Smart Optical Networks Enabled by Performance Monitoring and Machine Learning

by

Sandra ALADIN

MANUSCRIPT-BASED THESIS PRESENTED TO ÉCOLE DE TECHNOLOGIE SUPÉRIEURE IN PARTIAL FULFILLMENT FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
Ph.D.

MONTREAL, JULY 15, 2026

ECOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

BOARD OF EXAMINERS

THIS THESIS HAS BEEN EVALUATED

BY THE FOLLOWING BOARD OF EXAMINERS

Dr. Christine Tremblay, Thesis Supervisor
Electrical Engineering Department, École de technologie supérieure

Dr. Lena Wosinska, Thesis Co-supervisor
Electrical Engineering Department, Chalmers University of Technology

Dr. Nadjia Kara, President of the Board of Examiners
Software and IT Engineering Department, École de technologie supérieure

Dr. Aris Leivadeas, Member of the jury
Software and IT Engineering Department, École de technologie supérieure

Dr. Massimo Tornatore, External Independent Evaluator
Department of Electronics, Information, and Bioengineering, Politecnico di Milano

THIS THESIS WAS PRENSENTED AND DEFENDED

IN THE PRESENCE OF A BOARD OF EXAMINERS AND PUBLIC

ON JUNE 29, 2026

AT ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to my supervisor, Professor Christine Tremblay, whose guidance, insightful feedback, and support have shaped every stage of this doctoral journey. Her scientific vision, high expectations, and continual encouragement have deeply influenced both this work and my development as an independent researcher. I am truly grateful for the trust she placed in me and for the many conversations that pushed me to think more critically and rigorously.

I am equally deeply grateful to my co-supervisor, Professor Lena Wosinska, whose valuable discussions, constructive comments, and technical expertise have greatly strengthened this work. The complementary perspectives and guidance provided by both of my supervisors have been essential in bringing this thesis to completion.

I gratefully acknowledge the financial support provided by the Natural Sciences and Engineering Research Council of Canada (RGPIN 2019-03972), the ÉTS Program for the Development of International Research Collaborations and European Commission's Digital Europe Programme (DIGITAL) under grant agreement #101127973, and the 5G TACTIC "5G Trusted And seCure network servICes" project. Without these fundings this research would not have been possible. I also wish to thank the collaborators and colleagues whose discussions, feedback, and encouragement contributed to this work. Their support made the challenges of research more manageable and the achievements along the way far more meaningful.

Beyond academia, I wish to express my heartfelt thanks to my family for their constant support. To my parents, brothers and sister, thank you for your unconditional love. Your encouragement has been my greatest source of strength. I am equally grateful to my husband and my boys for their patience, understanding, and support throughout the demanding years of this doctoral journey. Their presence and reassurance have been a source of strength and motivation.

This thesis carries not only the weight of years of research, but also the love and strength of all those who walked beside me. To everyone who stood beside me, lifted me up, and believed in

me even when I struggled to believe in myself, my gratitude runs deeper than these words can express. This accomplishment is as much yours as it is mine.

Réseaux optiques intelligents rendus possibles par la surveillance des performances et l'apprentissage automatique

Sandra ALADIN

RÉSUMÉ

Le bon fonctionnement d'un réseau optique repose sur une gestion efficace de la qualité de transmission (*Quality of Transmission*, QoT) tout au long de la durée de vie opérationnelle de chaque connexion optique. Les défaillances matérielles interrompent immédiatement le service, tandis que les défaillances logicielles provoquent une dégradation progressive pouvant passer inaperçue jusqu'à ce que le service soit affecté. Dans les deux cas, elles se traduisent par des violations de la QoT. La gestion de la QoT regroupe l'ensemble des fonctions ou tâches visant à garantir qu'une connexion optique respecte en permanence les exigences de qualité de service. Ces fonctions incluent notamment l'évaluation de la faisabilité d'une connexion à partir de la QoT estimée avant son établissement, la détection de dégradations anormales de QoT durant l'opération, ainsi que la prédiction de son évolution afin de permettre des interventions proactives avant que le niveau de performance ne devienne insuffisant.

La gestion des défaillances dans les réseaux optiques (*Optical Network Failure Management,* ONFM) traite les défaillances dans les réseaux à travers six cas d'utilisation identifiés dans la littérature : l'estimation de la QoT, la détection des défaillances, la prédiction des défaillances, l'identification des cause profondes, la localisation des défaillances et l'estimation de leur ampleur. Bien que l'apprentissage automatique (*Machine Learning*, ML) se soit révélé efficace pour chacun de ces cas d'usage pris individuellement, les solutions existantes les traitent séparément. De plus, les principes qui relient les tâches de gestion de la QoT, à savoir le choix des approches ML appropriées pour chacune d'entre elles et la manière dont elles peuvent se renforcer mutuellement via une intégration inter-tâches, restent peu explorés. Cette thèse traite trois tâches de gestion de la QoT associées à trois cas d'utilisation de l'ONFM, à savoir l'évaluation de la faisabilité de connexions optiques par estimation de la QoT, la détection d'anomalies pour la détection de défaillances, et la prédiction de la QoT pour la prédiction de défaillances.

Pour l'évaluation de la faisabilité des connexions optiques, donc avant leur établissement, un modèle d'apprentissage supervisé léger et interprétable est développé pour l'estimation de la QoT des connexions optiques non établies. Afin de pallier la différence entre les données d'entraînement synthétiques ou simulées et les environnements de déploiement réels où les jeux de données sont rares, des techniques d'adaptation de domaine intégrées à une sélection de caractéristiques basée sur l'intelligence artificielle explicable (*eXplainable Artificial Intelligence*, XAI) sont mises en œuvre. Cela permet d'obtenir des performances robustes avec un nombre limité d'échantillons du domaine cible, tout en conservant l'efficacité computationnelle et l'interprétabilité requises pour les opérations de planification de réseaux optiques à grande échelle. Les classificateurs Random Forest (RF), combinés à la technique de sélection automatisée de caractéristiques Boruta-SHAP (*Boruta-SHapley Additive exPlanation*) et à l'apprentissage par transfert, atteignent une précision de 98,64% à 99,65%

pour l'estimation de la QoT sur quatre jeux de données, tout en offrant une efficacité de calcul supérieure (réductions du temps d'entraînement de 52,78% à 70,68%, et des temps d'inférence de 1,1 à 2,4 µs par connexion optique) ainsi qu'une interprétabilité basée sur SHAP. L'adaptation de domaine guidée par la sélection de caractéristiques atteint une précision de 86% avec seulement 50 instances du domaine cible.

Pour la détection d'anomalies pendant l'exploitation du réseau, des méthodes complémentaires d'apprentissage non supervisé sont appliquées afin d'identifier diverses anomalies dans des jeux de données multivariés de surveillance des performances optiques (*Optical Performance Monitoring*, OPM), sans s'appuyer sur des données de défaillance étiquetées. Une comparaison de méthodes non supervisées telles que (*Density-Based Spatial Clustering of Applications with Noise,* DBSCAN), (*One-Class Support Vector Machine,* OCSVM), et (*Isolation Forest*, IF) sur quatre années de données de réseau en production démontre que DBSCAN offre des performances supérieures (précision de 0,54, score de silhouette de 0,82) tout en détectant des dégradations subtiles que la surveillance traditionnelle basée sur des seuils ne parvient pas à détecter.

Pour la prédiction de la QoT, un cadre intégrant la détection des points de changement (*Change Point Detection*, CPD) et des anomalies collectives (*Collective Anomaly Detection*, CAD) est introduit à l'aide de modèles profonds séquence-à-séquence multivariés pour la prédiction de la QoT. En intégrant des indicateurs de points de changement et d'anomalies collectives dans la prédiction de séries temporelles, cette approche permet de calculer des fenêtres de réaction exploitables qui favorisent une gestion proactive du réseau. Le cadre de prédiction de la QoT basé sur un réseau à mémoire à long et court terme (*Long Short-Term Memory*, LSTM) tenant compte des points de changement et des anomalies collectives, permet une réduction de l'erreur quadratique moyenne (*Root Mean Square Error*, RMSE) de 28,21% et de l'erreur absolue moyenne (*Mean Absolute Error*, MAE) de 42,36% par rapport à la prédiction de référence, offrant des fenêtres de réaction cohérentes de 1 à 10 jours avant la survenue d'une défaillance. Une validation systématique sur plusieurs jeux de données supplémentaires, architectures de prédiction (*Temporal Convolutional Network*, TCN et LSTM), techniques de CAD non supervisées, et longueurs de séquence d'entrée, confirme la robustesse de l'approche. Les résultats montrent des améliorations constantes des prédictions (réductions moyennes de RMSE et MAE entre 10-25%) ainsi qu'une couverture des défaillances de 100% avec une fenêtre de réaction positive stable pour toutes les sensibilités de détection testées.

Ces résultats démontrent que la sélection de modèles ML adaptés au contexte propre à chaque tâche permet d'obtenir des solutions robustes pour la gestion de la QoT. Ils montrent également que l'intégration d'indicateurs CPD-CAD dans les modèles de prédiction de la QoT améliore significativement la précision des prédictions tout en offrant des capacités d'alerte précoce pour des réseaux optiques en exploitation.

Mots-clés : Détection d'anomalies, détection de points de changement, adaptation de domaine, IA explicable, gestion des défaillances des réseaux optiques, gestion proactive des réseaux, qualité de transmission, fenêtre de réaction, prédiction de séries temporelles, apprentissage par transfert

Smart Optical Networks Enabled by Performance Monitoring and Machine Learning

Sandra ALADIN

ABSTRACT

Reliable optical network operation depends on effective quality of transmission (QoT) management across the operational lifetime of each lightpath. Hard failures cause immediate service disruption, while soft failures produce gradual degradation that may go undetected until service is impacted. Both ultimately manifest as QoT violations. QoT management encompasses the functions or tasks required to ensure that each lightpath meets its quality-of-service requirements. These functions include assessing the feasibility of a lightpath based on its estimated QoT before establishment, detecting abnormal QoT degradations during operation, and forecasting the evolution of QoT to enable proactive intervention before service quality is compromised.

Optical network failure management (ONFM) addresses network failures with six identified use cases in the literature: QoT estimation, failure detection, failure prediction, root-cause identification, failure localization, and failure magnitude estimation. Although machine learning (ML) has proven effective for individual ONFM use cases, existing solutions have addressed them separately. Moreover, the relationships among the QoT management tasks, the selection of appropriate ML approaches for each task, and the ways in which these tasks can reinforce one another through cross-task integration, remain underexplored. This thesis addresses three QoT management tasks corresponding to three ONFM use cases: lightpath feasibility assessment through QoT estimation, anomaly detection for failure detection, and QoT forecasting for failure prediction.

For lightpath feasibility assessment, performed prior to lightpath establishment, a lightweight and interpretable supervised learning model is developed for QoT estimation of unestablished lightpaths. To address the difference between synthetic or simulated training data and real deployment environments with scarce datasets, domain adaptation techniques integrated with feature selection based on explainable artificial intelligence (XAI) are implemented. This enables robust performance with limited target domain samples while maintaining the computational efficiency and interpretability required for large-scale optical network planning operations. Random Forest (RF) classifiers with Boruta-Shapley additive explanation (Boruta-SHAP) automated feature selection and transfer learning achieve 98.64%-99.65% accuracy for QoT estimation across four datasets while providing superior computational efficiency (reductions of training time by 52.78%-70.68%, inference times of 1.1-2.4 μs per lightpath) and SHAP-based interpretability. Feature-selection-guided domain adaptation achieves 86% accuracy with only 50 target domain instances.

For anomaly detection during network operations, complementary unsupervised learning methods are applied to identify diverse anomalies within multivariate optical performance monitoring (OPM) datasets without relying on labeled failure data. A comparison of unsupervised methods such as density based spatial clustering applications with noise

(DBSCAN), one-class support vector machine (OCSVM), Isolation Forest (IF), on 4 years of production network data, demonstrates that DBSCAN achieves superior anomaly detection performance (precision 0.54, silhouette score 0.82) while detecting subtle degradations traditional threshold-based monitoring fails to observe.

For QoT forecasting, a framework that integrates change-point detection (CPD) and collective anomaly detection (CAD) is introduced using multivariate sequence-to-sequence (seq2seq) deep learning models for QoT forecasting. By integrating change point and collective anomaly indicators into time-series prediction, the approach enables the computation of actionable reaction windows that support proactive network management. A change point and collective anomaly-aware long short-term memory (LSTM)-based forecasting framework achieves 28.21% root mean square error (RMSE) reduction and 42.36% mean absolute error (MAE) reduction compared to baseline forecasting, providing consistent reaction windows of 1-10 days before failure events. A systematic validation across diverse additional datasets, forecasting architectures (temporal convolutional network (TCN) and LSTM), unsupervised CAD techniques, and input sequence lengths confirm the robustness of the approach. The results show consistent forecasting improvements (averaged reductions of RMSE and MAE of roughly 10-25%), along with 100% failure coverage with a stable positive lead time across all detection sensitivity.

These results demonstrate that selecting ML models according to the specific context of each task yields robust solutions for QoT management. Moreover, they show that integrating CPD-CAD indicators into the QoT forecasting framework significantly improves both prediction accuracy and early-warning capability in production optical networks.

Keywords: Anomaly detection, change point detection, domain adaptation, explainable AI, lead time, optical network failure management, proactive network management, quality of transmission, reaction window, time-series forecasting, transfer learning

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS AND ACRONYMS

ACF	Autocorrelation function
ADF	Augmented Dickey-Fuller
AI	Artificial intelligence
ANN	Artificial neural network
ASE	Amplified spontaneous emission
AUC	Area under the curve
BER	Bit error rate
CAD	Collective anomaly detection
CD	Chromatic dispersion
CNN	Convolutional neural network
CONUS	Core optical network of the United States
CORAL	Correlation alignment
CPD	Change point detection
CUSUM	Cumulative sum
DA	Domain adaptation
DBSCAN	Density-based spatial clustering of applications with noise
DDFO	Domain discrepancy feedback optimization
DGD	Differential group delay
DL	Deep learning
DMB	Dataset mixing baseline
DNN	Deep neural network

DSP	Digital signal processing
DWDM	Dense wavelength division multiplexing
EDFA	Erbium-doped fiber amplifier
ED-LSTM	Encoder-decoder long short-term memory
EGN	Enhanced Gaussian noise
FA	Feature augmentation
FEC	Forward error correction
FMF	Few-mode fiber
FWM	Four-wave mixing
GGN	Generalized Gaussian noise
GN	Gaussian noise
GNPy	Gaussian noise simulation in Python
GPR	Gaussian process regression
GPT	Generative pre-trained transformer
GRU	Gated recurrent unit
GSNR	Generalized signal-to-noise ratio
HCCS	High correction count seconds
HDBSCAN	Hierarchical density-based spatial clustering of applications with noise
HHI	(Fraunhofer) Heinrich Hertz institute
IF	Isolation forest
ILOS	Imminent loss of signal
IoT	Internet of things
ISRS	Inter-channel stimulated Raman scattering

ITU-T	International telecommunication union - telecommunication standardization sector
KNN	K-nearest neighbors
KPI	Key performance indicators
KPSS	Kwiatkowski-Phillips-Schmidt-Shin
LIME	Local interpretable model-agnostic explanations
LLM	Large language model
LOS	Loss of signal
LSTM	Long short-term memory
MAE	Mean absolute error
MDM	Mode division multiplexing
ML	Machine learning
MLP	Multilayer perceptron
MMD	Maximum mean discrepancy
MSE	Mean squared error
NLI	Nonlinear interference
NSERC	Natural sciences and engineering research council of Canada
OCSVM	One-class support vector machine
OFC	Optical fiber communications conference
ONFM	Optical network failure management
OPM	Optical performance monitoring
OPR	Optical power received
OSNR	Optical signal-to-noise ratio

OTN	Optical transport network
PCA	Principal component analysis
PDL	Polarization-dependent loss
PLATON	Planning tool for optical networks
PLI	Physical layer impairment
PMD	Polarization-mode dispersion
PR-AUC	Precision-recall area under the curve
QMIN	Minimum Q-factor
QoS	Quality of service
QoT	Quality of transmission
R^2	Coefficient of determination
RBF	Radial basis function
ReLU	Rectified linear unit
RF	Random forest
RFE	Recursive feature elimination
RFECV	Recursive feature elimination with cross-validation
RMSE	Root mean square error
RNN	Recurrent neural network
ROADM	Reconfigurable optical add-drop multiplexer
SBS	Stimulated Brillouin scattering
SDB	Source domain baseline
SDN	Software-defined networking

Seq2seq	Sequence-to-sequence
SHAP	Shapley additive explanations
SLA	Service level agreement
SMF	Single-mode fiber
SNR	Signal-to-noise ratio
SPM	Self-phase modulation
SRS	Stimulated Raman scattering
SSFM	Split-step Fourier method
SVM	Support vector machine
TCN	Temporal convolutional network
TDB	Target domain baseline
TED	Traffic engineering database
TL	Transfer learning
TSNN	Telefónica Spain national network
UAS	Unavailable seconds
VARMAX	Vector autoregressive moving average with exogenous variables
WDM	Wavelength division multiplexing
XAI	Explainable artificial intelligence
XGBoost	Extreme gradient boosting
XPM	Cross-phase modulation

LIST OF SYMBOLS AND UNITS OF MEASUREMENT

dB	decibel
dBm	decibel-milliwatt
ps	picosecond
ps/nm	picosecond per nanometer
nm	nanometer
GHz	gigahertz
km	kilometre
GBd	gigabaud
Gb/s	gigabit per second
Tb/s	terabit per second
s	second
μs	microsecond

INTRODUCTION

Modern optical fiber networks form the backbone of global telecommunications infrastructure, carrying exponentially increasing data traffic driven by cloud computing, video streaming, social media, and emerging applications like Internet of Things (IoT). As network capacity demands continue to rise, network operators face growing challenges in efficiently planning network evolution, ensuring reliable quality of service (QoS), and maintaining optimal performance across increasingly complex infrastructures (Cao et al., 2026). Traditional network management approaches, relying on manual processes, threshold-based alarms, and reactive interventions, struggle to cope with the scale, complexity, and dynamic nature of contemporary optical networks (Gu, Yang & Ji, 2020).

Machine learning (ML) has the ability to transform optical network management by enabling automated decision-making, intelligent pattern identification, and predictive capabilities that can enhance network operations (Musumeci et al., 2019a; Ayassi et al.,2022; Panayiotou, Michalopoulou & Ellinas, 2023). To turn these capabilities into operational benefits, ML needs to be connected to the specific quality of transmission (QoT) management challenges that determine network reliability. QoT management encompasses a set of operational tasks through which the lightpaths QoT is estimated, monitored, and predicted throughout their lifecycle. These tasks include assessing whether a candidate lightpath is expected to meet its required QoT prior to deployment, detecting QoT degradations during operation, and forecasting future QoT degradation before service is affected. Effectively applying ML to these tasks requires a clear understanding of their distinct objectives and the corresponding data and operational constraints, including limited labeled failure data, stringent computational budgets, interpretability requirements to foster operator trust, and heterogeneity of equipment and monitoring data across network domains. Musumeci and Tornatore (2025) present an overview of six key use cases in optical network failure management (ONFM), each focusing on specific objectives along the lifetime of an optical network, its lightpaths, or the associated optical components. This thesis addresses three ONFM use cases, selected for their direct relevance to QoT management across distinct stages of a lightpath's lifetime. The first is QoT estimation, addressed through an interpretable and efficient framework for lightpath feasibility assessment. The second is failure detection, addressed through unsupervised anomaly detection on multivariate optical performance monitoring (OPM) data. The third is failure prediction, addressed through an anomaly and regime-shift-aware QoT forecasting framework. While ML has proven effective for each individual task, applying it efficiently necessitates careful consideration of each task's distinct characteristics, such as available data sources, labeling constraints, computational requirements, and interpretability. Accordingly, we present three contributions with ML approaches customized to each ONFM-based task's unique challenges and opportunities. For QoT estimation, where labeled data are scarce in newly deployed networks, we combine supervised learning with feature selection and domain adaptation to transfer knowledge from data-rich to data-poor network domains in an interpretable and computationally efficient manner. For anomaly detection, where labeled failure events are rare or unavailable, we employ unsupervised learning on multivariate OPM data to detect QoT deviations without relying on failure labels. For QoT forecasting, where soft failures develop gradually through drifts and regime shifts, we propose a sequence-to-sequence (seq2seq) time series forecasting framework that integrates change point detection (CPD) and collective anomaly detection (CAD) to anticipate QoT degradation with sufficient lead time for proactive intervention.

This chapter introduces the research context and motivation, establishes the research questions and objectives, summarizes the thesis contributions, and provides an overview of the thesis organization.

0.1 Context and motivation

0.1.1 The growth of optical networks

Global internet traffic is growing fast, driven by bandwidth-intensive applications and the proliferation of connected devices. Global data traffic continues to grow at an unprecedented rate, with projections estimating the total volume of data generated worldwide will approach several hundreds of zettabytes by 2027 (International Data Corporation, 2023, as cited in Ciena, 2023). This places substantial pressure on telecommunications infrastructure. Optical fiber networks, which have high bandwidth capacity enabled by dense wavelength division multiplexing (DWDM) technology, are the only practical solution for delivering these massive data volumes across metropolitan, national, and worldwide distances (Musumeci et al., 2019a; Keiser, 2021).

Modern optical networks are complex systems in which several wavelength channels can propagate simultaneously across common fiber infrastructure via numerous reconfigurable optical add-drop multiplexers (ROADMs) and optical amplifiers (Keiser, 2021). The QoT for each end-to-end optical connection or lightpath that transports traffic is determined by a variety of factors, including route length, the number of intermediate nodes, fiber characteristics, channel power and wavelength allocation, modulation format, and interactions among co-propagating channels. Managing these networks to achieve service level agreements (SLAs) while increasing capacity utilization and reducing operational costs involves keeping the physical-layer performance of all active lightpaths stable under constantly changing conditions (Ramaswami, Sivarajan & Sasaki, 2010; Mukherjee, 2006; Musumeci et al., 2019a; Agrawal, 2010). A key part of this challenge is maintaining good QoT throughout each lightpath's lifetime, since QoT degradation directly affects both service reliability and operational efficiency. As aforementioned, the thesis addresses three ONFM-based QoT management tasks. Challenges linked to each of the tasks are described below.

0.1.2 Challenges in QoT estimation

Accurate QoT estimation is central to lightpath provisioning, as operators must determine whether candidate lightpaths will meet required QoT levels before deployment (Cao et al., 2026; Yan et al., 2017). Traditional approaches rely on analytical models that simplify physical layer behavior or on detailed simulations that offer higher accuracy but become computationally expensive at large scale. QoT estimation models typically rely on synthetic data generated from these analytical or simulation tools. This creates a domain gap between training and deployment environments: synthetic training data may not fully capture all physical layer effects, device imperfections, and parameter deviations encountered in production optical networks. Among others, these impairments include polarization-mode dispersion (PMD) and polarization-dependent loss (PDL), amplifier noise-figure variations, ROADM filter cascading penalties, arising from simplifications in analytical models or assumptions made in physical-layer simulations.

A further challenge is the scarcity of topology-, network- and transceiver-related data. Privacy and security concerns limit access to detailed network data, and early deployments typically include only a limited number of active lightpaths, restricting the availability of real-world samples for training. These constraints make it difficult to build robust supervised models using field data (Bergk, Shariati, Safari, & Fisher, 2022). Domain adaptation techniques address this by enabling models trained on synthetic data to adjust to small datasets, reducing distribution mismatches and improving generalization under data-constrained conditions (Rottondi et al., 2021; Cai et al., 2024; Musumeci et al., 2022). At the same time, operators require not only accurate estimations but also interpretable explanations for why a lightpath is expected to succeed or fail its QoT target (Ayoub et al., 2023). Therefore, the QoT estimation task demands ML methods that combine accuracy, computational efficiency, domain-adaptation capabilities, and explainability to support trustworthy, scalable QoT estimation.

0.1.3 Challenges in anomaly detection

Once lightpaths are established and carrying traffic, they must be continuously monitored to ensure they operate within acceptable performance levels throughout their lifetime. OPM systems collect measurements of key QoT parameters such as optical signal-to-noise ratio (OSNR), optical power received (OPR), chromatic dispersion (CD), PMD, and pre-forward error correction bit error rate (pre-FEC BER), typically at regular intervals ranging from every 15 minutes to once per day for each active lightpath.

Conventional monitoring relies on static threshold-based alarms and faces significant limitations. Fixed thresholds cannot adapt to the unique characteristics of individual lightpaths or differentiate normal fluctuations from true anomalies (Musumeci et al., 2019b; Boutaba et al., 2018). In this thesis, normal behavior refers to the dominant and recurring patterns observed in OPM parameters during healthy lightpath operation. These patterns are learned directly from data rather than predefined a priori. Anomalies refer to deviations from this learned normal behavior in observed OPM data. Optical networks faults are generally classified into hard failures, which cause immediate service disruption, and soft failures, which develop gradually and may progress into hard failures if left unaddressed (Musumeci & Tornatore, 2025). Chandola, Banerjee and Kumar (2009) present three main types of anomalies: point anomalies, where a single observation deviates significantly from the rest of the data; contextual anomalies, where an observation is anomalous only within a specific context; and collective anomalies, where a sequence of correlated observations deviates from expected behavior (Huang, Yun & Hu, 2023). In optical networks, such patterns correspond to soft failures, described as gradual QoT degradations that do not immediately cause service interruption but indicate emerging issues. These anomalies often appear only through the combined pattern of several correlated parameters while each individual metric may still fall within its acceptable threshold (Ahmed, Mahmood & Hu, 2016). This highlights the need for a multivariate strategy rather than a univariate one, to capture the combined pattern that is often missed by threshold-based monitoring systems.

A major obstacle for ML-based anomaly detection in operational networks is the scarcity of labeled training data. Operators rarely maintain detailed records specifying which time periods or which lightpaths exhibited anomalous versus normal behavior. While major outages may be documented, subtle degradations, soft failures, and slow-developing issues often go unlogged (Musumeci et al., 2019b; Tremblay et al., 2023). This scarcity of labels makes supervised learning impractical and motivates the use of unsupervised methods capable of identifying unusual patterns without relying on labeled datasets (Ahmed et al., 2016). Since no single unsupervised method can be expected to outperform all others across all anomaly types and data distributions, a systematic evaluation of different techniques with hyperparameter tuning is required to select the appropriate ones and understand their complementary strengths for different anomaly types.

0.1.4 Challenges in QoT forecasting

Modern network operations increasingly demand predictive capabilities that support proactive management rather than reactive troubleshooting. Instead of addressing issues only after they begin to degrade performance or disrupt service, proactive strategies aim to anticipate emerging problems and enable preventive action. For instance, if a forecasting model indicates that a lightpath's OSNR is likely to fall below acceptable thresholds within the next several hours, operators can schedule maintenance, prepare backup routes, or apply mitigation measures before service quality is affected.

A central requirement for effective failure prediction is the availability of adequate reaction windows, which is the interval between the time a prediction is issued and the expected occurrence of the predicted event. When a model forecasts degradation only a few minutes in advance, the reaction window is too short for any actions beyond automated rerouting. In contrast, predictions with one day or more of lead time would allow operators to coordinate field technicians, arrange replacement components, and plan maintenance windows (Behera, Panayiotou & Ellinas, 2023a; Murphy, Lavignotte & Lepers, 2024). Forecasting methods must therefore balance prediction horizon and accuracy, as extending the forecast further into the future typically increases uncertainty.

OPM data exhibits rich and complex temporal behavior, including long-term trends caused by component aging, periodic variations driven by daily temperature cycles, abrupt shifts due to configuration changes, and intermittent anomalies from transient interruptions. Capturing these diverse temporal patterns to project future performance requires advanced time-series forecasting techniques. Additionally, detecting early warning signals such as change points marking transitions between operating regimes or sequences of increasingly different conditions, provides interpretable indicators that systematic degradation may be developing.

Together, these challenges across QoT estimation, failure detection, and failure prediction motivate the three research contributions of this thesis.

0.2 Research questions and objectives

The challenges outlined in the previous section lead to three key research directions in ML-based optical network failure management. Each research direction targets a specific ONFM use case with ML methods aligned with its data characteristics and operational constraints. The following research questions and objectives are organized around these three directions.

RQ1: How can QoT be accurately estimated before lightpath establishment by leveraging large-scale synthetic training data through transfer learning to address data scarcity in new deployments, while ensuring operator trust through interpretable, explainable artificial intelligence (XAI)-based feature selection? What ML techniques can best balance accuracy with computational efficiency for QoT estimation?

RQ2: How can multivariate OPM data and different unsupervised learning paradigms (density-based, boundary-based, isolation-based) be leveraged to detect anomalies across correlated parameters in optical networks, without requiring labeled training samples?

RQ3: How can change point and collective anomaly detection bring context to multivariate seq2seq QoT forecasting models to enhance prediction performance and provide interpretable early warning signals with sufficient reaction windows for proactive network management? Does the enhanced forecasting framework generalize across independent production network datasets, alternative forecasting architectures, and input sequence lengths, and does it provide meaningful proactive failure detection?

The following objectives turn each research question into a concrete contribution. For each task, the objective outlines the method developed, the reasoning behind the choice, and - in the third contribution - the level of validation needed to show that the approach generalizes beyond a single dataset or configuration. First, this thesis introduces an automated, interpretable, and efficient ML framework for QoT estimation at lightpath provisioning, combining XAI-based feature selection with transfer learning to better align synthetic training data with real-world deployment conditions. Second, it implements and evaluates complementary unsupervised anomaly detection methods to identify different anomalies in multivariate OPM data without labeled training sets, highlighting their complementary strengths across anomaly types. Third, it builds a sequence-to-sequence forecasting framework integrating change points and anomalies for proactive failure prediction with actionable lead times or reaction windows. Additionally, it demonstrates the framework's generalizability across three additional production network datasets, two deep learning (DL) architectures, five anomaly detection methods, and two input sequence lengths, while presenting a preliminary evaluation of failure coverage, lead time and false alarm rates.

0.3 Thesis contributions

This thesis brings various novel contributions to the field of optical network management by introducing three contributions addressing tasks based on three ONFM use cases identified by Musumeci and Tornatore (2025). The contributions, presented across three peer-reviewed articles, are organized according to ML approaches adapted to each ONFM-based task's data context and operational constraints. They are summarized below.

Article 1: addressing generalizability requirements in QoT estimation

Aladin, S., Wosinska, L., & Tremblay, C. "Automated, interpretable and efficient ML models for real-world lightpaths' quality of transmission estimation". *Published in IEEE Open Journal of the Communications Society, November 2025,* vol. 6, pp. 9785–9801. https://doi.org/10.1109/OJCOMS.2025.3635533.

This article addresses RQ1 by developing an automated, interpretable, and efficient framework for lightpath feasibility assessment. It shows that random forest (RF) models used for domain adaptation integrated with Shapley additive explanation (SHAP) explainability can achieve robust data efficiency, computational efficiency, and interpretability, while addressing the domain shift between synthetic datasets and real-world deployment conditions. This challenges the idea that deep neural networks (DNNs) are required for high performance, demonstrating that lightweight, explainable models can deliver comparable results while maintaining operator trust.

Article 2: Tackling scarcity of labeled failure data constraints

Aladin, S., Wosinska, L., & Tremblay, C. “Detecting anomalies in the optical layer using unsupervised machine learning”. In *2024 Optical Fiber Communications Conference and Exhibition (OFC)* (pp. 1-3). Th3I.4.

This article addresses RQ2 through a systematic comparison of three complementary unsupervised methods: density-based spatial clustering applications with noise (DBSCAN), one-class support vector machine (OCSVM), and isolation forest (IF), for detecting anomalies in production OPM data. Results show that each method captures different types of degradation, and that using multiple approaches offers broader coverage. Effectiveness is validated on four years of field QoT data without requiring labeled training sets.

Article 3: Moving from QoT prediction to actionable intervention timing

Aladin, S., Wosinska, L., & Tremblay, C. “Change Point and Anomaly-Aware QoT Forecasting Framework for Proactive Network Management” in *IEEE Networking Letters*, vol. 8, pp. 224-228, 2026, doi: 10.1109/LNET.2026.3691249.

This article addresses RQ3 by integrating change point detection (CPD) and collective anomaly detection (CAD) into a multivariate seq2seq forecasting framework to detect QoT degradations and to estimate actionable reaction windows. Instead of producing raw QoT predictions that require further interpretation, the framework provides the lead time available for intervention, allowing operators to act before failures occur. This transforms QoT forecasting from a purely predictive task into an operational decision-support capability, bridging the gap between fault detection and the urgency of corrective actions.

Chapter 6 presents a systematic validation of the CPD-CAD-aware QoT forecasting framework introduced in Chapter 5 (Article 3) by evaluating two DL architectures, five collective anomaly detection techniques on three additional lightpath datasets from the same production network. This confirms that the enhanced framework shows consistent forecasting improvements with consistent positive lead time in all scenarios similarly to the results achieved with one single dataset in Article 3.

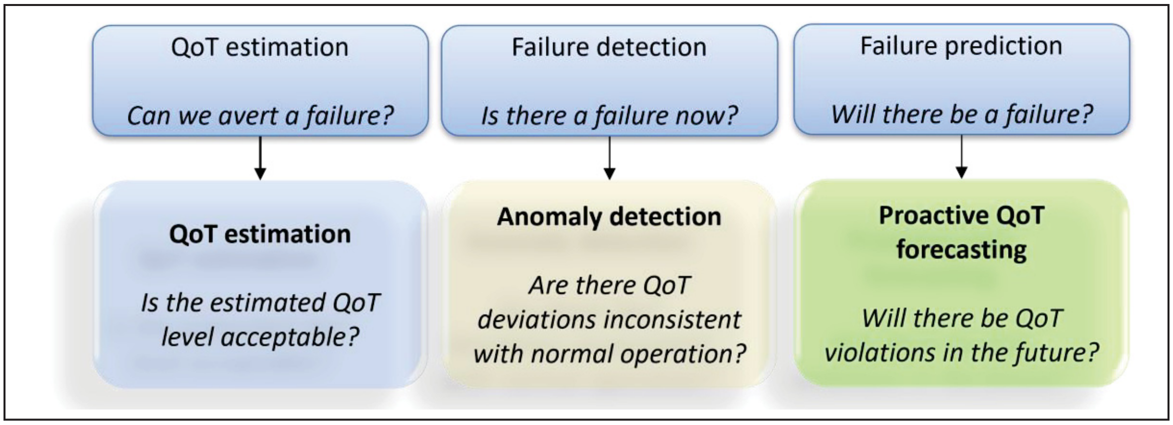


Figure 0.1 Overview of the thesis contributions addressing three of the optical network failure management use cases: QoT estimation, failure detection and failure prediction

Figure 0.1 shows the overview of the three contributions proposed in this thesis and their evolution adapted from the ONFM use cases aforementioned. Chapter 2 explains the role of each task, the reasoning behind its associated contribution, and how these contributions are tied together.

The mapping between research questions, ML paradigms, data contexts in the contributions, as well as the thesis chapters in which they are presented are summarized in Table 0.1.

Table 0.1 Mapping research questions to ONFM use cases, ML paradigms, data contexts, and thesis chapters

	RQ1 **Lightpath feasibility assessment**	**RQ2** **Anomaly detection**	**RQ3** **QoT forecasting**
ONFM use case	QoT estimation	Failure detection	Failure prediction
ML paradigm	Supervised learning XAI Domain adaptation	Unsupervised learning	Sequence modeling CPD + CAD
Data	Publicly available synthetic labeled QoT data	Multivariate time series field OPM datasets	Multivariate time series field OPM datasets
Thesis chapter	Chapter 3	Chapter 4	Chapters 5 and 6

0.4 List of publications

As mentioned in previous sections, the thesis is based on the three peer-reviewed publications presented in the previous section, as well as one additional chapter. Additional contributions on QoT management but not included in the thesis are also listed below.

0.4.1 Publications not included in the thesis

The publications listed below are not incorporated into the thesis as individual chapters. They reflect additional contributions made or co-authored by the candidate. They are included to provide a complete record of the research activities and to contextualize the thesis contributions within the Network Technology Lab's continuing research directions.

QoT estimation for lightpath feasibility assessment

Tremblay, C., Aladin, S., & Wosinska, L., (2026). Towards automated and interpretable ML for quality of transmission estimation in optical networks. In *Advanced Photonics Congress (APC) 2026* (Optica Publishing Group), Long Beach, California, July 26–30. [Invited]
This invited article presents a condensed version of Contribution 1, describing a lightweight, interpretable, and automated ML framework for QoT estimation. The proposed approach combines a RF classifier with Boruta-SHAP-based feature selection and domain adaptation techniques. Evaluated on four QoT datasets from the Fraunhofer Heinrich Hertz Institute (HHI), the framework achieves accuracy comparable to DL models while requiring lower computational resources and demonstrating strong cross-network generalization capability.

Allogba, S., Aladin, S., & Tremblay, C. (2022). Machine-learning-based lightpath QoT estimation and forecasting. *Journal of Lightwave Technology*, 40(10), 3115–3127. https://doi.org/10.1109/JLT.2022.3160379
The article provides a tutorial on ML methods for lightpath QoT estimation and short-term QoT forecasting, illustrating how both conventional models and advanced neural networks can

be applied to these tasks. It also presents a preliminary transfer learning study addressing the scarcity of field data.

Aladin, S., Tran, A. V. S., Allogba, S., & Tremblay, C. (2020). Quality of transmission estimation and short-term performance forecast of lightpaths. *Journal of Lightwave Technology*, 38(10), 2807–2814. https://doi.org/10.1109/JLT.2020.2975179
This article extends the work presented at ECOC 2019 in two directions: first, by introducing an artificial neural network (ANN)-based QoT estimator with feature engineering; and second, by incorporating encoder-decoder long short-term memory (ED-LSTM) and gated recurrent unit (GRU) forecasters alongside the LSTM model for prediction horizons up to 96 hours. The results show that GRU and LSTM modestly outperform the naive forecasting approach at longer horizons.

Diaz-Montiel, A. A., Aladin, S., Tremblay, C., & Ruffini, M. (2019). Active wavelength load as a feature for QoT estimation based on support vector machine. In *ICC 2019 — 2019 IEEE International Conference on Communications* (pp. 1–6). IEEE. https://doi.org/10.1109/ICC.2019.8761369
The article introduces active wavelength load and spectral position as input features for a multi-class support vector machine (SVM) model that predicts OSNR-based modulation format feasibility. The model is trained and evaluated using synthetic data generated by an emulator that captures Erbium-doped fiber amplifier (EDFA) power excursions. It achieves promising performance across different topological configurations.

Tremblay, C., & Aladin, S. (2018). Machine learning techniques for estimating the quality of transmission of lightpaths. In *2018 IEEE Photonics Society Summer Topical Meeting Series (SUM)* (pp. 237–238). IEEE. https://doi.org/10.1109/PHOSST.2018.8456791
This article is a condensed version of the work reported at OFC in 2018, comparing three classifiers using the GN-based cognitive QoT tool and synthetic BER data. The main contributions of this work are further developed and described in the subsequent publication.

Aladin, S., & Tremblay, C. (2018). Cognitive tool for estimating the QoT of new lightpaths. In *Optical Fiber Communications Conference and Exposition (OFC) 2018* (Optica Publishing Group), paper W2A.62.

This article proposes a QoT estimation tool that incorporates nonlinear impairments into the analytical OSNR formulation to generate a synthetic BER knowledge base. It then evaluates three classifiers, namely RF, SVM, and KNN, for lightpath classification. While SVM achieves the highest classification accuracy, it requires longer computation time compared with the other classifiers.

QoT forecasting

Allogba, S., Aladin, S., & Tremblay, C. (2021). Multivariate machine learning models for short-term forecast of lightpath performance. *Journal of Lightwave Technology*, 39(22), 7146–7158. https://doi.org/10.1109/JLT.2021.3110513

The study introduces multivariate LSTM and GRU models that incorporate OPM parameters and environmental features to improve short-term signal-to-noise ratio (SNR) forecasting over prediction horizons of up to 96 hours. The proposed models outperform univariate recurrent neural networks (RNNs) and persistence baselines. The work also demonstrates the applicability of transfer learning for short-term SNR prediction, showing effective generalization across lightpaths sharing the same fiber link and route.

Aladin, S., Allogba, S., Tran, A. V. S., & Tremblay, C. (2020). Recurrent neural networks for short-term forecast of lightpath performance. In *Optical Fiber Communication Conference (OFC) 2020* (Optica Publishing Group), paper W2A.24. https://doi.org/10.1364/OFC.2020.W2A.24

This study investigates gated RNN architectures, namely ED-LSTM, LSTM with attention mechanisms, and GRU. Using 18 months of field BER data, it analyses the impact of forecast horizon and observation window size on prediction accuracy. The results show that the benefits of RNN-based models over the naive baseline are modest and that prediction performance degrades rapidly beyond 12-hour horizons.

Tremblay, C., Allogba, S., & Aladin, S. (2019). Quality of transmission estimation and performance prediction of lightpaths using machine learning. In *45th European Conference on Optical Communication (ECOC 2019)* (pp. 1–3). IET. https://doi.org/10.1049/cp.2019.0757
This article presents an SVM-based QoT estimator trained on an expanded synthetic knowledge base, achieving promising classification performance. It also introduces an LSTM model trained on 13 months of field data to forecast SNR values from 2 to 24 hours ahead. The proposed LSTM model outperforms a naive baseline that assumes the SNR at a given time remains equal to the SNR at the previous time step.

0.5 Thesis outline

This thesis is organized to follow the progression from research context to methodology, to individual contributions and to synthesis. It outlines the research context and challenges, presents the research questions and objectives, and it identifies the open research questions. The three research articles, together with the systematic validation study are then presented in dedicated chapters, each positioned within the shared conceptual approach introduced in Chapter 2. The concluding chapters integrate the findings and relate them to the broader objective of ML-based QoT management.

Chapter 1 surveys the existing literature and identifies the open research questions that motivate the contributions. Chapter 2 outlines the research methodology, introduces the data sources and experiments environments, and the conceptual approach used in the thesis. Chapters 3 through 5 present the three peer-reviewed articles: Chapter 3 addresses QoT estimation for lightpath provisioning using an interpretable and efficient supervised learning framework with domain adaptation; Chapter 4 examines failure detection through a systematic comparison of unsupervised anomaly detection methods on production OPM data; Chapter 5 focuses on failure prediction using a CPD- and CAD-aware seq2seq forecasting framework that provides actionable reaction windows. Chapter 6 offers a systematic validation of the forecasting framework across three additional field lightpath datasets, two deep-learning architectures, and five CAD techniques, along with a preliminary operational evaluation of failure coverage, lead time, and false alarm rates. Chapter 7 synthesizes the results across all contributions and discusses their broader implications. Chapter 8 concludes the thesis and proposes directions for future research.

Collectively, these chapters show that progress toward proactive optical network management depends not on increasingly complex standalone models, but on the structured alignment of ML techniques with the requirements of the specific ONFM-based task they address. The next chapter reviews the state of the art in ML for smart optical network, establishing the technical and methodological foundations for the contributions that follow and identifying the specific research questions this thesis addresses.

CHAPTER 1

LITERATURE REVIEW

This chapter offers a critical analysis of prior work related to ML techniques for optical network management. It evaluates common methodological approaches, highlights their limitations, and positions the thesis contributions within the existing area of study.

The chapter is organized around the three ONFM-based tasks addressed in the thesis. It begins with a review of QoT estimation methods for lightpath provisioning, covering analytical modeling, simulation-based techniques, and ML-based solutions, focusing on domain adaptation and explainability. It then analyses research on anomaly detection in OPM data, with particular attention to the transition from threshold-based monitoring to unsupervised ML methods capable of operating without labeled failure data. Finally, it surveys QoT forecasting and proactive failure management strategies, including time-series methods and early-warning methodologies.

Across these areas, the chapter focuses on challenges such as domain shift, limited labeled data, interpretability requirements, and operational feasibility. It concludes by synthesizing the key research limitations that motivate the three contributions of this thesis.

Section 1.1 reviews QoT estimation for lightpath provisioning, highlighting how supervised learning with domain adaptation and explainability supports automated lightpaths feasibility assessment before their establishment. Section 1.2 explores anomaly detection for failure detection, where unsupervised learning techniques detect QoT degradation in production OPM data without labels. Section 1.3 addresses QoT forecasting for proactive failure prediction, showing how sequence modeling, change-point and collective anomaly detection, enable proactive failure management. Section 1.4 provides a synthesis of the ML concepts across the three ONFM-based tasks, highlighting shared challenges, recurring themes, and research limitations that the thesis addresses.

The organization around three distinct ONFM-based tasks, reflects both the operational reality that QoT estimation, anomaly detection and QoT forecasting occur at different stages of a lightpath's lifetime and involve different data contexts, and the methodological principle that paradigm selection should be guided by task requirements rather than the latest trends in ML algorithms.

1.1 QoT estimation for lightpath provisioning

Accurate QoT estimation during lightpath provisioning is crucial for reliable lightpath provisioning. Operators must assess whether candidate lightpaths will meet required QoT levels before allocating resources. This section reviews the evolution from traditional analytical and simulation-based methods to modern ML approaches, with particular attention to domain adaptation techniques that address the scarcity of data and explainability methods that strengthen operator confidence in automated QoT estimation.

1.1.1 Technical background: QoT fundamentals

QoT is a key concept in optical networks that allows to evaluate whether a transmitted signal can be correctly recovered at the receiver (Zhang, Li, Tang, Xin & Huang, 2022). Main QoT indicators include the OSNR, which compares signal power to accumulated amplified spontaneous emission (ASE) noise; the generalized signal-to-noise ratio (GSNR), which expresses the effective SNR of an optical channel by evaluating the ASE noise and the nonlinear interference (NLI) as a single value at the receiver, the BER, which reflects the likelihood of bit errors in the received data; and the Q-factor, a linear metric related to BER that summarizes overall transmission performance (Zhang et al., 2022).

A lightpath's QoT deteriorates as physical-layer impairments accumulate along its route. Linear impairments arise from fiber attenuation and component losses that necessitate amplification, CD that broadens optical pulses, PMD caused by fiber birefringence, and filtering effects introduced by ROADMs. At higher optical powers, nonlinear effects become prominent. Self-phase modulation (SPM), cross-phase modulation (XPM), and four-wave mixing (FWM) are all driven by the intensity-dependent variations in the fiber's refractive index, known as Kerr nonlinearity. The stimulated Raman scattering (SRS) and stimulated Brillouin scattering (SBS) are scattering-based nonlinearities. The combined impact of these impairments ultimately determines whether a lightpath can support error-free transmission at the intended data rate and modulation format (Keiser, 2021; Zhang et al., 2022).

During lightpath provisioning, QoT must be estimated before a lightpath is deployed. In new network deployments, operators usually have very few field data with lightpath characteristics for training ML models. As a result, ML-based QoT estimation built for such context relies on predictive methods that use link, system, and channel parameters such as route, wavelength, modulation format, and transmitted power along with transfer-learning strategies that allow models trained on synthetic data to adapt effectively to real-world environments with limited lightpath data.

1.1.2 Analytical and simulation-based QoT estimation techniques

Traditional QoT estimation mainly uses analytical models to approximate physical layer linear and nonlinear impairments. The two main types of analytical models are the comprehensive ones such as split-step Fourier method (SSFM) (Sinkin, Holzlöhner, Zweck, & Menyuk, 2003) and the simplified ones such as the Gaussian noise (GN) model (Poggiolini et al., 2014) which approximates nonlinear interference accumulation in coherent optical systems. Some extensions of the GN models such as enhanced GN (EGN) relax several simplifying assumptions to achieve higher accuracy (Carena et al., 2014). Generalized GN (GGN) variants account for frequency-dependent attenuation or gain and include stimulated Raman scattering (SRS) effects, making them suitable for modeling wideband wavelength division multiplexing (WDM) transmission (Cantono et al., 2018). Advanced analytical models, such as the inter-channel stimulated Raman scattering (ISRS) GN model, demonstrate remarkable efficiency, with Semrau, Killey, and Bayvel (2019a) achieving GSNR estimation using closed-form formulas with experimental validation in microseconds. Semrau, Sillekens, Killey, and Bayvel (2019b) extended this work with modulation-format-aware corrections, while Souza, Costa, Pedro, and Pires (2023) further confirmed that the closed-form ISRS-GN model achieves efficient QoT estimation suitable for real-time lightpath provisioning. Despite these advantages, these analytical models often require manual recalibration when applied to different network environments, or operating conditions. Wass, Thrane, Piels, Jones, and Zibar (2017) generated datasets using SSFM simulation and applied Gaussian process regression (GPR) to predict BER in WDM systems, demonstrating that supervised learning trained on SSFM-based data can effectively approximate nonlinear propagation behavior. Similarly, Rottondi, Barletta, Giusti, and Tornatore (2018) proposed using GN model simulation to generate synthetic training data for ML-based QoT prediction of unestablished lightpaths. Building on this foundation, the Gaussian Noise simulation in Python (GNPy) is an open source QoT estimation tool developed for lightpath analysis and network planning (Ferrari et al., 2020). It provides QoT estimation for WDM systems across diverse fiber types and network topologies. As with any analytical model, however, its accuracy depends on the precision of the physical-layer parameters provided as inputs to the tool. Many of these parameters are not consistently available, and deployed links are generally more heterogenous than laboratory testbeds, combining different fiber types and equipment from different vendors. Another limitation is scalability: in large-scale lightpath provisioning scenarios, where thousands of potential lightpath configurations must be evaluated, full simulations can quickly become impractical. Together, these accuracy and efficiency constraints motivate the development of faster data-driven methods that preserve high predictive accuracy while reducing reliance on complete and precisely characterized physical-layer parameters.

