K=1)	0.549 ± 0.046	0.402 ± 0.024	0.716 ± 0.042	0.713 ± 0.033
MC Dropout (T=50)	0.223 ± 0.009	0.168 ± 0.005	0.953 ± 0.005	0.812 ± 0.040
MDN (K=2, π_{abn} detection, ours)				
MDN-MLP	0.108 ± 0.006	0.073 ± 0.004	0.991 ± 0.001	0.973 ± 0.010
MDN-KAN	0.100 ± 0.006	0.067 ± 0.004	0.991 ± 0.001	**0.976** ± 0.001
MDN-ResNet	0.110 ± 0.005	0.074 ± 0.006	0.989 ± 0.001	**0.984** ± 0.012
MDN-Transformer	**0.085** ± 0.004	**0.060** ± 0.004	**0.993** ± 0.001	0.952 ± 0.022
MDN-ResKAN	**0.083** ± 0.005	**0.056** ± 0.005	**0.994** ± 0.001	**0.976** ± 0.059

models are trained and evaluated using the same data split. The results are reported as the mean and standard deviation over three random seeds.

For fair comparison, all neural models use comparable hidden dimensions, and all feature encoders output a 64-dimensional representation before the prediction or MDN head. For MDN-based models, two Gaussian components are used for each measured output channel. The anomaly alarm threshold is selected on the validation set and then fixed for test-set evaluation.

The evaluation contains two parts. First, QoT estimation performance is evaluated on nominal test samples using RMSE, MAE, and R^2. This setting reflects whether the model can accurately capture the normal operating trend of the deployed optical network. Second, anomaly detection performance is evaluated on the full test set using recall and false-positive rate (FPR). This setting reflects whether rare low-Q samples can be identified while keeping the alarm burden low.

IV. RESULTS AND DISCUSSION

Table I summarizes the estimation and detection performance of all nine methods, grouped into deterministic baselines, uncertainty baselines, and MDN-based models. The estimation metrics are calculated on nominal samples, while the detection metrics are evaluated on the full test set. Fig. 2 further visualizes the predicted and measured mean Q factors for representative methods with a random seed of 42.

A. QoT Estimation Accuracy

For nominal samples, the K = 2 MDN models reduce RMSE compared with the deterministic NN baseline, showing that mixture-density modeling improves QoT estimation under imbalanced data from deployed optical networks. MDN-ResKAN achieves the best estimation performance, with 0.083 dB RMSE, 0.056 dB MAE, and $R^2 = 0.994$. Compared with the NN baseline, this corresponds to a 32.5% reduction in RMSE and a 16.4% reduction in MAE. MDN-Transformer gives comparable accuracy, with 0.085 dB RMSE, whereas the MLP, KAN, and ResNet encoders remain in the range of 0.100–0.110 dB RMSE.

These results indicate that the performance gain does not only come from a specific feature encoder. Instead, the consistent improvement of $K = 2$ MDN models suggests that the mixture-density formulation plays a key role in reducing the impact of rare low-Q samples on nominal QoT estimation. By separating nominal and low-Q response modes, the model can better preserve the nominal Q-factor trend, which is important for digital-twin-assisted provisioning and optimization in deployed optical networks.

B. Low-Q Anomaly Detection

The deterministic NN baseline provides a strong threshold-based detector, achieving 0.929 recall by applying the 6.0-dB Q-factor threshold to its estimation. In comparison, all $K = 2$ MDN models achieve higher recall, indicating fewer missed low-Q events. MDN-ResKAN reaches 0.976 recall, corresponding to a 5.1% improvement over the NN baseline. Meanwhile, the false-positive rate of MDN-