1.1.3 ML for QoT estimation

ML offers a promising path to overcoming the scalability limits of analytical models and simulation-based tools. ML models can be trained offline on large datasets, either generated through simulations or collected from production networks, to be used for rapid inference during provisioning phase. This enables the capability of achieving accuracy similar to simulation-based models with computational costs close to those of analytical models.

Several studies have explored ML application to the QoT estimation task using regression-based or classification-based approach. Support vector machine was investigated in our previous studies (Aladin & Tremblay, 2018; Tremblay & Aladin, 2018), demonstrating that kernel-based techniques can achieve competitive accuracy with small training datasets. Although these studies established the ability of lightweight models to achieve promising results in QoT estimation of lightpaths before their establishment, they did not address the computational efficiency, the interpretability nor the generalization aspects, essential for practical deployment.

DL-based QoT estimation has expanded rapidly. Khan et al. (2020a) demonstrated that neural networks can generalize to previously unseen optical networks. Manso, Vilalta, Muñoz, Casellas, and Martínez (2021) integrated ML-based QoT estimation with GNPy and software defined networking (SDN) controllers, showing scalability for cloud-native transport networks. Amirabadi, Kahaei, Nezamalhosseini, Arpanaei, and Carena (2023) extended DNN models to both single-mode and few-mode fiber systems, addressing mode division multiplexing (MDM) scenarios. These models can achieve high accuracy, but they come with significant limitations: high computational costs during both training and inference, large data requirements, and limited interpretability due to their black-box nature. While these works highlight the capability of neural networks, they rarely investigate computational efficiency relative to lightweight models or integrate techniques for interpretability. Moreover, Ibrahimi et al. (2021) conducted a systematic comparison of regression algorithms for QoT estimation of unestablished lightpaths, offering guidance on model selection based on prediction accuracy and computational complexity. Their findings showed that simpler models can perform competitively under certain conditions, challenging the assumption that increasingly complex neural networks always yield superior results.

A further research direction is emerging beyond individual QoT estimators: the integration of QoT estimation into agentic frameworks for autonomous network management. Zhang et al. (2025) introduced a large language model (LLM)-based agent built on generative pre-trained transformer (GPT-4) that combines domain knowledge with optical network tools. The proposed agent addresses four tasks, namely QoT estimation, performance analysis, optimization, and physical-layer parameter calibration. The authors evaluated the agent's performance using a predefined scoring scheme. This work exemplifies a broader trend in which LLMs serve as coordinators that invoke specialized estimation and optimization models to address specific network management tasks.

Compared with simulation techniques, ML models offer similar inference time once trained offline, making them appealing for integration into automated planning tools. Their predictive accuracy, however, depends heavily on the quality of the training data and the choice of relevant features. Moreover, increasing model complexity does not always yield proportional accuracy improvements, especially when datasets are limited or structured by simplified simulation assumptions.

Overall, existing work shows that ML-based QoT estimation is a practical and scalable alternative to purely analytical methods, particularly when rapid evaluation of many configurations is required. Still, real-world deployment challenges such as ensuring model interpretability and enabling adaptation across heterogeneous network environments remain important challenges, which we address in the following sections.

1.1.4 Explainability and interpretability in QoT estimation

Network operators need more than accurate ML-based QoT estimation models, they require insight into which network parameters influence the model's decisions. Black-box models that output numerical results without explaining their reasoning face significant barriers to adoption in production environments, where incorrect lightpath provisioning decisions can lead to service disruptions and financial loss. The absence of interpretability is widely recognized as a major obstacle in optical networks, with operators often favoring simpler, transparent models even at the cost of performance (Khan, 2024a; Ayoub et al., 2023). This reflects broader challenges in critical infrastructure, where difficulties to understand model behavior remain one of the major bottlenecks to advanced ML models adoption (Hassija et al., 2024).

In optical networks, operators require not only accurate predictions but also clear insight into *why* a lightpath is expected to succeed or fail. Knowing which parameters among length of the lightpath, number of spans, central carrier frequency, modulation format, or number of links that compose the lightpath most influence a prediction helps validate model behavior and supports informed decision-making, especially when predictions are borderline.

Lundberg and Lee (2017) introduced SHAP as a unified framework for model interpretability, based on cooperative game theory, that assigns each feature a contribution value for a specific prediction. Applied to QoT estimation, SHAP can highlight which network parameters most influence a lightpath's predicted QoT, offering transparent, per-instance explanations across diverse ML models.

Beyond individual explanations, global interpretability analysis helps operators understand which parameters generally matter most across many predictions. Tree-based models such as RF naturally provide these rankings, combining interpretability with the strong accuracy and computational efficiency valuable for operators using these models within planning tools. Global feature importance from tree-based models summarizes which features matter most overall, while operators also need instance-level reasoning. SHAP provides these local attributions by showing which features drive each prediction and in what direction, enabling validation of lightpath provisioning decisions. Together, global and local explanations provide the transparency required for production deployment.

Although complex models often achieve high accuracy, they typically sacrifice transparency. Recent work suggests this trade-off is not absolute: lightweight interpretable models can match or even outperform DNNs in some settings while remaining easier to understand. For QoT estimation, interpretability is often as critical as accuracy.

The importance of interpretability extends across telecommunications and other critical systems. Barredo Arrieta et al. (2020) emphasized that interpretability and transparency are essential when artificial intelligence (AI) influences high-impact decisions. In optical networks, Musumeci and Tornatore (2025) identified interpretability as an important consideration for ML-based management, and Rottondi et al. (2018) and Ayoub et al. (2023)

demonstrated how feature importance analysis can clarify which parameters shape optical network performance. Rottondi et al. (2018) analyzed lightpath QoT estimation with multiple parameters, exploring how route characteristics, wavelength assignment, modulation format, and launch power contribute to estimation accuracy, and how multiple QoT metrics can be combined. This highlighted the importance of feature selection for both performance and interpretability, although it did not incorporate modern XAI techniques for feature importance assessment. Bergk et al. (2022) investigated dataset quality for ML-assisted QoT estimation, introducing visualization techniques to assess training data characteristics and identify biases or coverage limitations. Their work highlighted that understanding what a model learns requires transparency not only into the model itself but also into the structure and overall coverage of the underlying data, linking interpretability directly to data quality. More recently, Amirabadi, Nezamalhosseini, Kahaei and Carena, (2025) demonstrated the relevance of explainable and low-complexity QoT models for optical communication networks. They combined a DL QoT classifier for single-mode and few-mode fiber links with SHAP-based feature importance analysis and a comparison of computational costs against both ML baselines and traditional GN and EGN models. Their results showed that a limited number of physical parameters, including modulation format, span length and number of spans, consistently dominate feature rankings. They further demonstrated that learned models can reduce inference time by approximately two orders of magnitude compared with closed-form EGN computations. Overall, these findings support the thesis assumption that lightweight and interpretable models are well suited for real-time network control.

Despite the broader rise of XAI, QoT estimation research has only minimally adopted established interpretability tools such as SHAP or local interpretable model-agnostic explanations (LIME) techniques (Ribeiro, Singh, & Guestrin, 2016). Most studies present predictions as sufficient outputs without offering operators insight into local and global explanations of the model’s decisions. The lack of integrated explainability limits practical deployment, where operator trust depends on understanding how and why a model arrives at its predictions.

1.1.5 Transfer learning for data scarcity

Another key challenge in ML-based QoT estimation is the scarcity of lightpath data in newly deployed networks. Operators planning new regions or infrastructures often have little to no lightpath data for training, and although simulations can generate large synthetic datasets, real deployments may differ in equipment, topology, and vendor-specific characteristics that simulations do not fully capture. Transfer learning and domain adaptation mitigate this issue by enabling models trained on large synthetic datasets to generalize to real deployment environments, even when only small amounts of data are available due to new deployment conditions or privacy and security concerns that limit data sharing.

Transfer learning enables models trained in a source domain to perform effectively in a related target domain by leveraging shared knowledge. Pan and Yang (2010) categorize transfer learning into instance-based, feature-representation-based, parameter-based, and relational-knowledge approaches. In optical networking, simulation data with abundant labels typically serves as the source domain, while real networks with limited or no labels, form the target domain. By adapting models to the characteristics of real deployments, transfer learning improves generalization even with minimal target domain data.

The paradigm has been widely explored in optical networking due to its practical relevance. Azzimonti, Rottondi, Giusti, Tornatore, and Bianco (2021) identified transfer learning as essential for bridging the gap between models trained on simulation data and real networks, especially when only a small amount of real data is available. Liu, Chen, Proietti, and Yoo (2021a) proposed an evolutionary transfer learning approach using a genetic algorithm to optimize the architecture and weight sets transferred across domains, demonstrating that selective transfer significantly reduces training data requirements while maintaining accuracy above 95%. Pesic, Lonardi, Seve, Rossi, and Zami (2020) highlighted the importance of unbiased training datasets, demonstrating that careful dataset construction improves transfer effectiveness.

Domain adaptation is a transfer learning technique that addresses distribution shift between training and deployment datasets. It allows ML models trained in a source domain to generalize to a new target domain both sharing the same feature space (Azzimonti et al., 2021). In QoT estimation, this helps address the common situation in which abundant synthetic simulation data is available, but operational data from new networks are scarce.

Recent work has introduced increasingly sophisticated domain adaptation strategies for QoT estimation. Usmani et al. (2022a, 2022b, 2022c) conducted extensive studies on transfer learning for QoT computation in 400ZR-based networks, demonstrating that models trained on C-band wavelength ranges can be transferred to extended C-band 400ZR networks with strong performance. Gu, Qin, Zhou, Zhang, and Ji (2023) proposed sample-distribution-matching techniques to align source and target feature distributions, improving adaptation quality. Cai et al. (2024) developed a domain adversarial adaptation for few-shot QoT estimation framework that achieved effective cross-domain transfer with limited target-domain supervision.

The success of transfer learning depends heavily on neural architecture design and is also sensitive to how precisely knowledge is transferred between models. Standard transfer-learning methods typically operate at the layer level, freezing or retraining whole layers, which can be too coarse to capture the varying relevance of individual weights across domains. Zhou, Gu, Zhang, and Ji (2023, 2024) explored a more granular technique that uses particle swarm optimization to select specific neurons to freeze or retrain, rather than applying a uniform decision to an entire layer. When applied to QoT estimation, their findings highlight that neuron-level method outperforms conventional layer-level transfer learning in adaptation accuracy.

Beyond simple model transfer, hybrid approaches combine transfer learning with other concepts. Azzimonti et al. (2020, 2021) compared active learning, selectively labeling the most informative target samples, with transfer learning for QoT estimation using small datasets. Tariq et al. (2023) introduced iterative transfer learning strategies that refine models through multiple adaptation cycles. Usmani, Khan, Mehran, Ahmad, and Curri (2024) integrated knowledge distillation with transfer learning for QoT estimation, compressing large teacher models into smaller student models while maintaining comparable accuracy across network domains.

Zuo et al. (2025) introduced a domain discrepancy feedback optimization (DDFO) method designed to improve QoT estimation when only a limited number of target-domain samples are available. Their approach uses a dual-branch neural network with shared weights to jointly learn from the source and target domains. An adaptive alignment layer quantifies covariance-based discrepancies between feature distributions, and the resulting discrepancy term is incorporated into the loss function. This enables the model to actively reduce domain mismatch during training rather than relying on static preprocessing techniques.

The strong focus on transfer learning for QoT estimation comes from a practical challenge: new network deployments often lack enough real training data, so models must rely on synthetic simulations to fill this gap. Most existing methods still assume access to some labeled data in the target domain for fine-tuning and reducing this labeling requirement while maintaining effective adaptation remains an active research area.

Despite notable progress, important challenges persist, such as integrating interpretability, designing lightweight models for resource-constrained planning, and validating adaptation methods across different network topologies.

Table 1.1 summarizes the ML-based QoT estimation solutions reviewed in this section, comparing them in terms of modeling approach, explainability, domain adaptation, and whether computational efficiency. The comparison shows that studies addressing domain adaptation rarely incorporate interpretability capabilities, whereas works focusing on explainability generally do not consider domain adaptation for data-scarce target domains. Furthermore, most approaches rely on DNN architectures, and no reported solution simultaneously achieves high accuracy, computational efficiency, interpretability, and adaptability.

Table 1.1 Summary of ML-based QoT estimation approaches and comparison with the proposed solution

Reference	Approach	Explainability	Domain adaptation	Efficiency evaluation	Key limitation
Aladin & Tremblay (2018); Tremblay & Aladin (2018)	Cognitive QoT classification of unestablished lightpaths (K-NN, RF, SVM)	No	No	Yes	No XAI; generalization not addressed
Rottondi et al. (2018)	Feature analysis for lightpath QoT estimation (route, wavelength, modulation, power)	No	No	Yes (training and per-instance classification times)	No explainability; no adaptation
Khan et al. (2020a)	DNN generalizing to unseen networks	No	No	No	Black-box model; no explainability; no generalization
Ibrahimi et al. (2021)	Comparison of regression algorithms	No	No	Yes (training time)	No explainability, or adaptation integration
Manso et al. (2021)	ML-based estimation integrated with GNPy and SDN controllers	No	No	No	Focus on scalability; interpretability not investigated
Bergk et al. (2022)	Dataset-quality assessment and visualization for ML-assisted QoT estimation	No, but data-level transparency	No	No	Model-level explainability not addressed
Amirabadi et al. (2023)	DNN for single-mode and few-mode (MDM) systems	No	No	Yes (runtime comparison)	Black-box model; no XAI or adaptation
Azzimonti et al. (2020, 2021); Pesic et al. (2020)	Transfer and active learning with small datasets; unbiased dataset construction	No	Yes	No	Interpretability not integrated with adaptation
Liu et al. (2021a)	Evolutionary transfer learning (genetic algorithm)	No	Yes	No	Interpretability not addressed

Table 1.1 Summary of ML-based QoT estimation approaches and comparison with the proposed solution (continued)

Reference	Approach	Explainability	Domain adaptation	Efficiency evaluation	Key limitation
Usmani et al. (2022a-2022c, 2024); Tariq et al. (2023)	Transfer learning for 400ZR networks; active transfer learning (2023); parameter-efficient transfer (2024)	No	Yes	Yes (2024: training and prediction times)	Explainability not integrated with transfer learning
Gu et al. (2023); Zhou et al. (2023, 2024)	Sample-distribution matching; neuron-level transfer	No	Yes	Yes (2024: training and inference times)	Interpretability not explored; ensemble transfer increases training cost
Cai et al. (2024)	Domain-adversarial adaptation for few-shot QoT estimation	No	Yes	No	Requires limited target-domain supervision; interpretability not addressed
Zuo et al. (2025)	Domain discrepancy feedback optimization	No	Yes	Yes (inference times)	DNN-based; explainability not addressed
Amirabadi et al. (2025)	Low-complexity DL classifier with SHAP analysis (SMF/FMF)	Yes (SHAP)	No	Yes (inference time)	No adaptation to data-scarce target domains
Zhang et al. (2025)	LLM-based agent (GPT-4) coordinating QoT estimation and optimization tools	No	No	No	Agentic coordination framework; not a lightweight, interpretable estimator
This thesis (Contribution 1)	Lightweight RF combining automated SHAP-based feature selection with domain adaptation across networks	Yes (global + local SHAP)	Yes	Yes (training and inference times)	Jointly addresses accuracy, computational efficiency, interpretability, and adaptability

1.1.6 Open research questions in QoT estimation

The literature review highlights substantial progress in ML-based QoT estimation, particularly in the development of transfer learning techniques that address data scarcity in new network deployments. However, several limitations continue to restrict practical deployment. Although transfer learning methods have advanced considerably and explainability tools such as SHAP are well established, their integration remains limited. Operators require models that can both adapt from large synthetic datasets to environments with minimal real data and provide transparent, interpretable predictions. How explainability interacts with domain adaptation performance remains largely unexplored.

Recent work has also focused heavily on DNN architectures, often assuming that high accuracy necessitates increased model complexity. In contrast, lightweight and inherently interpretable models such as RF, which offer lower computational overhead, have received comparatively little attention within transfer learning frameworks. Whether such models can deliver competitive accuracy while maintaining efficiency and transparency remains an open question.

More broadly, existing research tends to optimize individual objectives such as accuracy through DL, efficiency through analytical models, adaptability through transfer learning, or interpretability through simpler methods, without demonstrating that these objectives can be achieved simultaneously. In practice, operators require solutions that combine competitive accuracy, computational efficiency for large-scale planning, adaptability to environments with limited data availability, and transparent decision-making.

Taken together, these observations indicate that while ML-based QoT estimation is feasible and transfer learning offers strong potential for data-constrained scenarios, important challenges remain. These include the integration of explainability with transfer learning, the exploration of lightweight models for domain adaptation, the validation of methods under genuinely data-scarce deployment conditions, and the development of approaches that balance accuracy, efficiency, adaptability, and interpretability within a unified framework. Addressing these limitations requires demonstrating that appropriately designed lightweight, interpretable models equipped with transfer learning capabilities can achieve competitive performance across multiple dimensions rather than excelling in only one sacrificing others.

1.2 Anomaly detection for failure detection

Once lightpaths are established and carrying traffic, continuous monitoring is essential to ensure they maintain acceptable performance over time. Traditional threshold-based monitoring is simple to implement but cannot adapt to lightpath-specific characteristics, often missing subtle degradations or triggering false alarms from normal fluctuations. ML-based anomaly detection can potentially offer a more powerful alternative by identifying complex patterns linked to performance issues, yet it faces a major challenge in production networks where labeled anomaly data is scarce. This section reviews the fundamentals of optical performance monitoring, existing anomaly detection strategies, and the shift toward unsupervised ML methods capable of operating effectively without extensive labeled datasets.

1.2.1 Technical background: optical performance monitoring

OPM refers to the techniques used to measure QoT parameters in production optical networks. Dong et al. (2016) provide a broad overview of OPM technologies for both direct detection and coherent systems, covering different monitored parameters and discussing future challenges for flexible and elastic optical networks. Modern coherent receivers now expose rich signal information through their digital signal processing (DSP) subsystems, enabling software-based monitoring without dedicated hardware. Commonly monitored parameters include optical power transmitted and received, OSNR estimated from the received signal, CD, differential group delay (DGD) from PMD, and pre-FEC BER as a direct indicator of the QoT.

In production networks, OPM data is collected periodically, typically every 15 minutes to 24 hours and stored in network management systems. Large networks with hundreds of active lightpaths therefore generate extensive multivariate time-series data, capturing how performance metrics evolve over time due to environmental conditions, component aging, and network reconfigurations.

1.2.1.1 Traditional threshold-based monitoring.

Conventional network monitoring relies on static parameter thresholds (OSNR below a fixed limit or OPR outside a defined window). Early work, such as Skoog et al. (2006) using SVM to classify impairment type from eye-diagram features, showed the potential of such approach in impairment detection, yet threshold-based methods remain dominant because of their simplicity and interpretability. However, they suffer from key limitations: thresholds must be manually tuned, they fail to account for lightpath-specific behavior, they cannot detect multivariate anomalies where individual parameters remain within limits, and they miss gradual soft failures that degrade performance before thresholds are crossed.

The combination of dense multivariate OPM data and the limitations of threshold-based monitoring motivates the use of ML methods capable of detecting complex, subtle anomalies without requiring manual threshold engineering for every parameter and lightpath.

1.2.2 Anomaly detection fundamentals

Anomaly detection, also known as outlier detection, focuses on identifying observations that deviate significantly from expected behavior. In optical networks, such deviations may signal performance degradation, component faults, configuration errors, or environmental disturbances that require operator attention.

1.2.2.1 Taxonomy of anomalies.

Chandola, Banerjee and Kumar (2009) introduced a widely used taxonomy distinguishing three main anomaly types. Point anomalies are defined as individual data instances that are clearly abnormal, such as a single OSNR reading far below the normal range. These are the easiest to detect and are typically captured by threshold-based monitoring. Contextual anomalies describe values that are unusual only within a specific context. For example, dispersion levels that are acceptable on long-haul links but abnormal on metro links, or performance drops that are expected during maintenance windows but concerning otherwise are considered contextual anomalies. Detecting these requires contextual awareness. Collective anomalies are groups of observations that appear normal individually but form an abnormal pattern together, such as a slow, sustained rise in pre-FEC BER over several days. These reflect temporal or structural deviations rather than isolated points.

Traditional threshold-based monitoring primarily captures point anomalies but leaves contextual and collective anomalies, often early indicators of emerging issues, undetected.

1.2.2.2 Supervised, unsupervised, and semi-supervised anomaly detection methods.

Anomaly detection approaches differ based on the availability of labeled data. Supervised methods train on labeled examples of both normal and anomalous data to learn decision boundaries that separate these classes (Chandola et al., 2009; Pang, Shen, Cao & Van Den Hengel 2021). While effective when labels exist, they are difficult to apply in optical networks where anomalies are rare, inconsistently documented, and often lack precise temporal extent of the affected periods (Abdelli et al., 2022a; Musumeci & Tornatore, 2025). Unsupervised methods require no labels, instead learn the structure of normal operations and flag deviations (Pang et al., 2021). This aligns well with operational networks, where normal data dominates and anomalies are scarce. Semi-supervised methods typically learn from normal data only, assuming unlabeled data is mostly normal, and detect deviations from this learned normal model (Pang et al., 2021). Limited anomaly labels, when available, are used to refine detection.

1.2.2.3 Challenges in anomaly detection.

Applying anomaly detection to production optical networks introduces challenges that extend beyond those encountered in generic network anomaly detection settings. Although general surveys such as Chandola et al. (2009) and Ahmed et al. (2016) classify anomaly detection techniques and discuss common issues in high dimensional or imbalanced datasets, optical networks impose additional domain-specific constraints.

First, OPM data encompasses several variables. Each lightpath is described by multiple physical-layer indicators such as Q-factor, optical power received, and BER, and production networks may contain several active lightpaths (Dong et al. 2016). This high dimensional structure creates computational and modeling challenges where the curse of dimensionality can make it difficult to reliably distinguish normal from abnormal behavior as the feature space dimension increases.

Second, anomalies in optical networks often evolve gradually over time rather than appearing as abrupt events. Soft failures typically manifest as progressive degradation over hours or days, requiring models capable of capturing sequential patterns rather than treating anomalies as isolated points (Musumeci et al., 2019b; Musumeci & Tornatore, 2025). This distinguishes operational anomaly detection from static outlier analysis and motivates approaches that incorporate sequential or temporal modeling (Behera, Panayiotou & Ellinas, 2023a).

Third, anomaly types in optical networks are diverse. They may arise from sudden component failures, slow aging effects, spectral interference, configuration changes, or environmental fluctuations. These phenomena exhibit different statistical signatures, making it challenging for a single detection strategy to perform consistently across all scenarios (Pang et al., 2021; Musumeci & Tornatore, 2025).

Fourth, labeled fault data is scarce. In production environments, comprehensive ground-truth labeling of anomalous events is rarely available, particularly for subtle degradations (Musumeci et al., 2019b; Abdelli et al., 2022a; Musumeci & Tornatore, 2025). This scarcity favors unsupervised or semi-supervised approaches that learn normal behavior and detect deviations without relying on extensive labeled anomaly samples (Pang et al., 2021; Abdelli et al., 2022a).

Finally, anomaly detection systems must generate interpretable alerts. Unlike offline analytics, network monitoring supports real-time decision-making, and operators must understand why a lightpath has been flagged as anomalous to validate alarms and initiate corrective actions. Consequently, transparency is essential for practical deployment (Musumeci & Tornatore, 2025; Barredo Arrieta et al., 2020).

Collectively, these challenges underscore the need for anomaly detection frameworks designed for multivariate, temporal, and operational characteristics of optical networks.

1.2.3 Unsupervised learning methods for anomaly detection

Given the scarcity of labeled anomalies in operational optical networks, unsupervised learning has become increasingly important for intelligent monitoring. Four major paradigms namely density-based, boundary-based, isolation-based, and reconstruction-based provide complementary ways to detect anomalies without labeled training data.

- **Density-based: DBSCAN and HDBSCAN**

Density-based spatial clustering of applications with noise (DBSCAN), introduced by Ester, Kriegel, Sander, and Xu (1996), identifies clusters based on point density rather than minimizing distance to centroids as in K-means. Points in dense regions form clusters, while those in sparse regions are flagged as noise or outliers. For OPM data, DBSCAN can detect lightpaths with unusual parameter combinations that fall outside the dense clusters representing normal behavior. It handles irregular cluster shapes and does not require specifying the number of clusters, though it is sensitive to parameter choices and struggles with datasets containing clusters of varying densities.

Hierarchical density-based spatial clustering of applications with noise (HDBSCAN) builds on DBSCAN by constructing a hierarchy of density-based clusters and then selecting the most stable groups across different density thresholds (Campello, Moulavi, Zimek & Sander 2015). This approach allows it to identify clusters with varying densities, making it well-suited for OPM data that exhibit diverse anomaly patterns. It preserves DBSCAN's strengths, such as handling arbitrarily shaped clusters and avoiding a predefined number of clusters yet comes with higher computational cost and reduced interpretability in high-dimensional scenarios.

- **Boundary-based: One-Class SVM**

Schölkopf, Platt, Shawe-Taylor, Smola, and Williamson (2001) proposed the OCSVM to learn the boundary of normal data when only normal samples are available. By mapping data into a high-dimensional feature space via kernel functions, OCSVM constructs a boundary that encloses the normal region; points outside this boundary are treated as anomalies. For optical networks, OCSVM can model the multivariate "normal operating envelope" defined by OPM parameters. Its strengths include flexible nonlinear boundary modeling, though it can be computationally expensive and sensitive to kernel and hyperparameter choices.

- **Isolation-based: Isolation Forest**

Isolation Forest, introduced by Liu, Ting, and Zhou (2008), takes a fundamentally different approach: it isolates anomalies rather than modeling normal behavior. Random decision trees recursively partition the feature space, and anomalies being rare and different require fewer splits to isolate. The average path length across trees yields an anomaly score. This method scales well to high-dimensional OPM data, requires minimal tuning, and is computationally efficient. Its main limitation is reduced effectiveness when anomalies form dense clusters rather than appearing as isolated points.

- **Reconstruction-based: Autoencoder**

Autoencoders were first proposed for unsupervised representation learning (Rumelhart, Hinton & Williams, 1986) and later became a popular tool for dimensionality reduction (Hinton & Salakhutdinov, 2006). They learn compact latent features by attempting to reconstruct their inputs. When trained exclusively on normal operating conditions, they capture typical system behavior, allowing anomalies to be flagged through unusually high reconstruction errors. This makes them a strong fit for OPM data, where complex nonlinear interactions exist among many performance indicators. Autoencoders handle high-dimensional data well and do not require labeled anomalies, though they can be sensitive to hyperparameter choices, susceptible to overfitting, and less interpretable than simpler approaches.

- **Complementary strengths**

These approaches excel in different scenarios: DBSCAN and HDBSCAN identify outliers far from dense clusters, with HDBSCAN better handling datasets of varying density; OCSVM learns robust boundaries around normal operating regions; Isolation Forest efficiently detects structurally unusual observations; and Autoencoders capture complex multivariate dependencies through reconstruction error.

1.2.4 Anomaly detection in optical networks

The optical networking community has increasingly adopted ML for anomaly detection and performance monitoring, moving from supervised methods that require labeled data toward unsupervised and semi-supervised techniques better aligned with production network environments.

Chen, Li, Proietti, Zhu, and Yoo (2019) proposed a self-taught anomaly detection framework that first clusters OPM data using unsupervised learning, then applies limited supervised labeling to interpret those clusters. This hybrid strategy reduces labeling requirements while still incorporating operator expertise, and it outperforms purely supervised or purely unsupervised methods when a small number of reliable labels are available. Liu, Wang, Zhang, Wang, and Zhang (2021b) addressed anomaly detection under severe class imbalance, a common challenge in optical networks where failures are rare. Their semi-supervised approach adapts to the scarcity of anomaly labels and improves detection performance compared to methods that ignore imbalance, highlighting the importance of modeling real-world data distributions. Behera et al. (2023a) developed a framework for detecting and localizing soft failures, also known as gradual degradations, that do not immediately trigger alarms but signal emerging problems. By combining anomaly detection with localization, the method identifies both the presence and underlying location of soft failures, enabling proactive maintenance before outages occur.

Du, Côté, Barber, and Liu (2021) proposed another approach forecasting loss of signal in optical networks using ML, marking a transition from reactive anomaly detection to forecasting failures based on historical patterns. By extracting anomalies in historical telemetry preceding loss of signal, these models can issue early-stage alerts that facilitate pre-emptive operational interventions. Furthermore, Allogba, Yameogo, and Tremblay (2022b) focused on early detection of anomalies in lightpath SNR using feature-engineering techniques such as time-series decomposition and statistical feature selection. By identifying subtle deviations before they escalate, their approach provides operators with intervention windows before service disruption. Tremblay et al. (2023) extended anomaly detection to include identifying the underlying causes of performance degradation. Their work shows that ML can not only detect anomalies but also help pinpoint the responsible components or conditions, supporting more effective troubleshooting. Wan et al. (2021) leveraged correlations among network-level key performance indictors (KPIs), detecting anomalies when normal relationships between parameters break down, even if individual parameters remain within acceptable ranges. This approach reflects the inherently multivariate nature of OPM data.

Khan et al. (2024) examined the practical limitations of supervised failure management by highlighting the severe class imbalance typically observed in production optical networks. In real-world deployments, critical failures occur rarely, resulting in limited failure samples that can compromise the reliability of supervised learning models. They investigated how model-centric and data-centric approaches can improve neural network-based failure identification under realistic data scarcity conditions. More recently, Ali et al. (2026) presented the first systematic side-by-side comparison of pre-, in-, and post-processing techniques for mitigating class imbalance in failure detection and identification. Using an experimental dataset, they showed that post-processing approaches, such as threshold adjustment, provided greater improvements for failure detection, whereas generative AI-based methods achieved better performance for multi-class failure identification.

Table 1.2 presents a summary of the anomaly detection approaches reviewed in this section, organized by learning paradigm. The comparison shows that most existing works either depend on labeled failure data, through supervised or semi-supervised learning, or evaluate a single anomaly detection approach. Comparisons of multiple unsupervised methods on the same production dataset remain absent from the literature.

Table 1.2 Summary of ML-based anomaly detection approaches for optical networks and comparison with the proposed solution

Reference	Learning paradigm	Approach	Key limitation
Chen at al. (2019)	Hybrid (unsupervised + supervised)	Self-taught framework with data labeling in a clustering module to train a DNN for online detection	Single density-based detection; simulation data
Liu et al. (2021b)	Semi-supervised	Autoencoder detection trained on normal data only, under severe class imbalance	Single reconstruction-based detection
Wan et al. (2021)	Hybrid (supervised + unsupervised clustering)	Correlative prediction of network indicators, with anomaly detection in the predicted series	Network-level data and single density-based detection
Du et al. (2021)	Supervised	Forecasting ILOS 1-7 days ahead using different techniques	Rule-based failure labels for supervised anomaly prediction
Allogba et al. (2022b)	Supervised with statistical labeling	Interquartile range (IQR)-based anomaly definition and early detection of lightpath SNR anomalies	Labels derived from a statistical IQR rule; supervised anomaly prediction
Behera et al. (2023a)	Supervised	Soft-failure evolution forecasting for proactive hard-failure detection + failure localization	Simulated data with EDFA gain degradation
Tremblay et al. (2023)	Supervised (multiclass)	Root-cause classification of failures on field data	Rule-based labeling from alarms and domain knowledge for root-cause classification
Khan et al. (2024b)	Supervised	Model-centric and data-centric strategies for failure identification under class imbalance	Experimental data for supervised failure identification

Table 1.2 Summary of ML-based anomaly detection approaches for optical networks and comparison with the proposed solution (continued)

Reference	Learning paradigm	Approach	Key limitation
Ali et al. (2026)	Supervised (+ generative AI)	Comparison of pre-, in-, and post-processing class imbalance mitigation	Experimental datasets for supervised failure identification
This thesis (Contribution 2)	Unsupervised	Comparison of DBSCAN, OCSVM and IF techniques on a multivariate production OPM dataset	Anomaly detection validated on production data without labels

1.2.5 Open research questions in anomaly detection

This literature review demonstrates strong progress in ML-based anomaly detection for optical networks, with increasing recognition that the scarcity of labeled anomalies makes unsupervised and semi-supervised approaches essential. Still, several important challenges remain:

Although a variety of unsupervised algorithms, such as clustering-based, boundary-based, and isolation-based techniques has been applied to OPM data, systematic evaluations across different paradigms on the same production network datasets remain scarce. This makes it challenging to determine which detection approaches are best suited to the wide range of anomaly behaviors found in production optical networks, where point deviations, gradual degradations, and collective anomalies can appear very differently across the multivariate OPM feature space.

Validating anomaly detection methods in production optical networks is difficult because operational data is proprietary and testing in live systems carries risk. As a result, many studies rely on synthetic or simulation datasets, controlled testbeds, or short field trials rather than long-term evaluations in production networks. Although controlled testing is useful for demonstrating specific failure scenarios, it does not fully reflect the range of anomaly behaviors, realistic noise characteristics, and subtle degradation patterns seen in real operational data. This gap between controlled research testing and deployment in actual networks limits confidence in how well these methods will generalize to real-world conditions.

OPM produces multivariate time-series data such as Q-factor, OPR, CD, and PMD, whose relationships provide important diagnostic information. However, many detection approaches still rely on single-parameter thresholds or treat parameters independently, missing the value of multivariate relationships. Methods that do use multiple parameters often depend on supervised learning with labeled failures or on deep-learning models that require significant computational resources. The potential of unsupervised multivariate methods that can capture parameter correlations while still supporting real-time processing remains underexplored.

Different unsupervised learning paradigms rely on distinct detection principles: density-based methods flag points in low-density regions, boundary-based methods learn the limits of normal behavior, and isolation-based methods identify observations that are easy to separate from the rest. These differences suggest that each paradigm may be better suited to certain anomaly types, for example, abrupt point anomalies versus gradual degradations or collective anomalies where individual parameters can appear normal but their combination is abnormal. Yet little research systematically examines these complementary strengths or offers practical guidance on choosing methods based on network characteristics and operational needs.

Beyond accuracy, deploying anomaly detection in production optical networks requires attention to computational efficiency, interpretability, and integration with existing operational workflows. Supervised methods can perform well when labeled data is available, but they require ongoing updates as network conditions change and new failure types appear. Unsupervised methods offer more flexibility but are less well understood in terms of performance across different anomaly types and operational environments. Study that evaluates unsupervised methods on real production data considering both detection performance and practical deployment requirements would provide valuable guidance for operational adoption.

1.3 QoT forecasting for failure prediction

Proactive network management reflects the transition from reacting to failures after they occur to anticipating and preventing issues before service quality deteriorates. QoT forecasting provides the predictive capabilities required for this shift, allowing operators to identify emerging degradation with enough lead time to implement preventive measures. This section reviews the evolution from traditional time-series forecasting techniques to modern deep-learning-based approaches, with particular attention to change-point and collective anomaly detection as early-warning mechanisms for signaling degradations ahead of critical failures.

1.3.1 Technical background: proactive vs. reactive QoT management

Traditional optical network operations rely on a reactive model: operators intervene only after problems surface as alarms, service degradation, or total failures. Although reactive measures such as rerouting traffic, replacing components, or adjusting operating parameters, can restore service, they essentially occur *after* disruption, risking SLA violations and impact on customers. Emergency interventions made under time constraints may not be as effective as planned maintenance during scheduled windows. And the inability to anticipate failures limits strategic resource allocation and proactive optimization.

1.3.1.1 Proactive management and reaction windows

Proactive management seeks to predict issues early enough to act before service is affected, creating a "reaction window" for preventive intervention. International standards such as Recommendation ITU-T G.827 (International Telecommunication Union, 2003), Recommendation ITU-T G.826 (International Telecommunication Union, 2002), and Recommendation ITU-T G.8201 (International Telecommunication Union, 2011) define availability and error-performance objectives for optical paths, highlighting the importance of predictive capabilities. Network operators offer services with different availability requirements, ranging from about 99.9% ("three nines") for standard enterprise services to

99.999% ("five nines") for carrier-grade telecommunications (Ramaswami et al., 2010; Gour, Ishigaki, Kong & Jue, 2021). These availability targets directly influence how thresholds are set for identifying critical QoT degradations. For example, any degradation that could lead to more than roughly 26 seconds of downtime per month, the limit under a five-nines SLA, would be treated as a critical event that must be predicted in advance.

An essential requirement for effective proactive QoT management is having a reliable reaction window - or lead time - defined as the interval between an early indication of degradation and the moment a service disruption occurs. Whereas reactive methods initiate intervention only after thresholds are breached, proactive approaches depend on forecasting models that can anticipate performance decline far enough in advance to allow operators to act. In optical networks, Murphy, Lavignotte, and Lepers (2023) explicitly frame fault prediction around this notion of lead time, emphasizing that predictive accuracy must be paired with sufficient advance warning time for operational response. In practice, a meaningful reaction window enables operators to reroute traffic, adjust resource configurations, or schedule maintenance before customers experience disruption. As a result, proactive management systems must be evaluated not just on forecasting precision but also on the length and reliability of the lead time they deliver, since this early-warning time ultimately determines their operational value.

1.3.1.2 Challenges in predictive management

Forecasting future network conditions is more complex than detecting current anomalies. Prediction accuracy declines as the forecast horizon increases, and optical networks exhibit diverse temporal behaviors: long-term trends for aging components, periodic environmental cycles for temperature fluctuations, abrupt configuration changes, and intermittent disturbances linked to environmental events (Allogba, Aladin & Tremblay, 2021; Yaméogo et al., 2020). Capturing these patterns to extrapolate future behavior requires advanced time-series forecasting. Moreover, distinguishing harmless fluctuations from early signs of systematic degradation demands methods that can recognize subtle but meaningful temporal shifts.

1.3.2 Traditional time-series forecasting methods

Classical time-series forecasting has a long history across economics, meteorology, and operations research, and these foundations also inform early approaches to network performance prediction.

1.3.2.1 Statistical approaches

Box, Jenkins, Reinsel, and Ljung (2015) consolidated the core framework for autoregressive integrated moving average (ARIMA) modeling, which represents a time series through linear combinations of past values (autoregressive component), past forecast errors (moving average component), and differencing to achieve stationarity (integrated component). Seasonal ARIMA extends this to periodic patterns. Hyndman and Athanasopoulos (2021) provided a modern, practical treatment of forecasting methods including exponential smoothing, state space models, and ARIMA variants, with emphasis on practical implementation and model evaluation.

However, ARIMA and related statistical techniques face notable limitations for QoT forecasting. Their linearity assumptions often fail to capture the nonlinear behavior of physical layer impairments, component aging, and environmental effects. Model specification requires expert tuning of AR and MA orders, and extending these methods to multivariate OPM data with many correlated parameters becomes complex. Most importantly, classical models struggle with the diverse temporal patterns seen in optical networks because they are designed for stationary or slowly-varying processes.

These limitations have motivated the shift toward ML-based forecasting methods that can automatically learn complex temporal dynamics without relying on explicit linear assumptions or manual model design.

1.3.3 Deep learning for time-series forecasting

DL architectures especially RNNs and temporal convolutional networks (TCNs) have shown strong ability to learn complex temporal dependencies in time-series forecasting across many domains.

- **Long Short-Term Memory (LSTM) Networks**

Hochreiter and Schmidhuber (1997) introduced LSTMs to overcome the vanishing gradient limitations of earlier RNNs. Their key innovation is the gated cell structure: input gates control what new information enters the cell state and output gates determine what influences the prediction. Gers, Schmidhuber, and Cummins (2000) subsequently added forget gates, enabling the network to actively reset cell state and remove outdated information. Together, these three gating mechanisms enable LSTMs to capture long-term dependencies, handle variable-length sequences, and learn which historical patterns matter for forecasting.

For optical networks, LSTMs naturally model the temporal structure of OPM data, maintain internal state across long historical context, and can stack multiple layers to capture patterns at different time scales.

- **Temporal Convolutional Networks (TCNs)**

Bai, Kolter, and Koltun (2018) showed that convolutional architectures, when designed with causal and dilated convolutions, can match or outperform recurrent models for sequence tasks. TCNs use causal convolutions to ensure predictions rely only on historical data, and dilation expands the receptive field efficiently to capture long-term dependencies without recurrent connections.

TCNs offer practical advantages: full parallelization across time steps, stable gradients, and better scalability for long sequences. Their ability to capture multi-scale temporal patterns aligns well with the diverse dynamics of optical networks, from short-term fluctuations to long-term trends.

- **Deep learning for QoT forecasting**

DL has been increasingly adopted for QoT and performance forecasting. Mezni, Charlton, Tremblay, and Desrosiers (2020) demonstrated multi-step forecasting in operational networks using a 1D convolutional neural network trained on field BER data, while Chouman, Djukic, Tremblay, and Desrosiers (2021) applied multilayer perceptron (MLP) and LSTM models specifically to lightpath QoT prediction. Hedayatnejad, Pei, Boertjes, Demeter, and Desrosiers (2025) explored multi-head attention transformers (Transfomer-XL) for span-loss forecasting, leveraging an architecture whose attention mechanism can highlight the most influential historical patterns. Du et al. (2021) used different ML techniques to forecast loss of signal, illustrating the potential for predicting failures before they occur.

- **Foundational QoT forecasting work by the thesis author**

Prior work by the thesis author and collaborators established early foundations for ML-based QoT forecasting. Tremblay et al. (2019) combined QoT estimation with performance prediction; Aladin et al. (2020b) demonstrated short-term forecasting using recurrent networks; Allogba et al. (2021) extended this to multivariate models incorporating multiple OPM parameters; and Allogba, Aladin, and Tremblay (2022a) provided a comprehensive treatment of ML-based QoT estimation and forecasting. Aladin et al. (2020a) validated LSTM and GRU architectures for short-term lightpath performance prediction.

This body of work confirmed the promise of DL while highlighting challenges such as accuracy degradation over longer horizons, difficulty distinguishing meaningful degradation from benign fluctuations, and the need for interpretable early-warning signals.

- **Uncertainty quantification**

Acknowledging that single point forecasts are inadequate for operational contexts that must account for risk, two parallel studies introduced uncertainty-aware QoT prediction for active lightpaths. Di Cicco, Talpini, Ibrahimi, Savi, and Tornatore (2023) developed a multi-input, multi-output seq2seq LSTM enhanced with variational dropout, enabling the model to produce probabilistic forecasts with near-optimal empirical coverage on the Microsoft wide area network (WAN) backbone dataset. Moreover, Yousefi et al. (2023) evaluated MLP, LSTM, and N-Beats forecasters on a production dataset, selecting MLP as the baseline for uncertainty estimation via Monte Carlo dropout, using quantile regression as a reference method and calibration applied to mitigate the miscalibration typical of Bayesian approximations. Collectively, these contributions position probabilistic QoT forecasting as a well-established research direction, shifting operational practice toward calibrated confidence intervals that support risk-aware decision-making.

1.3.4 Change point detection and collective anomaly detection

Proactive management benefits not only from forecasting future values but also from early-warning indicators that reveal when systematic degradation starts to develop. Change point detection and collective anomaly detection serve this role by identifying temporal patterns that signal regime shifts or progressive deterioration.

- **Change point detection fundamentals**

Change points mark moments when a time series undergoes a significant shift, such as changes in mean, variance, trend, or other statistical properties. In optical networks, these shifts may correspond to configuration updates, the onset of component degradation, or environmental changes. Detecting change points provides early indications that the network has entered a new behavioral regime.

Truong, Oudre, and Vayatis (2018, 2020) introduced *ruptures*, a widely used Python library implementing algorithms such as binary segmentation, dynamic programming, and kernel-based methods for offline change-point detection. Flynn and Yoo (2019) demonstrated kernel-cumulative sum (kernel-CUSUM) for distribution-free change detection without parametric assumptions, while Wei and Xie (2022) extended this to online detection with memory-efficient recursive computation for real-time monitoring. Xu, Song, Wu, Wang, and Zhou (2025) surveyed deep-learning-based change-point detection, highlighting how neural models can learn regime transitions directly from data.

- **Change-point detection in optical networks**

Mosayebi and Lampe (2024) applied change point detection specifically to detect EDFA malfunction in optical networks, using cumulative sum (CUSUM) and window-limited generalized likelihood ratio (GLR) on sliding-window SNR data to reliably detect the early onset of gradual gain reduction, showing that these techniques serve as valuable early-warning indicators for preventing connectivity disruptions.

- **Collective anomaly detection**

Unlike point anomalies, collective anomalies represent sequences that are anomalous when viewed together (Chandola et al., 2009). Gupta, Gao, Aggarwal, and Han (2014) examined this category in the temporal data context. These patterns, also known as subsequence outliers, are anomalous because of their joint behaviour over time, with each individual observation not necessarily deviating from normal behaviour. Detecting such patterns therefore requires analyzing temporal structure rather than individual observations. A gradual increase in pre-FEC BER, with each individual reading still within thresholds, is a typical collective anomaly in optical networks.

Change points and collective anomalies are related but distinct: change points identify *when* behavior shifts, while collective anomalies highlight sequences that deviate from normal patterns. Together, they form a powerful early-warning framework first signaling a regime transition, then revealing whether the new regime is anomalous, which may indicate imminent faults.

- **Integration with forecasting**

Combining forecasting with change point and collective anomaly detection yields more interpretable early-warning systems. Numerical predictions alone provide limited insight into specific forecasts that are concerning. Integrating CPD and CAD with forecasting address this by combining outputs from the three methods through interpretable rules to enhance decision-making in proactive network management. This interpretability helps operators understand both the prediction and the underlying evidence prompting the alert.

Table 1.3 summarizes the QoT forecasting and early-warning approaches reviewed in this section. The comparison shows that forecasting accuracy has been the primary objective throughout the literature. It also reveals that no existing study explicitly evaluates or derives operational reaction windows from its predictions to provide network operators with sufficient lead time for effective intervention.

Table 1.3 Summary of QoT forecasting and early-warning approaches and comparison with the proposed solution

Reference	Forecasting technique	CPD-CAD integration	Reaction window	Key limitation
Mezni et al. (2020)	1D-CNN multi-step forecasting of field SNR (from BER data), horizons up to 24 h	No	No	Error-metric evaluation only (RMSE, MAPE, R^2)
Chouman et al. (2021)	MLP and LSTM forecasting of lightpath QoT and minimum QoT (vs. naive and linear regression)	No	No	DNN advantage limited to the minimum-QoT task; no early-warning indicators
Du et al. (2021)	Supervised classification (RF, XGBoost, BRITS RNN) forecasting imminent loss of signal 1-7 days ahead	No	No	Requires labeled failure data; no CPD/CAD integration
Tremblay et al. (2019); Aladin et al. (2020a, 2020b); Allogba et al. (2021, 2022a)	SVM-based QoT estimation; LSTM, ED-LSTM and GRU forecasting (univariate and multivariate, horizons up to 96 h)	No	No	Marginal gains over naive baselines; accuracy degrades over longer horizons; no early-warning signals
Di Cicco et al. (2023)	Multi-channel seq2seq LSTM with variational dropout (Bayesian probabilistic forecasts)	No	No	Anomaly detection illustrative only; no reaction-window analysis

Table 1.3 Summary of QoT forecasting and early-warning approaches and comparison with the proposed solution (continued)

Reference	Forecasting technique	CPD-CAD integration	Reaction window	Key limitation
Yousefi et al. (2023)	MLP (selected over LSTM, N-Beats) with Monte Carlo dropout; quantile regression as reference	No	No	Calibration focus; DL did not outperform naive forecasts; no regime-shift detection
Hedayatnejad et al. (2025)	Transformer-XL (multi-head attention) single- and multi-step span-loss forecasting	No	No	No degradation indicators; no reaction-window analysis
Mosayebi & Lampe (2024)	CUSUM / window-limited GLR change-point detection for EDFA malfunction	CPD only	No	Detection only; no forecasting integration
This thesis (Contributions 3-4)	CPD-CAD-aware sequence-to-sequence forecasting (LSTM- and TCN-based) integrating change point and collective anomaly detection	Yes	Yes	Systematic validation across multiple lightpath datasets and forecasting architectures; operational evaluation using failure coverage, lead time, and false alarm rate)

1.3.5 Open research questions in QoT forecasting for proactive network management

The literature review shows significant progress in ML-based QoT forecasting for optical networks, particularly through DL architectures and time-series prediction models. However, several limitations remain.

Change point detection and QoT forecasting are typically addressed as separate tasks. Although change point detection techniques have been applied to identify regime shifts in monitoring data, and forecasting models have been developed to predict future QoT values, their integration remains limited. Most forecasting approaches focus on minimizing numerical prediction error without explicitly incorporating regime shifts that indicate emerging degradation. Consequently, the potential of change points to signal emerging degradation and to trigger proactive intervention before threshold violations occur remains underutilized.

Gradual performance degradation in optical networks generally manifests as collective temporal anomalies rather than isolated point deviations. Point anomaly detection has received extensive attention. In contrast, methods to detect collective anomaly patterns in multivariate OPM time series, which serve as early indicators of gradual QoT degradations, remain insufficiently developed. Methods that detect such degradation patterns and integrate them into predictive modeling remain insufficiently developed.

Despite the predictive capabilities of DL models, interpretability remains limited in forecasting contexts. Operators require actionable early warning indicators such as regime shifts or collective anomaly patterns, not only the numerical forecasts of future QoT values.

Furthermore, forecasting studies frequently evaluate performance using metrics based on the errors only, without systematically assessing whether predictions provide sufficient reaction windows for practical intervention. Demonstrating that forecasting models provide sufficient advance warning before service-impacting thresholds are reached is critical for validating their relevance in optical network operations.

Finally, existing research tends to tackle anomaly detection, change point detection, and forecasting as separate components. Comprehensive frameworks that combine collective anomaly detection, change point detection, and QoT degradation forecasting within a coherent architecture remain limited. Furthermore, no study has systematically validated CPD-CAD-aware forecasting across multiple independent network datasets, forecasting architectures, and detection sensitivity levels to assess whether the performance benefits generalize beyond a single experimental configuration.

These gaps motivate the development of an integrated forecasting framework that combine QoT degradation indicators with predictive models to enable proactive QoT management, as well as the systematic validation of the framework across multiple independent network datasets, forecasting architectures, and detection sensitivity levels to assess whether the performance benefits generalize beyond a single experimental configuration.

1.4 Synthesis, recurrent themes, and open research problems

The preceding sections reviewed ML approaches for three ONFM-based tasks: QoT estimation for lightpath provisioning, anomaly detection for failure detection, and QoT forecasting for failure prediction. This section synthesizes the key findings, examines themes that are present in multiple tasks, and states research gaps that motivate the three thesis contributions.

1.4.1 Summary of progress by ONFM-based task

QoT estimation for lightpath provisioning

ML-based QoT estimation has advanced considerably, demonstrating that ML models can replace computationally intensive simulations with fast inference while maintaining strong accuracy. Transfer learning has become highly popular, reflecting the reality that new network deployments suffer from data scarcity. However, much of the literature continues to prioritize accuracy over interpretability and computational efficiency. DNNs dominate despite their black-box nature and resource demands. The integration of explainability techniques into domain adaptation approaches remains limited, and few studies achieve a balanced combination of accuracy, efficiency, adaptability, and interpretability.

Anomaly detection for failure detection

Anomaly detection has evolved from simple threshold-based techniques to more sophisticated ML-based methods for pattern recognition. The scarcity of labeled anomalies in operational networks has motivated a strong shift toward unsupervised ML techniques. Yet systematic comparisons across different techniques remain limited. Most studies evaluate a single technique rather than assessing multiple approaches to understand their respective strengths for anomalies in multivariate time-series data. Validation on production networks is further constrained by proprietary data restrictions.

QoT forecasting for failure prediction

Forecasting has seen rapid progress with the adoption of DL, particularly LSTMs and TCNs, demonstrating the feasibility of multi-step QoT prediction. However, the integration of change point detection, collective anomaly detection, and forecasting into unified early-warning frameworks remains underexplored. Most work emphasizes accuracy metrics without assessing whether predictions provide sufficient lead time or reaction window for meaningful preventive action. Operators also require interpretable signals that clarify which patterns triggered concerns, rather than forecasts performance only.

1.4.2 Recurrent themes

1.4.2.1 Task-appropriate data contexts and paradigm selection

The data available for each ONFM-based task shapes the appropriate ML paradigm. QoT estimation relies on synthetic labeled lightpath datasets, motivating supervised learning with domain adaptation. Anomaly detection relies on unlabeled production OPM data, motivating unsupervised learning methods. QoT forecasting relies on sequential OPM time series, motivating DL sequence modeling. This progression in data type and availability across tasks provides a structured foundation for paradigm selection.

1.4.2.2 Interpretability across ONFM-based tasks

Across all three tasks, operators require explainable model outputs to ensure trust, validation, and operational accountability. Yet interpretability is often addressed only after accuracy has been optimized, rather than being incorporated as a foundational design principle. The appropriate interpretability mechanism differs by task: SHAP-based feature selection for QoT estimation, anomaly indicators for anomaly detection, and change point or collective anomaly indicators for QoT forecasting.

1.4.2.3 Computational efficiency and scalability

Large scale production networks, with multiple lightpaths generating continuous monitoring data, require ML solutions that are computationally efficient and scalable. In QoT estimation, models must evaluate thousands of candidate lightpaths quickly. In anomaly detection, models must process OPM data in real time without delay. QoT forecasting must scale across extensive lightpath populations. However, research often prioritizes accuracy over efficiency. Complex DNNs may yield minimal accuracy improvements at the cost of significant computational overhead, limiting their practical deployment. Meanwhile, lightweight and interpretable alternatives such as RF or efficient unsupervised methods receive comparatively less attention despite offering potentially competitive accuracy with superior efficiency.

Balancing predictive performance with operational efficiency constitutes a recurring design consideration throughout this thesis, albeit with different emphasis depending on the task. Indeed, computation times are presented in the QoT estimation and anomaly detection tasks, respectively. For the QoT forecasting task, the CPD-CAD module is implemented using online techniques, which limit computational cost by processing data sequences one sample at a time.

1.4.2.4 Generalization across diverse configurations

Optical networks vary widely in topology, equipment, technology, and operating conditions. Effective ML models must accommodate multi-vendor equipment, different fiber types, varying modulation formats, and heterogeneous operating conditions without overfitting to specific training scenarios. In QoT estimation, transfer learning enables models trained on synthetic data in source domain to run reliably in a different target domain with limited data. In contrast, for anomaly detection and QoT forecasting, generalization is evaluated through systematic validation across multiple independent datasets for the CPD-CAD-aware forecasting framework. Here, the emphasis is on robustness and consistency rather than explicit transfer mechanisms. Although the methods differ, both approaches aim to build deployable solutions capable of operating reliably across diverse network conditions.

1.4.3 Specific open research problems

1.4.3.1 P1: Data scarcity, computational efficiency and interpretability constraints in QoT estimation

Existing ML models for QoT estimation typically address accuracy, domain adaptation and interpretability separately. No study has demonstrated that these objectives can be achieved simultaneously within a single framework. Methods that combine automated feature selection, domain adaptation, and explainability, while maintaining accuracy and efficiency are still underdeveloped.

1.4.3.2 P2: Labeled failure data scarcity in anomaly detection

Production optical network telemetry collects multivariate monitoring data across several parameters. However, labeled failure samples are scarce because real failures do not happen often and it is difficult to capture complete fault signatures when they do occur. As a result, many existing anomaly detection methods rely on supervised or semi-supervised techniques that need labeled data, or they use single unsupervised methods without systematic comparison. What is missing is a systematic study of how different unsupervised learning approaches can be used to detect different types of anomalies, including collective anomalies that single-metric threshold monitoring cannot capture, across correlated parameters without requiring labeled training sets. The field also lacks comprehensive evaluations that compare these paradigms to clarify their respective strengths when used with field OPM.

1.4.3.3 P3: QoT prediction without actionable reaction windows

Forecasting methods typically predict future QoT values at specific horizons or identify faulty components through localization, but they do not indicate *when* operators must intervene. Recent hierarchical frameworks that combine forecasting, detection, and localization achieve high accuracy in identifying faulty components, yet they still provide only predicted values or binary failure classifications that require operator interpretation. The lead time to failure or reaction window quantifies the time available for intervention. Solutions that explicitly detect QoT degradation and estimate the remaining intervention time, thereby converting predictions into actionable maintenance decisions, are very rare. To the best of our knowledge, the integration of change point detection and collective anomaly detection with QoT forecasting to compute operational reaction window has not yet been developed.

1.4.3.4 P4: Lack of systematic validation of the proposed CPD-CAD-aware forecasting framework

Our recent work presented in Chapter 5, combining change point detection and collective anomaly detection with QoT forecasting shows promise for proactive network management, but the evaluations are based on a single field lightpath dataset, using one forecasting model and one fixed input sequence length. This narrow setup leaves a key empirical question unresolved: whether the reported improvements reflect inherent strengths of the CPD-CAD integration or simply the characteristics of a particular dataset, degradation pattern, or experimental configuration.

To the best of our knowledge, the literature offers no systematic multi-dataset assessment of a CPD-CAD-aware forecasting solution. We identify three aspects of generalizability that remain unexplored. First, it is unknown how sensitive CPD-CAD enhancements are to the choice of forecasting architecture, such as LSTMs, TCNs, or other sequence models. Second, the influence of the sequence length on forecasting accuracy and early-warning indicator has not been evaluated across different datasets. Third, and most important for network optimization, the proactive versus reactive nature of detection has not been systematically considered. Existing studies report accuracy metrics like root men square error (RMSE), mean absolute error (MAE) or coefficient of determination (R^2) but do not assess whether detections occur before degradation begins (positive lead time) or only after failures occur, even though this distinction is crucial for real-world deployment.

As a result, the field lacks a systematic validation framework that tests CPD-CAD-aware forecasting across multiple production network datasets, multiple forecasting architectures, and multiple anomaly-detection methods, while also evaluating operational metrics such as failure coverage, lead time, and false alarm rates alongside standard accuracy metrics. Without such evidence, the broader applicability of CPD-CAD-based forecasting for proactive QoT management remains an open question.

These four unsolved research problems collectively identify the boundaries of the existing literature and define the research space addressed in this thesis. The following chapter presents the research methodology and conceptual approach adopted for each contribution and explains how the contributions fill these gaps.

CHAPTER 2

RESEARCH METHODOLOGY AND CONCEPTUAL APPROACH

This chapter outlines the research methodology, the research strategy, and the conceptual approach that guide the development of the three contributions and the extended validation study that constitute this thesis. As a manuscript-based thesis, it comprises three peer-reviewed articles addressing complementary aspects of QoT management linked to ONFM use cases. An additional chapter, not structured as a peer-reviewed article, extends the third contribution through a systematic multi-dimensional validation study that establishes the generalizability of the proposed forecasting framework across different lightpath datasets from a production network and diverse configurations. Although each article focuses on different tasks, this chapter explains the reasoning behind the research design and details the conceptual steps that establish the link between the three contributions.

2.1 Choice of methodological approach

The articles proposed in the thesis address each specific gaps identified in previous chapter while contributing to a shared research objective: advancing ML-based QoT management in optical networks. The three ONFM-based tasks, namely QoT estimation, anomaly detection, and QoT forecasting, are significantly distinct in their data requirements and operational constraints. Nevertheless, they share the QoT as the performance objective and a chronological progression to form a coherent thesis. The progression reflects increasing data availability and complexity from synthetic data for QoT estimation before lightpath establishment, to multivariate field OPM data enabling anomaly detection and supporting QoT degradation forecasting.

The central methodological choice is the paradigm selection based on the tasks. Each task explores the ML approach that best suits its data context and operational requirements. This choice is grounded in the no free lunch theorem (Wolpert & Macready, 1997), which states that no single algorithm can be expected to perform optimally across all problem types. Moreover, the three ONFM use cases present fundamentally different data contexts. QoT estimation at lightpath provisioning relies on synthetic labeled data and require generalization under domain shift. Failure detection in production networks requires identifying anomalies without labeled training sets. And failure prediction, in the context of gradual QoT degradations forecasting, requires modeling temporal dependencies in time series OPM data to produce actionable early-warning signals. Three organizing principles provide coherence between the three tasks: the shared QoT performance objective, the paradigm selection based on each task, and the chronological progression of the tasks addressed.

Beyond these organizing principles, insights from the anomaly detection task, addressed in Chapter 4 (Article 2), are integrated with a change point module to feed into a forecasting architecture in the framework proposed in Chapter 5 (Article 3). This forms a technical connection enabled by the sequential development of the models and the shared data source.

2.2 Research strategy

The main objective of this thesis is to develop and validate ML-based solutions for three complementary tasks. These tasks address the open research problems identified in the previous chapter: QoT estimation approaches fail to generalize across network environments due to the domain gap between synthetic training data and real deployments. Anomaly detection methods rely on labeled failure samples that are rarely available in production environments. QoT forecasting methods report accuracy metrics without actionable lead time for proactive maintenance, and the validation of the CPD-CAD-aware forecasting framework across diverse network configurations remains to be addressed. The research questions formulated in Chapter 0 target each of these open problems and the objectives lead to the three contributions.

Chapter 3 (Article 1) addresses the first objective linked to the QoT estimation task. It builds an automated, interpretable, and computationally efficient framework, combining Boruta-SHAP feature selection with domain adaptation to reduce the domain shift between synthetic training data and real-world environments. Chapter 4 (Article 2) covers the second objective by implementing and evaluating three unsupervised anomaly detection methods on four years of production OPM data. Chapter 5 (Article 3) with an extension study in Chapter 6, address the third objective by building a CPD-CAD-aware seq2seq forecasting framework that turn QoT predictions into actionable reaction windows, and test its generalizability across three additional field lightpath datasets, two DL architectures, five CAD methods, and two input sequence lengths.

Two data sources support the contributions, selected to match each task's operational requirements. For QoT estimation, publicly available synthetic datasets, generated by the Fraunhofer institute modeling tool provide large-scale labeled QoT samples (Bergk et al., 2022). Domain adaptation is presented as the component to address the scarcity of field lightpath data in real-world scenarios (Azzimonti et al., 2021). For anomaly detection and QoT forecasting, multivariate OPM datasets collected from a production optical network provide real-world time series reflecting operational variability, environmental effects, and component aging. Full dataset descriptions and experimental environments are provided within each respective chapter.

In real-world optical network operations, QoT estimation, anomaly detection and QoT forecasting are not strictly sequential tasks. Anomaly detection and QoT forecasting run continuously and in parallel after a lightpath is established. The ordering of the three contributions in the thesis is a research design choice: QoT estimation is addressed first, establishing the QoT as the performance objective across all three contributions. Anomaly detection is addressed second to introduce the OPM data context and the anomaly indicators are a direct methodological prerequisite for the framework proposed in Chapter 5 (Article 3). QoT forecast is addressed last because it depends on both the OPM data context introduced in the second contribution and the anomaly indicators produced by the method. Each task therefore builds on methodological insights of the preceding one, and the three objectives form a coherent research progression.

2.3 Conceptual approach for ML-based QoT management in smart optical networks

Figure 2.1 provides an overview of the conceptual approach, illustrating the three contributions along the lightpath lifecycle: QoT estimation (Article 1) is performed in the pre-establishment phase, whereas anomaly detection (Article 2) and proactive QoT forecasting (Article 3) are applied during the operational phase, provided that the estimated lightpath QoT level is acceptable. Solid arrows represent the lightpath lifecycle, while the dashed arrow denotes the methodological integration between Articles 2 and 3, which is described in Section 2.4. For each contribution, the figure summarizes the data sources and ML paradigms. The following subsections describe the conceptual steps and methodological choices for each contribution.

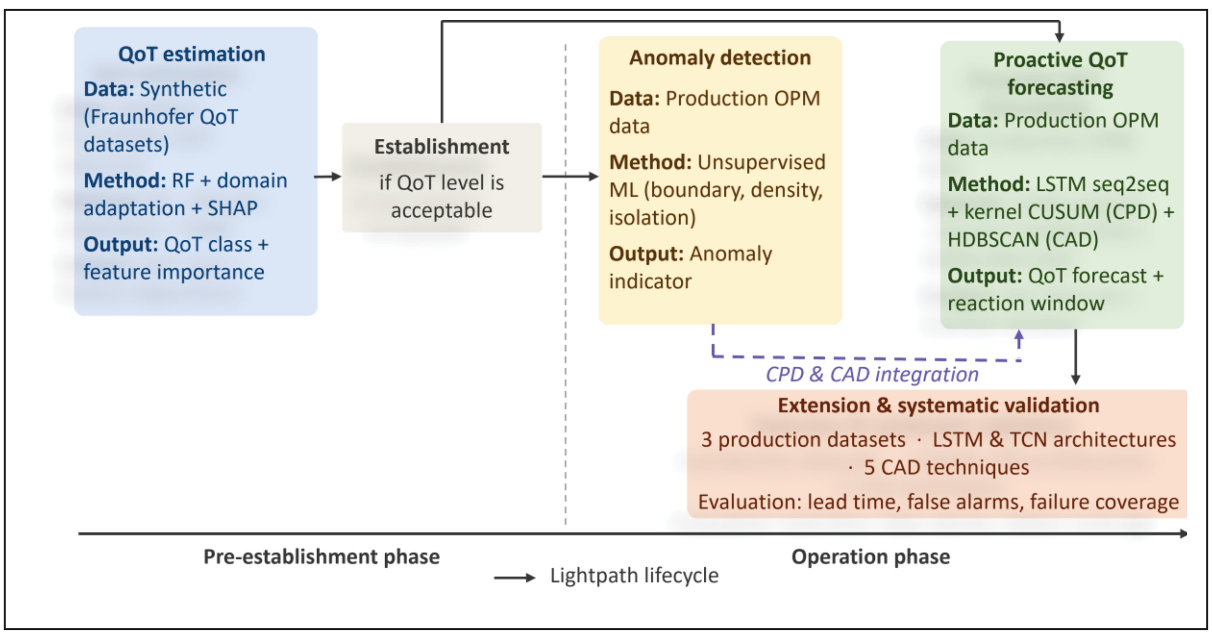


Figure 2.1 Work process overview: contributions with their data, methods, and outputs; CPD and CAD integration from anomaly detection to QoT forecasting tasks; Extension and systematic validation

2.3.1 QoT estimation for lightpath feasibility assessment

The first contribution addresses QoT estimation for lightpath provisioning, where the ONFM question is "*can we avert a failure*", in other terms, whether a candidate lightpath will meet QoT requirements before deployment. The main challenge is that models trained on synthetic data are expected to generalize to real networks environments with different physical layer characteristics and limited samples. Supervised learning techniques are not suited to resolve this problem linked to domain shift. Table 2.1 presents the research questions with the existing limitations and the proposed approach.

Table 2.1 QoT estimation task: research question, existing limitations, and proposed approach

RQ1	Limitations	Proposed contributions
How can QoT estimation balance accuracy, cross-domain generalization, interpretability, and computational efficiency for lightpath provisioning when only synthetic training data and limited target domain samples are available?	Models trained on synthetic data fail to generalize across domains; DL approaches for transfer leaning achieve accuracy but sacrifice interpretability and computational efficiency.	Boruta-SHAP feature selection guided domain adaptation using a RF model achieves promising accuracy with 50 target samples; SHAP explanations provide interpretable classification decision.

The conceptual approach combines Boruta-SHAP automated feature selection with domain adaptation in a supervised learning classification framework using RF technique. Boruta-SHAP identifies the most relevant features for QoT classification and generates SHAP-based explanations that make the feasibility decisions interpretable to operators. Integrated with domain adaptation, the framework refines the model for a target network using only a small number of target samples. The lightweight RF models are used to ensure the approach remains computationally efficient for large-scale lightpath provisioning, without sacrificing accuracy as compared to the ANN models.

This contribution demonstrates that accuracy, cross-domain generalization, interpretability, and computational efficiency can be achieved together within a QoT estimation framework.

2.3.2 Anomaly detection

The second contribution addresses failure detection in optical networks, where the ONFM question is "*is there a failure now?*". The main challenge for this task is that field OPM data are multivariate and unlabeled. This is because operators rarely keep records of when slow-developing issues begins. Table 2.2 show the research question, the limitations, and the proposed approach for the anomaly detection task.

Table 2.2 Anomaly detection task: research question, existing limitations, and proposed approach

RQ2	Limitations	Proposed contribution
How can complementary unsupervised learning methods detect different anomaly types in multivariate OPM data without labeled training data?	Supervised methods require labeled failure samples which are rarely available in production networks; no single unsupervised method can optimally capture anomalies.	Comparison of density-based, boundary-based, and isolation-based techniques using four years of OPM data. The density-based technique achieves promising performance in detecting QoT degradations.

Three complementary unsupervised techniques, namely DBSCAN, OCSVM, Isolation Forest are used with four years of OPM data to develop anomaly detection models. A systematic evaluation using both unsupervised metrics and supervised metrics based on the alarms metrics available in the datasets, allows to see how each technique perform in detecting QoT degradations.

This contribution establishes an unsupervised multivariate anomaly detection on field OPM data. It also lays the methodological foundation for the third task, where the point anomaly detection is extended to the collective type and combined with change point detection to integrate regime shift and anomaly insights into the QoT forecasting framework proposed in Article 3.

2.3.3 QoT forecasting and validation

In the third contribution for QoT forecasting, the ONFM question addressed is "*will there be a failure?*". The main challenge is turning forecasting performance metrics into output that is operationally meaningful. This requires distinguishing between two different concepts. The forecast window specifies how much historical OPM data the model uses as input. It controls what the model can learn from past QoT signal patterns and is a key design parameter for predictive accuracy (Aladin et al. 2020b; Allogba et al. 2021). The reaction window or lead time, on the other hand, is the interval between the earliest detected anomaly or change point and the actual failure event. It measures operational usefulness by quantifying how much time is available for intervention before service quality is affected (Murphy et al., 2023; Murphy et al., 2024). The two concepts are distinct and not necessarily correlated: a model may achieve high predictive accuracy for a given forecast horizon yet still offer insufficient lead time if the CPD and CAD indicators that define the reaction window are detected too close to or after the failure event. This contribution proposes seq2seq models trained with different forecast windows to improve prediction accuracy, while CPD and CAD modules allow generating early warnings that determine the reaction window. Together, these components provide operators with enough lead time for proactive intervention. Table 2.3 shows the research question, existing limitations, and proposed contribution.

The approach integrates CPD and CAD into a multivariate seq2seq forecasting framework using LSTM architectures. CPD identifies regime shifts in the OPM time series, while CAD identifies collective anomaly patterns in the input sequences. Both are incorporated as additional input signals to the forecasting model, providing degradation context that improves prediction accuracy and enables computation of lead time before QoT violation. This represents the link between the anomaly detection and the QoT forecasting tasks, demonstrating that the two tasks are mutually reinforcing and their integration produces measurable performance improvements.

Table 2.3 QoT forecasting task: research question, existing limitations, and proposed approach

RQ3	Limitations	Proposed contributions
How can CPD and CAD integrate into seq2seq forecasting to improve QoT prediction accuracy and provide actionable reaction windows?	Forecasting models report accuracy metrics without actionable lead time.	CPD and CAD integration into QoT forecasting models achieves RMSE and MAE reduction; consistent positive reaction windows before failure enable proactive maintenance
Does the CPD-CAD-aware framework generalize across diverse datasets, DL architectures, CAD techniques and sequence lengths?	Generalizability of CPD-CAD-aware forecasting across diverse production network configurations remains unestablished.	Consistent RMSE and MAE reductions across 3 datasets, 2 architectures, 5 CAD techniques, and 2 sequence lengths; 100% failure coverage with positive lead times.

Chapter 6 presents an extension of the CPD-CAD-aware forecasting framework introduced in Chapter 5 (Article 3) through a systematic validation study. This study investigates whether the reported improvements generalize across different datasets, model architectures and design choices. The validation is conducted across four dimensions. First, the framework is evaluated on three additional field datasets corresponding to lightpaths with different route characteristics and QoT dynamics from the one considered in Chapter 5. Second, the seq2seq LSTM model is compared with a TCN model to determine whether the framework's improvements depend on a specific recurrent architecture or extend across different model families. Third, the CAD stage is treated as a tunable design component rather than a fixed element: five CAD techniques are evaluated within the framework to identify the most suitable approach for detecting QoT deviations in each dataset. Fourth, two sequence lengths are investigated to assess the impact of input sequence length on forecasting performance.

Beyond conventional error metrics, Chapter 6 introduces an initial operational assessment that evaluates the framework from a network operator's perspective. This assessment considers practical indicators such as the lead time, namely the reaction window between an early CPD-CAD indicator and the subsequent QoT violation, the false alarm rate of QoT deviation indicators, and the coverage of imminent loss of signal (ILOS) events in the test datasets.

2.4 Integration of the three ONFM-based QoT management tasks

The three contributions of the thesis address distinct QoT management tasks while sharing a common performance objective: maintaining lightpath QoT. Their integration is based on a set of functions operating at successive stages of the lightpath lifecycle within a QoT management system.

The interaction between the QoT estimation task (Article 1) and the two subsequent tasks occurs through the lightpath establishment decision. Prior to establishment, the QoT estimation framework predicts whether a candidate lightpath can meet the required QoT level based on lightpath and network parameters. In contrast, anomaly detection relies on OPM data that become available only after the lightpath has been deployed. Therefore, the interaction between these tasks is sequential and conditional: a lightpath is established only if its estimated QoT is acceptable, after which its QoT can be monitored at the receiver to enable the subsequent tasks.

The anomaly detection task (Article 2) analyzes the monitored OPM data to identify anomalous QoT deviations associated with abnormal operating conditions. Building on this monitoring capability, the QoT forecasting task (Article 3) shifts the objective from reactive detection to proactive management by predicting the future evolution of lightpath QoT over a given forecast horizon, thereby providing advance warning for potential service degradation.

In contrast to the interaction between the QoT estimation and anomaly detection tasks, the link between the latter and the QoT forecasting task is methodological in nature. The anomaly detection approach developed in Contribution 2 is integrated into the forecasting framework of Contribution 3 through CPD and CAD indicators, which enrich the seq2seq models and enable the computation of a reaction window for proactive maintenance.

Together, the three contributions address QoT management tasks associated with the three ONFM use cases identified by Musumeci and Tornatore (2025) and interact at two distinct levels. At the lightpath lifecycle level, the QoT-based establishment decision links the pre-establishment QoT estimation task to the two operational tasks. At the methodological level, the anomaly detection approach is integrated into the QoT forecasting framework.

CHAPTER 3

AUTOMATED, INTERPRETABLE AND EFFICIENT ML MODELS FOR REAL-WORLD LIGHTPATHS' QUALITY OF TRANSMISSION ESTIMATION

Sandra Aladin [a], Lena Wosinska [b], Christine Tremblay [a]

[a] Department of Electrical Engineering, École de technologie supérieure,
1100 Notre-Dame Ouest, Montréal, Québec, Canada H3C1K3

[b] Department of Electrical Engineering, Chalmers University of Technology,
412 96 Gothenburg, Sweden

Article published in *IEEE Open Journal of the Communications Society*, vol. 6, pp. 9785-9801, November 2025, doi: 10.1109/OJCOMS.2025.3635533

3.1 Abstract

Fast and accurate estimation of lightpaths' quality of transmission (QoT) is crucial for ensuring quality of service (QoS) and seamless operation in real-world optical networks. ML algorithms are promising tools for QoT estimation of lightpaths before their establishment. In multi-domain optical networks, where learned QoT estimation models must be transferred between heterogeneous environments with limited target data, deep neural networks (DNNs) have shown promising results. However, DNN-based transfer learning (TL) approaches using fine-tuned artificial neural networks (ANNs) and convolutional neural networks (CNNs), offer limited interpretability. Consequently, little insight into the decision-making process is provided, and large labeled datasets as well as high computational resources are required, limiting their suitability for real-time, large-scale deployment in production networks. To address these challenges, we propose a novel lightweight and interpretable TL framework that integrates the Boruta-SHAP algorithm for automated feature selection (FS) together with two domain adaptation (DA) techniques: Feature Augmentation and Correlation Alignment. In

contrast to the existing approaches based on DNN, our strategy leverages interpretable and efficient ML models to enhance the adaptability across diverse datasets in real-world network environments. We show that our random forest (RF)-based models achieve better performance than the ANN-based models, without sacrificing the classification accuracy. The FS via Boruta-SHAP allows for reducing dimensionality as well as training and inference times up to 70.68%, and 41.64 %, respectively. Our proposed framework outperforms DA baseline models achieving 99.35% accuracy improvement in domain shift. Moreover, it offers 86% accuracy with a 99.83% reduction in the size of the target domain.

3.2 Introduction

Reliable optical networks are crucial for the rapid development of emerging online services such as the Internet of Things, 5G/6G applications, and edge computing. These technologies require high bandwidth and low latency connectivity while demanding excellent transmission performance. Coherent optical networks with flexible modulation and tunable transceivers enable reliable high-speed and long-distance data transmission. However, these technologies entail complex network architectures and dynamic changes of parameters. Quality of transmission (QoT) estimation of lightpaths before they are established is an important task to guarantee the required connection reliability performance. Analytical solutions, based on mathematical models, have been developed to estimate linear and nonlinear impairments (NLI) of lightpaths before their establishment. Two commonly used methods are the split step Fourier method (SSFM) and the Gaussian Noise (GN) model. The SSFM has been proven to. accurately estimate the QoT but it suffers from the long computation time and high complexity. Authors in Pilori, Cantono, Carena, and Curri (2017) developed an optical fiber simulator based on SSFM. The proposed fast fiber simulator software (FFSS) was tested using graphical processing units (GPU) computation on a standard desktop computer and a more robust server. Their simulations showed good accuracy and speed improvement in estimating the signal-to-NLI ratio on the server compared to a baseline result obtained by central processing unit (CPU) computation. Moreover, authors in Wass et al. (2017) used the same SSFM to generate data set through simulation to test a predictive models proposed in their study. On the other hand, the GN model, although less accurate, offers faster computation than the SSMF method. In Ferrari et al. (2020), the authors presented the validation of an open-source network simulation application that is called Gaussian noise simulation in Python (GNPy). The tool is based on the GN model and is used to generate QoT data in many research studies (Khan et al., 2020a; Manso et al., 2021; Usmani et al., 2021, 2022a, 2022b; Naik, Jeyachitra, & Doss, 2023). Due to its favorable trade-off between computational efficiency and accuracy, the GN model is more commonly employed than the SSFM for analytical QoT estimation of unestablished lightpaths. However, the accuracy of analytical QoT estimation methods is fundamentally constrained by the quality of input parameters - such as span loss, dispersion, and amplifier noise figures - whose precision and availability are often limited in real-world network environments.

Recently, QoT estimation based on ML gained high attention. In ML-based approaches, algorithms learn the complex mapping between input features - such as candidate lightpath characteristics and network configuration parameters - and the corresponding QoT outcomes. Thus, estimating the QoT of unestablished lightpaths involves leveraging these features to predict the feasibility and expected performance of the connection. By comparing the predicted QoT against the receiver sensitivity threshold, the viability of a potential lightpath can be assessed, which is an essential step for efficient service provisioning. An underestimated QoT may result in underutilized resources, while overestimating it can lead to connection failures and violations of service level agreements (SLAs). ML classifiers based on K-nearest neighbors (KNN), random forest (RF), and support vector machine (SVM) were proposed in Aladin and Tremblay (2018), Tremblay and Aladin (2018), and Allogba et al. (2022a). Additionally, a QoT estimation model using artificial neural network (ANN) was introduced in Aladin et al. (2020a). All these models were trained with synthetic Bit-Error-Rate (BER) data generated by using the additive Gaussian white noise (AGWN) model. While the results demonstrated the potential of ML techniques for efficient and automated lightpath provisioning, these studies relied exclusively on proprietary QoT datasets generated by individual research groups. Consequently, the challenges related to reproducibility and independent validation of results persist.

In 2022, the Fraunhofer Heinrich Hertz Institute (HHI) has made available a collection of four QoT datasets based on the Core Optical Network of the United States (CONUS) and Telefónica Spain National Network (TSNN) topologies, intended for evaluating QoT estimation models (Bergk et al., 2022). The use of the publicly accessible QoT datasets provides significant benefits, enabling researchers to validate their findings, benchmark models against existing methods, and compare performance effectively. So far different approaches using open QoT datasets to estimate the QoT for unestablished lightpaths have been explored. However, most proposed solutions rely on complex ML techniques that primarily focus on improving predictive performance, often overlooking critical aspects such as computation time (during both training and inference) and model interpretability. While interpretability is essential for building trust in model's decisions, inference time is directly linked to decision-making speed and can significantly impact lightpath provisioning in real-world applications. For network operators, short training times are crucial for adapting quickly to changing network conditions in dynamic, high-capacity optical networks. In addition, during a real-time provisioning, low-latency inference is essential to ensure SLA compliance and optimal resource utilization. Therefore, minimizing ML model training and inference times not only enhances operational efficiency but also supports the deployment of cognitive management systems in real-time optical network environments. Additional strategies such as transfer learning (TL), active learning, and meta-learning have been proposed to enable practical deployment of ML-based QoT estimation models in real-world networks (Ouyang, Chen, Liu, Chen, & Zhu, 2024). These approaches address the scarcity of field QoT datasets, as effective ML-based QoT estimation models require both high-quality and large volume of data. Active learning optimizes the selection of informative samples, meta-learning extracts transferable knowledge across tasks, and TL leverages prior knowledge from related domains to improve learning efficiency and generalization in new scenarios. Many TL solutions are based on neural networks pre-trained on source domain data and fine-tuned on target domain data (Cai et al., 2024; Zhou et al., 2023, 2024; Paudyal, Shen, Yan, & Simeonidou, 2021; Liu et al., 2021a; Usmani et al., 2022c, 2024; Gu et al., 2023; Pesic et al., 2020; Azzimonti et al., 2020; Tariq et al., 2023). While these models can achieve high accuracy, their deployment in production networks is hindered by computational complexity, large labeled data requirements (Usmani et al., 2024), and limited generalizability due to heterogeneous optical system configurations, as highlighted in Usmani et al. (2022a). As a result, models often require full retraining for each new system, leading to significant time and resource overhead.

In this work, we propose a novel TL framework that includes a feature selection stage before applying DA techniques with lightweight classifiers, improving model efficiency and generalizability across different optical network scenarios. We adopt an incremental approach. In the first step, we select the best-performing low-complexity classifier through a comparative performance analysis of three ML techniques - KNN, SVM, and RF - with a more advanced deep learning ANN-based approach. In the second step, we propose applying a feature selection strategy using the recursive feature elimination with cross-validation (RFECV) and Boruta-SHAP where SHAP (SHapley Additive exPlanations) is a game-theoretic approach used for feature importance. A comparative evaluation of model accuracy and computation times across the four datasets enables finding the most effective feature selection method. The feature selection methods automatically identify the most relevant features for the RF-based solution for each dataset. In the last step, leveraging the RF classifiers, we implement a transfer learning strategy that combines domain adaptation (DA) techniques with feature space derived from the automated feature selection process in the preceding phase.

This paper is organized as follows: Section 3.3 summarizes the related works on lightpath QoT estimation. Section 3.4 describes the methodology, as well as the dataset and proposed framework for automated feature selection and domain adaptation guided by reduced features. Section 3.5 details the results with a comparative analysis of the different models as well as the performance achieved by our proposed domain adaptation guided by feature selection framework in an inter-network scenario. Finally, Section 3.6 offers some conclusive remarks as well as future research directions.

3.3 Related works

In this section, we review prior research on ML-based lightpath QoT estimation, starting with general regression and classification approaches, followed by studies using the Fraunhofer

QoT datasets. We then examine efforts to enhance model trustworthiness through explainable artificial intelligence (XAI) and to improve adaptability for addressing data scarcity challenges in real-world deployment scenarios.

ML applications in management of optical networks have gained interest in recent years. Leveraging recent advancements in technologies and data accessibility, ML uses algorithms to enable cognition in optical networks. Four main categories of ML techniques are used in the field, i.e., supervised learning, semi-supervised learning, reinforcement learning and unsupervised learning. In supervised learning models are trained using labeled data to map input parameters to correct outputs. It includes classification and regression tasks. Semi-supervised learning trains models using a combination of labeled and unlabeled data. Reinforcement learning requires an agent to learn through trial-and-error interactions within an environment and make decisions, which can lead to rewards or penalties. Finally, unsupervised learning allows to find hidden patterns by using clustering and dimensionality reduction.

ML-based lightpath QoT estimation can be achieved through a regression or a classification strategy. Supervised learning is the most used approach to predict categorical or continuous QoT indicators. Cruzes (2024) presents an overview of studies on the topic categorized into three main groups: traditional regression methods, classification approaches and deep learning-based regression architectures such as RNNs and ANNs. Numerous studies have proposed both regression (Paudyal et al., 2021; Ibrahimi et al., 2021; Khan, Bilal, & Curri, 2020b; Amirabadi et al., 2023; Wang, Cai, Lau, Li, & Nadeem Khan, 2023) and classification approaches (Rottondi et al., 2018; Caballero et al., 2012; Morais & Pedro, 2018; Panayiotou, Savva, Shariati, Tomkos, & Ellinas, 2019; Diaz-Montiel, Aladin, Tremblay, & Ruffini, 2019) for estimating the QoT of unestablished lightpaths. However, as previously noted, these solutions were developed using proprietary datasets, limiting the reproducibility and independent validation of their results.

Most research on ML-based QoT estimation has focused primarily on the predictive performance of the models. In Fawaz et al. (2024), the authors employed a regression approach to predict the signal-to-noise ratio (SNR) of synthesized lightpaths, using data generated via the enhanced Gaussian Nonlinear (EGN) model. This method is similar to the one proposed in this paper. However, their solution suffers from the comparability issue discussed in Bergk et al. (2022), while it is addressed in the solution presented in this paper.

Moreover, researchers have often adopted complex ML techniques such as deep learning, which offers improved performance but come at the cost of interpretability and transparency in the model's decision-making process prior to deployment. In Ayoub et al. (2023), the authors addressed interpretability and transparency challenges by proposing an Extreme Gradient Boosting (XGBoost) model for lightpath QoT estimation which incorporates XAI techniques. They used SHAP as a framework for feature selection to identify a subset of features sufficient for accurate QoT estimation.

Recognizing the benefits of reproducibility, validation, and fair comparison, many researchers have leveraged the publicly available Fraunhofer QoT datasets to propose a variety of solutions for lightpath QoT estimation. Studies such as Bergk et al. (2022); Safari, Shariati, Bergk, and Fischer (2021a, 2021b); Hashemi, Safari, Shariati, and Fischer (2021) introduced complex models based on ANN, deep convolution neural networks (DCNN), vertical federated learning (VFL) and secure multi-party computation (SMPC) techniques. In contrast, works like Natalino, Mohammadiha, and Panahi (2023) and Natalino, Panahi, Mohammadiha, and Monti (2024) proposed an AI/ML-as-a-service framework for optical network automation, utilizing the Automated ML (AutoML) libraries that favor simpler models such as RF. However, RF was not selected as part of the final ensemble applied in their framework. The evaluation of the framework's inference performance revealed persistent challenges, particularly regarding inference speed, explainability and model trustworthiness.

Although the authors in Ayoub et al. (2023) also utilized the Fraunhofer datasets to propose an XGBoost-based solution for QoT estimation - incorporating explainable XAI techniques - they employed a heuristic-based feature selection approach and did not address computational efficiency in their proposed solution. Furthermore, XGBoost is generally considered more complex than RF, due to its greater number of tuning parameters, more complex architecture, higher computational cost, and lower interpretability. While RF trains decision trees in parallel and aggregates predictions through averaging, XGBoost builds tree training sequentially using gradient boosting and regularization step, which - although potentially more accurate - can hinder optimization and interpretability.

In production networks, adaptability - alongside interpretability and computational efficiency - is crucial for overcoming the scarcity of labeled QoT data. Many deep learning-based TL solutions have been proposed to adapt models from source domains with abundant labeled data to target networks with limited datasets. Although these approaches can achieve high prediction accuracy, their substantial computational complexity often limits their practical deployment in real-time or resource-constrained environments. Usmani et al. (2024) proposed a hybrid framework combining knowledge distillation (KD) with TL to transfer knowledge from a large, complex model to a smaller more efficient one, thereby reducing computational complexity, storage overhead, and dependence on large-scale training datasets. In Rottondi et al. (2021), a TL approach leveraging three standard ML algorithms, i.e., SVM, RF, and logistic regression (LR), was evaluated using two DA methods for transferring knowledge between two network domains, with results showing that DA outperformed simple mixing of source and target data. In Azzimonti et al. (2020), a comparative study of DA and active learning with Gaussian Processes (GPs) showed that active learning achieved similar QoT estimation performance improvements while requiring fewer labeled samples than the domain adaptation, highlighting its effectiveness in data-scarce scenarios.

In TL, DA facilitates knowledge transfer across network domains, enabling models trained on large source datasets to perform effectively in target domains with limited data. This capability is vital for the practical deployment of QoT estimation solutions in real-world optical networks where labeled data are scarce. Our approach combines automated feature selection with DA to enhance the generalization of lightweight, interpretable models.

In this work, we address three key aspects of ML-based lightpath QoT estimation with the goal of facilitating real-world deployment by focusing on model complexity, performance, benchmarking, interpretability as well as adaptability. First, we present RF as a meaningful, interpretable, and computationally efficient alternative for lightpath QoT estimation. Rather than introducing yet another baseline, our focus is on simpler ML models that have been relatively overlooked in the prior works, especially in comparison with more complex methods such as ANNs and XGBoost (Bergk et al., 2022; Ayoub et al., 2023). Although RF is a well-established ML technique, our contribution lies in demonstrating its effectiveness and practicality in the QoT estimation context, highlighting how its architectural simplicity translates into both operational efficiency and performance advantages. This contribution is significant as it provides practical insights into the operational feasibility and model selection criteria for QoT estimation tasks in real- world optical networks. Second, in contrast to the prior studies, we employ a fully automated and interpretable feature selection methodology, which enhances model transparency, reduces manual intervention in the modeling process, and supports scalable deployment across diverse network environments. Our implementation of RFECV and Boruta-SHAP represents a novel advancement over the heuristic-based method used in Ayoub et al. (2023), providing a replicable framework for other researchers working with publicly available datasets. Third, we introduce a feature-based DA approach that enhances alignment between source and target domains by leveraging automated feature selection to identify domain-invariant and relevant features prior to applying TL. This contrasts with the source domain baseline (SDB) approach described in Azzimonti et al. (2020), Di Marino, Rottondi, Giusti, and Bianco (2020), Rottondi et al. (2021) and proposed in Bergk et al. (2022), to address the limited availability of datasets related to topology, transceivers, and network status. Additionally, our proposed solution enables more stable transfer between domains by only sharing relevant features while applying the DA techniques.

To support multi-domain infrastructure deployments, network operators need fast, scalable and accurate decision-making for real-time lightpath provisioning, as well as model interpretability for easier troubleshooting and refinement. Keeping in mind the above mentioned preferences, in this paper we demonstrate the advantages of simpler models in terms of performance, interpretability, and adaptability across varied optical network architectures. Hence, the main motivation and impact of our contribution is the relevance for network operators, which has not yet been sufficiently addressed by other existing works.

3.4 Methodology

This section outlines the methodology, beginning with a description of the dataset, model selection, and comparative assessment strategy. We then detail the automated feature selection methods and DA techniques applied in this paper. Finally, we present the procedures for data preprocessing, feature selection, and feature-guided DA.

Figure 3.1 shows the overall process of our proposed QoT estimation approach with subprocesses such as data pre-processing, model selection, automated feature selection including the visualization of how the important features impact the classification of lightpath instances. The figure also illustrates the implementation of DA guided by the Boruta-SHAP-based feature selection, and the performance analysis of our proposed solution in inter-networks context.

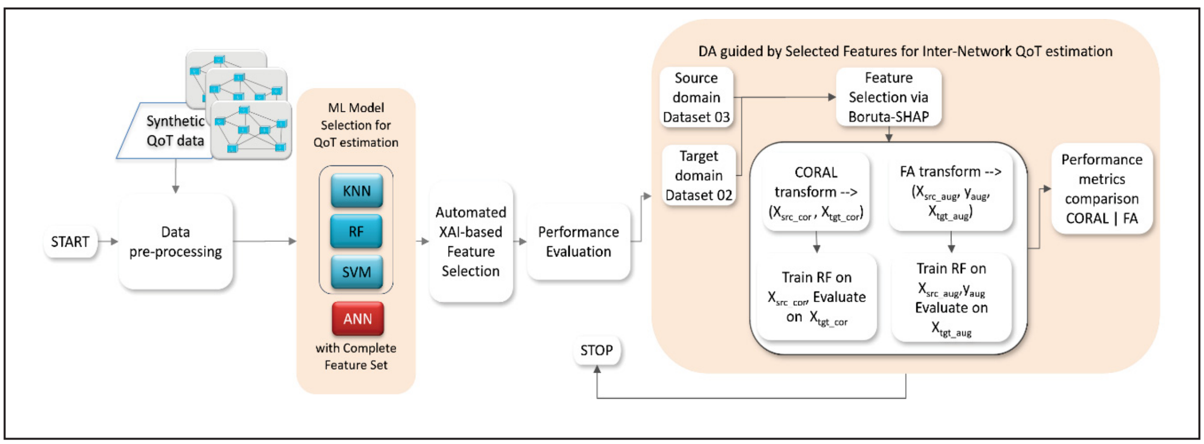


Figure 3.1 Overall process to implement automated, interpretable and adaptable lightpath QoT estimation models. Xsrc_cor and Xtgt_cor are source and target subsamples transformed with CORAL; Xsrc_aug, yaug, and Xtgt_aug are source subsamples with labels and target subsamples transformed by feature augmentation

3.4.1 Dataset description

This study is conducted using the QoT datasets that are made available to the research community in Bergk et al. (2022). The synthetic datasets are generated using the Fraunhofer Institute's planning tool for optical networks (PLATON). More information about the configuration and procedures to generate the data are available in Bergk et al. (2022). A QoT estimation (QOTE) module uses the Gaussian-Noise-based nonlinear channel model proposed by Poggiolini (2012) to calculate the Optical Signal-to-Noise (OSNR) on each link which helps estimating the QoT of lightpaths. The QOTE module estimates the QoT of new lightpaths through a provisioning procedure triggered by a traffic request. Based on the estimated QoT values, the new lightpath can be either established or blocked. QoT metrics of established and rejected connection requests are documented in PLATON along with numerous statistics in its Traffic Engineering Database (TED).

The simulation scenarios behind the gathered QoT data are based on elastic optical networks using the C-band. With a fixed-grid flex-rate network scheme and a fixed channel spacing of 37.5 GHz, the configuration allows for 96 channels carrying various data rates. Nominal central frequencies range from 192.2 to 195.7625 THz with the frequency grid anchored to 194 THz (Bergk et al., 2022). Possible modulation formats include polarization-multiplexed binary phase-shift keying (PM-BPSK), polarization-multiplexed quadrature phase-shift keying (PM-QPSK), and higher-order quadrature amplitude modulation (QAM) schemes such as PM-8QAM, PM-16QAM, PM-32QAM, and PM-64QAM. Physical link parameters are set as follows: the EDFA noise figure to 5 dB, the amplifier gain to 16 dB, the fiber attenuation coefficient to 0.2 dB/km, the fiber dispersion to 17 ps/nm/km, the fiber nonlinear coefficient to 1.3 1/W/km, the noise bandwidth to 12.5 GHz, and the reference wavelength to 1550 nm. The symbol rate is set to 28 GBd and the channel launch power to -3 dBm. All links consist of one or more 80-km spans of ITU-T G.652 SMF. Erbium-Doped Fiber Amplifiers (EDFAs) are used to compensate for span loss. A Hard Decision Forward Error Correction (HD-FEC) is implemented for error correction with a required minimum pre-FEC BER of 3.8 $\times 10^{-3}$. More details about the simulations to build the four datasets can be found in Bergk et al. (2022), included in the documentation provided with each dataset. With almost 1.2 million samples each, these datasets provide a wide range of simulated networking scenarios. Each dataset contains QoT metrics such as the OSNR, SNR, and BER as well as a binary class label indicating QoT compliance. Although the GN model used in PLATON includes neither inter-channel stimulated Raman scattering (ISRS) nor modulation format correction, these datasets include consistent structure across multiple simulated network scenarios and offer rich feature representation with topology-related, network-related and transceiver-related characteristics to support both network-wide and lightpath-based QoT estimation tasks, enabling diverse ML applications.

The lightpath dataset comprises 35 scalar features with different numbers of instances. Each subset includes 31 lightpaths' characteristics with 4 simulation-related meta-data attributes. Four target variables are associated to each instance: a binary class label based on a predefined BER threshold (BER_{th} = 0.0038), the BER, the OSNR and the SNR of the lightpath. Binary class labels can be positive (y = 1) or negative (y = 0). More information about the features and labels can be found in Bergk et al. (2022).