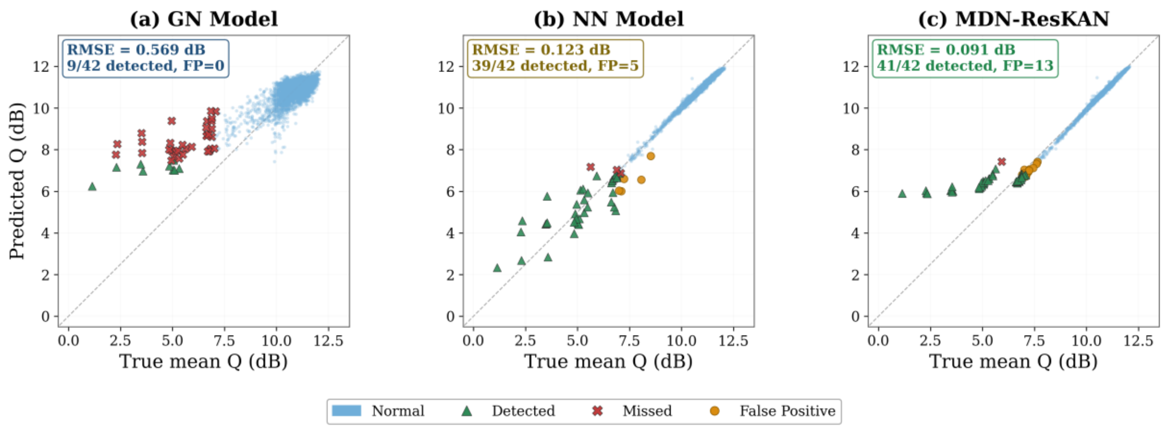

Fig. 2: Predicted vs. true mean Q factor (seed = 42, $\tau = 0.18$): (a) GN model, (b) NN model, and (c) MDN-ResKAN.

ResKAN remains below 0.5%, corresponding to about 13 nominal test samples falsely flagged as abnormal on average.

From an operational perspective, this trade-off is desirable because missed low-Q events may lead to unreliable provisioning or delayed failure response, whereas a small number of false alarms can be further filtered by network-control policies. Therefore, the lower-mean mixture weight provides a practical risk indicator rather than only an auxiliary classification output. This property is particularly useful for imbalanced field-deployed data, where low-Q events are rare but critical for service assurance.

C. Effect of Mixture Modeling

The MDN models perform better than the uncertainty baselines. The Single-Gaussian model fits each Q-factor output with one distribution, resulting in 0.549 dB RMSE and 0.713 recall. MC Dropout uses predictive variance as the anomaly score, but still gives 0.223 dB RMSE and 0.812 recall. In contrast, all $K = 2$ MDN models keep RMSE at or below 0.110 dB and recall above 0.95.

This result shows that predictive uncertainty alone is not sufficient for low-Q anomaly detection. The variance estimated by MC Dropout mainly reflects model uncertainty, while the lower-mean mixture weight in the proposed method directly corresponds to the probability of a degraded-Q response mode. Therefore, the two-component mixture is more suitable for imbalanced QoT data, where the key challenge is not only estimating uncertainty but also distinguishing nominal and degraded transmission regimes.

V. Conclusions

We proposed an MDN-ResKAN QoT estimation model for imbalanced data from deployed optical networks. By modeling the conditional Q-factor distribution with two Gaussian compo-nents, it reduces the influence of rare low-Q samples on nominal QoT estimation while providing an intrinsic anomaly score. On the WDM dataset, MDN-ResKAN reduces nominal-sample RMSE by 32.5%, achieves anomaly recall above 97%, and keeps the FPR below 0.5%. These results indicate that mixture-density modeling improves QoT estimation reliability under imbalanced data from deployed optical networks.

Acknowledgment

This work was supported by Shanghai Pilot Program for Basic Research-Shanghai Jiao Tong University (21TQ1400213) and Open Project in 2025 (Project2: Research on Mecha-nism of Ultra-High-Speed Large-Capacity All-Optical Switching for Satellite Network) from State Key Laboratory of Satellite Network and Shanghai Key Laboratory of Satellite Network.

References

[1] Q. Zhuge, X. Liu, Y. Zhang, M. Cai, Y. Liu, Q. Qiu, X. Zhong, J. Wu, R. Gao, L. Yi, and W. Hu, “Building a Digital Twin for Intelligent Optical Networks [Invited Tutorial],” Journal of Optical Communications and Networking, vol. 15, no. 8, pp. C242–C262, 2023, doi: 10.1364/JOCN.483600.