For each dataset, we adopt the same strategy as in Bergk et al. (2022) to have class-balanced subsets of 100,000 instances. We keep the same names for Datasets 01, 02, and 04 which derive from the CONUS optical network as well as Dataset 03 which is based on the TSNN. Table 3.1 shows a complete list of the features used as the input to our models.

Table 3.1 Description of features representing lightpath instances in the datasets with their measurement units

Feature	**Description**
Path Len (km)	Path length
Avg Link Len (km)	Average link length
Min Link Len (km)	Minimum link length
Max Link Len (km)	Maximum link length
Num Links	Number of links
Num Spans	Number of spans
Freq (THz)	Central frequency
Mod Order	Cardinality of the modulation format
LP Linerate (Gb/s)	Lightpath line rate
Source Node Degree	Source node degree
Destination Node Degree	Destination node degree
Min Link Occ	Minimum link occupation
Max Link Occ	Maximum link occupation
Avg Link Occ	Average link occupation

Table 3.1 Description of features representing lightpath instances in the datasets with their measurement units (continued)

Feature	Description
Std Link Occ	Standard deviation of the link occupation
Sum Link Occ	Sum of the spectral occupation of the links the lightpath traverses
Max BER	Maximum BER of interfering lightpaths
Min BER	Minimum BER of interfering lightpaths
Avg BER	Average BER of interfering lightpaths
Min Mod Order (L)	Minimum cardinality of the modulation format (left)
Max Mod Order (L)	Maximum cardinality of the modulation format (left)
Min Mod Order (R)	Minimum cardinality of the modulation format (right)
Max Mod Order (R)	Maximum cardinality of the modulation format (right)
Min LP Linerate (L) (Gb/s)	Minimum lightpath line rate (left)
Max LP Linerate (L) (Gb/s)	Maximum lightpath line rate (left)
Min LP Linerate (R) (Gb/s)	Minimum lightpath line rate (right)
Max LP Linerate (R) (Gb/s)	Maximum lightpath line rate (right)
Min BER (L)	Minimum BER (left)
Max BER (L)	Maximum BER (left)
Min BER (R)	Minimum BER (right)
Max BER (R)	Maximum BER (right)

3.4.2 Model selection and comparative assessment strategy

As mentioned in the Introduction, we first implement twelve models leveraging the four subsets of dataset 01, 02, 03, and 04 using the three classical ML techniques KNN, SVM and RF. After comparing their classification accuracies, the RF models outperform the two other models for dataset 01, 02, and 04, while the SVM model performed better with the dataset 03. Table 3.2 shows a summary of the results obtained for each scenario.

To ensure a fair comparative evaluation with the first formulation of Occam's Razor principle as stated in Domingos (1999), we compare models based on both classification performance and computational efficiency within a consistent execution environment. As in Bergk et al.

(2022), we implement ANN models for comparison with the RF classifiers. We use the *SciKeras* library version 0.13.0 with the module *scikeras.wrappers* to develop the classifiers.

Table 3.2 Classification accuracy (%)

Dataset	**KNN Train/Test**	**RF Train/Test**	**SVM Train/Test**
01	88.10/84.60	99.90/98.50	98.20/97.80
02	97.90/97.20	99.90/99.60	99.60/99.50
03	97.30/96.70	99.90/99.20	99.30/99.30
04	95.70/94.80	99.80/99.50	99.50/99.30

In the hyperparameter optimization step, we consider 100 epochs with 128/256 hidden units, the learning rate in {0.001, 0.01}, 0.1/0.2 for the model dropout rate and the batch size in {50, 100}. The rectified linear unit (ReLU) and the hyperbolic tangent (tanh) are the two activation functions tested for the hidden layer. For the RF models, we use the *scikit-learn* library version 1.7.1. In the hyperparameter tuning for the RF classifiers, we take two values for number of estimators {50, 75}, as well as for the maximum depth {27, 30} and three values for the maximum leaf nodes {66, 72, 80}. We keep the default values for the other parameters. We consider a 5-fold cross-validation using *GridSearchCV* from the *scikit-learn* library for both techniques.

Table 3.3 Performance comparison of RF and ANN models

Dataset	**ANN**		**RF**	
	Time (s) Train/Test	***Accuracy (%) Train/Test***	***Time (s) Train/Test***	***Accuracy (%) Train/Test***
01	884.76/1.45	98.45/98.10	85.86/0.21	99.12/98.59
02	881.27/1.43	99.57/99.60	26.65/0.12	99.79/99.64
03	887.61/1.27	99.18/99.34	63.30/0.20	99.41/99.24
04	872.62/1.34	99.45/99.45	35.40/0.14	99.75/99.50

The performance results in terms of classification accuracy and computation time for the new ANN and RF models are shown in Table 3.3. Only the classification performances are available for the models presented in Bergk et al. (2022). Our proposed RF models outperform both their and our ANN models for dataset 01, 02 and 04. However, our ANN model for dataset 03 performs better than both the ANN proposed in Bergk et al (2022) and our RF models (with classification accuracy of 99.34%, 99.03% and 99.24% respectively).

Moreover, as expected, the RF models consistently have faster computation times than the ANN models in all scenarios. The computation time presented in Table 3.3 for each scenario is achieved on a workstation using an Intel® Core™ i5-8250U 1.6 GHz CPU and 16 GB RAM. These findings reveal that RF-based models significantly reduce training time (by 10 to 33 times) and inference time (by 6 to 11 times) while maintaining classification accuracy.

We also perform a comparison of our RF-based models with the XGBoost models. To validate the replication of the solution proposed in Ayoub et al. (2023), we use a similar approach with the whole set of features for Datasets 03 and 04, including the simulation- related meta-data attributes, and without any pre-processing of the class-imbalanced instances in the two datasets. The replicated XGBoost models perform similarly to Ayoub et al. (2023) with accuracies of 99.60% and 99.84%, compared to 99.70% and 99.80% for Datasets 03 and 04, respectively. Moreover, to allow a fair comparison between our simpler RF-based solution and the proposed XGBoost-based models, we use our class-balanced sub-samples with the selected features from Ayoub et al. (2023). The RF-based models achieve almost similar accuracies with differences of 0.01% and 0.02% for Datasets 03 and 04, respectively, making it a formidable choice to assess how a simple approach compares to a complex one in a context of lightpath's QoT estimation, leveraging the available QoT datasets proposed in Bergk et al. (2022).

3.4.3 Automated feature selection methods and domain adaptation techniques

In Ayoub et al. (2023), the authors applied the recursive feature elimination (RFE) method to select important features from the datasets used to build their classifiers. But this method is only recommended when prior knowledge of the optimal number of features is available. We implement an automated approach using the following two methods.

RFECV is a feature selection method that automatically determines the most relevant features for a specific model. An RFE selector is used on different cross-validation splits. The number of selected features is set to the number of features that maximize the cross-validation score, obtained by evaluating the performance for different numbers of selected features and averaged across folds. This algorithm uses a supervised learning estimator that returns information about feature importance.

Boruta-SHAP is proposed here in order to leverage the robustness of the Boruta method and the interpretability of SHAP values. Boruta is a tree-based method working for tree models and other classification models that output feature importance measure. The algorithm shuffles original features to create shadow attributes at each iteration. The algorithm iteratively compares the importance of the features with the importance of the shadow attributes. Features are consecutively dropped or admitted when their importance compared to the shadow ones is worse or better, respectively. The algorithm stops either when all features are covered, or it reaches a specified limit. SHAP values help to determine feature importance based on principles of cooperative game theory. They allow a better understanding of the model's decisions by measuring the contribution of each feature to the classifiers' performance.

In Fawaz et al. (2024), the authors used another automated version of the techniques, namely BoostRFE, and BoostBoruta, applying SHAP values, but the solutions were designed for gradient boosting models such as light gradient boosting (LGBM) and XGBoost. With our RF-based models, we explore the two automated methods for the feature selection, i.e., RFECV and Boruta-SHAP, which both are compatible with RF.

The fully automated feature selection process retains only statistically significant features, eliminating the need to manually specify their number. This brings a notable improvement over the manual approaches used in the previous studies, e.g., in Ayoub et al. (2023).

We evaluate two DA methods as in Azzimonti et al. (2020), Di Marino et al. (2020), and Rottondi et al. (2021).

Feature Augmentation (FA) encodes the domain of a sample by augmenting its feature vector. In particular, the length of the original feature vector **x** is tripled with a rule that depends on the domain. This way, both commonalities between the two domains and the unique characteristics of each domain are captured. While for sample coming from the source domain S, the triple feature vector is defined as $x' = \langle x, x, 0 \rangle$, for sample coming from the target domain T, the triple feature vector is defined as $x' = \langle x, 0, x \rangle$. It is a supervised DA technique.

Correlation Alignment (CORAL) transforms the features in the source domain to match the second-order statistics of the features in the target domain. This allows minimizing the domain shift. A transformation is applied that re-colors the whitened features of the samples in S with the covariance computed from the feature distribution of the samples in T. The transformed data is used to train the model. This method does not require any information about the labels of samples in T; thus, it is considered an unsupervised DA technique.

In this study, we implement the source domain baseline (SDB) where the model is trained only with source domain data, for comparison with the two DA techniques. The other baseline scenarios such as target domain baseline (TDB) or dataset mixing baseline (DMB) are not considered in our assessment.

To improve the model's ability to find patterns, FA adds domain-specific information, such as statistical summaries or domain markers, to the original feature collection. CORAL applies a linear adjustment to align the second-order statistics (covariance) of the source and target feature distributions, minimizing domain shift (Rottondi et al., 2021).

3.4.4 Data preprocessing, automated feature selection and the feature-selection-based domain adaptation

The four gathered class-balanced subsets of data used in this work include 50,000 instances of each class, forming four sub-samples of 100,000 instances as in Bergk et al. (2022). The *StandardScaler* class from the *sklearn* package in Python is used to remove the mean and scale the data to unit variance as a way of standardizing the features. It is worth noting that while data scaling benefits techniques such as KNN, SVM, and ANN, it is intentionally not applied to the RF technique in order to preserve the interpretability of feature importance. Data scaling can improve gradient convergence in ANN, facilitate the construction of decision boundaries in SVM, and enhance distance calculations in KNN for more accurate neighbor selection. In contrast, RF does not rely on distance-based or gradient-based computations and therefore does not benefit from feature scaling. Furthermore, scaling the data can obscure the real-world meaning of feature values, making SHAP-based explanations less intuitive.

1) Automated Feature Selection

Reducing the number of features when building ML models simplifies the resulting models. This can lead to faster training and prediction time as well as easier interpretation and reduced risk of overfitting. Two methods, compatible with RF, are explored for an automated feature selection to have a better insight to how the parameters affect the lightpaths' QoT. We applied the RFECV and Boruta-SHAP methods to the four datasets and built RF-based classifiers using the reduced data sets obtained with the two methods.

After the model selection process, we perform the automated feature selection process. We start with the RFECV method, where we apply the scikit-learn ML library in Python, using the *RFECV* class. We pass our previously created RF-based estimator into the RFE with cross-validation instance to find the optimal number of features for each of the four class-balanced sub-samples. We keep the number of folds (cv parameter) to 5, its default value. The most relevant features are obtained by fitting the RFECV model. We then transform the training and test data using the RFECV model with the '*transform()*' method. Next, we train the final model using the transformed dataset with selected features to perform final classification of the instances in the feature-reduced test dataset. For the Boruta-SHAP implementation, we install the package and import the method. Similarly to the RFECV implementation, we pass our RF-based estimator into the BorutaSHAP instance. We set the importance measure parameter to '*shap*', and classification to '*True*'. We transform the training and test datasets in this case by dropping the less relevant features: we use the attribute 'features_to_remove' from the *BorutaSHAP* class which tracks the rejected features during the selection process. And for the final steps, as with the RFECV method, we train the model with the selected features then we predict the class for the transformed test dataset. These final steps allow us to compare the two strategies and establish our choice. While RFECV is inherently more efficient when used with large datasets, our findings show that simpler models built using class-balanced sub-samples can achieve better overall classification accuracy and reduced computation time. This makes them a more practical option for real-world applications where both efficiency and performance are critical.

To interpret the ML models built with reduced feature sets, we install the SHAP framework and use its explainer interface. The interface supports both local and global explanations: local explanations allow to visualize and interpret the model's decision for a specific instance, while global explanations reveal the contribution of each feature to the models' predictions. In our study, we use 'beeswarm' for global explanations and 'forceplot' plots for local explanations.

2) Selected-Feature-Based DA

We propose a DA framework based on the Boruta-SHAP feature selection performed in previous step. We integrate the Boruta-SHAP-based feature selection with the DA transformation into an efficient, explainable, and adaptable RF-based solution, addressing computational complexity and opacity in DL-based models while lowering the requirement for large amounts of training data in the development of supervised learning models. The pseudo-code of our approach is shown in Algorithm 3.1. We determine the source and target domains based on the worst-case scenario for domain adaptation inter-networks, namely with Dataset 03 from TSNN as the source domain and the Dataset 02 from CONUS as the target domain. Then we assess the two methods FA and CORAL, commonly used for transfer learning in QoT estimation to test our selected-feature-based models' generalization between different networks. In this case, we consider the scenario where the source domain is gathered from the small network TSNN, and the target domain from the more comprehensive network CONUS. We pick three different target sample sizes of 50, 100, and 1000 instances. After the first transformation performed using the Boruta-SHAP-based feature selection, we perform the second transformation using the CORAL method, then we fit the RF-based model using the transformed data. We predict the class for new data instances from the target domain. For the FA method, the outputs from the first transformation are used along with the labels to build

Algorithm 3.1 Feature-selection-based domain adaptation

```
Input:
    • Source dataset: X_src, y_src
    • Target dataset: X_tgt, y_tgt
    • Target sample sizes: TS = {50, 100, 1000}
Output: {accuracy, precision, recall, f1_score, auc,
computation time}
Feature selection phase
    Initialize Boruta-SHAP selector
    Selected_features ← BorutaSHAP.fit_transform (X_src, y_src)

    X_src_sel ← X_src[:, selected_features]
    X_tgt_sel ← X_tgt[:, selected_features]
Domain adaptation phase
    for each num in TS do:
        | Generate target samples based on the size
        | X_tgt_num, y_tgt_num
        | Method 1 : CORAL Domain Adaptation
        |     Initialize CORAL transformer COR
        |     Fit the COR model using (X_src_sel X_tgt_num)
        |
        |     Transform the two data samples
        |     Train the RF-based model using transformed source samples and labels
        |
        |     Predict classes for transformed target samples
        |
        | Method 2 : FA Domain Adaptation
        |     Generate domain indicators for source and
        |     target domains
        |     Concatenate the source samples with the source domain indicators and the
        |     target samples with the target domain indicators
        |
        |     Combine the obtained augmented data samples
        |
        |     Combine the source and target domain labels
        |
        |     Train the RF models using the augmented data and the combined labels
        |
        |     Predict classes for transformed target samples
        |
        | Jump to performance evaluation
    end for
Performance evaluation
    for each method m ∈ {CORAL, FA} do:
        | Compute performance_metrics_m
        | Store metrics in results[num][m]
    end for
```

the augmented feature vectors. The RF-based model is then fit using the new augmented data, and the resulting model is used to predict the class for unseen instances from the target dataset.

Moreover, we implement the SDB scenario, where only source domain data is used to build the model, with the reduced feature set. While the feature sets were arbitrarily chosen to build the models in Rottondi et al. (2021), our solution leverages a more comprehensive feature set, where the automated feature selection Boruta-SHAP allows transferring only relevant attributes.

We test our models using different sizes of target domain datasets and we present different performance parameters, along with the classification accuracy, to have better insight on the models' decision.

3.5 Numerical results and comparative analysis

This section highlights the performance of automated feature selection and evaluates the effectiveness of the proposed feature-guided DA framework in an inter-network scenario.

3.5.1 Boruta-shap-based feature selection

The two aforementioned feature selection methods, i.e., Boruta-SHAP and RFECV, are used to select relevant features from the four datasets. In the context of lightpath QoT estimation, both the classification accuracy and the training time are good parameters to evaluate models' performance.

As shown in Table 3.4 the Boruta-SHAP method achieves better performance than the RFECV method with slightly higher classification accuracy and faster training time for Datasets 02 (99.68% and 8.89 seconds against 99.67% and 10.3 seconds) and 04 (99.56% and 10.01 seconds against 99.55% and 15.62 seconds), while the RFECV method performs better for Datasets 01 (98.59% and 7.14 seconds against 98.55% and 17.98 seconds) and 03 (99.29% and 7.82 seconds against 99.25% and 14.91 seconds). We pick the Boruta-SHAP method for the rest of this study because it provides superior global and local interpretability compared to the alternative RFECV method used in this context of explainable models for lightpath's QoT estimation using the Fraunhofer QoT datasets. Indeed, while RFECV focuses on eliminating features based on the model's internal feature importance, Boruta-SHAP uses the SHAP values for consistent feature attributions and interpretability.

Table 3.4 Performance comparison of RFECV and Boruta-SHAP

Dataset	**Boruta-SHAP**		**RFECV**	
	Time (s) Training	***Accuracy (%)***	***Time (s) Training***	***Accuracy (%)***
01	17.98	98.55	7.14	98.59
02	8.89	99.68	10.3	99.67
03	14.91	99.25	7.82	99.29
04	10.01	99.56	15.62	99.55

The resulting sets of features obtained with Boruta-SHAP are of 14 features, 10 features, 17 features, and 10 features for Dataset 01, 02, 03, and 04 respectively. The rankings of selected features for the four datasets are displayed in Table 3.5.

With the class-balanced sub-samples, we can see that the same more important attributes appear for the four datasets with similar ranking for the first four features, with the exception for Dataset 03 where the sum of the spectral occupation of the links traversed by the lightpath (S*um Link Occ*) is ranked 3rd while it is ranked 7th for the other 3 datasets. And the path length (*Path Len*) is ranked 7th for Dataset 03 while it is third for all other datasets. This can be supported by the path length distributions per modulation format and data rate displayed in Bergk et al. (2022) for the two network topologies CONUS and TSNN. The first four important features for the sub-samples gathered from the CONUS topology are the cardinality of the modulation format (*Mod Order*), the number of spans *(Num Spans*), the path length *(Path Length*), and the average link occupation (*Avg Link Occ*). And for Dataset 03 from TSNN topology, the four most important features are Mod order, Num Spans, Sum Link Occ, and minimum link occupation (Min Link Occ). This confirms that datasets from CONUS exhibit greater variation in Path Len than the one from TSNN, and the distributions of Sum Link Occ and Min Link Occ are more heterogeneous in Dataset 03. Despite using class-balanced subsamples of the datasets and a fully automated feature selection process, we obtained similar results for the top three most important features across Dataset 03 (Mod Order, Num Spans, Sum Link Occ) and Dataset 04 (Mod Order, Num Spans, and Path Len), aligning closely with the outcomes reported by the solution proposed in Ayoub et al. (2023). This highlights one of the key advantages of using publicly available datasets: it increases our confidence in the outcomes of the proposed solution. More importantly, the process enables us to focus on the most informative features, reducing feature-space noise and improving computational efficiency. Additionally, leveraging only relevant features enhances model interpretability.

Table 3.5 Selected features from the most to the least significant in model’s decision for each dataset

Rank	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
D01	Mod Order	Num Spans	Path Len	Avg Link Occ	Freq	Min Link Occ	Sum Link Occ	Max Link Occ	Max BER (R)
D02	Mod Order	Num Spans	Path Len	Avg Link Occ	Freq	Min Link Occ	Sum Link Occ	Min Mod Order (L)	Max BER (R)
D03	Mod Order	Num Spans	Sum Link Occ	Min Link Occ	Freq	Min Mod Order (R)	Path Len	Max BER (R)	Avg Link Occ
D04	Mod Order	Num Spans	Path Len	Avg Link Occ	Min Link Occ	Freq	Sum Link Occ	Min BER (R)	Max BER (R)
	10	**11**	**12**	**13**	**14**	**15**	**16**	**17**	
D01	Min BER (R)	Avg Link Len	Min Mod Order (R)	Min Link Len	Min Mod Order (L)				
D02	Min BER (L)								
D03	Min BER (R)	Max Link Len	Min BER	Std Link Occ	Min Mod Order (L)	Max BER (L)	Min BER (L)	Max LP Linerate (L)	
D04	Source Node Degree								

Figure 3.2 and Figure 3.3 show the SHAP summary plots for the models built with the reduced feature sets for global explanation purposes. The figures show only the most relevant features, amongst the 31 attributes, for lightpath QoT estimation using the four datasets. These two figures display bee swarm plots, with features ordered by their impact on the classifier's

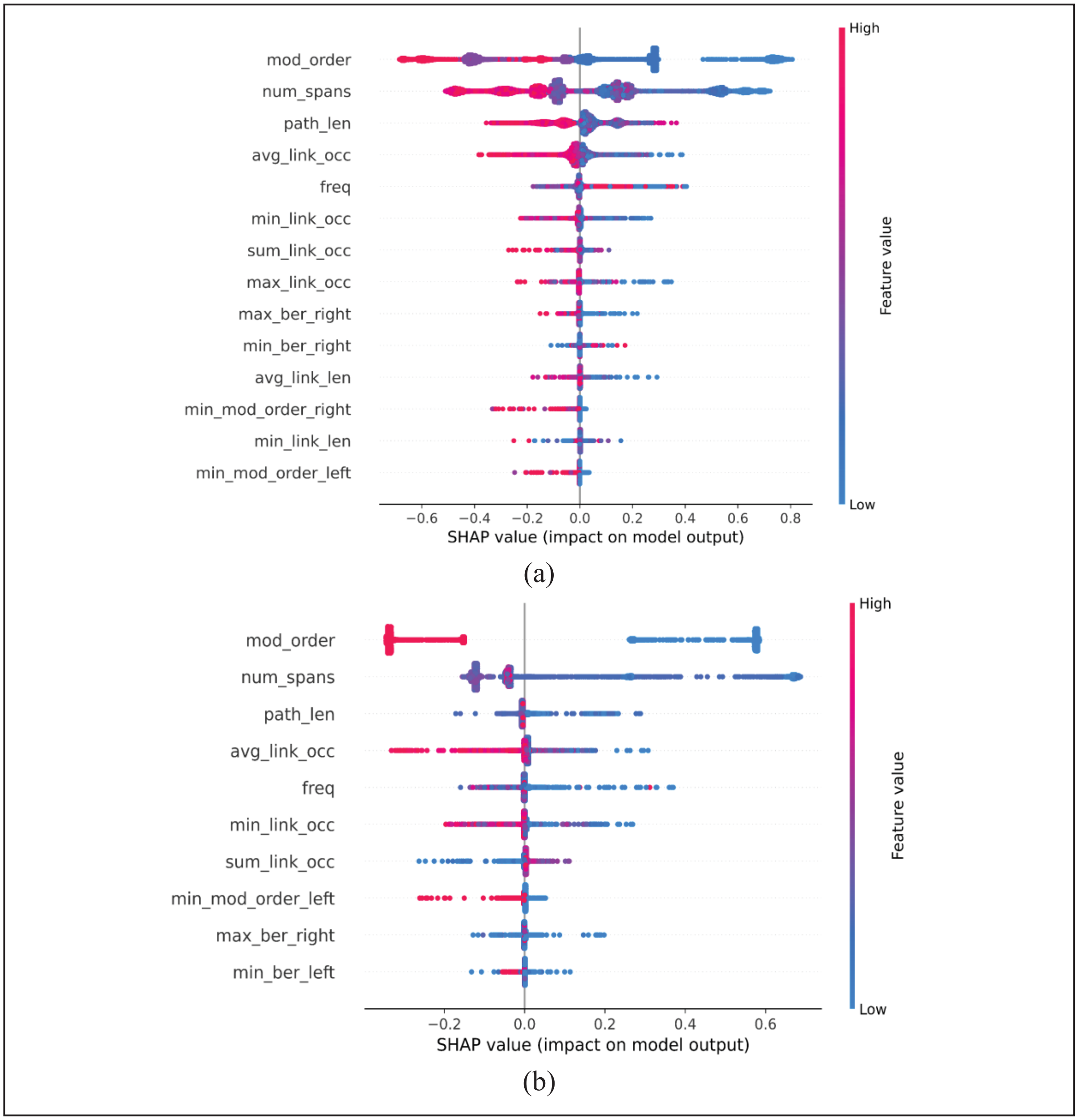


Figure 3.2 SHAP summary plots for the RF-based models using: (a) Dataset 01 and (b) Dataset 02

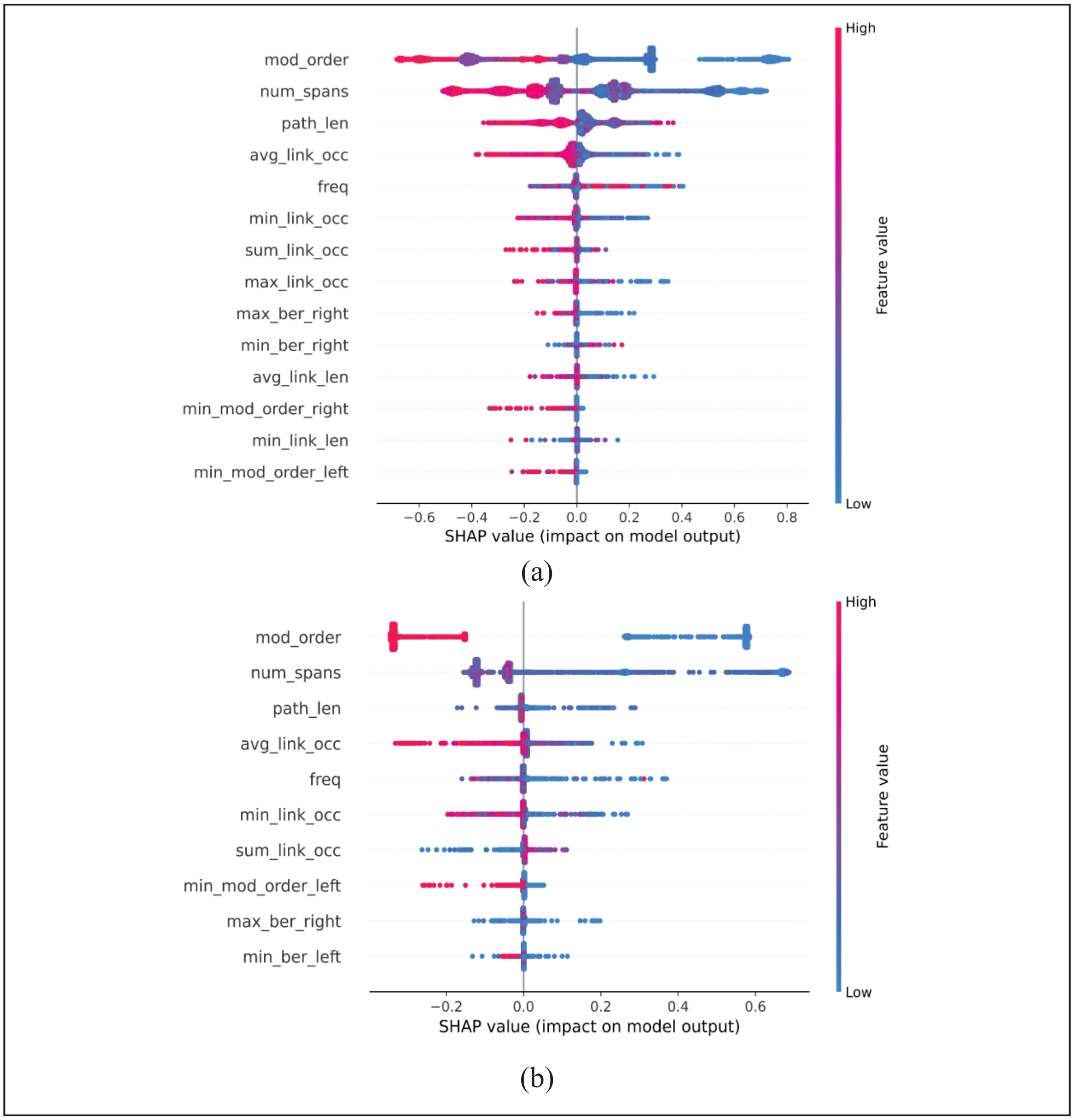


Figure 3.3 SHAP summary plots for the RF-based models using: (a) Dataset 03 and (b) Dataset 04

decisions as measured by SHAP values. Each dot represents an individual observation, with color indicating whether the corresponding feature is high or low relative to other instances in the dataset. Negative SHAP values indicate a negative impact on the model's decision, while positive values correspond to a positive impact. For all four datasets, high values of *Mod Order* contribute negatively to the classification. In other words, a higher modulation format leads the model to assign a negative class label (class 0) to the corresponding instance. Similarly, high values of *Num Spans* exhibit a similar pattern across all datasets, with the exception of Dataset 02 shown in Figure 3.2 (b). This aligns with the observation that, in this specific subsample, all instances with a class 0 label have low *Num Spans* values ranging from 23 to 39, while the instances with a positive class label (class 1) have *Num Spans* values between 1 to 102. Another interesting observation in Figures 3.2 (a), 3.2 (b), and 3.3 (b), concerns the feature *Path Len*, which ranks as the third most important feature across all datasets, with the exception of Dataset 03 in Figure 3.3 (a). The visualization proposed in Bergk et al. (2022) shows *Path Len* values ranging from 84 to 1382 km for Dataset 03 while the feature values range from 24 to more than 7000 km for the three other datasets. Although this feature is ranked seventh in Dataset 03, its contribution to the model's decisions remains consistent across Dataset 01 and Dataset 04. High values of *Path Len* are associated to class 0 while low values tend to lead the model toward class 1, as seen in Figure 3.2 (a), Figure 3.3 (a) and (b). For the model built with Dataset 02, we can see in Figure 3.2 (b) that high values of *Path Len* have no significant impact on the model's decision. This is consistent with the observation that most of the instances in the subsample with high values of *Path Len* (values greater than 2640 km) are labeled as class 1, as clearly shown in Figure 4 published in Bergk et al. (2022).

In real-world scenarios, this Boruta-SHAP-based feature selection integration allows not only to identify the relevant features for QoT estimation of unestablished lightpaths while preserving the classification performance, but with the knowledge provided on the impact of important features on each model, network operators can develop domain-specific strategies to address different SHAP patterns for different network topologies. Also, Figures 3.2 and 3.3 can help in prioritizing features with high impact on the models in provisioning decisions.

The *shap* package in Python allows displaying local explanation about the classifier's decision built with selected features using the Boruta-SHAP method. Figure 3.4 shows examples of misclassified instances for Dataset 01 for a false positive and a false negative case respectively, namely local explanation examples. Figure 3.4 (a) illustrates a false positive case in which the modulation order, the path length and the number of spans increase the model's confidence in predicting class 1. For this particular instance, only the *Avg Link Occ* contributes negatively to the model's prediction score. A lightpath instance, characterized by a *Path Len* of 3462 km, *Num Spans* of 48 and *Mod Order* of 8, supports a prediction of "class 1". However, an *Avg Link Occ* of 66.22 reduces the model's confidence in this prediction. Notably, the ground truth label for this instance is class 0. Figure 3.4 (b) shows a false negative case in which *Mod Order* of 8 increases the model's confidence in predicting class 1. But the values for *Num Spans* of 51, *Avg Link Occ* of 53.22, *Path Len* of 3686.93 km, and *Freq* of 193.88 THz all contribute negatively to the model's prediction score, leading to a "class 0" label. The ground truth label in this case is class 1. The length of the bar, representing a feature in the figure, indicates the magnitude of its impact on the classifier's decision. The *Mod Order* is the most influential feature contributing to the false positive shown in Figure 3.4 (a), while the *Num Spans*, *Avg Link Occ*, and *Path Len* have a greater impact on the false negative shown in Figure 3.4 (b). This transparency in the decision-making process enables a deeper understanding of model misclassification and supports further analysis and strategies to complement the automation provided by ML. A step added to the lightpaths' provisioning process where the knowledge provided by the Boruta-SHAP feature selection can be used to review lightpaths' requests presenting specific feature values as the ones shown in Figures 3.4 (a) and (b). Reviewed requests can be added back to the training set to improve the ML models.

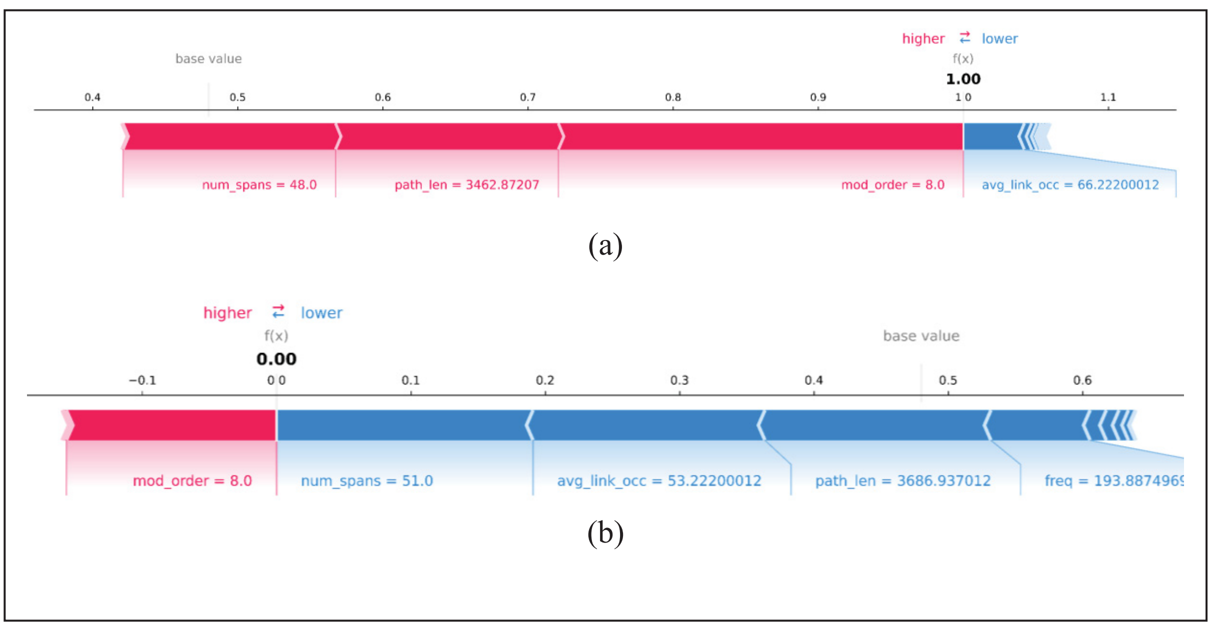


Figure 3.4 Examples of misclassification from Dataset 01: (a) False positive and (b) False negative

The RF-based model constructed with the 14 selected features outperforms the model built with all 31 features for Dataset 01, with accuracies of 98.64% and 98.6% respectively. And the classifier built with 10 features from Dataset 02 performs slightly better than the 31-feature classifier, with 99.65% and 99.64% respectively. The model built with 17 features from Dataset 03 achieves a classification accuracy of 99.27%, which is better than the performance of the model with all features, which is 99.24%. Lastly, we observe a difference of 0.02% between the model built with 10 features from Dataset 04 and the one built with 31 features (99.52% and 99.5%, respectively).

Table 3.6 Performance of the RF-based models using complete vs. reduced features sets

Dataset	**Training Time (s)**		**Inference Time (ms)**	
	Complete	***Reduced***	***Complete***	***Reduced***
01	21.98	**10.38**	73.80	**71.80**
02	11.44	**4.43**	47.80	**33**
03	16.24	**11.14**	60.80	**41.90**
04	21.18	**6.21**	71.80	**41.19**

Although there is only a slight improvement in the classification accuracy for each model, while running on a workstation using an Intel® Core™ i7-9700T 2.00GHz CPU and 16 GB RAM, they show clear reductions in the training times: from 21.98 to 10.38 seconds for Dataset 01, from 11.44 to 4.43 seconds for Dataset 02, from 16.24 to 11.14 seconds for Dataset 03, and from 21.18 to 6.21 seconds for Dataset 04. This represents reductions of 52.78%, 61.28%, 31.40%, and 70.68%, respectively. Moreover, inference times decrease from 73.8 to 71.8 milliseconds for Dataset 01, from 47.8 to 33 milliseconds for Dataset 02, from 60.8 to 41.9 milliseconds for Dataset 03, and from 71.8 to 41.9 milliseconds for Dataset 04. This represents reductions of 2.71%, 30.96%, 31.09%, and 41.64% for reduced feature sets of Dataset 01, 02, 03, and 04 respectively.

Table 3.6 shows a summary of the training time and inference speed of the classical RF-based classifiers built with the complete sets of data and with the reduced feature sets for comparison. These findings show that, with automated feature selection, simpler RF-based models can achieve accuracy comparable to full-feature models while significantly reducing training and inference times. This highlights the practical advantages of using lightweight, interpretable models in time-sensitive applications such as real-time QoT estimation without sacrificing classification performance.

The inference times in Table 3.6 are achieved with 30,000 test samples. Consequently, the times from 33 to 72 ms presented are the total inference times for the whole test datasets. Thus, on a workstation using an Intel® Core™ i7-9700T 2.00GHz CPU and 16 GB RAM, the average inference times per lightpath vary from 1.1 μs to 2.4 μs. These values are acceptable for practical real-world scenarios if we consider the target of well below one second per lightpath set for the computational time by the Open Optical & Packet Transport-Physical Simulation Environment (OOPT-PSE) group within the Telecom Infra Project (TIP) (Ferrari et al., 2020).

It is worth noting that the order of magnitude of the average inference times per lightpath achieved by our RF-based classifiers is comparable to those reported in Fawaz et al. (2024), where the models were evaluated using a more powerful environment (T4 GPU as the primary hardware accelerator, 13 GB of RAM, and approximately 80 GB of disk space) and a smaller, in-house dataset of 10,000 instances. The CatBoost model with a FS using the model's built-in feature importance ranking in Fawaz et al. (2024), achieves the lowest inference time of 0.49 ms for a 2000-instance test dataset (average 0.24 μs per lightpath), while the lowest inference time achieved by our RF-based model is 33 ms for a 30,000-instance test dataset (in average 1.1 μs per lightpath).

3.5.2 RF-based transfer learning guided by selected features

In this section, we describe the case where we apply a TL approach by utilizing the selected features to construct four source models, each corresponding to one of the datasets. As mentioned in Section 3.4.2, we apply a feature-representation transfer strategy, where the features selected from each source domain with the pre-trained RF models are used to transform the target domain. We perform a transformation of the target domain to find shared representation in the data, as explained in the section about the DA in shallow or classical learning presented in Singhal, Walambe, Ramanna, and Kotecha (2023). This feature transformation method corresponds to the SDB presented in Rottondi et al. (2021). We then evaluate each pre-trained RF model against the three other datasets with reduced feature sets. The outcome of this study is shown in Table 3.7.

Table 3.7 Classification accuracies of the RF-based models trained on source datasets on target datasets

Target	**Dataset01**	**Dataset02**	**Dataset03**	**Dataset04**
Source				
Dataset01	-	**99.33**	98.62	99.17
Dataset02	**70.21**	-	49.93	65.39
Dataset03	**60.35**	49.93	-	57.98
Dataset04	98.19	**99.50**	98.30	-

Our results confirm that Dataset 01 is the most comprehensive of the four subsamples, with classification accuracies above 99% on Dataset 02 and Dataset 04

In addition, with the selected features the proposed RF-based models achieve 98.62% on Dataset 03, which is better than the performance obtained in Bergk et al. (2022). Moreover, models built with Dataset 02 and Dataset 03 achieve classification accuracies below 71% on each of the other datasets. It is worth noting that our RF-based model trained with Dataset 04 achieves an accuracy of 98.19% on Dataset 01, while the model trained with Dataset 01 achieves 99.17% accuracy on Dataset 04. Similarly to the results in Bergk et al. (2022), the model trained with Dataset 01 achieves more than 99% accuracy in classifying test instances from Dataset 04, even though the BPSK modulation format does not appear in Dataset 01. However, we could see an increase in classification accuracy (from 73.3% to 98.19%) for the transfer learning on the target domain Dataset 01 using the model trained with Dataset 04, as compared to the performance obtained in Bergk et al. (2022). This may be because of the applied feature selection process which, with the class-balanced subsamples, allows to find new feature combinations to gain better insight on how to apply feature-based transfer learning as well as which attributes are important for QoT estimation in optical networks.

These results show that RF can face generalization challenges when DA is performed across heterogeneous domains using the SDB approach, in contrast to the conclusions reported in Rottondi et al. (2021). Nonetheless, the results obtained using the FA and CORAL feature-selection-based methods exhibit a significant improvement in classification accuracy for one of the worst-case scenarios presented in Table 3.7. This worst-case scenario corresponds to TL from source Dataset 03, representing the smaller TSNN network, to target Dataset 02, associated with the larger CONUS network.

Our novel TL framework integrates the feature selection and domain adaptation into a unified algorithm. We use the Boruta-SHAP technique to build a feature importance mask that captures the most relevant features from the source domain while providing insights on classification decision made by the model. By using the mask to transform both the source and target domain data, we ensure that adaptation is performed only over the most meaningful and statistically stable features. Following the first transformation using the Boruta-SHAP to select only relevant features, the solution further adjusts the joint source and target data to minimize domain shift, allowing the model to learn transferable relationships. Rather than applying the two components independently, our approach combines them into an integrated transformation that aligns domains, improves generalization on small target samples, and provides interpretability of classification outcomes.

Overall better accuracies of 86%, 98%, and 99.10% are achieved with FS-FA as compared to 82%, 97%, and 99% obtained with the standard feature augmentation technique for target domains of 50, 100, and 1000 instances respectively as shown in Table 3.8.

We test our proposed feature-selection-based feature augmentation (FS-FA) DA framework with the same source and target domain sizes used with the SDB approach. The proposed solution achieved a classification accuracy of 99.54% with Dataset 03 as the source domain and Dataset 02 as the target domain. That is an increase of 99.35% in classification accuracy when we compare it to the 49.93% achieved with the source domain baseline approach.

The results show that increasing in the number of samples in the target domain allows the combined (FS-FA) framework to achieve better performance, with a slight increase in training time ranging from 10.42 to 10.68 and 10.91 seconds for 50, 100, and 1000 instances, respectively. The same trend is observed for the inference time which varies from 10.1 to 10.9,

Table 3.8 Performance metrics of the DA models with Dataset 03 as source domain and Dataset 02 as target domain

	Accuracy (%)	***Precision (%)***	***Recall (%)***	***F1-Score (%)***	***AUC (%)***	***Train/Test Time (s)/(ms)***
Target domain size = 50						
CORAL	*68.00*	*73.33*	*73.33*	*73.33*	*69.33*	*46.19/10.2*
FA	*82.00*	*95.65*	*73.33*	*83.01*	*94.75*	*12.482/10.1*
FS-CORAL	*82.00*	*83.87*	*86.66*	*85.24*	*89.50*	*29.04/10.1*
FS-FA	*86.00*	*96.00*	*80.00*	*87.27*	*94.16*	*10.42/10.1*
Target domain size = 100						
CORAL	*80.00*	*82.00*	*78.84*	*80.39*	*88.10*	*44.41/11.9*
FA	*97.00*	*98.03*	*96.15*	*97.08*	*99.41*	*12.94/18.3*
FS-CORAL	*55.00*	*53.93*	*92.30*	*68.08*	*78.14*	*28.02/11.9*
FS-FA	*98.00*	*100.00*	*96.15*	*98.03*	*99.75*	*10.68/10.9*
Target domain size = 1000						
CORAL	*84.30*	*84.70*	*81.27*	*82.95*	*91.42*	*44.60/18.1*
FA	*99.00*	*99.78*	*98.08*	*98.92*	*99.89*	*13.02/10.1*
FS-CORAL	*64.70*	*57.72*	*92.97*	*71.23*	*78.47*	*28.34/14.9*
FS-FA	*99.10*	*100.00*	*99.03*	*99.03*	*99.84*	*10.91/15.9*

and 15.9 milliseconds for target domain sizes of 50, 100, and 1000 instances, respectively. For a better view of trade-offs between sensitivity and specificity, we calculate the area under the curve (AUC) for each scenario. A similar trend is observed for target domain sizes of 50, 100, and 1000 instances with AUC values of 94.16%, 99.75%, and 99.84% respectively. The proposed framework achieves faster training times in the three scenarios and faster inference times in average, along with the combined feature selection with CORAL with 12.3 seconds as compared to 12.8 and 13.4 milliseconds for standard FA and CORAL respectively.

Moreover, while the model built with the CORAL method exhibits reduced training times after combining it with the feature selection (FS-CORAL) compared to its counterpart without feature selection, the results indicate that its performance degrades as the size and the complexity of the associated feature-space of the target domain increases. The classification performance for the CORAL method decreases from 80% to 55% and from 84.3% to 64.7% for the model without feature selection and the feature-selected-based model with target domain sizes of 100 and 1000 respectively.

Ultimately, we obtain overall better results while applying the proposed (FS-FA) DA framework using RF-based model than when applying the standard DA methods (CORAL or FA) without prior feature selection. Thus, we conclude that the proposed FS-FA framework significantly outperforms standard DA methods, which use FA and CORAL without applying the FS method, in classification performance as well as computational efficiency for QoT estimation of lightpaths before their establishment. We perform additional tests using Dataset 01 as the source domain and subsamples from the three other datasets (Dataset 02, 03, and 04) as target domains. An ablation style comparison of the standard DA techniques (standard FA) with our proposed framework (FS-FA) shows consistent improvements in the computation times with the small subsamples of 50 instances. The training time decreases from 17.41 to 10.20 seconds and the inference time from 8.97 to 8.49 ms for DA between Dataset 01 and Dataset 02. The same trend is seen for DA between Dataset 01 to Dataset 03 and Dataset 04 with the training time decreasing from 17.48 to 10.16 seconds and 17.0 to 10.39 seconds, and the inference time reductions from 9.97 to 8.97 ms and 9.47 to 8.99 ms, respectively, without sacrificing the classification accuracy. Our Boruta-SHAP-based method yields shorter computation times while maintaining the same classification accuracy of 98% for both DA approaches, when using the standard feature augmentation and the Boruta-SHAP-based feature augmentation schemes, with Dataset 01 as the source domain and Datasets 02, 03, and 04 as target domains.

In contrast to the hybrid solution integrating KD with TL proposed in Usmani et al. (2024) where knowledge is transferred from a bigger "teacher" ANN model to a smaller "student" ANN model, our proposed approach leverages RF-based classifiers with Boruta-SHAP-based feature selection providing model transparency with lower complexity. Our TL framework automatically identifies important features while in previous works the authors manually build feature vector with characteristics of the considered lightpath as well as the ones of the neighboring channels (Rottondi et al., 2021), and other scholars filter samples by manually selecting length and number of adjacent channels features (Gu et al., 2023).

Our framework achieves reduction in both training and inference times while preserving the classification accuracy with 99.83% reduction in the size of the target domain. The Boruta-SHAP-based feature augmentation (FS-FA) achieves reduction of the training times by 41.41%, 41.89%, and 38.88%, as well as reduction of the inference times by 5.35%, 10.03%, and 5.07% for Dataset 01 to Dataset 02, 03, and 04, respectively. Although the authors of Usmani et al. (2024) used a workstation with a 12th Generation Intel® Core™ i7 12700F 2.10 GHz CPU and 32GB RAM, which is superior to our workstation with an Intel® Core™ i7-9700T 2.00GHz CPU and 16 GB RAM, their proposed KD+TL model, with their 6,000-instance target data samples, achieves training time of 6.7 seconds and prediction time of 0.09 seconds while our framework achieves training and inference times of 2.51 and 0.04 seconds when we use a 12,000-instance subsample from Dataset 01 as source domain and a 6,000-instance subsample from Dataset 03 as target domain. Moreover, while a direct comparison with Natalino et al. (2024) is not entirely feasible due to differences in the experimental environments, datasets, and ML methods, our lightweight RF-based model nonetheless achieves a better inference time - by an order of magnitude - than the AI/ML-as-a-service framework reported therein. In Natalino et al. (2024), the average inference time achieved by the ANN model, tested with a 10,000-instance subsample gathered from Dataset 01 on an

Ubuntu 22.04 AMD Ryzen Threadripper 3960X 24-Core Processor is 40 ms while their proposed AI/ML-as-a-service framework achieves 553 ms per inference, due in part, to the ensemble of models implemented in the framework. We demonstrate that our framework achieves similar order of magnitude for both the training and the inference times when compared to Usmani et al. (2024) for domain adaptation. When compared to Natalino et al. (2024), our lightweight RF-based model achieves better inference times than the proposed AI/ML-as-a-service framework. This shows the benefits of lightweight RF-based models compared to ANN-based models for QoT estimation of unestablished lightpaths.

3.6 Conclusions and future research direction

We develop an interpretable and efficient QoT estimation framework tailored in particular for the real-world use cases, where the training datasets are scarce and the training time is a critical factor. A comparative analysis using the four publicly available datasets from the Fraunhofer Institute with different representative techniques clearly demonstrates the advantages of our approach. In particular, we show effectiveness of classical RF-based models in estimating the QoT of unestablished lightpaths, offering enhanced efficiency and interpretability while preserving classification accuracy.

We propose an efficient automated approach to select the most important features, guiding the models in classifying test instances across each class-balanced sub-dataset. Our approach reduces the dimensionality of the training data (from 31 features to 14, 10, 17, and 10 for Datasets 01, 02, 03, and 04, respectively), thereby significantly lowering processing time while preserving, or even slightly improving, classification accuracy. Moreover, the combination of SHAP values with the Boruta technique, referred to as the *Boruta-SHAP method,* provides deeper insights into how important features influence the model's classification decisions for test instances. Basic visual error analysis helps uncover potential explanations for the two misclassification cases presented in Figure 3.4 (a) and (b). The RF-based models, built with reduced feature sets, show clear reductions in the training times of 52.78%, 61.28%, 31.40%, and 70.68% for the four sub-samples obtained from Datasets 01, 02, 03, and 04, respectively.

Additionally, the inference times also show reductions of 2.71%, 30.96%, 31.09%, and 41.64% for the four sub-samples, respectively. This demonstrates the practical applicability of RF-based models in real-world scenarios.

Furthermore, we implement a feature-selection-driven feature augmentation (FS-FA) domain adaptation framework to evaluate the cross-domain performance of models trained on one dataset and tested on others. In this approach, Boruta-SHAP-based feature selection is applied prior to the feature transformation step for transfer learning. Depending on the subsample used as the source domain, a pre-trained models achieve accuracies ranging from 49.93% to 99.50% with an SDB approach. Our framework reaches a classification accuracy of 99.54% with the complete 30,000 samples target domain used with the SDB approach and of 86% with a small target domain of 50 samples. Moreover, our framework shows a reduction in training time of 77.44%, 16.52%, and 64.11% for the small target dataset, when compared to the standard CORAL, the standard FA, and the FS-CORAL methods, respectively. The framework also exhibits fast inference time of 10.10 milliseconds, which makes it an efficient solution for deployment in real-world optical networks.

As a future research direction, we plan to explore more advanced DA strategies to improve the generalizability of QoT estimation models across diverse network scenarios. And we will perform simulations with other target domain sizes to determine appropriate sizes for real-world implementations.

CHAPTER 4

DETECTING ANOMALIES IN THE OPTICAL LAYER USING UNSUPERVISED MACHINE LEARNING

Sandra Aladin [a,b], Lena Wosinska [b], Christine Tremblay [a]

[a] Department of Electrical Engineering, École de technologie supérieure,
1100 Notre-Dame Ouest, Montréal, Québec, Canada H3C1K3

[b] Department of Electrical Engineering, Chalmers University of Technology,
412 96 Gothenburg, Sweden

Article published in *Optical Fiber Communications Conference and Exhibition (OFC)*, San Diego, CA, USA, March 2024, pp. 1-3, Th3I.4.

4.1 Abstract

We propose an unsupervised machine learning (ML) approach using field data for the detection of optical layer anomalies. We show how multivariate ML models can forecast hard failures by detecting soft failures.

4.2 Introduction

The rapid development of emerging technologies and online services, such as real-time video and cloud services, leads to the drastically increasing capacity demand and infrastructure complexity as well as a need for high network flexibility. In such a context, any network failure can lead to massive loss of data and customers churn due to non-compliance with the service level agreements. It is thus particularly important to ensure high quality of service (QoS) for all applications requiring high bandwidth and/or ultra-low latency and jitter.

Machine learning (ML) has been proposed to enable proactive fault management in optical networks. Network faults can be categorized into hard failures linked to immediate service disruptions, and soft failures which are linked to gradual quality of transmission (QoT) degradations in optical networks. The main objective of an effective ML-based fault management system is to detect, identify and localize soft failures in order to prevent them from deteriorating to hard failures. In Chen, Li, Proietti, Zhu & Yoo (2019), the authors proposed a hybrid unsupervised and supervised ML approach for anomaly detection in optical networks. Their framework analyzes patterns of experimental optical performance monitoring (OPM) data using an unsupervised data clustering module. Then, the learned patterns are fed to a supervised data regression and classification module for online anomaly detection. The authors in Liu et al. (2021b) and Behera et al. (2023a) , similarly to aforementioned authors, used experimental imbalanced data with a semi-supervised approach to detect failures in optical networks, and synthetically generated data with encoder-decoder long short-term memory (ED-LSTM) to model and predict the evolution of soft failures for several days with the aim to trigger proactive maintenance and component replacement actions before hard failures occur. In Du et al. (2021) and Allogba et al. (2022b), the authors used optical performance monitoring data gathered from production networks. The first study showed that it was possible to use ML techniques to predict the loss of signal with good precision 1 to 7 days before they occur. Similarly, the second study presents an ML-based model to detect performance degradations in optical networks. In both studies, the authors used a supervised learning approach to detect QoT degradations in the OPM data, which required a labeling phase to perform loss of signal forecast and anomaly detection tasks.

In this study, we propose the use of three unsupervised learning techniques to detect anomalies from the OPM data. Then we use the labeling principles based on the alarm parameters in the dataset as in Du et al. (2021), to visualize and compare the results of the ML-based anomaly detection models.

4.3 Methodology

4.3.1 Anomalies in OPM data

Figure 4.1 shows the OPM metrics for one lightpath with the anomalies defined from alarms over 4 years. In the context of this study, anomalies represent instances of OPM parameters that do not conform to the majority of values recorded over a collection period. Point anomalies are single data points that are significantly different from the rest of the data. Contextual anomalies are instances that deviate from the other data points in a particular context, while being considered normal in a different context. A group of data points which, taken individually may be considered normal, but deviate considerably from the rest of the data points when gathered together, represent the collective anomalies (Mejri et al., 2024). A point-anomaly detection approach was adopted for the algorithms implementation as opposed to a collective or a contextual anomaly detection approach.

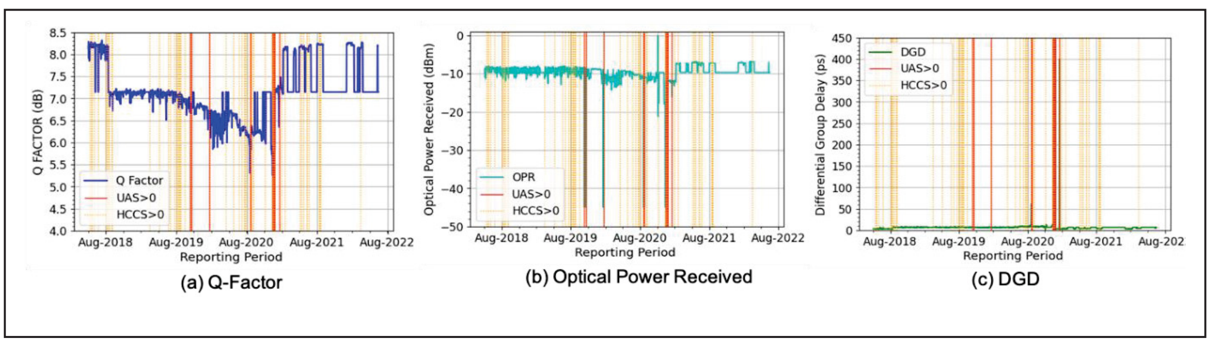


Figure 4.1 Anomalies observed in the OPM metrics for one lightpath during 4 years. UAS: unavailable seconds; HCCS: high correction count seconds; OPR: optical power received; DGD: differential group delay

4.3.2 Dataset description and preprocessing

For this study, we used a dataset of daily binned PMs collected from equipment in the production network of a North American Service Provider over 4 years. The PM metrics are pulled from specific ports of these devices linked to specific facilities. We gathered instances for the optical transport module services among others (Ethernet, OC-n, …), encapsulated by the optical transport network. Amongst these parameters, the Q-Factor for the signal quality, the optical power received (OPR) and the differential group delay (DGD) were the ones used to build our anomaly detection models. Moreover, the high correction count seconds (HCCS) and the unavailable seconds (UAS) alarm parameters were used to identify ground truth anomalies in the OPM time series (Du et al., 2021; Tremblay et al., 2023; Pei et al., 2022). While the equipment's unavailability duration in seconds is reported as UAS, the duration in seconds where the number of errors to be corrected by forward error correction exceeds some threshold is referred to as HCCS. We structured the data as multivariate time series, using the collection date of each record. Missing data was replaced using a moving average using a 7-day sliding window. Then we normalized the data using the preprocessing module in scikit learn (Pedregosa et al., 2011). A correlation analysis using the *stats* module from the *SciPy* library revealed that the correlation between the OPR and DGD parameters with the Q-Factor is statistically significant. The Pearson coefficients calculated for the OPR and DGD parameters with regards to the Q-Factor were 0.48 and -0.18, respectively.

4.3.3 Unsupervised learning-based anomaly detection

We implemented three anomaly detection models utilizing the density-based spatial clustering of applications with noise (DBSCAN), one-class support vector machine (OCSVM) and isolation forest (IF) techniques. Three performance metrics commonly used with clustering techniques have been used to compare the three models. The silhouette score (SS) allows to validate consistency within clusters of data. It ranges from -1 to +1, with higher values indicating a well-matched instance to its own cluster. The Calinski-Harabasz index (CH), also known as the variance ratio criterion, is the ratio of the sum of the dispersion between-clusters and within-cluster for all clusters (where dispersion is defined as the sum of distances squared). Higher CH index means well-separated clusters. The Davies-Bouldin index (DB) measures the average similarity between clusters. The lower the index value the better the separation between clusters (https://scikit-learn.org/stable/modules/clustering.html#clustering).

The DBSCAN algorithm is an unsupervised learning technique that can find arbitrary shaped clusters and clusters with noise. It uses the number of samples in a neighborhood for a point to be considered as a core point (*min_samples*) and the maximum distance between two samples for one to be considered as in the neighborhood of the other (*eps*). The unsupervised learning algorithm OCSVM is a variation of the regular SVM. It can learn the boundary for normal data points where instances outside the boundary are considered to be anomalies. The isolation forest technique uses a repeated recursive process to randomly select a feature from the dataset and create partitions of the data based on a split value, until anomalies are isolated. The OCSVM and the IF models use the parameters "*nu*" and *contamination,* respectively. These parameters correspond to the proportion of outliers in the data set.

4.4 Results and comparative analysis

We used the Silhouette score to find the optimal *min_samples* and *eps* parameters for the DBSCAN model within value ranges [5-80] and [0.05-0.9] respectively. And the proportion of outliers was set to 4% for the OCSVM and IF algorithms as per the ratio obtained by dividing the number of instances with HCCS > 0 or UAS > 0 by the total number of instances in the data set. Table 4.1 shows how well the DBSCAN model outperformed the OCSVM and IF models, based on the *Precision* score as well as the unsupervised learning performance metrics as in Wan et al. (2021) and Chouhan, Atulkar, & Nagwani (2022). The low *Recall* score might be explained by the fact that the dataset is highly imbalanced in terms of anomalies compared to normal instances. Moreover, the training time for IF, OCSVM and DBSCAN on a 12.7 GB RAM Virtual Machine with Nvidia's T4 GPUs was 0.37 s, 0.02 s and 0.05 s, respectively, with OCSVM slightly faster than DBSCAN.

Table 4.1 Performance scores for the anomaly detection models

	DBSCAN	**OCSVM**	**IF**
Precision	**0.54**	0.26	0.32
Recall	0.18	0.23	**0.28**
F1-score	0.27	0.24	**0.30**
SS	**0.82**	0.69	0.68
CH	**363.65**	78.61	201.82
DB	**1.10**	2.71	1.49

Figure 4.2 a) shows the clusters created by the DBSCAN algorithm. The yellow cluster represents normal instances while the dark purple points with triangles are the anomalies detected by the model. Figure 4.2 b) displays a 3D scatter of the normal data in yellow with the labeled anomalies in dark purple.

Figure 4.3 presents a 60-day period where the system was unavailable for more than 1 hour (December 19th, 20th 2020 and January 16th 2021), with the detected anomalies displayed for the different features used. The orange vertical lines represent the HCCS within [1, 60] s (transient loss of signal (LOS)) and the red vertical lines are days when UAS > 3600 s (imminent loss of signal (ILOS)) as in Du et al. (2021). We can see that the DBSCAN model detected 14 anomaly points including the 3 ILOS events for the period, but did not detect 4 LOS events (Dec 15th, 21st, 22nd 2020 and January 27th 2021).

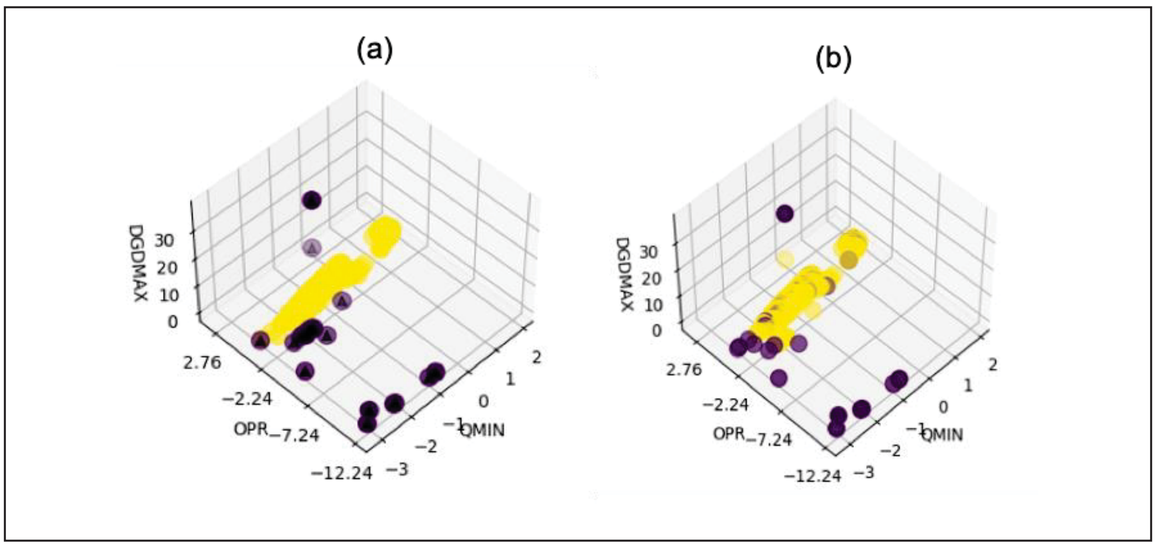


Figure 4.2 (a) Detected anomalies by DBSCAN; (b) Labeled anomalies

These undetected cases may be related to alarms raised for events not linked to the PM metrics considered in this study. On the other hand, 9 days during which anomalies were detected by the DBSCAN model don't have any alarms recorded. On December 12th, 14th 2020 for example, there were two occurrences of Q-Factor degradation, with very low (-45 dBm) OPR values (indicated by A and B), while the DGD remained stable for both days. This shows that the model can find combinations of PMs abnormal, while PMs might not all be anomalies when taken one by one.

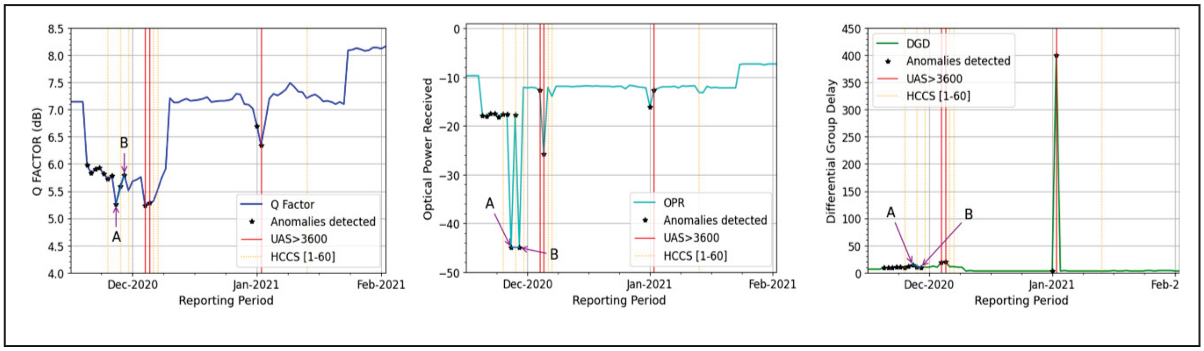


Figure 4.3 Anomalies detected on the 3 PM metrics (Q-Factor, OPR and DGD) by DBSCAN

4.5 Conclusions

We propose an unsupervised ML approach for anomaly detection. Our study demonstrates the benefits of unsupervised learning's ability to process unlabeled data and create clusters, thus allowing anomaly detection in multivariate time series without time consuming hand-labeling task on the data. In future work, we will develop a strategy allowing to capture gradual degradations in lightpath QoT considering multiple parameters simultaneously.

4.6 Acknowledgements

This work was supported by the Natural Sciences and Engineering Research Council of Canada (RGPIN 2019-03972), the ÉTS Program for the Development of International Research Collaborations and the EUREKA cluster CELTIC-NEXT project AI-NET PROTECT funded by the Swedish Innovation Agency.

CHAPTER 5

CHANGE POINT AND ANOMALY-AWARE QOT FORECASTING FRAMEWORK FOR PROACTIVE NETWORK MANAGEMENT

Sandra Aladin [a], Lena Wosinska [b], Christine Tremblay [a]

[a] Department of Electrical Engineering, École de technologie supérieure, 1100 Notre-Dame Ouest, Montréal, Québec, Canada H3C1K3

[b] Department of Electrical Engineering, Chalmers University of Technology, 412 96 Gothenburg, Sweden

Article published in *IEEE Networking Letters*, vol. 8, pp. 224-228, May 2026.
https://doi.org/10.1109/LNET.2026.3691249

5.1 Abstract

Reliable forecasting of optical performance monitoring (OPM) parameters under non-stationary conditions is a key enabler of proactive optical network management. In this letter, we propose a forecasting enhancement framework that integrates change point and collective anomaly detection scores as auxiliary features into a sequence-to-sequence QoT forecaster. Results on production OPM data show consistent improvements in forecasting accuracy over baseline models that do not integrate the auxiliary features. As a result, the proposed framework provides a consistently positive reaction window prior to QoT violations, with observed durations ranging from 1 to 10 days across failures, enabling proactive network operation without introducing additional control loop complexity.

5.2 Introduction

Optical networks are fundamental to modern communication systems in the era of data-intensive applications and real-time connectivity. As the demand for uninterrupted service continues to grow, enhanced reliability performance becomes critical, since even brief disruptions can make a very high impact by affecting many users as well as leading to significant data loss and system failures (Wang et al., 2022b). Conventional threshold-based monitoring typically identifies issues only after substantial performance degradation has occurred, preventing early intervention and exposing networks to unplanned downtime and expensive repairs. Proactive network management provides the ability to overcome this limitation by enabling reliable predictions of quality of transmission (QoT) degradations prior to hard failures or violations of service-level agreement (SLA).

Machine learning (ML) has been proposed as a promising solution for improving management and recovery processes in optical networks (Musumeci & Tornatore, 2025). Time-series forecasting models based on recurrent neural networks (RNNs) architectures, such as long short-term memory (LSTM) and gated recurrent unit (GRU), have shown promising results for short-term QoT prediction, using signal-to-noise ratio (SNR), Q-factor, and related metrics over short to multi-step time horizons. (Allogba et al., 2021). However, their performance degrades when the underlying data distribution changes. In practice, major system shifts are often preceded by small gradual QoT degradations caused by component aging, polarization effects, and fiber impairments.

Recent work has proposed an embedding architecture that integrates link-level information to improve ML-based QoT estimation of unestablished lightpaths (Lechowicz et al., 2025). However, these methods focus on instantaneous QoT estimation under known configurations and do not address non-stationary or multi-step QoT forecasting, both critical for proactive network operation.

To address non-stationarity in optical network monitoring, change point detection (CPD) and anomaly detection (AD) have received considerable attention. CPD discovers regime shifts, while anomaly detection highlights deviations from normal behavior. KernelCPD, from the *ruptures* library, accesses the complete dataset for offline CPD (Truong et al., 2018, 2020) while the kernel cumulative sum (Kernel-CUSUM), a maximum mean discrepancy (MMD)-based sequential test, may detect distributional shifts without requiring any assumptions about the data distribution or the need to explicitly model how failures occur (Flynn & Yoo, 2019).

Kernel-CUSUM can capture both rapid and gradual changes in network patterns within high-dimensional data while being sensitive to subtle distributional shifts (Wei & Xie, 2022). Few studies specifically address CPD in optical networks (Mosayebi & Lampe, 2024).

Soft failures are characterized by degraded system performance or non-critical errors, (e.g., due to components aging) that often allow continued operation but may lead to service disruption if not detected early (Kruse, Kuhl, Dochhan & Pachnicke, 2024). In contrast, hard failures are abrupt and unexpected events (e.g., fiber cuts). Unsupervised and semi-supervised learning approaches, such as autoencoders, generative adversarial networks (GANs), vision transformers, and clustering methods, have been applied to soft failure detection (Musumeci & Tornatore, 2025; Aladin et al., 2024). More recent studies explore uncertainty-aware forecasting and attention-based architectures (Di Cicco et al., 2023; Hedayatnejad et al., 2025; Yousefi et al, 2023), while event-oriented approaches formulate preventive maintenance as a discrete failure prediction problem (Du et al., 2021). These methods rely on collective anomaly detection (CAD), which analyzes temporal observation sequences to identify abnormal patterns rather than individual instances.

Proactive and predictive network management has also attracted increasing interest within the networking community. Learning-based and predictive models have been proposed to support proactive decision making by providing information-rich outputs accessible to network controllers serving as a key enabler for the automation of optical network operation and management tasks (Musumeci et al., 2019a). In optical networks, ML approaches have largely treated QoT estimation, anomaly detection and QoT forecasting as separate tasks. However, most focus on detection or prediction accuracy, with limited attention to how improved forecasting translates into actionable reaction windows for proactive network operation.

In this letter, we propose (i) a CPD-CAD-aware QoT forecasting framework that, to the best of our knowledge, is the first to explicitly integrate change point and anomaly indicators as auxiliary inputs to a multivariate sequence-to-sequence (seq2seq) LSTM model for optical QoT forecasting; (ii) a performance assessment showing the forecasting improvements obtained on production OPM data with aggregate MAE reductions of up to 16.7% and R^2 improvement from negative to positive values; and (iii) strictly positive reaction windows of

1-10 days preceding all observed QoT violations, providing actionable lead time for proactive network operation. This reaction window establishes a concrete link between physical-layer forecasting and predictive network operation, enabling proactive management decisions to be considered without introducing additional control-loop complexity.

The next sections introduce and evaluate our proposed QoT forecasting enhancement framework for proactive network management. Finally, we present our conclusions.

5.3 CPD-CAD-aware QoT forecasting framework

Figure 5.1 illustrates the proposed CPD-CAD-aware QoT forecasting framework. Monitored QoT metrics are first processed to identify regime shifts and degradations. The resulting CPD-CAD-aware feature representation is subsequently input to a multivariate seq2seq forecasting model, which forecasts future QoT signals under non-stationary conditions. The predicted QoT signals, along with reaction windows computed from detected change points and anomalies, feed into the decision support module, enabling early control actions via the network controller, and subsequently supporting post-hoc validation and continuous model refinement through the performance feedback loop.

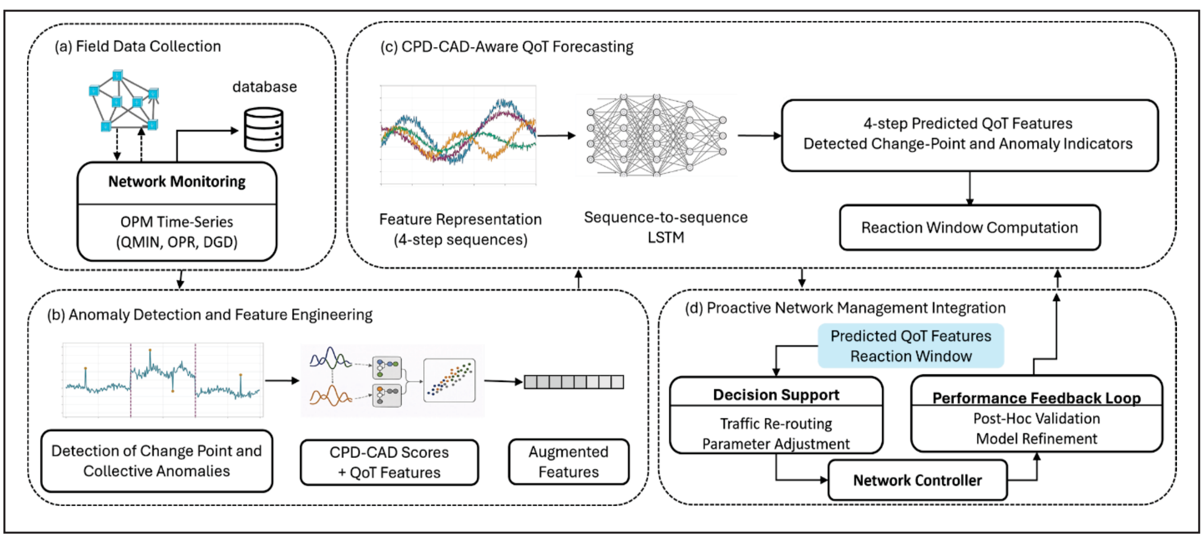


Figure 5.1 Conceptual workflow of the proposed CPD-CAD-aware QoT forecasting framework including: (a) OPM data collection in a production network; (b) the change point and collective anomaly detection module; (c) the CPD-CAD-aware QoT forecasting module; (d) the proactive network management integration via a network controller

5.3.1 Dataset description and problem formulation

We consider multivariate OPM time series collected from an optical channel of a production optical transport network over nearly 4 years. Each daily time step contains three QoT metrics: minimum Q-factor (QMIN), optical power received (OPR), and differential group delay (DGD). Alarm parameters such as unavailable seconds (UAS) and high correction count seconds (HCCS) are available but only used for weak supervision and evaluation.

Given an input sequence $X_{t-L+1:t} \in \mathbb{R}^{L\times 3}$, where L denotes the observation window, the proposed model aims to forecast future QoT values over a given horizon while remaining robust to variable operations situations such as regime shifts and anomalous patterns.

Transient loss of signal (LOS) and imminent loss of signal (ILOS) are annotated following the strategy in Du et al. (2021). A QoT violation is defined as a sustained degradation event in the monitored QoT time series, corresponding to service outages characterized by UAS exceeding 1800 s. The dataset is highly imbalanced, with only four ILOS events over approximately 1,502 samples, reflecting typical production networks and motivating the use of an unsupervised CAD approach.

Figure 5.2 shows the non-stationary time series of the QMIN parameter (Augmented Dickey-Fuller (ADF) Statistic: -1.653 with p-value: 0.455, and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) Statistic: 0.890 with p-value: 0.01), with a degradation period around December 2020 and January 2021 which we consider to be predictable.

Although defined at the physical layer, the degradation of these QoT parameters has system-level impacts as it may escalate into service disruptions. Modeling their temporal evolution is therefore directly tied to proactive reliability management rather than mere physical-layer monitoring. Because degradations patterns can be caused by environmental changes or component aging, integrating CPD and CAD enables the detection of changes in feature dynamics under non-stationary conditions. In this work, a regime shift denotes a sustained change in the statistical behavior (mean, variance, or trend) of the QoT signal, while a collective anomaly captures short-term correlated deviations outside the normal operating ranges. Together, these elements provide sensitivity to both gradual and abrupt QoT degradation, enabling more robust QoT forecasting

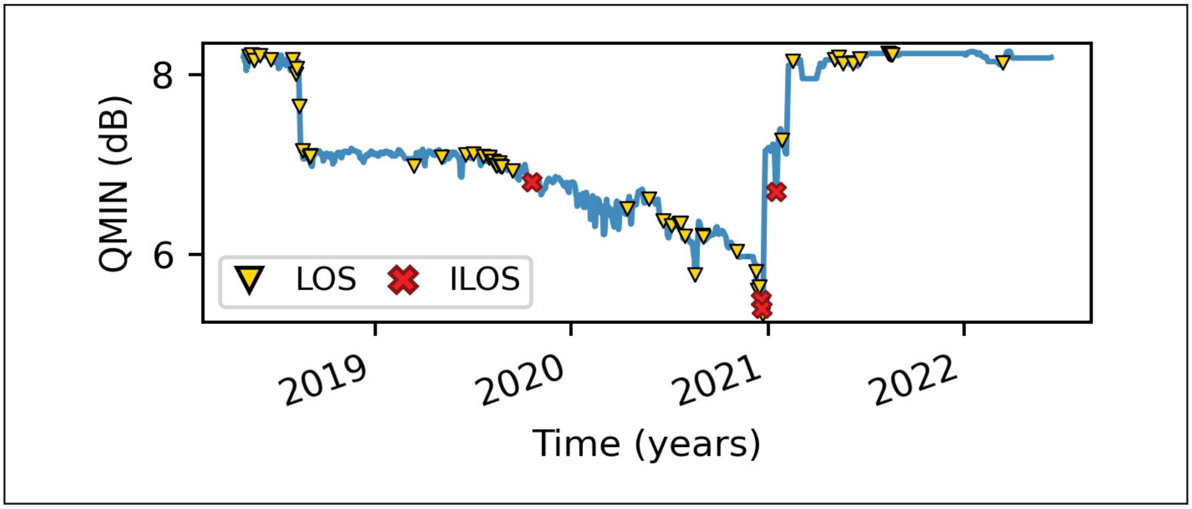


Figure 5.2 Time series of OPM parameter QMIN and alarms over a 4-year period for one channel

5.3.2 Change point and collective anomaly awareness

To address fluctuating regime shifts and subtle degradations in dynamic QoT time series, we leverage both CPD and CAD techniques. An offline KernelCPD, with RBF kernel and the default number of segments $n\_bkps = 10$, is first applied to the training data to detect regime shifts and calculate CPD scores that reflect regime changes. In addition, an unsupervised hierarchical density-based spatial clustering of applications with noise (HDBSCAN) CAD model is trained on historical data using grid search with $min\_cluster\_size \in \{10, 15, 20\}$ and $min\_samples \in \{10, 15\}$, to detect anomalous multivariate sequences and produce CAD scores. During evaluation, CPD and CAD scores are computed sequentially on test data using an online Kernel-CUSUM with the detection threshold set to $\mu + 2\sigma$ of MMD statistics computed on the training set, and the same CAD model retrained incrementally on observed historical data, without access to future sequences.

Rather than using CPD and CAD as standalone modules, the proposed framework uses their outputs as auxiliary features: a continuous CPD score = $(1 + d)^{-1}$, where d is the number of sequence steps to the nearest detected change point to provide regime shift insight as a normalized signal in range [0, 1]; and a binary CAD score (0 = normal, 1 = collective anomaly). Both are added to the QoT feature vector to provide the forecaster with additional insight into regime shifts and anomalous patterns. The specific CPD and CAD techniques are selected for their ability to work in an unsupervised manner and to provide indicators suitable for integration into forecasting models.

5.3.3 CPD-CAD-aware QoT forecasting

The forecasting model includes an encoder with two stacked LSTM layers (64 and 32 units, dropout 0.2) and a decoder using a RepeatVector followed by a 32-unit LSTM layer and a TimeDistributed Dense output layer. Inputs consist of 4-day sequences with three QoT features expanded to five when CPD and CAD indicators are included. The CPD and CAD scores augment the original QoT feature vectors as follows. Let $\boldsymbol{X_t} = (x_{t-L+1}, \ldots, x_t) \in \mathbb{R}^{L\times 3}$ denote the input sequence of $L = 4$ daily OPM observations. The CPD score $s_i^{cpd} = \frac{1}{1+d_i}$,where $d_i = min_{cp}|i - cp|$ is the distance from sequence i to the nearest change point cp, and the binary HDBSCAN anomaly score $s_i^{cad} \in \{0, 1\}$ are appended at each timestep to form $\boldsymbol{X_t'} \in \mathbb{R}^{L\times 5}$, processed by the seq2seq LSTM to produce $\boldsymbol{\widehat{Y}}_{t+1:t+L} \in \mathbb{R}^{L\times 3}$. All inputs are normalized using MinMaxScaler technique and the output shares the same sequence length processed by the seq2seq architecture.

LSTMs are selected for their ability to model complex temporal dependencies and their established effectiveness as RNN tool to manage multivariate time series (Hall & Rasheed, 2025). From a network operations perspective, this enhanced input representation enables the forecasting model to condition its predictions on the current regime, yielding more informative forecasts for proactive control decisions.

5.3.4 Reaction window computation

Beyond forecasting accuracy, the approach reliably computes a reaction window before a QoT violation. The reaction window is defined as the interval between a detected QoT degradation and the observed subsequent failure, representing the time available for preventive network actions.

In this work, the reaction window is computed as the difference between the failure timestamp and the last CP or CA detection before the event and collected directly from the framework output and assessed along with the enhanced forecasting model's prediction robustness, rather than as an independent detection metric.

5.3.5 Proactive network management integration

Although the proposed framework does not include a control policy or a closed-loop mechanism, the predicted QoT features and computed reaction windows feed into the decision support module as actionable signals for the network controllers, enabling proactive actions such as traffic rerouting, protection switching, or system parameter adjustments (e.g., modulation format, amplifier gain). Moreover, the framework supports the integration of a performance feedback loop in which predicted OPM parameters and their corresponding reaction windows can be leveraged to validate the outputs and retrain the model as follows: Predicted degradations can be compared against subsequent QoT observations, alarms, and logged maintenance events to evaluate the consistency between predicted and observed network behavior. Any discrepancies such as undetected degradations or systematic overestimation of OPM parameters' fluctuations can serve as indicators for model retraining or threshold recalibration using newly collected data. The reaction window also provides a temporal context that can be used for ranking predictions by operational relevance, enabling selective inclusion of events in future training datasets and facilitating continuous model improvement as network conditions evolve.

5.3.6 Scope and assumptions

The proposed framework enhances QoT forecasting under dynamic conditions and estimates the associated reaction window, while incorporating a feedback loop for post-hoc validation and model refinement. The design and evaluation of specific network control actions are beyond the scope of this study. The framework operates without failure labels during training. Feedback from subsequent QoT metrics, alarms, and maintenance logs can be used offline for post-hoc validation, threshold tuning or model retraining. Notably, such feedback is not used in the present evaluation, as all reported results are derived exclusively from field OPM data.

5.4 Experimental validation

Numerical evaluations are conducted on field OPM data, comparing the CPD-CAD-aware LSTM against a baseline LSTM trained solely on QoT features. The framework is evaluated based on forecasting accuracy and its ability to identify reaction windows before QoT violations occur.

5.4.1 Forecasting performance evaluation

We evaluate the CPD-CAD-aware forecasting model on multivariate OPM data from a production optical network, using a chronological train/validation/test split (61/2/37) to reflect realistic conditions in production environment and avoid leakage. Both baseline and enhanced models use the same seq2seq LSTM architecture and hyperparameters, ensuring that any performance improvements come only from integrating the CPD and CAD indicators.

Table 5.1 lists the hyperparameters used via grid search for forecasting models optimization in both scenarios. Under identical conditions, the best results for both the baseline LSTM and the E-LSTM models are achieved with the following hyperparameters values: 64 hidden units with a dropout rate of 0.2 to mitigate overfitting, Adam optimizer, a learning rate of 0.001 and a batch size of 32.

Table 5.1 Hyperparameter Values

Hyperparameter	Tested Values
Hidden units	32, 64
Learning rate	0.0001, 0.0005, 0.001
Batch size	32
Dropout	0.1, 0.2, 0.3
Optimizer	Adam, RMSprop

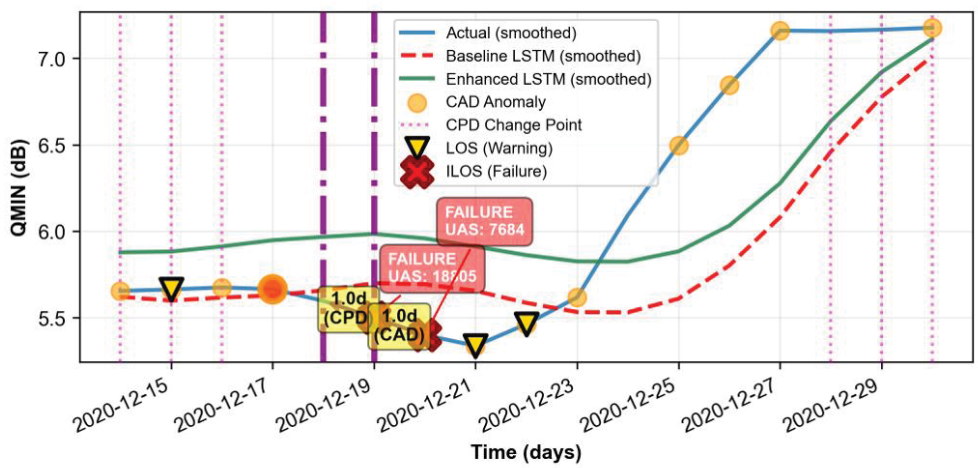
Table 5.2 summarizes the forecasting performance for QMIN, OPR, and DGD over a 4-day horizon. The CPD-CAD-aware enhanced LSTM (E-LSTM) model consistently outperforms the baseline model, achieving RMSE and MAE reductions of up to 28.21% and 42.36%, respectively. For QMIN, the coefficient of determination improves from 0.477 to 0.731, indicating better tracking of QoT variations under dynamic conditions. For OPR and DGD, the enhanced model achieves lower RMSE and better variance capture, with more evenly distributed prediction errors. These improvements reflect the model's ability to remain robust when QoT time series deviates from the stationary behavior seen during training. With the multivariate seq2seq architecture optimizing a single joint loss across all three QoT features, E-LSTM achieves aggregated MAE and RMSE reductions from 1.330 to 1.108 and from 9.937 to 9.888 respectively, while also improving R^2 from -0.031 to 0.139. This confirms that CPD-CAD integration improves the multivariate LSTM forecaster despite the OPR MAE increase.

Table 5.2 Comparative performance evaluation of the forecaster per Feature

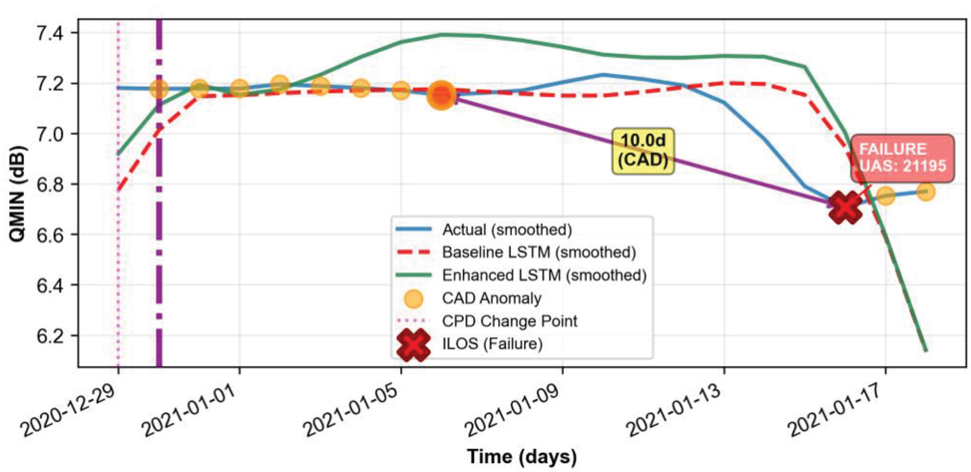
	FEATURES	RMSE	MAE	R^2
LSTM	QMIN (dB)	0.319	0.203	0.477
	OPR (dBm)	2.336	**1.441**	-0.573
	DGD (ps)	17.048	2.345	0.001
E-LSTM	QMIN (dB)	**0.229**	**0.117**	**0.731**
	OPR (dBm)	**2.142**	1.729	**-0.323**
	DGD (ps)	**16.991**	**1.477**	**0.008**

5.4.2 Reaction windows analysis

Beyond improving the forecasting accuracy, the proposed framework identifies reaction windows preceding QoT violations. These windows are obtained by identifying the interval between the first detected CA or CP and the subsequent QoT violation marked as ILOS.



(a)



(b)

Figure 5.3 E-LSTM forecasting results (QMIN) with reaction windows for: (a) an ILOS on December 19, 2020, preceded by a CP (purple vertical line) detected on December 18, and an ILOS on December 20, preceded by a CA (dark orange circle) and a CP detected on December 19; (b) an ILOS on January 16, 2021, preceded by a CA detected on January 6

Figure 5.3 (a) and (b) illustrate representative QMIN degradations and the corresponding reaction windows computed by the CPD-CAD-aware LSTM model. In Figure 5.3 (a), two ILOS events on December 19 and December 20, 2020 are preceded by CP and CA detections one day before, yielding reaction windows of one day. In contrast, Figure 5.3 (b) shows an ILOS event on January 16, 2021 preceded by a CA on January 6, providing a reaction window of ten days and reflecting a more gradual degradation behavior. Red boxes show ILOS durations in seconds, while yellow boxes indicate the detected event type (CP or CA) with the corresponding reaction window before each ILOS.

Figure 5.4 shows the distribution of reaction-window durations across all three failure events. All windows are strictly positive, ranging from one to ten days with a median of one day and a mean of four days, indicating consistent actionable early-warning capability under non-stationary conditions. Despite variability across events, the interquartile range confirms that detected degradations always precede ILOS events. The reaction window is evaluated as an element of the proposed forecasting enhancement approach. The information is provided as a decision-support signal that may be leveraged for proactive network operation. In addition, the structured outputs comprising the reaction window and the predicted OPM parameters can be fed into a performance feedback loop as aforementioned.

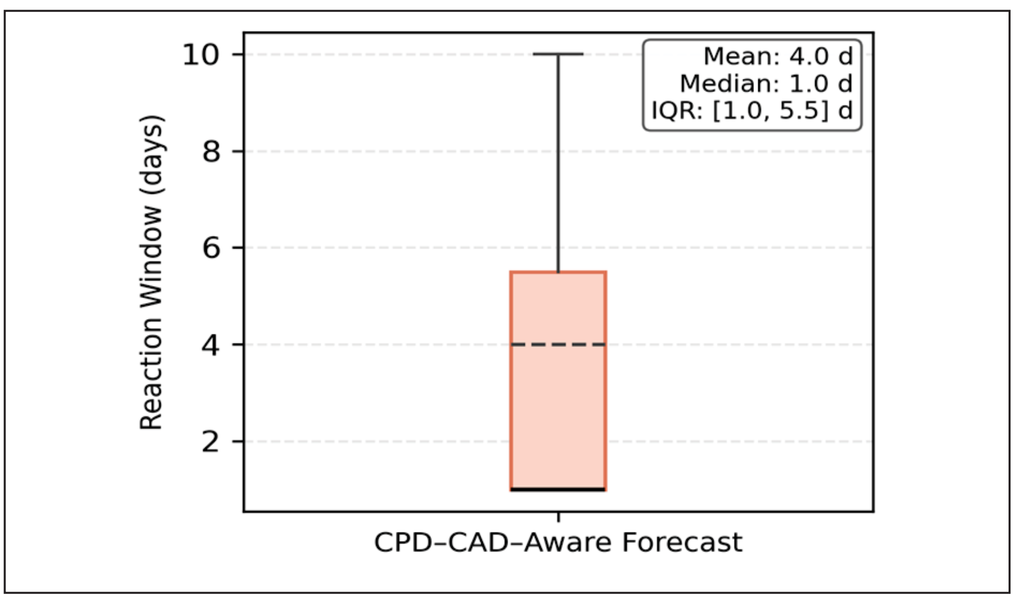


Figure 5.4 Reaction window distribution across failure events

These results show the benefits of interpreting forecasting performance alongside the degradation regime present in the OPM time series. Unlike conventional ML models that treat detection and forecasting as separate tasks, CPD-CAD-awareness not only improves predictive consistency, as reflected in Table 5.2, but also yields reliable reaction windows across heterogeneous degradation patterns. This behavior is particularly important in evolving optical networks, where aging, reconfiguration, and environmental factors introduce non-stationary dynamics. The proposed framework thus shifts the focus from minimizing prediction error to detecting degradations early enough to prevent service disruption.

5.5 Conclusion

This letter introduces a CPD-CAD-aware QoT forecasting framework that improves prediction robustness under non-stationary network conditions. By integrating CPD and CAD indicators, the multivariate seq2seq model achieves consistent forecasting improvements including aggregate MAE reductions of up to 16.7% and improvements in R^2 from negative to positive values, outperforming the baseline models that rely only on raw monitoring features.

Beyond improving accuracy, the proposed CPD-CAD-aware forecasting framework yields strictly positive reaction windows of 1-10 days prior to all three failure events in the data. Representative examples and distribution analysis confirm that early detected degradations reliably precede service-impacting events, enabling actionable intervals for proactive operation.

Although explicit network control strategies are beyond the scope of this letter, a performance feedback loop is incorporated into the CPD-CAD-aware forecasting framework, enabling retrospective comparison of predicted OPM parameters and monitored values for post-hoc validation of QoT predictions and reaction windows. This process supports threshold tuning and continuous model refinement under evolving network conditions. Future work will extend validation to additional datasets and deployment scenarios.

CHAPTER 6

ENHANCING PROACTIVE QOT MANAGEMENT WITH CHANGE-POINT- AND COLLECTIVE-ANOMALY-ENABLED DEEP LEARNING FORECASTING

6.1 Introduction

6.1.1 Context and motivation

Optical transport networks (OTNs) form the backbone of today's communication systems, and growing demand for continuous, diverse services is pushing them to operate increasingly close to full capacity. Under these conditions, even brief service disruptions can result in major data losses and significant financial consequences (Wang et al., 2022b). Conventional threshold-based monitoring approaches typically identify issues only after noticeable performance degradation has already occurred, supporting a reactive maintenance model that leads to costly repairs and extended downtime. Preventive methods that rely on fixed thresholds or routine scheduled interventions are often expensive and insufficiently responsive to the dynamic nature of real-world optical networks. This highlights the need for predictive maintenance techniques capable of detecting early signs of QoT degradation before they evolve into service disruption.

ML techniques have been widely adopted to extract insights from the large volumes of OPM data produced by modern network infrastructures, enabling accurate failure prediction and early detection of anomalies (Musumeci & Tornatore, 2025). Time-series forecasting models based on RNNs, notably LSTM and GRU architectures, have shown strong performance in short-term prediction tasks (Tremblay et al., 2019; Aladin et al., 2020a; Allogba et al., 2022a; Allogba et al., 2021; Aladin et al., 2020b). However, their accuracy typically deteriorates when the underlying data distribution evolves over time, a pattern driven by gradual QoT variations induced by component aging, slow environmental drifts, polarization dynamics, and minor fiber-level impairments. These subtle degradations often precede more pronounced system failures, making their timely identification essential for proactive network management.

CPD algorithms can discover both abrupt and gradual regime shifts in time-series data and are therefore well suited to model such subtle changes in QoT signals. Offline techniques, such as KernelCPD from the *ruptures* library in Python, rely on full dataset access and are therefore unsuitable for real-time operation (Truong et al., 2018, 2020). Conversely, online CPD techniques process data sequentially and detect changes without referencing future observations. Among these, Kernel-CUSUM based on MMD is particularly attractive, as it detects shifts without assumptions about the data distribution or explicit modeling of failures mechanisms (Flynn & Yoo, 2019). Despite these benefits, CPD remains underutilized in real-world optical networks and to the best of our knowledge, no prior work has explicitly integrated it with QoT forecasting models.

Anomaly detection techniques have also been extensively studied for highlighting patterns that deviate from expected normal behaviors in OPM data. Early approaches rely primarily on point-wise algorithms that examined individual observations separately (Aladin et al., 2024; Abdelli et al., 2024a; Abdelli et al., 2024b; Sun et al, 2024; Musumeci et al., 2022). Although these algorithms work well for detecting individual outliers, they are poorly suited to the behavior of lightpaths in optical networks, where impairments frequently appear as gradual subtle degradations. CAD addresses this limitation by analyzing sequences of observations to uncover unusual patterns that may not be detected when examining individual data points. CAD techniques such as DBSCAN, HDBSCAN, Isolation Forest, OCSVM, and Autoencoder-based reconstruction can capture complex multi-variable interactions like those reflected in abnormal QoT fluctuations. However, these techniques are still typically applied as standalone anomaly detection tools rather than being integrated within a forecasting framework.

Although prior studies have shown that ML methods are effective for anomaly detection, fault identification, localization, and QoT prediction using recurrent and seq2seq architectures, most existing approaches address CPD, CAD, and forecasting as isolated tasks. As a result, state-of-the-art forecasting models often experience reduced accuracy when confronted with physical-layer perturbations too different from the range of values used to train the models. A limited number of works have begun to investigate integrated strategies. For example, Bayesian recurrent neural networks have been employed to generate predictive intervals for QoT forecasts, where interval violations are interpreted as anomalies, thereby combining forecasting for fault localization with anomaly detection (Di Cicco et al., 2023; Behera et al., 2023b; Abdelli et al., 2022a). Another study introduced a statistical hypothesis-testing framework to differentiate predictable from unpredictable soft failures by detecting real-time deviations in BER measurements, combining statistical indicators with online predictions from an Encoder-Decoder LSTM model (Lun et al., 2023). Abdelli et al. (2022b) proposed a multi-stage predictive maintenance pipeline that incorporates an attention-based GRU model for real-time performance degradation prediction of semiconductor lasers, a convolutional autoencoder for anomaly detection using the predicted values, and an attention-based model to forecast remaining useful life (RUL) triggered upon detection of abnormal behavior. However, none of these approaches directly integrate change point information or collective anomaly indicators within the forecasting architecture itself.