[2] D. Wang, Y. Song, Z. Zhao, X. Chen, and M. Zhang, “Digital Twin of Optical Networks: A Review of Recent Advances and Future Trends,” Journal of Lightwave Technology, vol. 42, no. 12, pp. 4233–4251, 2024, doi: 10.1109/JLT.2024.3401419.

[3] P. Poggiolini, “The GN Model of Non-Linear Propagation in Uncompensated Coherent Optical Systems,” Journal of Lightwave Technology, vol. 30, no. 24, pp. 3857–3879, 2012, doi: 10.1109/JLT.2012.2217729.

[4] X. Yang, C. Sun, G. Charlet, M. Tornatore, and Y. Pointurier, “Digital-twin-based active input refinement for insertion loss estimation and QoT optimization in C and C+L networks,” Journal of Optical Communications and Networking, vol. 16, no. 12, pp. 1261‑1272, 2024, doi: 10.1364/JOCN.537734.

[5] R. Ayassi, A. Triki, N. Crespi, R. Minerva, and M. Laye, “Survey on the Use of Machine Learning for Quality of Transmission Estimation in Optical Transport Networks,” Journal of Lightwave Technology, vol. 40, no. 17, pp. 5803–5815, 2022, doi: 10.1109/JLT.2022.3184178.

[6] S. Allogba, S. Aladin, and C. Tremblay, “Machine-Learning-Based Lightpath QoT Estimation and Forecasting,” Journal of Lightwave Technology, vol. 40, no. 10, pp. 3115–3127, 2022, DOI: 10.1109/JLT.2022.3160379.

[7] C. Delezoide, H. Akbari, P. Ramantanis, F. Boitier, and P. Layec, “On the Detection, Modeling, and Mitigation of Anomalous QoT Fluctuations in Optical Networks,” Journal of Optical Communications and Networking, vol. 16, no. 5, pp. B35–B44, 2024, DOI: 10.1364/JOCN.514779.

[8] Q. Qiu, X. Liu, Y. Zhang, L. Yi, W. Hu, and Q. Zhuge, “TrustOPT: A Trusted Online Optimization Strategy for Reliable Autonomous Optical Networks with Field-Trial Demonstration,” Journal of Optical Communications and Networking, vol. 18, no. 1, pp. A44–A53, 2026, DOI: 10.1364/JOCN.572249.

[9] F. Musumeci and M. Tornatore, “Failure management in optical networks with ML: a tutorial on applications, challenges, and pitfalls [Invited],” Journal of Optical Communications and Networking, vol. 17, no. 8, pp. C144–C155, 2025, DOI: 10.1364/JOCN.551910.

[10] Q. Wang, Z. Cai, and F. N. Khan, “Lifelong QoT prediction: an adaptation to real-world optical networks,” Journal of Optical Communications and Networking, vol. 16, no. 11, pp. 1159–1169, 2024, DOI: 10.1364/JOCN.531851.

[11] C. Natalino, B. Shariati, P. Safari, J. K. Fischer, and P. Monti, “Analysis and Mitigation of Unwanted Biases in ML-based QoT Classification Tasks,” in Optical Fiber Communication Conference (OFC) 2024, Technical Digest Series, San Diego, CA, USA, 2024, paper M1H.3, DOI: 10.1364/OFC.2024.M1H.3.

[12] J. Chen, Y. Yu, and Y. Liu, “Physics-Guided Mixture Density Networks for Uncertainty Quantification,” Reliability Engineering & System Safety, vol. 228, p. 108823, 2022, DOI: 10.1016/j.ress.2022.108823.

[13] Z. Liu, Y. Wang, S. Vaidya, F. Ruehle, J. Halverson, M. Soljačić, T. Y. Hou, and M. Tegmark, “KAN: Kolmogorov-Arnold Networks,” in Proceedings of the International Conference on Learning Representations, 2025.