This chapter addresses this limitation by presenting a unified framework that integrates CPD and CAD directly into the multivariate QoT forecasting models first introduced in Chapter 5 and substantially expanded here. By enriching seq2seq LSTM and TCN forecasters with Kernel-CUSUM CPD scores and CAD scores as additional input features, the framework supports adaptive prediction under evolving operating conditions and delivers early-warning capabilities through reaction windows preceding QoT violations. While Chapter 5 demonstrated the feasibility of this approach on a single lightpath dataset from a production network using an LSTM model with a fixed input sequence length, the present chapter significantly broadens that initial study across three field lightpath datasets, forecasting architectures, five CAD methods, and additional sequence lengths, thereby providing a rigorous and generalizable assessment of the framework's performance.

6.1.2 Research motivation and scope of the framework extension

A preliminary version of this framework was introduced in Chapter 5, where the feasibility of CPD- and CAD-aware QoT forecasting was demonstrated using a single field lightpath dataset, a four-day input window, HDBSCAN as the CAD method, and an LSTM-based forecasting model. That initial study established the core integration concept and yielded promising results, including a 28.21% reduction in RMSE and a 42.36% reduction in MAE relative to a baseline LSTM, as well as reaction windows ranging from 1 to 10 days prior to ILOS event. While these findings validated the feasibility of the approach, they were obtained under an experimental design that leaves four key challenges open.

The first concerns generalizability across network environments. The experiments in Chapter 5 relied exclusively on a single field lightpath dataset from a North American service provider's optical transport network. Such results may reflect characteristics specific to that optical channel, degradation types, and therefore cannot be assumed to generalize without broader experimental evaluation.

The second relates to the choice of forecasting architecture. Chapter 5 examined only an LSTM as the forecasting technique. TCNs, which employ dilated causal convolutions rather than recurrent connections, represent a fundamentally different paradigm for sequence modeling and have demonstrated competitive or superior performance in various time-series applications. Whether TCNs offer advantages over LSTMs within the CPD-CAD integration framework, particularly with respect to forecast accuracy and reaction windows, remains an open question.

The third involves the selection of CAD techniques. Chapter 5 employed HDBSCAN as a fixed CAD method without systematic comparison to alternatives. Because different CAD algorithms capture distinct forms of collective anomalous behavior, a model selection process evaluating multiple techniques may yield different results than reliance on a single CAD method.

The fourth is about sensitivity to the sequence length. Chapter 5 used a fixed four-day forecast window. Sequence length directly influences both forecasting accuracy and lead time to failure: shorter windows may fail to capture subtle regime shifts that CPD and CAD are designed to detect, whereas longer windows may introduce noise from historical segments. Whether 3-day or 7-day sequences yield improved performance, and whether such effects are consistent across datasets and architectures, has practical implications for deployment.

6.1.3 Chapter objectives and contributions

This chapter extends the preliminary framework introduced in Chapter 5 along the four key limitations identified above. Its contributions are as follows:

Contribution 1: Robustness evaluation using multiple datasets.

The CPD-CAD-aware forecasting framework is assessed across three additional field lightpath datasets, drawn from different routes and channels in the network. This evaluation, using three lightpath datasets from a production network, provides a basis for determining whether the performance improvements observed in Chapter 5 generalize across diverse contexts.

Contribution 2: Comparative evaluation of LSTM and TCN architectures.

Both LSTM and TCN are implemented as forecasting models within the same CPD-CAD integration pipeline and evaluated under identical datasets, metrics, and experimental conditions. This comparison establishes whether architectural choice influences forecasting accuracy or reaction window duration, thereby guiding architecture selection for future deployments.

Contribution 3: Systematic CAD model selection.

Rather than relying on a single CAD method, this chapter introduces a structured model selection and hyperparameter-tuning procedure that evaluates five CAD techniques, namely DBSCAN, HDBSCAN, Isolation Forest, OCSVM, and autoencoder-based reconstruction, within the unified CPD-CAD-aware framework. This systematic evaluation identifies which CAD methods perform most effectively across different datasets and network conditions.

Contribution 4: Sensitivity analysis of sequence length.

Input windows of three, and seven days are examined across both forecasting architectures and three additional datasets. The 3-day window evaluates the impact of reduced temporal context on early-detection capability, while the 7-day window aligns with the 7-day QoT-degradation forecast horizon reported in prior literature (Du et al., 2021). This analysis yields practical guidance on selecting input-sequence lengths for real-world CPD-CAD-aware forecasting deployments.

To the best of our knowledge, this study provides the first comprehensive evaluation of a unified CPD-CAD-aware forecasting framework for proactive optical network management that simultaneously examines robustness across datasets, CAD methods, deep-learning architectures, and sequence lengths.

The remainder of this chapter is organized as follows. Section 6.2 reviews related work on CPD, CAD, and QoT forecasting, with emphasis on integrated approaches and the specific paradigms expanded in this chapter. Section 6.3 formally defines the forecasting problem and experimental variables. Section 6.4 provides the description of the datasets used for the systematic validation of the CPD-CAD-aware forecasting framework, details the experimental environment, the labeling strategy and the input sequence configuration. Section 6.5 presents the change point detection approach adopted for regime shift detection. Section 6.6 presents the comparative study of the collective anomaly detection models. Section 6.7 details the integration of CPD and CAD insights into the forecasting models, while Section 6.8 presents the forecasting architectures, including the TCN architecture introduced in the framework, with information about the baseline models. Section 6.9 presents and analyzes results obtained for the different scenarios. Section 6.10 describes the pseudo-lead-time analysis with different evaluation considerations for comparison with the baseline models. Section 6.11 concludes with an integrated discussion.

6.2 Related work

Proactive failure management in optical networks increasingly depends on data-driven analysis of OPM parameters. Soft failures or gradual QoT degradations caused by component aging or environmental factors represent a significant and well-documented operational challenge (Lun et al., 2023; Sun et al., 2023; Silva et al., 2023; Behera et al., 2023a; Kruse et al., 2024).Three methodological areas are especially relevant for detecting these degradation patterns: change point detection, collective anomaly detection, and QoT forecasting. Although ML has been widely applied to anomaly detection and QoT prediction, explicit CPD remains comparatively underexplored, and the three paradigms are still mostly studied separately despite their complementary roles in early-warning and proactive maintenance workflows.

Given the structure of the proposed CPD-CAD-aware forecasting framework, this section reviews prior work on CPD, CAD, and QoT forecasting, with particular attention to the four aspects expanded in this chapter relative to the preliminary study in Chapter 5: the choice of deep-learning architecture, the evaluation of CAD techniques within forecasting pipelines, sensitivity to input-sequence length, and generalizability across three lightpath datasets from a production network.

6.2.1 Change point detection in optical networks

CPD techniques aim to identify moments when the statistical properties of a time series, such as its mean, variance, trend, or correlation structure, undergo significant shifts, whether abrupt or gradual. These techniques are well suited to optical networks, where degradations often appear as subtle, gradual changes rather than clear threshold violations, reflecting the slow dynamics of component aging and fiber impairments. Figure 6.1 shows an illustration of the detected change points indicating period where statistical properties of the signal change.

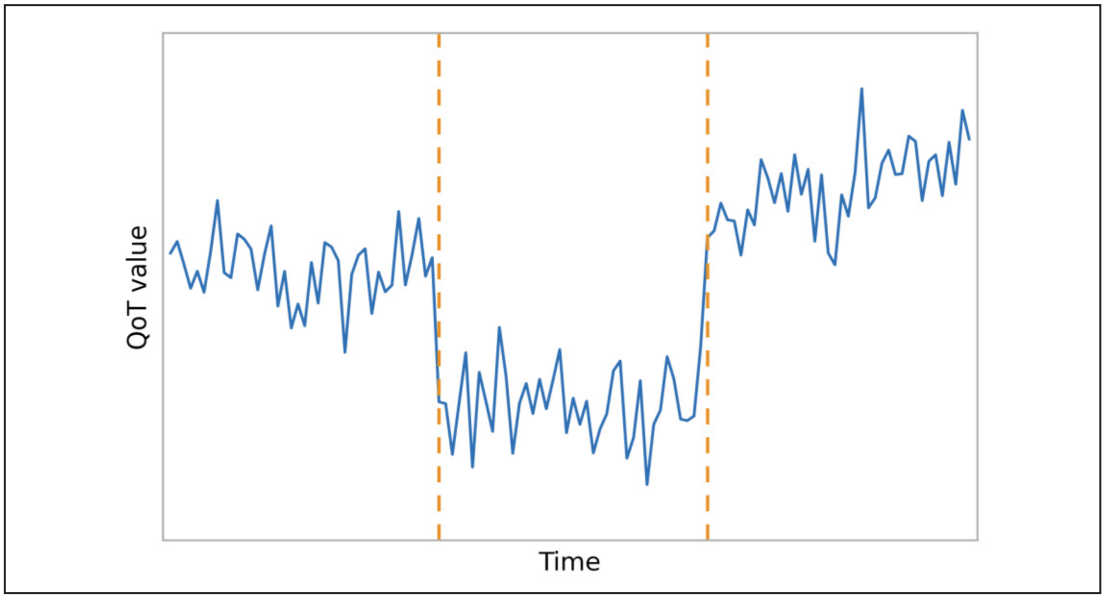


Figure 6.1 Illustration of change points (orange vertical lines) detected in a signal

Kernel-based approaches, particularly Kernel-CUSUM using MMD, have proven effective at detecting both rapid and gradual distributional shifts. Their non-parametric nature makes them especially appropriate for multivariate OPM datasets, whose distributions are shaped by complex physical-layer interactions that are difficult to model analytically. While offline CPD methods, such as those in the *ruptures* library, require full dataset access and are unsuitable for real-time use, online variants operate sequentially and are therefore viable for operational monitoring.

Only a limited number of studies have applied CPD explicitly to optical networks. One work proposes combining fast CPD with classification to detect regime transitions using sliding-window SNR features, highlighting the value of viewing performance monitoring as a shift-detection problem rather than a point-anomaly detection task (Mosayebi & Lampe, 2024). Xu et al. (2025) provide a broad survey of DL-based CPD techniques, highlighting their importance for timely decision-making in dynamic systems. However, CPD-aware forecasting, where CPD outputs serve as contextual inputs to predictive models rather than standalone diagnostics, remains largely unexplored. This chapter directly addresses that gap as an extension of the work proposed in Chapter 5.

6.2.2 Collective anomaly detection in optical networks

Early work in optical network fault management showed that statistical and ML models can detect, identify, and localize faults using OPM data (Vela et al., 2017; Furdek, Natalino, Giglio & Schiano, 2021; Mayer et al., 2021). These studies demonstrated that deviations in parameters such as BER, OSNR, optical power received, and transmitter power contain diagnostic information about lightpath health, and that automated methods can uncover patterns missed by threshold-based monitoring.

More recent research has shifted toward unsupervised and semi-supervised techniques, such as autoencoders, GANs, vision transformers, and clustering-based approaches, reflecting the scarcity of labeled failure data in production networks (Abdelli et al., 2024a; Abdelli et al., 2024b; Sun et al., 2024). These models better capture the multivariate interactions typical of real optical impairments, where degradations affect several correlated OPM parameters simultaneously. However, most of these approaches focus on instantaneous or short-window observations and do not explicitly model the temporal evolution of anomalies. As a result, they struggle to detect the gradual, sequence-level deviations that characterize early-stage optical degradation.

CAD addresses this limitation by analyzing sequences of observations to identify patterns that are anomalous only when considered together. Techniques such as DBSCAN, HDBSCAN, Isolation Forest, OCSVM, and autoencoder-based reconstruction each capture different forms of collective anomalous behavior, and their effectiveness depends on the temporal characteristics of the underlying degradation process. The preliminary study in Chapter 5 relied on a single CAD technique and did not examine robustness to model choice or sensitivity to forecasting sequence length, factors that are critical for multivariate seq2seq forecasting in predictive maintenance (Allogba et al., 2021). This chapter fills that gap through a systematic evaluation of five CAD techniques within the unified CPD-CAD-aware forecasting framework.

6.2.3 QoT forecasting in optical networks

QoT forecasting has been widely studied for predictive maintenance. Prior work has applied RNN and seq2seq models to predict SNR, Q-factor, and related metrics over single- and multi-step horizons (Tremblay et al., 2019; Aladin et al., 2020a; Allogba et al., 2022a; Allogba et al., 2021; Aladin et al., 2020b; Mezni et al., 2020; Chouman et al. 2021). LSTMs have been especially effective due to their gating mechanisms, which help capture long-term temporal dependencies that reflect degradation patterns unfolding over hours or days. More recent efforts have explored uncertainty-aware forecasting using Bayesian RNNs and attention-based models such as transformers to improve robustness and long sequence modeling (Di Cicco et al., 2023; Hedayatnejad et al., 2025). Yousefi et al. (2023) have proposed an approach for uncertainty estimation to produce prediction distributions and confidence margins by combining Monte Carlo dropout with an MLP model. Moreover, they used an uncertainty calibration error (UCE) to calibrate the uncertainty. Event-based approaches have also been proposed, framing loss-of-signal prediction as a classification task optimized for high precision (Du et al., 2021). The authors showed that such models can forecast loss of signal events up to seven days in advance, establishing a practical reference horizon that motivates the 7-day input window examined in this study.

TCNs, introduced by Bai et al. (2018) in a broad comparison of sequence-modeling architectures, offer a structurally different alternative to recurrent models. TCNs use dilated causal convolutions to process sequences in parallel, enabling stable gradients, efficient training, and explicit control over temporal receptive fields. Their receptive field is determined entirely by architectural parameters such as layers, kernel size, and dilation schedule, making the effective sequence length a transparent design choice. Despite these advantages, comparisons of LSTM and TCN for optical QoT forecasting, particularly within a CPD- and CAD-enhanced input space, have not been previously explored. This chapter provides the first such comparison across three production network datasets under consistent experimental conditions.

Most existing forecasting approaches treat prediction as an isolated regression or classification task, without considering regime shifts or collective anomalies in the input data. Training on data containing undetected regime changes can degrade performance, and these models typically offer no early-warning indicators beyond raw predictions. Chapter 5 addressed this limitation by incorporating CPD and CAD scores as additional inputs to LSTM-based forecasters, showing that these contextual signals improve accuracy and provide consistent early-warning performance. However, key questions remained regarding robustness across datasets, CAD model selection, architectural diversity, and sensitivity to input-sequence length.

With respect to sequence length, prior optical network forecasting studies have generally used fixed lookback windows, also called forecast windows. Broader time-series research shows that optimal window length can vary and that longer windows do not always improve performance; in some cases, shorter windows reduce noise and enhance generalization, especially for recurrent models (Aladin et al., 2020b; Allogba et al., 2021). For TCNs, the relationship between input length and receptive field adds another consideration: windows shorter than the receptive field underutilize the model, while excessively long windows may introduce unnecessary noise. This chapter evaluates 3-day, and 7-day input windows across both architectures and all datasets to characterize this sensitivity and provide support for practical deployment.

6.2.4 Multi-dataset generalizability in ML-based network management

Reproducibility and generalizability remain major challenges for ML-based optical network management. Most proposed methods are evaluated on a single dataset from one operator or monitoring setup, which demonstrates feasibility but cannot show whether the reported improvements reflect true methodological advantages or simply the characteristics of that specific network. This distinction is critical in practice: proactive network management systems deployed using data from an environment must perform reliably under different conditions, not only those seen during initial validation.

Evaluating a framework across three field lightpath datasets provides a direct way to test this. Applying the same training and evaluation procedures to independently collected datasets allows to determine whether performance improvements are consistent or dependent on context.

For CPD-CAD-aware QoT forecasting specifically, the key question is whether the performance improvements observed in Chapter 5, where CPD and CAD indicators improved forecasting accuracy and provided early warning, persist across datasets with different types of degradation, OPM distributions, and failure characteristics. It is possible that the improvements in Chapter 5 were influenced by the particular temporal structure of that dataset. Conversely, if similar improvements appear across diverse datasets, this would indicate that CPD and CAD indicators provide broadly useful predictive information.

This chapter addresses this question by evaluating the CPD-CAD-aware forecasting framework on three lightpath datasets from a production network under identical experimental conditions. For each dataset, performance is measured both in absolute terms and relative to a baseline without CPD-CAD integration, allowing the contribution of the detection modules to be isolated. Consistent relative improvements across datasets, regardless of differences in absolute difficulty, serve as the primary indicator of generalizability.

6.2.5 Open research questions addressed in this chapter

The literature review above highlights some open research questions that have not yet been addressed in prior work. Although deep-learning-based QoT forecasting is well established, and CPD and CAD methods are individually mature, their systematic integration as contextual inputs to forecasting models for proactive network management has only been explored in the preliminary work of Chapter 5. That study demonstrated feasibility but did not validate four key design challenges: generalizability across different datasets, the relative performance of LSTM versus TCN forecasting models, robustness to CAD model selection, and sensitivity to input-sequence length. No prior work has jointly examined these challenges within a unified CPD-CAD-aware forecasting framework using production OPM data.

These open research problems motivate the broader evaluation presented in this chapter. Here, the CPD-CAD-aware framework is tested across three lightpath datasets from a production network, incorporates five CAD techniques through a unified model selection process, compares LSTM and TCN architectures, and examines two input-sequence lengths. Together, these extensions advance CPD-CAD-aware forecasting toward a more robust and deployable proactive optical network management solution.

6.3 Problem formulation

This chapter extends the problem statement introduced in Chapter 5. It repeats the core forecasting objective which is predicting QoT degradation with enough lead time to enable proactive intervention, and then introduces the central generalization question: does the CPD-CAD-aware QoT forecasting framework maintain its performance across datasets that differ in degradation behaviors? The three control variables used in this study, namely data source, architecture, and sequence length, are defined here as the experimental factors.

6.3.1 Forecasting objective and reaction window definition

This chapter focuses on multivariate seq2seq QoT forecasting for proactive optical network management. At each time step, multiple OPM parameters, including Q-factor, OPR, and DGD, are observed for a given lightpath, as in Chapter 5. The forecasting model receives a recent window of these metrics and predicts their future evolution over a specified horizon. Both inputs and outputs are multivariate, allowing the model to learn joint temporal patterns across all monitored parameters rather than treating them independently.

In the CPD-CAD-aware formulation, the raw OPM inputs are augmented with two additional contextual signals: an online Kernel-CUSUM CPD score indicating recent distributional shifts, and a CAD score from the selected CAD method indicating whether the recent sequence exhibits abnormal behavior. These signals are appended to the OPM features at each time step, enabling the forecasting model to condition its predictions on the current regime and anomaly context.

Forecast accuracy is measured using MAE and RMSE. RMSE emphasizes large deviations and is therefore sensitive to sharp degradation events, while MAE provides a stable measure of average error. The coefficient of determination R^2, shows how much the input variables explain the variance of the output variable. The three metrics are reported per dataset and in aggregate.

Operational usefulness is assessed through the reaction window metric, which measures the time between the model's first degradation indicator and the actual occurrence of an ILOS event. A positive reaction window indicates advance warning, which is essential for preventive actions such as rerouting or maintenance scheduling.

6.3.2 Generalization problem statement

The preliminary study in Chapter 5 showed that CPD-CAD-aware forecasting can improve prediction accuracy and provide meaningful reaction window or lead times, but only for a single field lightpath dataset. The central question addressed in this chapter is whether those benefits generalize to other operational environments with different monitoring intervals, OPM distributions, and degradation patterns.

Generalizability is evaluated by applying the framework independently to three lightpath datasets from a production network under identical conditions: the same model architectures, training and validation procedures, CAD model selection process, and evaluation metrics. No parameters are shared across datasets. Performance is assessed both in absolute terms and relative to a baseline without CPD-CAD enhancement, allowing the impact of the detection modules to be isolated. Consistent relative improvements across all three datasets serve as the primary evidence of generalizability.

6.3.3 Experimental variables

The experimental design is structured around three independent variables, directly aligned with the four contributions outlined in Section 6.1.3.

Data source

Four field lightpath datasets are used. Channel 01 is the reference dataset from Chapter 5, while Channels 02-04 are newly introduced and represent distinct optical channels' characteristics. Each dataset corresponds to an individual lightpath. Full dataset details are provided in Section 6.4.1.

Forecasting architecture

Two deep-learning models, namely LSTM and TCN, are evaluated as forecasting architectures within the same CPD-CAD integration framework. For each architecture, a baseline version without CPD-CAD enhancement is also tested to isolate the contribution of the detection modules.

Input sequence length

Two sequence lengths are examined: 3 and 7 days. The 4-day window serves as the baseline from Chapter 5. The 3-day window tests whether a shorter temporal context still supports early-warning capability, while the 7-day window evaluates whether capturing a full weekly cycle improves performance, motivated by prior 7-day LOS forecasting results (Du et al., 2021).

CAD model selection

The CAD model selection process operates as an inner loop within this design. For each of the combinations of dataset, architecture, and sequence length, all five CAD techniques are evaluated on the validation split, and the best performing method is selected for final testing. The full experimental matrix and implementation details are provided in Section 6.4.

6.4 Extended experimental setup

This chapter builds on the framework introduced in Chapter 5. The experimental setup focuses only on the new elements: the three additional datasets, the expanded forecasting configurations with the two sequence lengths, and the CAD model selection procedure across five techniques. The CPD module, with the offline KernelCPD, online Kernel-CUSUM algorithm and their integration with the forecasting models, remains unchanged from Chapter 5 and is not repeated here. More details of the CPD implementation can be found in Chapter 5.

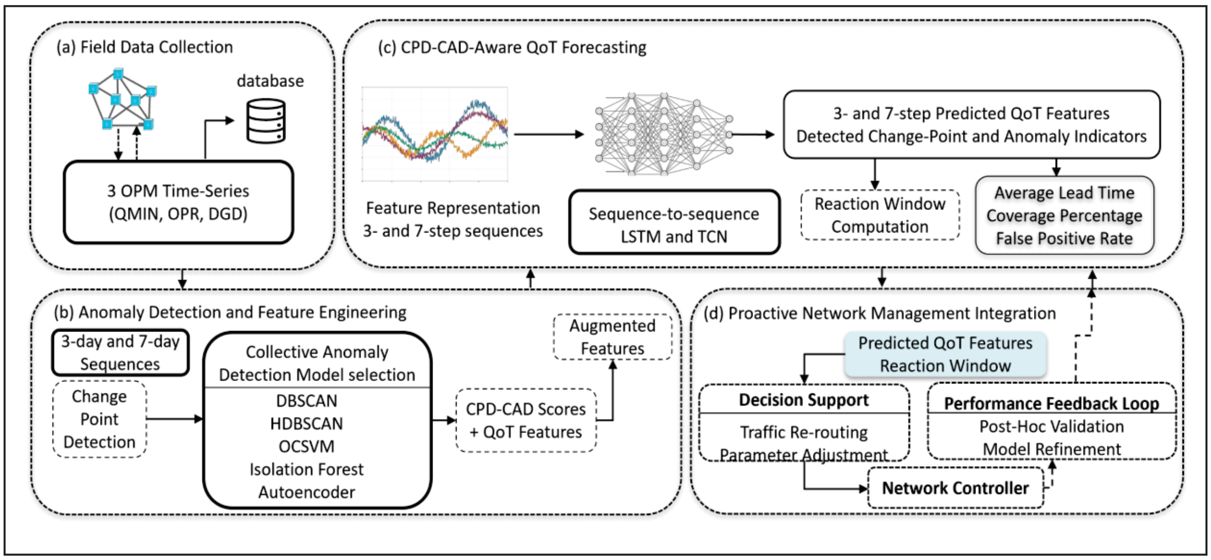


Figure 6.2 Experimental elements for the systematic validation of the CPD-CAD-aware forecasting framework

The extended experimental setup is structured around five dimensions. Four of these are the variables already introduced in Section 6.3.3. The fifth dimension adds an operational evaluation element. While Chapter 5 (Article 3) offers an estimate of reaction windows based on three failure events within a single field lightpath dataset, Chapter 6 develops this into a structured evaluation framework. It introduces two additional metrics: a coverage rate, capturing the fraction of failure events for which a strictly positive reaction window is obtained, and a false positive rate, measuring the proportion of predicted degradations that do not correspond to actual failure events. This evaluation is then extended across all three datasets to test how well the methods generalize beyond a single context. Together, these five dimensions represent a systematic broadening of the CPD-CAD-aware forecasting framework originally proposed in Article 3, as depicted in Figure 6.2.

6.4.1 Dataset description and temporal characterization

The lightpath datasets are collected from a production optical network, over approximately 4 years. They include daily records of network performance parameters and alarm indicators extracted at the receiver side. In this study, we focus on the following parameters:

- QMIN (dB): Minimum Q-factor represents the signal quality, approximately linear with the SNR with slope 1 in a dB/dB scale for several modulation formats.
- OPR (dBm): Minimum optical power received corresponds to the amount of light measured at the receiver end of the fiber. It can be influenced by fiber losses, connectors, splitters, and amplifiers while traveling through the optical channel.
- DGD (ps): Differential group delay represents the arrival time difference for two pulses traveling in different polarization modes of a fiber and at different velocities due to asymmetries and other random imperfections in the fiber.
- HCCS (s): High correction count seconds counts the number of seconds each day when the number of errors to be rectified using "forward error correction" approaches exceeds a specified threshold.
- UAS (s): Unavailable seconds counts the number of seconds when a facility was unavailable and unable of carrying out its purpose.

Four datasets corresponding to four lightpaths on different paths in the production network are considered, denoted as Channel 01 through Channel 04. Unless otherwise stated, results in this study are reported for Channel 02 to Channel 04, while Channel 01 was used in the preliminary study (Aladin, Wosinska & Tremblay, 2026).

Channel 02 dataset comprises instances from August 2nd 2018 to June 11th 2022. Channel 03 and Channel 04 datasets cover a period from November 2nd, 2018 to June 11th, 2022. Figure 6.3 shows an example of the QMIN signal for Channel 03 with one LOS event and three ILOS events for the period.

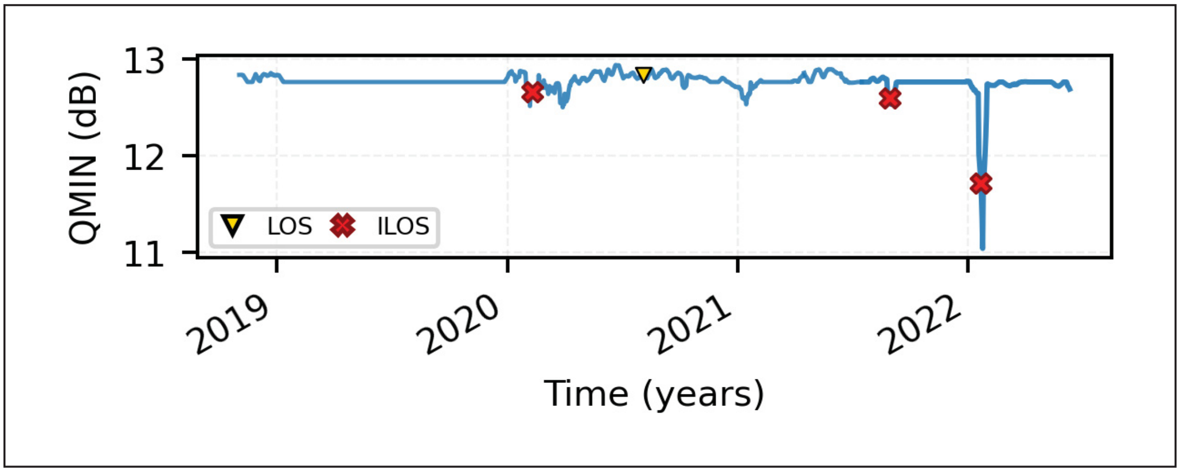


Figure 6.3 Time series of OPM parameter QMIN and alarms over a 4-year period for Channel 03

The QoT forecasting problem is modeled as a multivariate seq2seq prediction task. Given an input sequence $X_{t-L+1:t} \in \mathbb{R}^{L\times3}$, with $L \in \{3, 7\}$ and each time step comprising the QoT features QMIN, OPR, and DGD, the models predict their future evolution over the specified horizon. This formulation aligns with operational maintenance requirements and enables early detection of QoT degradations (Du et al., 2021). A chronological split of the data into training, validation and test sets is adopted to represent realistic deployment and prevent information leakage.

To evaluate whether QMIN, OPR, and DGD are appropriate inputs for seq2seq DL models and to motivate the CPD-CAD integration approach, the temporal characterization of each feature is analyzed across Channels 02 - 04 using the ADF and KPSS stationarity tests, and autocorrelation function (ACF) analysis. Table 6.1 summarizes the findings. For ADF, stationarity is indicated by rejection of the unit root null (p-value $<$ 0.05). For KPSS, stationarity is indicated by failure to reject the stationarity null (p-value $>$ 0.05). ACF lag-1 measures how quickly the influence of past observations fades. QMIN and DGD are stationary at Channels 02 and 03 under both tests. OPR shows inconclusive stationarity for these channels, consistent with gradual optical channel degradation, and is addressed through the Min-Max normalization applied in the forecasting configuration. At Channel 04, OPR becomes stationary while QMIN becomes inconclusive, reflecting signal characteristics specific to each route. The ACF lag-1 values for QMIN across datasets (0.14 at Channel 04, 0.26 at Channel 02, and 0.35 at Channel 03) further indicate that QMIN's intrinsic day-to-day predictability varies considerably by dataset, with Channel 04 exhibiting the weakest short-term temporal dependencies of the three. Despite these differences, individual stationarity results do not fully characterize the multivariate input space.

Table 6.1 Temporal characterization of each QoT features by Channel

Channel	**Feature**	**ADF Statistic with p-value**	**KPSS statistic and p-value**	**Stationarity**	**ACF Lag-1**
02	**QMIN**	-7.99 with p-value < 0.001	0.19 with p-value = 0.100	Stationary	0.26
	OPR	-9.08 with p-value < 0.001	0.48 with p-value = 0.046	Inconclusive	0.47
	DGD	-21.26 with p-value < 0.001	0.19 with p-value = 0.100	Stationary	0.37
03	**QMIN**	-6.31 with p-value < 0.001	0.42 with p-value = 0.067	Stationary	0.35
	OPR	-5.34 with p-value < 0.001	0.80 with p-value = 0.010	Inconclusive	0.74
	DGD	-6.97 with p-value < 0.001	0.39 with p-value = 0.079	Stationary	0.74
04	**QMIN**	-20.08 with p-value < 0.001	0.77 with p-value = 0.010	Inconclusive	0.14
	OPR	-18.08 with p-value < 0.001	0.19 with p-value = 0.100	Stationary	0.24
	DGD	-6.27 with p-value < 0.001	0.07 with p-value = 0.100	Stationary	0.75

6.4.2 ILOS definition, labeling strategy, and pseudo-lead-time

ILOS events are identified using a UAS threshold of 1800 seconds (30 minutes) whereas transient LOS indicates degradations on the channel for a few seconds. This 30-minute criterion, although different from the 1-hour threshold in Du et al. (2021), follows ITU-T guidelines for defining unavailability states and provides a balance between capturing severe degradations and preserving enough subtle early-warning instances for effective training and evaluation (International Telecommunication Union, 2003, 2002, 2011).

Designating failure using UAS ≥ 1800 s, results in approximately 2% daily unavailability. This is significantly higher than availability requirements of "three-nines - four-nines" (99.9-99.99%) for OTN systems (Ramaswami et al., 2010; Gour, et al., 2021). This value remains within the regime of severe degradation while providing a broader and more realistic scenarios of early-warning examples than the more extreme 3600 s threshold used in previous work such as Du et al. (2021) and Aladin et al. (2024) which corresponds to 4% unavailability and captures only the most catastrophic outages.

Using this alarm-based failure concept, we introduce a pseudo-lead-time metric to quantify early detection capability, as in Chapter 5. It is defined as the number of days between the first timestamp of an anomaly or change point sequence and the subsequent ILOS occurrence (UAS > 1800 s). This metric indicates how far in advance the framework flags QoT degradations relative to alarm-based indicators. From an operational perspective, the pseudo-lead-time represents the time available for network operators to take preventive measures before service disruptions.

Because failure occurrences in OTNs are rare and often reported only through high-level alarms, supervised learning is challenging. To enable evaluation of anomaly detection performance, a practical labeling strategy based on the alarm indicators UAS and HCCS is adopted, similarly to the process in Du et al. (2021).

Because labeled failure data are scarce in production optical networks, the framework does not use supervised training at any stage. Instead, a weak labeling strategy based on alarm indicators, following Du et al. (2021), is used only for evaluation. Using a 14-day sliding window, a sequence is labeled “abnormal” if any UAS or HCCS alarm occurs in the final 7 days; for 3-day models, the window is shortened to 6 days with a 3-day alarm horizon.

Given a sliding window of length $S = 14$ days, as used in Du et al. (2021), a sequence is labeled as “abnormal” if any UAS/HCCS alarm occurs within the last 7 days of the sequence as represented in Eq. (6.1):

$$y_t = \begin{cases} 1, & \text{if } \max(A_{t+1:t+L}) > 0 \\ 0, & \text{otherwise.} \end{cases} \tag{6.1}$$

Where y_t is the degradation label at timestamp t. $L = 7$ is the sequence length. And $A_{t+1:t+L}$ denotes the list of alarm indicators (UAS or HCCS) over the last 7 days of the 14-day sequence. The 3-day seq2seq forecasting models are built using the same process with $L = 3$ and $S = 6$ days, respectively.

This labeling strategy reflects the operational assumption that failures are typically preceded by anomalous behaviors. It provides a weak but meaningful sign of degradations before failures. These labels are never exposed to the CPD, CAD, or forecasting modules during training.

They are used only to compute precision-recall area under the curve (PR-AUC) during CAD model selection on the validation split and to evaluate pseudo-lead-times or reaction windows on the test split. PR-AUC is preferred due to the strong class imbalance in production networks, where abnormal sequences are far less frequent than normal ones.

The pseudo-lead-time is defined as the number of days between the first detected anomaly or change point and the subsequent ILOS event. For each ILOS occurrence, the closest preceding CPD detection or CAD anomaly is used to compute the lead time. This metric captures the framework’s early-warning capability by quantifying how far in advance the CPD-CAD module detects QoT degradation, directly indicating the time available for operators to take preventive action before service disruption.

Table 6.2 summarizes the three additional datasets used in this chapter across their key characteristics. The diversity in dataset duration, available features, and train/validation/test split ratios: it ensures that the generalizability evaluation covers a meaningful range of operational conditions rather than three datasets with similar characteristics.

Table 6.2 Summary of the three field lightpath datasets with the monitoring period, features, and the train/validation/test split ratio for 3-day and 7-day sequence lengths

Dataset	**Period**	**Features**	**Number of ILOS in Test Set**	**Split (Train/Validation/Test)**
Channel 02	Aug 2018 - Jun 2022	QMIN, OPR, DGD	3	3-DAY: 52/3/45 7-DAY: 51/4/45
Channel 03	Nov 2018 - Jun 2022	QMIN, OPR, DGD	2	3-DAY: 70/5/25 7-DAY: 68/7/25
Channel 04	Nov 2018 - Jun 2022	QMIN, OPR, DGD	2	3-DAY: 62/11/27 7-DAY: 58/15/27

6.4.3 Input sequence configuration

Two input sequence lengths are examined across all datasets and architectures: L=3 days and L=7 days. The 7-day window is chosen for two reasons: it captures the weekly periodicity commonly observed in network traffic, potentially informative for load-related optical stress, and it matches the 7-day forecasting horizon used by Du et al. (2021) for LOS prediction, enabling direct comparison with prior work. The 3-day window, however, evaluates how reducing temporal context affects forecasting accuracy and early-detection performance.

Input sequences and their prediction targets are constructed using a fixed-length sliding window that advances one step at a time. At each position, the model receives L consecutive daily OPM observations and predicts the following L time steps. This seq2seq setup aligns with Chapter 5 and with operational requirements for multi-step QoT forecasting. All sequences are generated after appending normalized CPD and CAD scores to the raw OPM features, ensuring that the augmented input is used consistently across all model variants. The only distinction between baseline and enhanced models is whether these additional features are included, allowing the specific contribution of CPD and CAD to be isolated.

Missing values in QMIN, OPR, and DGD are imputed using a rolling average matched to the sequence length (7-day average for 7-day windows, 3-day average for 3-day windows). For alarm indicators (UAS and HCCS), missing entries are replaced with zeros, a conservative assumption that avoids injecting artificial anomalies. All numerical features undergo Min-Max normalization prior to training to promote stable neural network optimization and maintain consistent scaling between QoT features and CPD-CAD scores.

Train/validation/test splits follow a strictly chronological order: training precedes validation, which precedes testing, with no overlap. Each dataset uses its own split boundaries to ensure that at least one ILOS event appears in the test set. Because the number of samples produced by the sliding window depends on L, the resulting split proportions differ slightly between the 3-day and 7-day configurations; these proportions are reported in Table 6.2. This temporal ordering prevents information leakage and mirrors real-world deployment, where models are trained on historical data and evaluated on future observations.

Forecasting performance is evaluated using three complementary metrics: RMSE, which emphasizes large errors and is therefore sensitive to sharp degradation events; MAE, which reflects the average prediction error across the evaluation period; and R^2, which measures the proportion of variance explained by the model. All models use identical data partitions. A vector autoregressive moving average with exogenous regressors (VARMAX) model from the *statsmodels* package provides a linear multivariate baseline. Enhanced LSTM and TCN models incorporating CPD and CAD features are denoted E-LSTM and E-TCN, respectively.

6.5 Change point detection for regime shift awareness

Change point detection is applied in two stages corresponding to the training and test phases of the framework. During training, an offline KernelCPD (Truong et al, 2020) with a radial basis function (RBF) kernel and a maximum of 10 breakpoints ($n\_bkps = 10$) is applied to the training set to identify regime shifts in the multivariate QoT feature space.

During evaluation, change points are detected incrementally using an online Kernel-CUSUM procedure based on the MMD statistic with and RBF kernel. The RBF kernel bandwidth parameter γ is estimated once using the median heuristic applied to at most 200 uniformly subsampled training observations and is held fixed throughout the online detection phase. At each test step i, a reference window of length equal to the sequence length $L \in \{3, 7\}$ and a test window of fixed length of 5 are extracted from the growing historical data with accumulated test observations up to step i. The detection threshold is calibrated on the training set as $\tau = \mu + 2\sigma$ where μ and σ are the mean and standard deviation of the squared MMD (MMD^2) statistics computed over adjacent training windows. A change point is detected at step i whenever the MMD^2 statistic exceeds τ. The raw MMD^2 scores are subsequently normalized to $[0, 1]$ via Min-Max scaling. This procedure ensures that no future test observations are accessed at any detection step.

The resulting online CPD scores are appended to the test QoT feature vectors alongside the online CAD score as described in Sections 6.6 and 6.7 to form the augmented test input sequences. The augmented test input vector is fed to the enhanced LSTM and TCN forecasters. The same CPD configuration is applied across all three datasets and for both sequence lengths with the reference window adapting to the sequence length while the test window and the threshold calibration procedure remain the same.

6.6 Comparative collective anomaly detection study

A major limitation of the preliminary study in Chapter 5 was its reliance on a single CAD method, namely the HDBSCAN, without a systematic comparison against alternative approaches. Because the CAD score is used directly as an input feature to the forecasting models, the choice of CAD technique influences not only which anomaly patterns are detected but also the quality and stability of the signal provided to the forecaster. A method that yields noisy, inconsistent, or poorly calibrated scores can ultimately hinder forecasting performance, even if it performs well as a standalone anomaly detector. This section introduces the five CAD techniques evaluated in this chapter, outlines the criteria used to compare them, and analyzes how the CAD method is selected for subsequent steps.

6.6.1 CAD techniques evaluated

All five CAD techniques are applied in a collective anomaly detection setting: instead of evaluating individual observations, each technique operates on a sliding window of consecutive OPM data instances and outputs a single scalar labeling that window. The input for each technique consists in 7-day or 3-day sliding window sequences constructed from the three monitoring features QMIN, OPR, and DGD. The offline step is performed using the training portion of the datasets. This score is then added to the model's input at each time step, as described below in Section 6.7. The five techniques are presented in Section 1.2.3 and the offline training process is as follows.

DBSCAN

The DBSCAN model is trained on the sequences in the training portion of the data. It produces binary anomaly scores, where any sequence assigned the cluster label -1 (anomaly) is given a score of 1, while other sequences receive a score of 0. To tune the hyperparameters, a grid search is conducted over $eps \in \{0.05, 0.3, 0.5, 0.7, 1.0, 1.5, 2.0\}$ and $min\_samples \in \{5, 10, 15, 20, 50, 80\}$, selecting the combination that allows achieving the highest silhouette score on the training data.

HDBSCAN

The scoring scheme for HDBSCAN model is similar to the DBSCAN's: any sequence labeled -1 (classified as anomaly) is assigned an anomaly score of 1. The model is fitted on the training sequences. The grid search explores $min\_cluster\_size \in \{5, 10, 15, 20\}$ and $min\_samples \in \{5, 10, 15\}$, selecting the parameter combination that maximizes the silhouette score.

Isolation Forest

Isolation forest isolates observations by recursively partitioning the feature space using random splits. Each sequence is evaluated using the average path length across the ensemble of isolation trees. Sequences that are isolated in fewer splits receive lower path lengths and are considered more anomalous. The anomaly score is represented as the negation of the $score\_samples$ output, so that larger values indicate stronger anomalies. Sequences for which the score exceeds the decision threshold are marked as anomalies. The model is fitted on the training sequences with a fixed random seed for reproducibility. The grid search covers $n\_estimators \in \{100, 200\}$, $contamination \in \{0.04, 0.15, 0.2\}$, and $max\_samples \in \{256, 512\}$, selecting the configuration that achieves the highest mean negative score.

One-Class SVM (OCSVM)

The OCSVM model learns a decision boundary in the feature space that encloses the majority of normal training sequences. It marked any sequence that falls outside the boundary as anomalies. The model is trained on the training portion of the data. The anomaly score for each sequence is defined as the negative values of the model's decision function, such that larger values indicate stronger anomalies. Sequences for which the decision function returns a negative value are flagged as anomalous. Hyperparameter tuning considers $kernel \in \{'rbf', 'poly'\}$, $gamma \in \{'scale', 'auto', 0.001, 0.01, 0.1\}$, and $nu \in \{0.01, 0.04, 0.1, 0.15, 0.2\}$, retaining the configuration that maximizes the mean negated decision-function value.

Autoencoder

The Autoencoder model trains a symmetric encoder-decoder architecture: the input is passed through three dense layers in the encoder using ReLU activations. The decoder mirrors the structure back, ending with a linear activation. A dropout rate of 0.2 is applied after the first dense layer in both the encoder and decoder. The model is trained on the sequences to minimize mean squared error (MSE) reconstruction loss with early stopping (patience = 5) applied during training. Each sequence's anomaly score is its per-sample MSE reconstruction error, and a sequence is marked as anomaly if its score exceeds the threshold defined by the contamination percentile of the training reconstruction errors. Hyperparameter tuning covers $encoding\_dim \in \{10, 15, 20\}$, $epochs \in \{50, 100\}$, $batch\_size \in \{32, 64\}$, and $contamination\ percentile \in \{0.9, 0.95\}$, selecting the configuration that minimizes the mean reconstruction error on the training portion of the data.

Table 6.3 summarizes the five CAD techniques, including their hyperparameters, and the values tested.

Table 6.3 Hyperparameters tested for the five CAD techniques across datasets

CAD Technique	Hyperparameter	Values
DBSCAN	eps	0.05, 0.3, 0.5, 0.7, 1.0, 1.5, 2.0
	min_samples	5, 10, 15, 20, 50, 80
HDBSCAN	min_cluster_size	5, 10, 15, 20
	min_samples	5, 10, 15
Isolation Forest	n_estimators	100, 200
	max_samples	256, 512
	contamination	0.04, 0.15, 0.2
One-Class SVM	kernel	'rbf','poly'
	nu	0.01, 0.04, 0.1, 0.15, 0.2
	gamma	'scale', 'auto', 0.001, 0.01, 0.1
AutoEncoder	encoding_dim	10,15,20
	epochs	50,100
	batch_size	32, 64
	contamination	0.9, 0.95

6.6.2 CAD model selection criteria

For each experimental configuration, defined by the dataset, architecture, and input sequence length, the optimal CAD technique is selected using the training split prior to any evaluation on the test split. The selection process integrates two complementary criteria that reflect the dual purpose of the CAD score in this framework: it must reliably characterize collective anomalous behavior in the OPM data, and it must provide a signal that enhances the forecasting model.

Each technique identifies its best hyperparameter configuration separately, using an unsupervised metric appropriate to the detection technique. For the density-based clustering techniques such as DBSCAN and HDBSCAN, the silhouette score is used to measure how well sequences fit within their assigned clusters relative to neighboring clusters. For Isolation Forest and OCSVM, the mean negated score from the model's scoring function is used while the mean reconstruction error is used for the Autoencoder.

The best configuration from each technique is assessed against the practical weak labels using the PR-AUC, which is appropriate given the strong class imbalance typical of field optical network failure data. The selected technique, achieving the highest PR-AUC on the training split is then applied unchanged to the test set.

As in Chapter 5, standalone CPD and CAD performance is not the focus of this chapter; their value is assessed indirectly through their contribution on the forecasting models' performance.

6.7 Integration with the forecasting models

The binary anomaly scores generated by the best CAD model on the training set are scaled to the range $[0, 1]$ and added as extra input features alongside the CPD scores for the enhanced LSTM and TCN forecasters. Let $\boldsymbol{X_t} = (x_{t-L+1}, \dots, x_t) \in \mathbb{R}^{L\times 3}$ denote the baseline input sequence of L daily OPM observations, where $L \in \{3, 7\}$. The CPD score $s_i^{cpd} = \frac{1}{1+d_i}$, where $d_i = min_{cp}|i - cp|$ is the distance from sequence i to the nearest change point cp, and the binary anomaly score $s_i^{cad} \in \{0,1\}$ produced by the selected CAD model, are appended at each timestep to form the augmented input $\boldsymbol{X_t'} \in \mathbb{R}^{L\times 5}$, which is processed by the seq2seq forecaster to produce $\widehat{\boldsymbol{Y}}_{t+1:t+L} \in \mathbb{R}^{L\times 3}$. All inputs, including the CPD and CAD scores, are normalized using MinMaxScaler technique prior to being fed to the forecasting models.

For the test portion, an online detection procedure incrementally retrains the best CAD model using the growing history that includes all training data plus the accumulated test observations. At each update, the anomaly prediction for the current test sequence is obtained from the retrained model normalized to $[0, 1]$, and added as the CAD score to the forecaster input. This setup ensures that the enhanced forecasting models receive anomaly insights without leaking future observations. This integration strategy is assessed across three field lightpath datasets and for both sequence lengths, allowing for a systematic assessment of the contribution of CPD and CAD integration to forecasting performance under varying temporal contexts and network operating conditions.

6.8 Forecasting architecture overview

This section outlines the two DL architectures used as forecasting architectures within the CPD-CAD-aware framework: the LSTM-based seq2seq model described in Chapter 5, and the TCN introduced here for comparison. Both architectures operate on the same augmented input, which consists in raw OPM metrics combined with CPD and CAD scores, and generate the same multivariate output sequence over the prediction horizon. The integration strategy is therefore identical for both models; only the internal mechanism for capturing temporal dependencies differs. This controlled setup ensures that any performance differences observed between the two architectures can be attributed to the architectures themselves rather than to variations in input processing, training strategy, or evaluation methodology.

6.8.1 LSTM configuration

The LSTM architecture used in this chapter follows the same seq2seq configuration developed and validated in Chapter 5. The encoder consists of two stacked LSTM layers with 64 and 32 units respectively, incorporating a dropout rate of 0.2. This configuration enables the network to capture long-term temporal dependencies in OPM metrics by propagating hidden and cell states through the sequences, allowing the gated recurrent units to selectively retain relevant historical patterns. The decoder mirrors this structure by using a RepeatVector layer followed by a 32-unit LSTM layer and a TimeDistributed Dense layer to generate the output sequences.

The baseline input comprises sequences of the three QoT features over two forecasting horizons $L \in \{3, 7\}$ days, extended to five features when CPD and CAD indicators are included. All inputs are normalized using a MinMaxScaler method, and the output shares the same sequence length as the input processed by the seq2seq models. Hyperparameters are optimized via grid search on each dataset's validation split. All hyperparameters are identical to those in Chapter 5 and remain fixed across datasets and sequence lengths to ensure comparability.

The choice of LSTM networks is motivated by their established capacity to model complex dependencies between past and future observations and their widespread adoption for multivariate time series forecasting (Hall & Rasheed, 2025).

6.8.2 TCN configuration

The TCN introduced in this chapter uses four stacked of dilated causal convolutional blocks, with exponentially increasing dilatation rates, giving the network a receptive field that grows with depth while strictly preserving temporal order. Each block contains two causal one-dimensional convolution (Conv1D) layers with the same dilation factor, followed by a batch normalization, a ReLU activation, dropout, and a residual connection that both stabilizes training and allows the network to learn refinements to a simpler baseline mapping (Bai et al., 2018). The output of the four blocks is aggregated via global average pooling and passed through a 128-unit Dense layer with ReLU activation before the final projection to the forecast horizon. Hyperparameters are optimized via grid search on each dataset, with an EarlyStopping mechanism (patience = 5) and a ReduceLROnPlateau learning rate scheduler (factor = 0.5, patience = 3). All models are compiled using the Adam optimizer and MSE loss.

By using dilated causal convolutions, TCNs can capture patterns that span hundreds of time steps without relying on recurrence, while still preserving the correct temporal order. Their fixed, explicitly defined receptive field makes it easy to control how much historical context the model sees, and residual connections help stabilize training even in deep architectures. Because TCNs process entire sequences in parallel rather than step-by-step, they train significantly faster than recurrent models like LSTMs, making them well-suited for large-scale or real-time forecasting tasks.

6.8.3 Baseline models

A VARMAX model from the *statsmodels* package is used as a linear statistical baseline. Unlike naive baseline such as last-value persistence or moving-average repetition commonly used in QoT forecasting, VARMAX explicitly models both the autocorrelation of each time series and the cross-correlation structure among the three OPM features QMIN, OPR, and DGD. This makes VARMAX a more informative baseline for evaluating multivariate seq2seq forecasting models (Lütkepohl, 2005). The model is configured using the attributes $VAR\ order = 2$, $moving\ average\ order = 1$, and $trend\ term = \text{‘}c\text{’}$ for a constant intercept. No CPD or CAD augmentation is applied to the input features for the VARMAX model, as it serves strictly as a linear baseline. It produces joint forecasts for the three features over the same prediction horizons and data splits as the DL models, providing a reference point for evaluating their added value.

Table 6.4 Hyperparameters tested for LSTM and TCN forecasting models

Model	Hyperparameter	Values Tested
LSTM	Hidden units	32, 64
	Learning rate	1e-4, 5e-4, 1e-3
	Batch size	32, 64
	Dropout	0.1, 0.2, 0.3
	Optimizer	Adam, RMSprop
TCN	Number of filters	32, 64, 128
	Kernel size	3, 5, 7
	Dilation factor	1, 2, 4, 8
	Dropout	0, 0.2, 0.5
	Learning rate	1e-4, 5e-4, 1e-3
	Batch size	32, 64
	Optimizer	Adam, RMSprop

Together, the VARMAX, baseline LSTM, and baseline TCN models trained on the original QoT features without CPD-CAD integration, constitute the baseline forecasters in the study. This setup enables two layers of comparison: the DL baselines reveal the added value of temporal representation learning over a linear multivariate model, while the enhanced LSTM and TCN variants isolate the specific contribution of CPD and CAD insights to forecasting performance. Table 6.4 summarizes the hyperparameter search spaces used for both DL architectures.

6.9 Results and analysis

6.9.1 Deep learning forecasting models implementation

The LSTM and TCN models are implemented using Python 3.0 with the TensorFlow 2.12.0 package. Hyperparameters for LSTM and TCN models are tuned independently for each dataset via grid search over the search values reported in Table 6.4. The VARMAX model is trained on the same chronological partitions and evaluated on the same test sets as the DL models. For conciseness in tables and figures, the CPD-CAD-aware LSTM is referred to as E-LSTM (for Enhanced LSTM) and the CPD-CAD-aware TCN as the E-TCN (for Enhanced TCN). The baseline models retain their regular acronyms (LSTM, TCN, and VARMAX). The train/validation/test split ratios are reported in Table 6.2.

6.9.2 Aggregated forecasting performance

Table 6.5 presents aggregated RMSE, MAE, and R^2 for all models across Channels 02, 03, and 04 for both sequence lengths. E-LSTM and E-TCN consistently outperform their respective baselines. E-LSTM achieves averaged RMSE and MAE reductions of approximately 10-25% and R^2 increases of up to 0.15 over the baseline VARMAX and approximately 2-20% over the baseline LSTM. These results confirm that integrating CPD and CAD scores improves forecasting robustness across all datasets and both sequence lengths.

Table 6.5 Aggregated multivariate forecasting performance across test sets of Channels 02 to 04 for 7-day and 3-day sequences

Model	Channel 02			Channel 03			Channel 04		
	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2
7-day									
VARMAX	2.354	1.247	-0.239	1.421	0.547	-0.147	3.278	1.732	-0.100
LSTM	2.282	1.189	-0.124	1.271	0.534	-0.053	2.643	1.353	0.092
E-LSTM	**2.201**	**1.077**	**-0.098**	**1.232**	**0.507**	**-0.051**	**2.545**	**1.284**	**0.129**
TCN	11.157	9.110	-44.527	2.709	1.692	-3.367	7.507	5.566	-10.366
E-TCN	9.297	6.970	-31.850	1.465	0.792	-2.051	6.499	4.742	-5.590
3-day									
VARMAX	2.341	1.245	-0.225	1.409	0.540	-0.156	3.185	1.684	-0.083
LSTM	2.251	1.147	**-0.133**	1.327	0.525	-0.129	2.446	1.308	0.033
E-LSTM	**2.209**	**1.057**	-0.307	**1.140**	**0.464**	**-0.073**	**2.245**	**1.049**	**0.144**
TCN	11.236	9.349	-34.069	1.646	0.790	-0.771	6.034	4.597	-10.256
E-TCN	7.429	5.892	-20.212	1.512	0.763	-0.681	4.872	3.627	-3.156

LSTM-based models consistently outperform VARMAX, demonstrating the value of DL for non-stationary OPM forecasting. The performance difference is most apparent on Channel 04, where LSTM achieves a positive R^2 of 0.092 at 7-day and 0.033 at 3-day, and E-LSTM reaches R^2 = 0.129 and 0.144 respectively. Channel 04 is the only dataset in the extended study where any model achieves positive aggregate R^2, suggesting that its OPM signals exhibit a degree of predictable temporal structure not present in Channels 02 and 03. On Channels 02 and 03, all models yield negative R^2, including VARMAX, a finding that is explained below.

TCN baseline models underperform VARMAX across all scenarios, with significantly elevated RMSE and MAE values, indicating that the TCN architecture in its baseline configuration is less suited to this data. The most extreme case appears in Channel 02 with the 7-day window, where the TCN baseline RMSE of 11.157 is nearly five times higher than the VARMAX baseline RMSE (2.354). This is likely driven by the fact that the TCN cannot selectively suppress pre-shift historical context in the way an LSTM's forget gate can. As a result, predictions become influenced by outdated operating regimes, leading to degraded accuracy. The CPD-CAD integration significantly reduces this limitation. On the 7-day window, E-TCN decreases RMSE relative to the TCN baseline by roughly 17% on Channel 02, 46% on Channel 03, and 14% on Channel 04. For the 3-day window, the corresponding reductions are 34% on Channel 02, 8% on Channel 03, and 19% on Channel 04. The CPD and CAD indicators help balance the TCN's architectural constraint by explicitly encoding regime shift and anomaly information. However, E-TCN does not reach the performance level of E-LSTM in any scenario, indicating that LSTM remains the more effective forecasting technique for the characteristics of these datasets.

Negative R^2 values are observed for Channel 02 and Channel 03 across all models, including VARMAX, indicating that QoT variability in these datasets contains abrupt spikes rather than temporal structure. This is expected behavior when ILOS events produce abrupt signal drops that no forecasting model attempts to predict. The QoT features in these datasets, particularly QMIN and OPR can exhibit abrupt drops during ILOS events. The framework is designed to anticipate degradation trends; thus, the predictions fail to follow the true QoT signal when the abrupt events occur. In this context, negative R^2 is an expected consequence of the forecasting problem formulation rather than an indication that the models perform poorly. More importantly, the RMSE and MAE reductions achieved by E-LSTM and E-TCN over their respective baselines are positive and consistent across all three datasets, and these relative improvements reflect the model's enhanced capacity to track joint feature dynamics under regime shifts, rather than the behavior of any individual feature.

6.9.3 Per-feature forecasting performance

Tables 6.6 and 6.7 break down the aggregated multivariate results into per-feature RMSE, MAE, and R^2 for LSTM and E-LSTM across all three datasets, using 7-day and 3-day input windows, respectively. This decomposition highlights which OPM parameters are most suitable for seq2seq deep-learning forecasting and where CPD-CAD integration provides the largest per-feature improvements.

Table 6.6 Per-feature forecasting performance for 7-day sequences achieved with LSTM and E-LSTM. With QMIN in dB, OPR in dBm and DGD in ps

Feature	*Channel 02*			*Channel 03*			*Channel 04*		
	RMSE	*MAE*	*R²*	*RMSE*	*MAE*	*R²*	*RMSE*	*MAE*	*R²*
LSTM									
QMIN	**0.182**	**0.080**	**-0.022**	0.320	**0.084**	-0.032	0.404	**0.188**	-0.259
OPR	**2.357**	1.408	**-0.156**	**0.230**	**0.151**	**-0.206**	**3.473**	**1.341**	**-0.008**
DGD	3.167	2.079	-0.193	2.166	1.368	0.079	2.954	2.531	0.546
E-LSTM									
QMIN	0.186	0.083	-0.067	**0.319**	0.087	**-0.021**	**0.397**	0.190	**-0.218**
OPR	2.368	**1.393**	-0.168	0.236	0.165	-0.268	3.521	1.476	-0.036
DGD	**2.977**	**1.749**	**-0.054**	**2.096**	**1.270**	**0.137**	**2.621**	**2.185**	**0.642**

Table 6.7 Per-feature forecasting performance for 3-day sequences achieved with LSTM and E-LSTM. With QMIN in dB, OPR in dBm and DGD in ps

Feature	*Channel 02*			*Channel 03*			*Channel 04*		
	RMSE	**MAE**	**R²**	**RMSE**	**MAE**	**R²**	**RMSE**	**MAE**	**R²**
LSTM									
QMIN	**0.191**	**0.093**	**-0.126**	**0.299**	**0.062**	**0.094**	0.396	0.217	-0.436
OPR	2.275	1.276	-0.087	**0.254**	**0.152**	**-0.472**	3.217	1.432	-0.097
DGD	3.161	2.071	-0.185	2.265	1.361	-0.011	2.728	2.277	0.632
E-LSTM									
QMIN	0.237	0.153	-0.733	0.307	0.071	0.048	**0.372**	**0.182**	**-0.265**
OPR	**2.244**	**1.086**	**-0.058**	0.259	0.211	-0.531	**3.149**	**1.183**	**-0.051**
DGD	**3.089**	**1.932**	**-0.131**	**1.934**	**1.109**	**0.262**	**2.251**	**1.781**	**0.749**

DGD is the only feature for which the models achieve better R^2 values across all datasets and sequence lengths. At the 7-day horizon, LSTM yields $R^2 = 0.079$ on Channel 03 and $R^2 = 0.546$ on Channel 04, while E-LSTM improves these to 0.137 and 0.642 respectively, indicating that CPD-CAD integration meaningfully enhances the model's ability to capture DGD variability. At 3-day, E-LSTM reaches $R^2 = 0.262$ on Channel 03 and $R^2 = 0.749$ on Channel 04, the highest R^2 values observed in the extended study. DGD's positive R^2 reflects its stronger temporal continuity: it evolves more smoothly than QMIN or OPR and exhibits temporal dependencies that seq2seq models can exploit effectively as reported in Section 6.4.1.

In contrast, QMIN and OPR generally produce negative R^2 values despite moderate absolute errors, consistent with the findings at the aggregate level. E-LSTM reduces MAE for these features in several scenarios. As an example, it decreases OPR's MAE from 1.408 to 1.393 dBm on Channel 02 using 7-day sequences and from 1.276 to 1.086 dBm on Channel 02 using 3-day sequences. A notable exception is QMIN on Channel 02 with 3-day sequences, where E-LSTM increases RMSE from 0.191 to 0.237 dB and decreases R^2 from -0.126 to -0.733. This degradation suggests that QMIN's short-term autocorrelation does not provide enough information for reliable seq2seq forecasting, particularly during the periods preceding a QoT violation where the feature's dynamics shift without affecting the global stationarity across all three OPM features. CPD-CAD integration provides its strongest improvements at Channel 04, where QMIN shows the weakest short-term dependency (ACF lag-1 = 0.14, Section 6.4.1), making the model most dependent on regime shift information provided by the CPD integration. Moreover, the integrated approach significantly improves DGD performance with DGD's MAE reduction from 2.071 to 1.932 ps, and the model may accept a modest decrease in QMIN accuracy as a trade-off. This interplay between per-feature behavior and multivariate training dynamics is an inherent property of multivariate seq2seq forecasting models rather than evidence of a systematic limitation in the CPD-CAD integration.

6.9.4 Sequence-length sensitivity analysis

We evaluate sequence length in two ways: how it affects overall forecasting accuracy and how it affects pseudo-lead-time performance. These are treated separately because they address different questions. Accuracy metrics reflect average predictive quality across the entire evaluation period, while pseudo-lead-times focus on the model's behavior right before ILOS events. A sequence length that improves accuracy may not improve early-warning performance, and the reverse may also hold.

At the aggregated multivariate level, 3-day and 7-day windows perform similarly. Both benefit from CPD-CAD integration, and the relative ranking of models stays the same under both lengths, with E-LSTM outperforming LSTM, E-TCN outperforming TCN and LSTM outperforming VARMAX. The magnitude of the improvements is also comparable.

At the per-feature level, a clearer trend appears: DGD forecasts improve more with the shorter 3-day window. For example, E-LSTM achieves higher R^2 for DGD on both Channel 03 and Channel 04 with 3-day sequences than with 7-day sequences. This matches the idea that DGD is easier to model when the input focuses on recent, more consistent behavior. QMIN and OPR dynamics are event-driven rather than temporally structured and show negative R^2 under both sequence lengths.

For pseudo-lead-times, sequence length has no effect. Across all three extended datasets and both architectures, the 3-day and 7-day models produce identical lead times: the CPD module identifies regime shifts roughly 28 days before ILOS events, and the CAD module detects collective anomalies approximately 1 day before QoT violation. Although the 7-day model generates fewer total CAD anomaly indicators, the closest detected anomaly to each ILOS event remains within the same 1-day window, leaving computed lead times unchanged.

This stability is important: it shows that early-warning performance depends on the CPD and CAD indicators themselves, not on the forecasting window length. Kernel-CUSUM is independent of the sequence length, while CAD captures anomaly patterns that immediately precede ILOS event and are detectable within both window sizes.

In practice, this means sequence length can be chosen based on accuracy needs and computational cost without affecting early-warning capability. The 3-day window offers better DGD performance, while the 7-day window provides slightly more stable overall results in some datasets. If DGD accuracy is the priority, 3-day sequence is preferable; for broader deployment across diverse scenarios, the 7-day window offers a modestly more conservative choice while preserving early-warning effectiveness.

6.9.5 Qualitative analysis: QMIN forecasting with QoT violations

Figure 6.4 shows the QMIN forecasting for Channel 02 with a 7-day sequence length, during both the normal operating period and the period leading to the ILOS events on January 21 and January 24, 2021. During normal operation, the baseline LSTM (red dashed) and E-LSTM (green solid) follow the overall QMIN evolution (blue solid) with similar accuracy. The closest change points and collective anomalies detected before the QoT violations occur are presented as bold purple vertical lines and dark orange circles respectively.

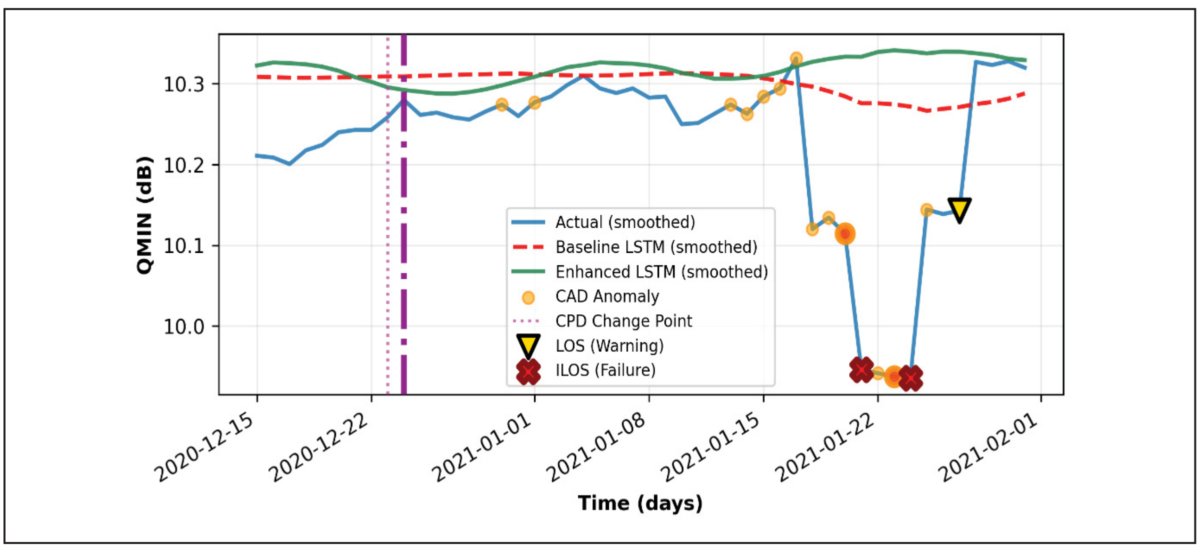


Figure 6.4 Smoothed QMIN timeline (blue solid) and 7-day-ahead forecasts for Channel 02, overlaying baseline LSTM (red dashed) and enhanced LSTM (green solid) models. CAD (yellow circles), CPD (purple vertical line), LOS (yellow triangles) and ILOS (red crosses) are indicated

After the CPD module detects a regime shift on December 23, 2020, E-LSTM adjusts its predictions slightly toward the new post-change-point indicator, whereas the baseline LSTM continues to follow its pre-shift pattern. This adaptation is driven by the elevated CPD score in the augmented input, which signals to E-LSTM that the statistical properties of the network, thus the underlying operating conditions have changed and that older historical context should be given less weight.

In the days leading up to the QoT violation, the CAD module identifies collective anomalies several days before the ILOS events, as indicated in the figure (from January 13 up to January 20, 2021). Neither model attempts to reproduce the abrupt QMIN drops on January 21 and January 24, which is expected: the forecasting task is designed to predict QoT evolutions, not to anticipate discrete failure events. In this period, the operational value comes from the CPD and CAD warnings rather than from the forecasted QoT feature itself.

Following the failure period, E-LSTM readjusts with the recovered QMIN signal more quickly than the baseline model, showing greater robustness to transient disruptions. This faster recovery is operationally useful when the network returns to normal operation after a maintenance intervention, as it allows predictions to stabilize around the new operating point more rapidly with CPD-CAD information included.

6.10 Pseudo-lead-time analysis

Beyond raw forecasting accuracy, the main operational value of the CPD-CAD-aware framework is its ability to issue early warnings ahead of ILOS events. Standard forecasting metrics such as RMSE and MAE capture average performance over the entire evaluation period, including long periods of normal operation where prediction is relatively easy. In contrast, the pseudo-lead-time analysis focuses specifically on the periods immediately before QoT violations that can affect service, and measures how far in advance each component of the framework provides actionable insights. This analysis is performed on all three datasets (Channels 02 to 04) for the 3-day sequence configuration.

To provide more insight on the practical benefits of the CPD-CAD-aware DL forecasting framework we introduce two complementary metrics, namely the coverage and the false alarm rate and we evaluate at three sensitivity levels corresponding to the 90th, 95th, and 99th percentiles of the training residual distribution. This threshold-based evaluation is applied consistently to the baseline LSTM and TCN models configurations across Channels 02 to 04 for both sequence lengths.

6.10.1 Evaluation considerations

For the enhanced models (E-LSTM and E-TCN), failure detection is driven by the CPD and CAD indicators included in the input. A change point or anomaly is detected at any time step where either score is greater than zero, and the average lead time for each ILOS event is computed as the minimum of the CPD-based and CAD-based lead times for each ILOS event, using the closest detection before the QoT violation. For the baseline models (LSTM and TCN), no CPD or CAD indicator is available, so detection relies on forecast residuals: the mean squared error across OPM features is computed for each test sequence and compared against the 90th, 95th, or 99th percentile of the training residual distribution. A detection is flagged whenever the test residual exceeds that threshold. This residual-based method is structured but reactive, where the forecast error reflects the model's inability to track an abnormal signal degradation, but it occurs only after the failure has already begun.

Coverage is defined as the percentage of ILOS events for which at least one detection occurs before the event. Average lead time is the mean number of days between the last detection preceding each ILOS event and the event itself; negative values indicate that detection occurred after failure onset.

False alarms per month count detections not followed by an ILOS event within three days, normalized by the length of the test period. False alarms are computed from forecast residuals for all models, reflecting the operational convention that the alarm signal available to a network operator is the model's prediction error rather than its internal CPD or CAD scores directly. Results are reported for the 3-day sequence scenario because pseudo-lead-time results are identical for both 3-day and 7-day sequence lengths (as shown in Section 6.9.4) the tables below present the results.

6.10.2 Results by dataset

Tables 6.8, 6.9 and 6.10 report the summary of the coverage, average lead time and false alarms per month for Channel 02, 03, 04 respectively.

Table 6.8 Summary of coverage percentage, average lead time and false alarms per month using Channel 02 dataset

Model	Threshold percentile	Coverage (%)	Avg lead time (days)	False alarms/ month	Type
LSTM	90%	100%	-3.3	1.51	Baseline
LSTM	95%	100%	+1.7	0.57	Baseline
LSTM	99%	0%	-	0.00	Baseline
TCN	90%	100%	-3.0	5.39	Baseline
TCN	95%	66.7%	-2.5	0.80	Baseline
TCN	99%	0%	-	0.00	Baseline
E-LSTM	90%	100%	+1.0	0.85	Enhanced
E-LSTM	95%	100%	+1.0	0.57	Enhanced
E-LSTM	99%	100%	+1.0	0.14	Enhanced
E-TCN	90%	100%	+1.0	1.32	Enhanced
E-TCN	95%	100%	+1.0	0.43	Enhanced
E-TCN	99%	100%	+1.0	0.14	Enhanced

Table 6.9 Summary of coverage percentage, average lead time and false alarms per month using Channel 03 dataset

Model	Threshold percentile	Coverage (%)	Avg lead time (days)	False alarms/ month	Type
LSTM	90%	100%	+4.0	3.52	Baseline
LSTM	95%	100%	+7.5	1.76	Baseline
LSTM	99%	100%	+76.0†	0.28	Baseline
TCN	90%	100%	-4.0	5.93	Baseline
TCN	95%	100%	-4.0	4.72	Baseline
TCN	99%	100%	-4.0	1.85	Baseline
E-LSTM	90%	100%	+1.0	3.06	Enhanced
E-LSTM	95%	100%	+1.0	2.13	Enhanced
E-LSTM	99%	100%	+1.0	0.93	Enhanced
E-TCN	90%	100%	+1.0	2.69	Enhanced
E-TCN	95%	100%	+1.0	1.39	Enhanced
E-TCN	99%	100%	+1.0	0.37	Enhanced

Table 6.10 Summary of coverage percentage, average lead time and false alarms per month using Channel 04 dataset

Model	Threshold percentile	Coverage (%)	Avg lead time (days)	False alarms/ month	Type
LSTM	90%	100%	-4.0	3.87	Baseline
LSTM	95%	100%	-4.0	2.94	Baseline
LSTM	99%	100%	-3.0	0.76	Baseline
TCN	90%	100%	-4.0	13.70	Baseline
TCN	95%	100%	-4.0	11.60	Baseline
TCN	99%	100%	-3.0	1.26	Baseline
E-LSTM	90%	100%	+1.0	3.61	Enhanced
E-LSTM	95%	100%	+1.0	2.61	Enhanced
E-LSTM	99%	100%	+1.0	0.76	Enhanced
E-TCN	90%	100%	+1.0	4.54	Enhanced
E-TCN	95%	100%	+1.0	1.85	Enhanced
E-TCN	99%	100%	+1.0	1.01	Enhanced

6.10.3 Analysis of the practical metrics

The key finding across all datasets is the clear difference in lead time behavior between baseline and enhanced models. On Channels 02 and 04, baseline models show negative lead times, meaning their residual-based detections occur only after the failure has already started, an inherent limitation of reactive, error-driven detection.

E-LSTM and E-TCN, by contrast, consistently achieve a positive 1-day lead time across all thresholds and datasets. Their performance is stable because CPD-CAD indicators reliably spike before each ILOS event, independent of threshold tuning. Enhanced models also maintain full coverage even when baselines fail. For example, baseline models lose all coverage on Channel 02 at the 99th percentile, while enhanced models retain 100%. On Channel 04, baseline models achieve 100% coverage but with negative lead times, confirming that detection occurs after failure occurrence rather than before it. The enhanced models again achieve 100% coverage with positive lead times.

False alarm rates reveal a further operational advantage of the enhanced framework. Baseline TCN on Channel 04 produces 13.70 false alarms per month at the 90th percentile threshold, roughly one false alarm every two days, making it operationally unusable at that sensitivity level. E-TCN on Channel 04 reduces this to 4.54 false alarms per month at the same threshold. At the 99th percentile, E-TCN further reduces the false alarm rate to 1.01 per month while maintaining 100% coverage and a 1-day lead time, a configuration that has no baseline equivalent: the 99th percentile baseline configurations either achieve 0% coverage (Channel 02) or negative lead times (Channels 03 and 04)

Two unusual baseline results require special attention. The Channel 03 baseline LSTM 76-day lead time at the 99th percentile is a sequence with high residual from an unrelated anomaly 76 days prior to the ILOS event, which survives the strict threshold when most other detections are suppressed. The Channel 03 baseline TCN produces negative lead times at all thresholds despite 100% coverage, which is consistent with its poor forecasting performance on this dataset, where the model underperforms consistently rather than detecting specific degradation before QoT violation.

6.10.4 Progressive early-warning structure

The 1-day lead time reported in the summary tables comes from the CAD detections since the framework takes the earlier of the CPD and CAD indicators for each event, and CAD consistently triggers closest to the QoT violation. Figure 6.5 (a) and (b) shows that CPD operates on a longer timescale: for example, the change point detected on December 24, 2020 provides a 28-day warning before the January 21, 2021 ILOS event, which is useful for planned actions such as maintenance or traffic migration. The collective anomaly detected 1 day before failure serves as a confirmation that the degradation is progressing toward a service-impacting event and triggers immediate protective actions like emergency rerouting or parameter adjustments.

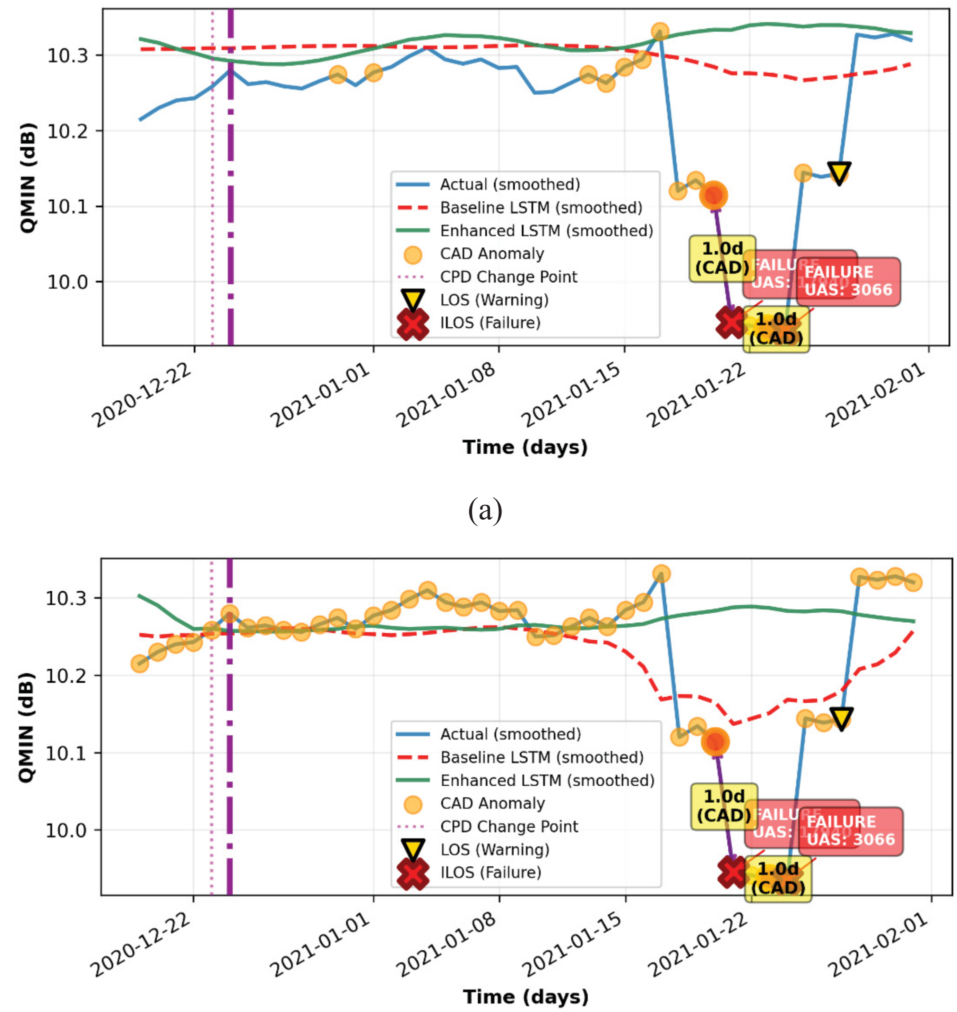


Figure 6.5 Pseudo-lead-times analysis of the CPD-CAD-aware forecasting framework using Channel 02 with (a) 7-day and (b) 3-day input sequences. Change points are detected several days before the ILOS events on January 21 and January 24, 2021, while collective anomalies are detected closer to failures, providing progressive early-warning indications

Together, the CPD and CAD indicators can form a progressive two-stage early-warning structure. The 100% coverage and stable positive lead times achieved by E-LSTM and E-TCN across all thresholds and datasets demonstrate that this structure is robust across network conditions and cannot be matched by residual-based detection alone.

6.11 Discussion and conclusion

6.11.1 Interpretation of key findings

Across three production network datasets, two neural architectures, multiple CAD methods, two sequence lengths, and an operational evaluation, four main conclusions emerge about the robustness and generalizability of the CPD-CAD-aware forecasting framework.

First, CPD-CAD integration consistently improves forecasting accuracy for both LSTM and TCN models across all datasets and window lengths. Averaged RMSE and MAE reductions of 10-25% with R^2 improvements up to 0.15 over the baseline VARMAX, and of approximately 2-20% over the baseline LSTM, show that regime shift and collective anomaly context provide insights not captured by raw OPM features alone. The consistency of these improvements across diverse datasets indicates that the approach is broadly applicable, not specific to a configuration.

Second, LSTM is the stronger forecasting architecture. TCN underperforms even the linear VARMAX baseline, likely due to its fixed convolutional receptive field being poorly suited to non-stationary OPM data. This limitation appears both in accuracy and in operational behavior: baseline TCN produces high false alarm rates (13.70 per month on Channel 04 at the 90th percentile threshold). CPD-CAD integration improves TCN performance substantially, but E-LSTM remains superior.

Third, sequence length has little effect on early-warning performance. Coverage, lead times, and false alarm rates are nearly identical for 3-day and 7-day windows. Sequence length can therefore be chosen based on accuracy and computational needs: 3-day windows improve DGD prediction, while 7-day windows offer slightly more stable aggregate performance.

Fourth, the early-warning behavior of enhanced models differs qualitatively from baselines. Residual-based detection in baseline models yields negative lead times on Channels 02 and 04, meaning that detections occur after failures occur. This reflects the reactive nature of residual detections. In contrast, E-LSTM and E-TCN achieve a consistent positive 1-day lead time and 100% coverage across all thresholds and datasets, with manageable false alarm rates. Their threshold-robustness comes from relying on CPD-CAD indicators rather than residual magnitude alone.

6.11.2 Contribution to QoT forecasting for failure prediction

This chapter extends the work introduced in Chapter 5 by demonstrating that CPD-CAD-aware forecasting generalizes beyond a single dataset. The framework is validated on three additional lightpath datasets from a production network, confirming that its benefits come from the methodology itself rather than from any one network configuration.

The two-stage warning structure is supported by both direct feature inspection and operational coverage results. On Channel 02, CPD provides approximately a 28-day warning and CAD a 1-day warning before failure. The coverage analysis shows that this pattern holds across all datasets and thresholds: enhanced models consistently achieve 100% failure coverage with positive lead times, while residual-only baseline models do not. This demonstrates that the early-warning capability is an inherent property of the CPD-CAD integration, not a consequence of favorable threshold selection.

The two detection periods align with operational practice: CPD enables planned interventions weeks in advance, while CAD supports immediate protective actions. Together they form a dual-timescale early warning structure for different intervention strategies. Both CPD and CAD operate without labeled failure data, which is consistent with the operational reality that labeled failure data are rarely available in production optical networks. This makes the framework practical to deploy in production networks.

6.11.3 Limitations

Some limitations of this extended study should be noted. All four lightpath datasets come from the same production network, limiting diversity in vendor equipment, and operating environments. Validation across multiple networks would strengthen claims of generalizability, but such datasets are rarely accessible in the public domain.

Because no ground truth labels exist for change points, CPD performance cannot be independently evaluated. The reported 28-day lead time is computed relative to the ILOS timestamp, not the true onset of the regime shift, which may occur earlier. Resolving this would require expert annotation or synthetic datasets with known change points.

The operational coverage and lead time analysis in Section 6.10 covers three datasets, two architectures, and two sequence lengths, but does not explore alternative threshold-calibration strategies or different false alarm cost assumptions that may apply to different network operators. The evaluation framework itself uses a fixed 3-day horizon for false alarm attribution, which may not reflect the operational realities of all deployment contexts. A comprehensive operational validation, including sensitivity to these design choices, is identified as a direction for future work.

Finally, all experiments use daily OPM data. The framework's performance at finer monitoring intervals, where QoT degradations may be captured with higher temporal resolution, remains to be evaluated.

6.11.4 Chapter summary

This chapter extends the CPD-CAD-aware QoT forecasting framework from Chapter 5 along five axes: validation on three additional lightpath datasets from a production network (Channels 02-04), comparison of LSTM and TCN architectures, systematic CAD model selection across five unsupervised techniques, analysis of 3-day and 7-day sequence lengths, and an analysis of QoT violations coverage, lead time, and false alarm rates across threshold sensitivity levels.

The main findings are clear. CPD-CAD integration improves forecasting across all datasets and sequence lengths, reducing the averaged RMSE and MAE by 10-25% and increasing R^2 by up to 0.15 over the baseline VARMAX, and by approximately 2-20% over the baseline LSTM. LSTM consistently outperforms TCN, whose baseline version underperforms VARMAX and produces high false alarm rates; CPD-CAD improves TCN substantially but does not close the gap. DGD is the only OPM feature with consistently positive R^2, benefiting most from the 3-day window. Operationally, enhanced models achieve 100% ILOS coverage with a stable positive 1-day lead time across all thresholds and datasets, while baseline models detect QoT violations reactively with negative lead times. The two-stage early-warning structure, with CPD at 28 days and CAD at 1 day before QoT violation, remains consistent across environments and sensitivity settings. All improvements are achieved without labeled failure data.

Overall, the framework moves from feasibility result using a single dataset to a validated, generalizable approach for proactive QoT management, demonstrating operational value of the CPD-CAD-aware forecasting framework beyond conventional forecasting accuracy metrics.

CHAPTER 7

DISCUSSION

This chapter brings together the contributions of the three published articles and Chapter 6 and shows how they collectively address three ONFM use cases for QoT management in smart optical networks. The discussion concentrates on the insights across tasks, methodological implications, and practical considerations relevant to real-world deployment.

7.1 Summary of key findings

This section synthesizes the key findings from each article and from the systematic validation study, establishing the practical basis for the cross-task discussion that follows.

7.1.1 Article 1: Automated, Interpretable and Efficient ML Models for Real-World Lightpaths' Quality of Transmission Estimation

Publication: S. Aladin, L. Wosinska, and C. Tremblay, "Automated, Interpretable and Efficient ML Models for Real-World Lightpaths' Quality of Transmission Estimation," *IEEE Open Journal of the Communications Society*, vol. 6, pp. 9785-9801, 2025.

Research question addressed: How can quality of transmission be accurately estimated before lightpath establishment by leveraging large-scale synthetic training data through transfer-learning to address data scarcity in new deployments, while ensuring operator trust through interpretable, XAI-based feature selection? What ML techniques can best balance accuracy with computational efficiency for QoT estimation?

Key findings: Article 1 examines QoT estimation for lightpath feasibility assessment under domain adaptation context. The study shows that lightweight RF models with Boruta-SHAP feature selection outperform ANN-based baselines while maintaining high classification accuracy across four QoT datasets. Feature selection significantly reduces dimensionality and computation time, with training time reduction by 52.78%-70.68% and inference time reduction to 1.1-2.4 μs per lightpath without sacrificing classification accuracy. The integration of the feature selection to domain adaptation further enhances cross-domain performance: using Dataset 03 as source domain and Dataset 02 as target domain, the proposed framework improves classification accuracy from 49.93% under the SDB to 99.54% with domain adaptation (a relative improvement of 99.35%) demonstrating that Boruta-SHAP feature selection allows effective distribution alignment between domains. Moreover, the framework achieves 86% accuracy when only 50 target domain samples are used. This represents a 99.83% reduction in required target samples relative to the full target domain training, enabling deployment in new networks with minimal lightpath data.

These findings show that accuracy, computational efficiency, and interpretability can be achieved simultaneously within a single QoT estimation framework. SHAP-based feature importance identifies modulation format, number of spans, path length and spectral occupation of the links the lightpath traverses as the most important features supporting model interpretability and operators' trust.

Annex I presents a sensitivity analysis evaluating how deviations between GN-model estimation and actual QoT values would influence classification performance and SHAP-based feature importance. The results show that both accuracy and features importance remain stable under BER perturbations consistent with realistic GN model errors.

7.1.2 Article 2: Detecting Anomalies in the Optical Layer Using Unsupervised Machine Learning

Publication: Aladin, S., Wosinska, L., & Tremblay, C. "Detecting anomalies in the optical layer using unsupervised machine learning". In *2024 Optical Fiber Communications Conference and Exhibition (OFC)* (pp. 1-3). Th3I.4

Research question addressed: How can multivariate OPM data and different unsupervised learning paradigms (density-based, boundary-based, isolation-based) be leveraged to detect anomalies across correlated parameters in optical networks, without requiring labeled training samples?

Key findings: Article 2 addresses multivariate anomaly detection using unsupervised learning applied to production OPM data, relying solely on unlabeled field data and alarm-based evaluation rather than supervised fault labels. The study investigates three unsupervised methods applied to four years of production OPM data. Among the tested methods, DBSCAN delivered the strongest clustering performance, achieving a Silhouette score of 0.82, compared with 0.69 for OCSVM and 0.68 for IF. In alarm-based evaluation, DBSCAN reached an F1-Score of 0.27, outperforming OCSVM (0.24) but not IF (0.30), and successfully detected all three imminent loss-of-signal (ILOS) events within the observed degradation window. Training times remained operationally practical across all three methods, with 0.05 s for DBSCAN, 0.02 s for OCSVM, and 0.37 s for IF, demonstrating that multivariate unsupervised detection can be deployed in an efficient way. The findings highlight that unsupervised learning enables soft-failure detection without the need for expensive labeling process, and that no single method dominates across all evaluation criteria.

7.1.3 Article 3: Change Point and Anomaly-Aware QoT Forecasting Framework for Proactive Network Management

Publication: Aladin, S., Wosinska, L., & Tremblay, C. “Change Point and Anomaly-Aware QoT Forecasting Framework for Proactive Network Management” in *IEEE Networking Letters*, vol. 8, pp. 224-228, 2026, doi: 10.1109/LNET.2026.3691249

Research question addressed: How can change point and collective anomaly detection bring context to multivariate seq2seq QoT forecasting models to enhance prediction performance and provide interpretable early warning signals with sufficient reaction windows for proactive network management?

Key findings: Article 3 addresses QoT forecasting for failure prediction by incorporating CPD and CAD scores into multivariate seq2seq LSTM forecasting models. Under identical hyperparameters, the CPD-CAD-enhanced LSTM consistently outperforms the baseline LSTM, achieving up to a 28.21% reduction in RMSE, a 42.36% reduction in MAE, and an increase in QMIN R^2 from 0.477 to 0.731. These improvements demonstrate that incorporating change point and collective anomaly awareness improves forecasting robustness under non-stationary conditions.

Beyond accuracy improvements, the framework computes actionable reaction windows. Across three failure events, reaction windows were strictly positive, ranging from 1 to 10 days, with a median of 1 day and a mean of 4 days, showing that forecasting performance translates into practical intervention time. This capability links physical-layer degradation forecasting with operational decision support, enabling proactive actions such as traffic rerouting, parameter adjustments, or preventive maintenance before service-impacting QoT violations occur.

Annex II reports 95% bootstrap confidence intervals for the RMSE, MAE, and R^2 improvements, confirming that all three gains are statistically significant across 10,000 resamples.

7.1.4 Chapter 6: Systematic validation and extension of the CPD-CAD-aware forecasting framework

Research question addressed: Does the enhanced forecasting framework generalize across independent field lightpath datasets, alternative forecasting architectures, and input sequence lengths, and does it provide operationally meaningful proactive failure detection?

Key findings: Chapter 6 extends the study presented in Article 3 by evaluating three additional field lightpath datasets (Channels 02, 03, and 04), two DL architectures (LSTM and TCN), five unsupervised collective anomaly detection methods with systematic model selection, and two input sequence lengths (3-day and 7-day windows). Before model development, a statistical characterization of the OPM features is conducted using augmented Dickey-Fuller stationarity tests and autocorrelation function analysis. These analyses confirm that the OPM time series exhibit regime shifts consistent with the change points identified by the CPD module. Autocorrelation patterns further show strong short-lag temporal dependencies in DGD, aligning with its consistently positive R^2 values in forecasting results, whereas QMIN and OPR display weaker autocorrelation structures, accounting for their event-driven forecasting difficulty. This statistical evidence supports the architectural choices of the framework: the CPD-CAD module explicitly addresses regime shifts and non-stationarity, the LSTM technique is better suited than the TCN technique for handling regime shifts in OPM signals, and the 3-day window enhances DGD forecasting by limiting the influence of pre-shift data.

The comprehensive forecasting evaluation demonstrates that CPD-CAD integration produces consistent performance improvements across all configurations, yielding averaged RMSE and MAE reductions of approximately 10-25% and increases in R^2 of up to 0.15 for the E-LSTM model over the baseline VARMAX, and 2-20% over the baseline LSTM. Across all configurations, LSTM outperforms TCN in all evaluated configurations; TCN baselines perform worse than even the VARMAX model, due to the interaction between their fixed convolutional receptive fields and the non-stationarity of the lightpath datasets. Among the three OPM features, DGD is the only one to achieve consistently positive R^2 values, benefiting most from the 3-day sequence, with R^2 reaching 0.749 for the enhanced LSTM on Channel 04. Input sequence length did not affect the early warnings: CPD detects regime shifts approximately 28 days before ILOS events, and CAD detects collective anomalies approximately one day before QoT violation under both the 3-day and 7-day configurations.

A preliminary operational assessment of failure coverage, lead time, and false alarm rates reveals a qualitative distinction between the enhanced and baseline models that is not captured by forecasting accuracy metrics alone. The enhanced models (E-LSTM and E-TCN) achieve 100% failure coverage and maintain a consistent positive lead time of 1 day across all threshold sensitivity levels and across all three datasets. In contrast, the baseline models produce negative average lead times in most settings, indicating that their residual-based detections are triggered only after the ILOS event had already begun and loose all coverage under the strictest threshold on Channel 02. These findings confirm that the CPD-CAD integration enables proactive, rather than reactive, detection and that this capability is robust across diverse datasets and detection sensitivity configurations.

7.2 Integration across tasks based on ONFM use cases

Collectively, the three articles and the additional chapter, share the QoT as the central performance objective throughout the tasks: QoT is estimated before lightpath deployment in Chapter 3 (Article 1), monitored through multivariate OPM data for failure detection in Chapter 4 (Article 2), and forecast to anticipate future violations for failure prediction in Chapter 5 (Article 3) and Chapter 6. This shared objective provides conceptual coherence across contributions despite difference in data types and modeling approaches.

Method selection reflects the data characteristics and operational constraints of each ONFM-based task: supervised learning with domain adaptation for synthetic lightpath data, unsupervised learning for unlabeled field OPM data, and seq2seq forecasting enhanced with change point and collective anomaly indicators for non-stationary behavior. This task-based approach shows that model selection is driven by data characteristics and operational requirements rather than algorithmic preference.

The three contributions are connected at two levels. At the technical level, Contributions 2 and 3, the collective anomalies detected for failure detection provide the CAD indicators which, combined with CPD indicators, are incorporated into the forecasting framework in Chapter 5 (Article 3). This leads to improvements in both prediction accuracy and reaction window reliability. Chapter 6 reinforces the generalizability of the QoT forecasting framework by showing that the benefits of CPD-CAD integration and the progressive early-warning pattern are consistent across three additional, independent production network datasets. At the conceptual level, Chapter 3 relates to Chapters 4 and 5 through the shared QoT as the performance objective, where effective provisioning reduces the risk of deploying poorly configured lightpaths that could generate anomalies or degrade faster than expected.

The difference in data sources across contributions reflects the operational reality that distinct ONFM-based tasks arise at different stages of a lightpath's lifetime and thus require different types of dataset. QoT estimation relies on synthetic lightpath data because field lightpath data are unavailable prior to deployment. Anomaly detection and QoT forecasting use field OPM time series as failure detection and prediction depend on observed degradations in production networks. This difference reflects a direct consequence of what each task requires and it is consistent with Musumeci and Tornatore (2025) characterization of each ONFM use case as having its own operational context and objectives.

7.3 Broader implications

The results show that lightweight models can perform as well as, or even better than, more complex ones when they are designed with operational constraints of the specific ONFM-based task in mind. In QoT estimation, RF models with Boruta-SHAP feature selection achieve performance comparable to, and in some cases better than, DNNs, while reducing inference time to microsecond levels. This finding challenges the common assumption, that increasing model complexity is the primary way to improve the QoT estimation.

Integrating CPD and CAD awareness enhances forecasting robustness under non-stationary conditions. The statistical analysis in Chapter 6 clarifies this effect: CPD integration addresses the regime shifts present in the OPM signals, and the LSTM architecture handles such non-stationarity more effectively than TCN. Moreover, computing reaction windows transforms forecasting models into actionable decision-support mechanisms for network operators.

The preliminary operational assessment in Chapter 6 adds an insight that extends beyond accuracy metrics. The consistent finding that residual-based detection is inherently reactive, typically triggering only after failure start, while CPD-CAD-aware forecasting remains consistently proactive across all datasets and threshold sensitivity levels demonstrate that the CAD model selection is as important as the forecasting architecture selection. A framework that reduces RMSE but fails to produce a positive lead time does not meaningfully advance proactive failure management.

Taken together, these findings suggest that operational impact in ML-based failure management depends less on model complexity and more on how well the selected approach aligns with the data characteristics and operational requirements of each ONFM-based task. When contributions address distinct tasks using methods matching their specific requirements, while sharing a common performance objective, insights can transfer across tasks. This is illustrated by the CAD integration between anomaly detection and QoT forecasting tasks CAD and CPD indicators enable more effective QoT forecasting.

7.4 Limitations and methodological constraints

While the results demonstrate the effectiveness of the proposed contributions, several limitations must be acknowledged.

First, the QoT estimation contribution relies on synthetic datasets generated from physical-layer simulation environments. Although domain adaptation significantly mitigates distribution shift, full validation against large-scale field lightpath datasets remains limited. The generalization performance under diverse vendor equipment configurations and heterogeneous fiber infrastructures therefore needs further investigation.

Second, anomaly detection performance for failure detection is evaluated using alarm indicators such as UAS and HCCS as weak supervision. These alarms do not provide precise ground-truth degradation onset timestamps and may reflect other events not fully captured by the selected OPM metrics. Consequently, evaluation metrics may underestimate or overestimate early detection capability in certain cases.

Third, although the systematic validation in Chapter 6 substantially extends the foundation of the CPD-CAD-aware forecasting framework, demonstrating consistent improvements across three additional field lightpath datasets, two forecasting architectures, five CAD techniques, and two sequence lengths, all datasets originate from a single network operator. This limits diversity of the generalizability assessment with respect to topology, vendor environments, geography, and operational practice. Broader multi-operator validation would strengthen claims of generalizability, though such datasets remain difficult to obtain.

Fourth, the operational evaluation in Chapter 6 relies on design parameters that are not optimized, including the 3-day sequence used for the assessment and the chosen threshold percentiles. A comprehensive operational validation with a sensitivity analysis of these parameters is still required.

Fifth, the proposed contributions do not include closed-loop control or automated intervention mechanisms. Reaction windows are computed as decision-support signals rather than directly triggering network reconfiguration actions. The operational impact of integrating the framework within live control-plane systems therefore remains to be validated.

Finally, while computational feasibility is explicitly addressed for QoT estimation and anomaly detection, large-scale deployment analysis of the forecasting framework across hundreds or thousands of lightpaths was not conducted.

These limitations do not undermine the contributions of the thesis but instead define clear boundaries for interpretation and highlight opportunities for further empirical validation and methodological development.

The three articles and the systematic validation chapter in relation to the research questions, considered together, demonstrate that promising results come from matching ML paradigms with the data context and operational constraints of each ONFM-based task. Also, robust and operationally meaningful solutions are achieved from integrating anomaly detection into QoT forecasting when the data context allows it. Building on this synthesis, the final chapter presents the overall conclusions of the thesis, highlights the main contributions, and outlines future research directions.

CONCLUSION

This thesis addresses a key challenge of applying ML to optical network management in a way that is operationally relevant and methodologically coherent across three ONFM-based QoT management tasks. Rather than considering QoT estimation for lightpath feasibility assessment, anomaly detection for failure detection, and QoT forecasting for failure prediction as separate problems, the work adopts an approach based on the ONFM-based tasks in which each contribution applies an ML paradigm matching the task's data context and operational constraints, while maintaining QoT as a central performance objective.

The findings show that effective deployment of ML in optical networks requires more than predictive accuracy. Robust generalization across network configurations, interpretability of model behavior, and the ability to compute actionable intervention time are equally essential components of practical and operationally meaningful QoT management solutions. Chapter 3 (Article 1) shows that accuracy, cross-domain generalization, interpretability, and computational efficiency are not competing objectives and can be achieved simultaneously within a QoT estimation framework. Likewise, Chapters 4 and 5 (Articles 2 and 3) demonstrate that using task-appropriate methods, with unsupervised learning for anomaly detection and sequence modeling for QoT forecasting, provide both robust detection performance and practical operational value.

8.1 Main contributions

The thesis makes three principal contributions, each addressing a task linked to distinct ONFM use case.

First, it introduces an interpretable domain adaptation framework for QoT estimation at lightpath provisioning, combining Boruta-SHAP automated feature selection with domain adaptation and RF-based classification. The framework demonstrates that accuracy, computational efficiency, and explainability can be pursued simultaneously rather than treated as competing objectives, and that domain adaptation with minimal target domain data enables deployment across network configurations. A preliminary sensitivity analysis demonstrating the robustness of these results to realistic GN model deviations is presented in Annex I.

Second, it provides a systematic assessment of unsupervised multivariate anomaly detection methods applied to four years of lightpath data from a production network. The results show that soft-degradation patterns are more effectively captured through multivariate modeling than through isolated threshold-based techniques and establish the anomaly detection foundation used in the third contribution.

Third, it proposes and systematically validates a CPD-CAD-aware forecasting framework that incorporates change point and collective anomaly indicators into seq2seq DL models. Chapter 5 (Article 3) demonstrates the feasibility of this approach on a single lightpath dataset, providing actionable reaction windows prior to QoT violations. Chapter 6 extends this contribution through systematic validation across three additional lightpath datasets, two forecasting architectures (LSTM and TCN), five collective anomaly detection techniques, and two input sequence lengths. A preliminary statistical analysis confirms the non-stationarity and autocorrelation structure that motivates the CPD-CAD integration, and a preliminary operational evaluation demonstrates that the framework achieves 100% failure coverage with consistently positive lead times across all configurations. Annex II provides 95% bootstrap confidence intervals for the forecasting performance metrics in Article 3, showing that the observed RMSE, MAE, and R^2 improvements are statistically robust across distinct resamples. Together, Chapter 5 (Article 3) and Chapter 6 advance the QoT forecasting contribution from a feasibility demonstration to a systematically validated approach to proactive failure prediction.

8.2 From reactive to proactive QoT management

A key finding of this research is showing that proactive QoT management requires more than improved forecasting models. It depends on the integration of detection and prediction methods, explicit modeling of QoT degradation, and the computation of available intervention time.

The progression achieved in this thesis, from lightpath feasibility assessment prior to deployment and QoT degradation detection during operation, to computation of reaction windows for proactive maintenance, represents a shift from reactive analytics toward proactive network intelligence. Each contribution extends the temporal aspect of decision-making: QoT estimation supports provisioning before a lightpath is established, anomaly detection identifies degradation already present in OPM data, and QoT forecasting anticipates degradations before it affects the service.

The systematic validation in Chapter 6 reinforces this progression: by demonstrating that CPD-CAD-aware forecasting is consistently proactive across diverse datasets. This shows that the proactive capability is a property of the CPD-CAD integration rather than results linked to a single dataset's characteristics. This finding indicates that the evolution toward autonomous optical network management relies not only on increasingly complex algorithms, but also on the structured alignment of ML methods, the data available for each task, and the operational requirements of the specific ONFM use case being addressed.

8.3 Future research directions

The limitations identified in Chapter 7 outline several directions for future research.

First, broader validation of QoT estimation through domain adaptation using heterogeneous field lightpath datasets would strengthen external generalization. Incorporating vendor diversity and varied outside plants would further test model robustness.

Second, large-scale, multi-network evaluation of the unsupervised anomaly detection and QoT forecasting frameworks would improve statistical confidence in anomaly detection coverage and reaction window stability across diverse degradation scenarios. The validation presented in Chapter 6, although based on three lightpath datasets from a single production network, highlights the need for cross-network evaluation as the next step toward establishing broader methodological generalizability.

Third, beyond evaluating the proposed methods with different optical-network scenarios, extending them to include data-center interconnects, metro networks, and multi-domain optical infrastructures, would assess their generality beyond the deployment context studied in the thesis. These scenarios differ in reach regimes, traffic dynamics, and operational practices. This would help determine how well the observed performance transfers across heterogeneous network segments.

Fourth, a comprehensive operational validation of the CPD-CAD-aware forecasting framework is still required. The evaluation in Chapter 6 demonstrates its proactive detection advantage but relies on design parameters, such as the sequence length or the three percentile levels used, that are not optimized. Future work should include sensitivity analyses of these parameters and evaluation at sub-daily OPM data granularity, which may enhance both forecasting accuracy and anomaly detection performance.

Fifth, integrating the CPD-CAD-aware forecasting framework into closed-loop network control systems represents an essential next step. Combining reaction windows with automated reconfiguration strategies, such as traffic rerouting, modulation format adaptation, or reinforcement-learning-based policy optimization, would shift the framework from decision support toward autonomous operation.

Sixth, extending the contributions to incorporate cross-layer information, including traffic patterns, routing dynamics, and control plane telemetry, could enrich degradation context and enhance forecasting robustness under complex network conditions.

Seventh, the methods developed in this thesis an end-to-end ML formulation, in which models are trained directly on monitoring data without relying on analytical physical layer models. Another future research direction includes hybrid input-refinement approaches, where analytical models provide estimates that are subsequently refined through ML-based calibration. This combination offers physical consistency of analytical models while allowing data-driven techniques to handle missing or uncertain parameters, making them particularly relevant in deployments where link, system and channel information is not consistently available.

Finally, scalable deployment studies assessing computational overhead across large number of lightpaths would further demonstrate feasibility in real network environments.

Collectively, these directions point toward the long-term goal of resilient, self-optimizing optical networks capable of anticipating and mitigating degradation before service disruption occurs.

8.4 Final reflection

Although ML has become increasingly prominent in optical network research, its application often remains limited to isolated tasks. This thesis shows that greater impact arises when QoT estimation, anomaly detection, and QoT forecasting are organized as complementary ONFM-based tasks guided by consistent performance objectives and structured task-appropriate methodological choices.

By maintaining QoT as a common metric, selecting ML paradigms matching the task-appropriate data context and operational requirements, and computing actionable reaction windows, the work contributes not only new models but also a coherent approach for proactive QoT management. The systematic validation across three independent lightpath datasets from a production network, demonstrates that the benefits of the CPD-CAD-aware forecasting framework arise from fundamental methodological properties rather than results linked to characteristics of any specific dataset.

As optical network infrastructures continue to grow in scale and complexity, aligning ML methods with the specific requirements of each QoT management tasks will be essential for delivering solutions that remain reliable across diverse configurations, transparent enough to support operator trust, and oriented toward proactive rather than reactive QoT management.

ANNEX I

SENSITIVITY ANALYSIS OF HOW DEVIATIONS BETWEEN GN MODEL ESTIMATION AND REAL QOT VALUES WOULD AFFECT THE MODELS PERFORMANCE AND FEATURE IMPORTANCE IN ARTICLE 1

The QoT estimation solution presented in the first article in Chapter 3 is trained and evaluated on synthetic datasets generated through the PLATON simulation framework, which relies on the GN model. Although the GN model has been extensively validated against simulations in the literature (Carena et al., 2014; Semrau et al., 2019a), it introduces simplifications that may cause deviations from real QoT measurements in production networks or from simulations. Specifically, the GN model neglects certain physical effects, including modulation-format-dependent nonlinear interference and inter-channel stimulated Raman scattering (ISRS), that can affect the accuracy of estimated QoT values.

This annex addresses the question raised by peer reviewers during the revision of Article 1: how sensitive are the trained RF classifiers and the SHAP-based feature importance rankings to realistic deviations between GN model estimations and real QoT values? The sensitivity analysis presented here assesses model robustness under BER perturbations applied to the training labels, providing evidence that the reported classification accuracy and feature importance results are stable under realistic analytical noise.

I.1 Experimental setup

I.1.1 Dataset

The sensitivity analysis uses a 12,000-instance class-balanced subsample of Dataset 01 from the Fraunhofer HHI QoT dataset collection. We keep a class-balanced sample to ensure equal representation of feasible and infeasible lightpath instances and avoids any bias in the accuracy and AUC estimates caused by class imbalance. Dataset 01 is used as it provides the largest sample from the available datasets and yields stable estimates across multiple experimental runs. The mean BER of the subsample is 0.004109.

I.1.2 Perturbation design

BER perturbations are applied to the label generation threshold to simulate deviations between GN model estimations and real QoT values. BER perturbations are expressed as absolute values, since the physical interpretation of absolute deviations is more directly meaningful for comparison with GN model error reported in the literature.

The perturbation ranges are chosen based on the values of GN model errors in optical settings provided by Carena et al. (2014), which reports typical NLI power estimation errors for practical transmission scenarios.

The relationship between BER and SNR deviations in this experiment correspond to PM-QPSK, which appears as the lowest-order modulation format present in the subsample. Using the PM-QPSK is a conservative choice because it has a steeper BER-vs-SNR slope than the higher-order formats at the same operating BER, so the same SNR deviation produces a larger BER perturbation than it would for PM-8QAM, PM-16QAM, PM-32QAM, and PM-64QAM. The BER perturbation levels, the corresponding approximated SNR deviations, and the GN model regimes are presented in Table-A I-1. The different values are based on SNR underestimation of 0.3 to 0.6 dB for standard single-mode fiber (SMF), pure silica core fiber (PSCF), and non-zero dispersion-shifted fiber (NZDSF). Additionally, the SNR deviation values go up to 0.8 dB for low dispersion fiber. All these values are reported in Carena et al. (2014).

Table-A I-1 BER perturbation levels and their corresponding GN model error regimes based on Carena et al. (2014)

Absolute BER deviation	**Approx. SNR deviation**	**GN model regime**
0.00073	0.2 dB	Within GN model accuracy range for SMF/PSCF/ NZDSF
0.0011	0.3 dB	Within GN model accuracy range for SMF/PSCF/ NZDSF
0.0022	0.6 dB	Within GN model accuracy range for SMF/PSCF/ NZDSF
0.0029	0.8 dB	Upper bound of realistic GN model error (low-dispersion fiber)
0.0047	1.3 dB	Stress test - beyond realistic GN model error

The 0.0029 threshold corresponds to the upper bound of what Carena et al. (2014) report as typical NLI estimation errors, specifically the 0.8 dB SNR underestimation reported for low-dispersion fibers such as LS fiber in the EGN model validation. The 0.0047 threshold is included as a stress test representing deviations of approximately 1.3 dB. This level is included to establish where the model's robustness begins to degrade.

I.1.3 Evaluation metrics and procedure

Two performance dimensions are evaluated:

Classification performance: mean accuracy and AUC are computed over 3 independent runs for each perturbation level. Computing the mean over multiple runs accounts for variability in the RF training process due to bootstrap sampling and feature selection.

Feature importance stability: the Spearman rank correlation coefficient (ρ) is computed between the SHAP-based feature rankings obtained under perturbed and unperturbed conditions. Spearman ρ is chosen over Pearson correlation because it measures rank-order agreement rather than linear correlation, which is the appropriate measure for evaluating whether the relative ordering of feature importances is preserved under perturbation. A Spearman ρ greater than 0.95 is considered indicative of high ranking stability.

I.2 Results

Table-A I-2 reports mean classification accuracy, AUC and SHAP rank stability at each perturbation level.

Table-A I-2 Classification performance and SHAP feature ranking stability under increasing BER perturbations.

BER perturbation (SNR deviation)	**Mean accuracy (%)**	**Mean AUC**	**Spearman ρ (SHAP)**
No perturbation (baseline)	98.55	0.9995	0.994
0.00073 (0.2 dB)	97.95	0.9995	0.993
0.0011 (0.3 dB)	97.5	0.9994	0.963
0.0022 (0.6 dB)	95.5	0.9925	0.951
0.0029 (0.8 dB)	93.9	0.988	0.931
0.0047 (1.3 dB)	89.9	0.965	0.994

Figures-A I-1 and I-2 illustrate the continuous degradations underlying these values. Figure-A I-1 covers the perturbation ranges from 0 to 0.0022 and Figure-A I-2 extends to the low-dispersion LS fiber and stress test with values ranging from 0 to 0.0047.

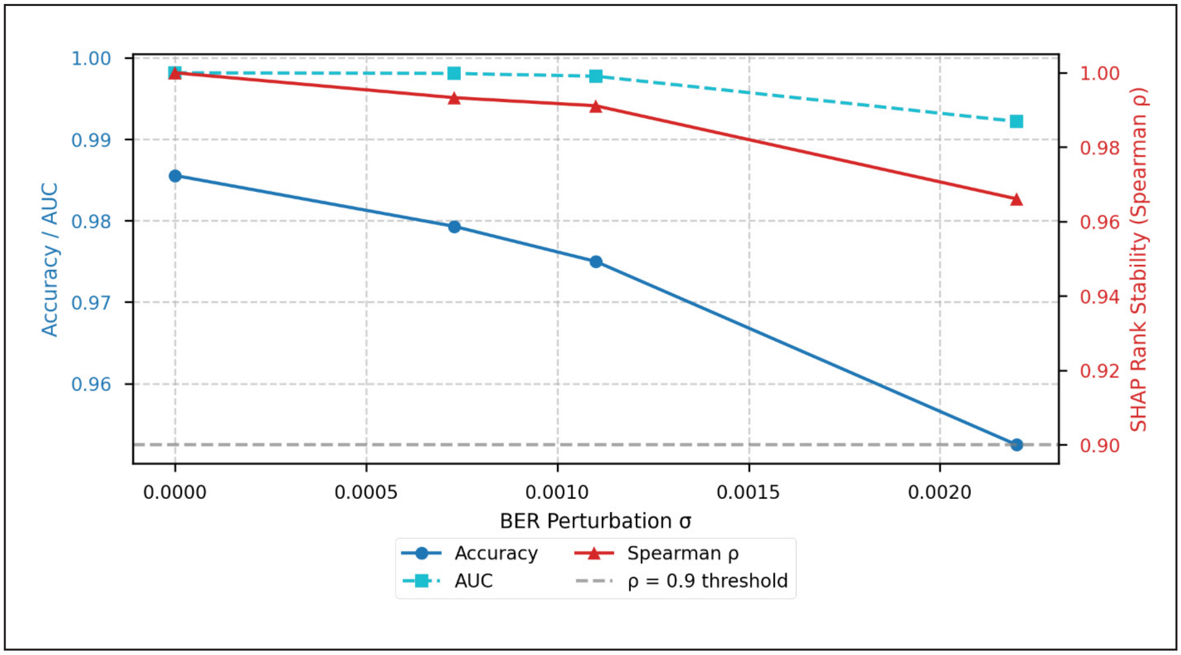


Figure-A I-1 Model and SHAP stability under low BER perturbation levels

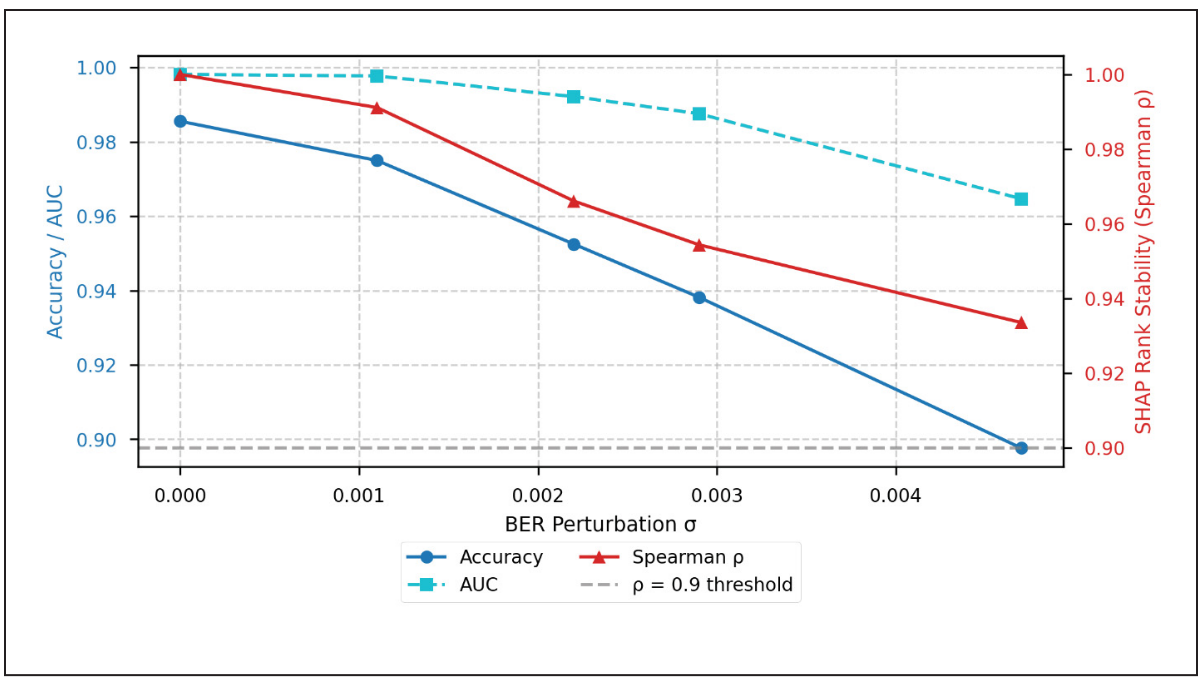


Figure-A I-2 Model and SHAP stability under extended range of BER perturbation levels

I.3 Interpretation

The results show that the RF-based QoT estimation model is robust to BER perturbations within the range of realistic GN model errors. For perturbations up to 0.0011 (approximately 0.3 dB SNR variation), which covers the typical accuracy range of the GN model as indicated by Carena et al. (2014), classification accuracy and AUC remain essentially unchanged, and SHAP feature rankings are highly stable (Spearman ρ greater than 0.99). This indicates that both the model's classification performance and its interpretability are robust to the level of analytical noise that would be expected in practice.

Performance degrades more when BER perturbations reach 0.0029 (approximately 0.8 dB SNR variation), where accuracy decreases by 4.7% and SHAP stability declines to $\rho = 0.951$. At the stress-test level of 0.0047 (approximately 1.3 dB SNR variation), accuracy reaches 89.9% and SHAP ranking stability declines to $\rho = 0.931$. This degradation establishes the boundary beyond which GN model inaccuracies would affect model performance, while confirming that the RF-based classifier operates reliably within realistic GN model error limits.

Figures-A I-1 and I-2 confirm that the reported classification results and feature importance rankings are robust to GN model typical error ranges. The AUC curve in Figure-A I-1 remains essentially flat for all the perturbation values, indicating that the classifier's ability to separate feasible from infeasible lightpaths is unaffected up to 0.6 dB SNR deviation. Figure-A I-2 extends the analysis to a 0.0047 BER perturbation, corresponding to a stress-test region beyond realistic GN model error. At 0.0029 (0.8 dB SNR deviation) accuracy remains at 93.9% and $\rho = 0.951$, still well above the 0.90 threshold. Taken together, Figures-A I-1 and I-2 show that the model's accuracy and interpretability are preserved within the GN model's known error ranges for all standard fiber types used in production optical networks.

These findings provide two practical implications. First, the classification results reported in Article 1 are not artifacts of the synthetic data generation process, they reflect real dependencies between lightpath configuration features and QoT outcomes that are preserved under realistic levels of analytical noise. Second, SHAP-based feature importances, particularly the top-ranked features (modulation format, number of spans, path length, and spectral occupation), remain consistent under perturbation, supporting the interpretability claims of the contribution.

ANNEX II

BOOTSTRAP CONFIDENCE INTERVALS FOR QOT FORECASTING PERFORMANCE METRICS REPORTED IN ARTICLE 3

The forecasting performance improvements for QMIN reported in Chapter 5 (Article 3), specifically the RMSE reduction of 28.21% and the MAE reduction of 42.36% achieved by the CPD-CAD-aware LSTM framework over the baseline LSTM, are computed on sequences in the test portion of the dataset. During the review process, the reviewers requested statistical significance measures or confidence intervals to confirm that these improvements are not due to the specific test samples and that the reported improvements are statistically meaningful.

This annex reports bootstrap 95% confidence intervals for the three primary forecasting performance metrics MAE, RMSE, and R^2, for both the baseline LSTM and the CPD-CAD-aware LSTM, along with confidence intervals for the improvement between them.

II.1 Bootstrap methodology

II.1.1 Method

Bootstrap confidence intervals are computed using the non-parametric bootstrap method. The bootstrap is appropriate here because the test set consists of multivariate OPM time-series sequences and no distributional assumption about the forecasting residuals is imposed. The method proceeds as follows: (1) from the available test sequences, $B = 10{,}000$ bootstrap resamples are drawn with replacement; (2) for each resample, the performance metric is computed for both the baseline and enhanced models; (3) the 2.5th and 97.5th percentiles of the resulting bootstrap distribution are used as the limits of the 95% confidence interval. The percentile method is used given the large number of resamples, which makes the percentile interval a reliable estimate of the sampling distribution.

II.1.2 Parameters

The choice of $B = 10{,}000$ resamples is standard practice for percentile bootstrap confidence intervals and provides stable interval estimates. It ensures that the 2.5th and 97.5th percentile limits are estimated with high precision. The resampling unit is the individual test sequence (multivariate feature vector) rather than individual time steps, which respects the temporal structure of the OPM data and avoids inflating the effective sample size by treating correlated observations as independent. The parameters are presented in Table-A II-1.

Table-A II-1 Bootstrap procedure parameters

Parameter	Value
Number of resamples (B)	10,000
Confidence level	95%
Interval method	Percentile bootstrap
Resampling unit	Test sequences (multivariate feature vectors)
Metrics evaluated	MAE, RMSE, R^2

II.2 Results

Table-A II-2 shows the three performance metrics MAE, RMSE, and R^2 for the baseline LSTM and the enhanced LSTM models wit the confidence intervals computed over 10,000 bootstrap resamples. The results are presented for the overall framework stability in a multivariate context. The improvements remain positive for all the metrics, showing supporting evidence of the robustness of the CPD-CAD-aware forecasting framework.

Table-A II-2 Bootstrap 95% confidence intervals for forecasting performance metrics

Metric	Baseline LSTM [95% CI]	Enhanced LSTM [95% CI]	Improvement [95% CI]
MAE	1.3295 [1.1105, 1.6009]	1.1076 [0.8888, 1.3787]	+0.2219 [+0.2103, +0.2346]
RMSE	9.9366 [5.1790, 13.9450]	9.8884 [5.0753, 13.9172]	+0.0482 [+0.0212, +0.1240]
R^2	0.3134 [0.1944, 0.6168]	0.3200 [0.1975, 0.6326]	+0.0066 [+0.0024, +0.0182]

II.3 Interpretation

The confidence interval for the MAE improvement is [+0.2103, +0.2346], meaning the entire interval is strictly above zero across all 10,000 bootstrap resamples. This confirms that the MAE reduction is statistically robust: the CPD-CAD-aware LSTM consistently achieves lower absolute forecasting error than the baseline LSTM, with the improvement representing a 16.7% reduction in MAE. The narrow width of this interval [0.0243 units] indicates low variability in the improvement estimation, supporting confidence in the reported gain.

The confidence interval for the RMSE improvement is [+0.0212, +0.1240]. This interval is also strictly positive, confirming that the RMSE improvement is statistically meaningful, though the wider interval reflects greater variability in the RMSE metric across test sequences. The RMSE is more sensitive to individual large errors than the MAE, which accounts for the higher dispersion in the bootstrap distribution. The improvement remains consistently positive across all resamples.

The confidence interval for the R^2 improvement is [+0.0024, +0.0182]. While the absolute magnitude of the R^2 improvement is small, the interval is strictly positive, confirming that the enhanced model explains a consistently greater proportion of variance in the QoT time-series than the baseline. The relatively modest R^2 values for both models (0.31 - 0.32) reflect the difficulty of forecasting non-stationary QoT signals, as discussed in Chapter 5 and confirmed by the validation study in Chapter 6.

The bootstrap confidence intervals confirm that all three performance improvements reported in Article 3 are statistically meaningful. The improvement lower bounds are strictly positive in all cases, and the results hold consistently across 10,000 independent resamples. The MAE improvement is the most robust and largest in magnitude. RMSE and R^2 improvements are smaller but consistently positive. These results provide statistical validation that the CPD-CAD integration produces promising forecasting improvements rather than results obtained for a specific dataset.

LIST OF BIBLIOGRAPHICAL REFERENCES

Abdelli, K., Cho, J. Y., Azendorf, F., Griesser, H., Tropschug, C., & Pachnicke, S. (2022a). Machine-learning-based anomaly detection in optical fiber monitoring. *Journal of Optical Communications and Networking*, *14*(5), 365-375. https://doi.org/10.1364/JOCN.451289

Abdelli, K., Grießer, H., & Pachnicke, S. (2022b). A machine learning-based framework for predictive maintenance of semiconductor laser for optical communication. *Journal of Lightwave Technology*, *40*(14), 4698-4708. https://doi.org/10.1109/JLT.2022.3163579

Abdelli, K., Lonardi, M., Gripp, J., Olsson, S., Boitier, F., & Layec, P. (2024a). Unsupervised anomaly detection and localization with generative adversarial networks. In *ECOC 2024: 50th European Conference on Optical Communication* (pp. 998-1001).

Abdelli, K., Lonardi, M., Gripp, J., Correa, D., Olsson, S., Boitier, F., & Layec, P. (2024b). Anomaly detection and localization in optical networks using vision transformer and SOP monitoring. In *2024 Optical Fiber Communications Conference and Exhibition (OFC)* (pp. 1-3), Tu2J.4.

Agrawal, G. P. (2010). *Fiber-optic communication systems* (4th ed.). John Wiley & Sons

Ahmed, M., Mahmood, A. N., & Hu, J. (2016). A survey of network anomaly detection techniques. *Journal of Network and Computer Applications*, *60*, 19-31. https://doi.org/10.1016/j.jnca.2015.11.016

Aladin, S., & Tremblay, C. (2018). Cognitive tool for estimating the QoT of new lightpaths. In *Optical Fiber Communications Conference and Exposition (OFC)* (pp. 1-3). Optica Publishing Group, M3A.3.

Aladin, S., Tran, A. V. S., Allogba, S., & Tremblay, C. (2020a). Quality of transmission estimation and short-term performance forecast of lightpaths. *Journal of Lightwave Technology*, *38*(10), 2807-2814. https://doi.org/10.1109/JLT.2020.2975179

Aladin, S., Allogba, S., Tran, A. V. S., & Tremblay, C. (2020b). Recurrent neural networks for short-term forecast of lightpath performance. In *Optical Fiber Communication Conference (OFC 2020)*, paper W2A.24. Optica Publishing Group. https://doi.org/10.1364/OFC.2020.W2A.24

Aladin, S., Wosinska, L., & Tremblay, C. (2024). Detecting anomalies in the optical layer using unsupervised machine learning. In *2024 Optical Fiber Communications Conference and Exhibition (OFC)* (pp. 1-3). Optica Publishing Group

Aladin, S., Wosinska, L., & Tremblay, C. (2025). Automated, interpretable and efficient ML models for real-world lightpaths' quality of transmission estimation. *IEEE Open Journal of the Communications Society*, 6, 9785-9801. https://doi.org/10.1109/OJCOMS.2025.3635533

Aladin, S., Wosinska, L., & Tremblay, C. (2026). Change Point and Anomaly-Aware QoT Forecasting Framework for Proactive Network Management. *IEEE Networking Letters*, vol. 8, 224-228, 2026, doi: 10.1109/LNET.2026.3691249

Ali, Y. M., Prilepsky, J. E., Sambo, N., Pedro, J., Hosseini, M. M., Napoli, A., Turitsyn, S. K., & Freire, P. (2026). From data to decision: A multi-stage framework for class imbalance mitigation in optical network failure analysis. *Journal of Optical Communications and Networking*, *18*(1), 42–58. https://doi.org/10.1364/JOCN.576774

Allogba, S., Aladin, S., & Tremblay, C. (2021). Multivariate machine learning models for short-term forecast of lightpath performance. *Journal of Lightwave Technology*, *39*(22), 7146-7158. https://doi.org/10.1109/JLT.2021.3110513

Allogba, S., Aladin, S., & Tremblay, C. (2022a). Machine-learning-based lightpath QoT estimation and forecasting. *Journal of Lightwave Technology*, *40*(10), 3115-3127. https://doi.org/10.1109/JLT.2022.3160379

Allogba, S., Yaméogo, B. L. M., & Tremblay, C. (2022b). Extraction and early detection of anomalies in lightpath SNR using machine learning models. *Journal of Lightwave Technology, 40*(7), 1864-1872. https://doi.org/10.1109/JLT.2021.3134098

Amirabadi, M. A., Kahaei, M. H., Nezamalhosseini, S. A., Arpanaei, F., & Carena, A. (2023). Deep neural network-based QoT estimation for SMF and FMF links. *Journal of Lightwave Technology*, *41*(6), 1684-1695. https://doi.org/10.1109/JLT.2022.3225827

Amirabadi, M. A., Nezamalhosseini, S. A., Kahaei, M. H., & Carena, A. (2025). Deep learning-based classification for QoT estimation in SMF and FMF links. *Journal of Lightwave Technology*, *43*(1), 120–136. https://doi.org/10.1109/JLT.2024.3451242

Ayassi, R., Triki, A., Crespi, N., Minerva, R., & Laye, M. (2022). Survey on the use of machine learning for quality of transmission estimation in optical transport networks. *Journal of Lightwave Technology*, *40*(17), 5803-5815. https://doi.org/10.1109/JLT.2022.3184178

Ayoub, O., Troia, S., Andreoletti, D., Bianco, A., Tornatore, M., Giordano, S., & Rottondi, C. (2023). Towards explainable artificial intelligence in optical networks: The use case of lightpath QoT estimation. *Journal of Optical Communications and Networking, 15*(1), A26-A38. https://doi.org/10.1364/JOCN.470812

Azzimonti, D., Rottondi, C., Giusti, A., Tornatore, M., & Bianco, A. (2020). Active vs transfer learning approaches for QoT estimation with small training datasets. In *Optical Fiber Communications Conference and Exhibition (OFC)* (pp. 1-3). Optica Publishing Group M4E.1.

Azzimonti, D., Rottondi, C., Giusti, A., Tornatore, M., & Bianco, A. (2021). Comparison of domain adaptation and active learning techniques for quality of transmission estimation with small-sized training datasets [Invited]. *Journal of Optical Communications and Networking*, *13*(1), A56-A66. https://doi.org/10.1364/JOCN.405510

Bai, S., Kolter, J. Z., & Koltun, V. (2018). *An empirical evaluation of generic convolutional and recurrent networks for sequence modeling*. arXiv. https://arxiv.org/abs/1803.01271

Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, *58*, 82-115. https://doi.org/10.1016/j.inffus.2019.12.012

Behera, S., Panayiotou, T., & Ellinas, G. (2023a). Machine learning framework for timely soft-failure detection and localization in elastic optical networks. *Journal of Optical Communications and Networking,* *15*(10), E74-E85. https://doi.org/10.1364/JOCN.490008

Behera, S., Panayiotou, T., & Ellinas, G. (2023b). Machine learning for real-time anomaly detection in optical networks. In *2023 23rd International Conference on Transparent Optical Networks (ICTON)* (pp. 1-4). IEEE. https://doi.org/10.1109/ICTON59386.2023.10207370

Bergk, G., Shariati, B., Safari, P., & Fischer, J. K. (2022). ML-assisted QoT estimation: A dataset collection and data visualization for dataset quality evaluation. *Journal of Optical Communications and Networking*, *14*(3), 43-55. https://doi.org/10.1364/JOCN.442733

Boutaba, R., Salahuddin, M. A., Limam, N., Ayoubi, S., Shahriar, N., Estrada-Solano, F., & Caicedo, O. M. (2018). A comprehensive survey on machine learning for networking: Evolution, applications and research opportunities. *Journal of Internet Services and Applications*, *9*(1), 16. https://doi.org/10.1186/s13174-018-0087-2

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time series analysis: Forecasting and control (5th ed.). Wiley

Caballero, A., Aguado, J. C., Borkowski, R., Saldaña, S., Jiménez, T., de Miguel, I., Arlunno, V., Durán, R. J., Zibar, D., Jensen, J. B., Lorenzo, R. M., Abril, E. J., & Monroy, I. T. (2012). Experimental demonstration of a cognitive quality of transmission estimator for optical communication systems. In *2012 38th European Conference and Exhibition on Optical Communications (ECOC)* (pp. 1-3). IEEE

Cai, Z., Wang, Q., Deng, Y., Zhang, P., Zhou, G., Li, Y., & Khan, F. N. (2024). Domain adversarial adaptation framework for few-shot QoT estimation in optical networks. *Journal of Optical Communications and Networking, 16*(11), 1133-1144. https://doi.org/10.1364/JOCN.530915

Campello, R. J. G. B., Moulavi, D., Zimek, A., & Sander, J. (2015). Hierarchical density estimates for data clustering, visualization, and outlier detection. *ACM Transactions on Knowledge Discovery from Data*, 10(1), 1-51. https://doi.org/10.1145/2733381

Cantono, M., Pilori, D., Ferrari, A., Catanese, C., Thouras, J., Augé, J.-L., & Curri, V. (2018). On the interplay of nonlinear interference generation with stimulated Raman scattering for QoT estimation. *Journal of Lightwave Technology*, *36*(15), 3131-3141. https://doi.org/10.1109/JLT.2018.2830816

Cao, D., Ren, C., Wu, Z., Liang, Y., Li, L., & Zhang, H. (2026). A survey of AI-driven QoT analysis in lightpath management. Digital Communications and Networks. https://doi.org/10.1016/j.dcan.2025.12.006

Carena, A., Bosco, G., Curri, V., Jiang, Y., Poggiolini, P., & Forghieri, F. (2014). EGN model of non-linear fiber propagation. *Optics Express*, *22*(13), 16335-16362. https://doi.org/10.1364/OE.22.016335

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. ACM Computing Surveys, 41(3), Article 15. https://doi.org/10.1145/1541880.1541882

Chen, X., Li, B., Proietti, R., Zhu, Z., & Yoo, S. J. B. (2019). Self-taught anomaly detection with hybrid unsupervised/supervised machine learning in optical networks. *Journal of Lightwave Technology, 37*(7), 1742-1749. https://doi.org/10.1109/JLT.2019.2902487

Chouhan, R. K., Atulkar, M., & Nagwani, N. K. (2022). An unsupervised attack detection approach for software defined networks. In *2022 International Conference on Augmented Intelligence and Sustainable Systems (ICAISS)* (pp. 1025-1030). IEEE. https://doi.org/10.1109/ICAISS55157.2022.10010577

Chouman, H., Djukic, P., Tremblay, C., & Desrosiers, C. (2021). Forecasting lightpath QoT with deep neural networks. In *Optical Fiber Communication Conference (OFC 2021)*, paper Th4J.5. Optica Publishing Group. https://doi.org/10.1364/OFC.2021.Th4J.5

Ciena. (2023). *It's time for a new interconnection model for the growing amount of data in motion*. https://www.ciena.com/insights/articles/2023/its-time-for-a-new-interconnection-model-for-the-growing-amount-of-data-in-motion

Cruzes, S. (2024). *Overview of quality of transmission estimation in optical networks*. TechRxiv. https://doi.org/10.36227/techrxiv.171340665.50082312/v2

Di Cicco, N., Talpini, J., Ibrahimi, M., Savi, M., & Tornatore, M. (2023). Uncertainty-aware QoT forecasting in optical networks with Bayesian recurrent neural networks. In *ICC 2023 - IEEE International Conference on Communications* (pp. 441-446). IEEE. https://doi.org/10.1109/ICC45041.2023.10278767

Di Marino, R., Rottondi, C., Giusti, A., & Bianco, A. (2020). Assessment of domain adaptation approaches for QoT estimation in optical networks. In *Optical Fiber Communications Conference and Exhibition (OFC)* (pp. 1-3). Optica Publishing Group.

Diaz-Montiel, A. A., Aladin, S., Tremblay, C., & Ruffini, M. (2019). Active wavelength load as a feature for QoT estimation based on support vector machine. In *ICC 2019 - IEEE International Conference on Communications (ICC)* (pp. 1-6). IEEE. https://doi.org/10.1109/ICC.2019.8761369

Domingos, P. (1999). The role of Occam's razor in knowledge discovery. *Data Mining and Knowledge Discovery*, *3*, 409-425. https://doi.org/10.1023/A:1009868929893

Dong, Z., Khan, F. N., Sui, Q., Zhong, K., Lu, C., & Lau, A. P. T. (2016). Optical performance monitoring: A review of current and future technologies. *Journal of Lightwave Technology*, *34*(2), 525-543. https://doi.org/10.1109/JLT.2015.2480798

Du, W., Cote, D., Barber, C., & Liu, Y. (2021). Forecasting loss of signal in optical networks with machine learning. *Journal of Optical Communications and Networking*, *13*(10), E109-E121. https://doi.org/10.1364/JOCN.423667

Ester, M., Kriegel, H., Sander, J., & Xu, X. (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *Knowledge Discovery and Data Mining*.

Fawaz, H., Arpanaei, F., Andreoletti, D., Sbeity, I., Hernández, J. A., Larrabeiti, D., & Ayoub, O. (2024). Reducing complexity and enhancing predictive power of ML-based lightpath QoT estimation via SHAP-assisted feature selection. In *2024 International Conference on Optical Network Design and Modeling (ONDM)* (pp. 1-6). IEEE. https://doi.org/10.23919/ONDM61578.2024.10582777

Ferrari, A., Filer, M., Balasubramanian, K., Yin, Y., Le Rouzic, E., Kundrat, J., Grammel, G., Galimberti, G., & Curri, V. (2020). GNPy: An open source application for physical layer aware open optical networks. *Journal of Optical Communications and Networking, 12*(6), C31-C40. https://doi.org/10.1364/JOCN.382906

Flynn, T., & Yoo, S. (2019). Change detection with the kernel cumulative sum algorithm. In *2019 IEEE 58th Conference on Decision and Control (CDC)* (pp. 6092-6099). IEEE. https://doi.org/10.1109/CDC40024.2019.9029854

Furdek, M., Natalino, C., Di Giglio, A., & Schiano, M. (2021). Optical network security management: requirements, architecture, and efficient machine learning models for detection of evolving threats [Invited]. *Journal of Optical Communications and Networking*, *13*(2), A144-A155. https://doi.org/10.1364/JOCN.402884

Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to forget: Continual prediction with LSTM. *Neural Computation*, *12*(10), 2451-2471. https://doi.org/10.1162/089976600300015015

Gour, R., Ishigaki, G., Kong, J., & Jue, J. P. (2021). Availability-guaranteed slice composition for service function chains in 5G transport networks. *Journal of Optical Communications and Networking*, *13*(3), 14-24. https://doi.org/10.1364/JOCN.403721

Gu R., Yang Z., Ji Y. (2020) Machine learning for intelligent optical networks: A comprehensive survey, Journal of Network and Computer Applications, Volume 157, 102576, ISSN 1084-8045, https://doi.org/10.1016/j.jnca.2020.102576.

Gu, Z., Qin, T., Zhou, Y., Zhang, J., & Ji, Y. (2023). Sample-distribution-matching-based transfer learning for QoT estimation in optical networks. *Journal of Optical Communications and Networking*, *15*(9), 649-663. https://doi.org/10.1364/JOCN.493053

Gupta, M., Gao, J., Aggarwal, C. C., & Han, J. (2014). Outlier detection for temporal data: A survey. *IEEE Transactions on Knowledge and Data Engineering*, *26*(9), 2250-2267. https://doi.org/10.1109/TKDE.2013.184

Hall, T., & Rasheed, K. (2025). A survey of machine learning methods for time series prediction. *Applied Sciences, 15*(11), Article 5957. https://doi.org/10.3390/app15115957

Hashemi, N., Safari, P., Shariati, B., & Fischer, J. K. (2021). Vertical federated learning for privacy-preserving ML model development in partially disaggregated networks. In *2021 European Conference on Optical Communication (ECOC)* (pp. 1-4). IEEE. https://doi.org/10.1109/ECOC52684.2021.9605846

Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K., Scardapane, S., Spinelli, I., Mahmud, M., & Hussain, A. (2024). Interpreting black-box models: A review on explainable artificial intelligence. *Cognitive Computation*, *16*(1), 45-74. https://doi.org/10.1007/s12559-023-10179-8

Hedayatnejad, M., Pei, Y., Boertjes, D., Demeter, D., Desrosiers, C., & Tremblay, C. (2025). Multi-step span loss prediction in optical networks using multi-head attention transformers. *IEEE Photonics Journal*, *17*(3), 1-8, Article 8500608. https://doi.org/10.1109/JPHOT.2025.3571220

Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, *313*(5786), 504-507. https://doi.org/10.1126/science.1127647

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735

Huang, X., Yun, P., & Hu, Z. (2023). A survey of collective anomaly detection on sequence dataset. *International Journal of Data Warehousing and Mining*, *19*(1). https://doi.org/10.4018/IJDWM.327363

Hyndman, R. J., & Athanasopoulos, G. (2021). Forecasting: Principles and practice (3rd ed.). OTexts. https://otexts.com/fpp3/

Ibrahimi, M., Abdollahi, H., Rottondi, C., Giusti, A., Ferrari, A., Curri, V., & Tornatore, M. (2021). Machine learning regression for QoT estimation of unestablished lightpaths. *Journal of Optical Communications and Networking, 13*(4), B92-B101. https://doi.org/10.1364/JOCN.410694

International Data Corporation. (2023). *Worldwide global DataSphere forecast, 2023–2027: It's a distributed, diverse, and dynamic (3D) DataSphere* (Doc. No. US50554523). IDC.

International Telecommunication Union. (2002). End-to-end error performance parameters and objectives for international, constant bit-rate digital paths and connections (Recommendation ITU-T G.826). https://www.itu.int/rec/T-REC-G.826-200212-I/en

International Telecommunication Union. (2003). Availability performance parameters and objectives for end-to-end international constant bit-rate digital paths (Recommendation ITU-T G.827). https://www.itu.int/rec/T-REC-G.827-200309-I/en

International Telecommunication Union. (2011). Error performance parameters and objectives for multi-operator international paths within optical transport networks (Recommendation ITU-T G.8201). https://www.itu.int/rec/T-REC-G.8201-201104-P

Keiser, G. (2021). *Fiber optic communications*. Springer Singapore. https://doi.org/10.1007/978-981-33-4665-9

Khan, F. N. (2024a). Non-technological barriers: The last frontier towards AI-powered intelligent optical networks. *Nature Communications*, *15*, Article 5995. https://doi.org/10.1038/s41467-024-50307-y

Khan, I., Bilal, M., Siddiqui, M., Khan, M., Ahmad, A., Shahzad, M., & Curri, V. (2020a). QoT estimation for light-path provisioning in unseen optical networks using machine learning. In *2020 22nd International Conference on Transparent Optical Networks (ICTON)* (pp. 1-4). IEEE. https://doi.org/10.1109/ICTON51198.2020.9203364

Khan, I., Bilal, M., & Curri, V. (2020b). Advanced formulation of QoT-estimation for un-established lightpaths using cross-train machine learning methods. In *2020 22nd International Conference on Transparent Optical Networks (ICTON)* (pp. 1-4). IEEE. https://doi.org/10.1109/ICTON51198.2020.9203334

Khan, L. Z., Pedro, J., Costa, N., Sgambelluri, A., Napoli, A., & Sambo, N. (2024b). Model and data-centric machine learning algorithms to address data scarcity for failure identification. *Journal of Optical Communications and Networking*, *16*(3), 369–381. https://doi.org/10.1364/JOCN.511863

Kruse, L. E., Kuhl, S., Dochhan, A., & Pachnicke, S. (2024). Experimental investigation of machine-learning-based soft-failure management using the optical spectrum. *Journal of Optical Communications and Networking,* *16*(2), 94-103. https://doi.org/10.1364/JOCN.500930

Lechowicz, P., Natalino, C., Arpanaei, F., Melin, S., Diaz, R., Lindgren, A., Larrabeiti, D., & Monti, P. (2025). Toward better QoT estimation: An ML architecture with link-level embedding layers. *IEEE Networking Letters,* *7*(2), 122-125. https://doi.org/10.1109/LNET.2025.3561336

Liu, C.-Y., Chen, X., Proietti, R., & Yoo, S. J. B. (2021a). Performance studies of evolutionary transfer learning for end-to-end QoT estimation in multi-domain optical networks [Invited]. *Journal of Optical Communications and Networking*, *13*(4), B1-B11. https://doi.org/10.1364/JOCN.415209

Liu, S., Wang, D., Zhang, C., Wang, L., & Zhang, M. (2021b). Semi-supervised anomaly detection with imbalanced data for failure detection in optical networks. In *2021 Optical Fiber Communications Conference and Exhibition (OFC)* (pp. 1-3), Th1A.24

Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation forest. In Proceedings of the 2008 Eighth IEEE International Conference on Data Mining (ICDM) (pp. 413-422). IEEE. https://doi.org/10.1109/ICDM.2008.17

Lun, H., Fu, M., Zhang, Y., Jiang, H., Yi, L., Hu, W., & Zhuge, Q. (2023). A GAN based soft failure detection and identification framework for long-haul coherent optical communication systems. *Journal of Lightwave Technology*, *41*(8), 2312-2322. https://doi.org/10.1109/JLT.2022.3227719

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In I. Guyon, U. von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems 30* (pp. 4765-4774). Curran Associates. https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer. https://doi.org/10.1007/978-3-540-27752-1

Manso, C., Vilalta, R., Muñoz, R., Casellas, R., & Martínez, R. (2021). Scalable for cloud-native transport SDN controller using GNPy and machine learning techniques for QoT estimation. In *2021 Optical Fiber Communications Conference and Exhibition (OFC)* (pp. 1-3). Optica Publishing Group, M1B.5.

Mayer, K. S., Soares, J. A., Pinto, R. P., Rothenberg, C. E., Arantes, D. S., & Mello, D. A. A. (2021). Machine-learning-based soft-failure localization with partial software-defined networking telemetry. *Journal of Optical Communications and Networking*, *13*(10), E122-E131. https://doi.org/10.1364/JOCN.424654

Mejri, N., Lopez-Fuentes, L., Roy, K., Chernakov, P., Ghorbel, E., & Aouada, D. (2024). Unsupervised anomaly detection in time-series: An extensive evaluation and analysis of state-of-the-art methods. *Expert Systems with Applications, 256*, 124922. https://doi.org/10.1016/j.eswa.2024.124922

Mezni, A., Charlton, D. W., Tremblay, C., & Desrosiers, C. (2020). Deep learning for multi-step performance prediction in operational optical networks. In *Conference on Lasers and Electro-Optics (CLEO: S&I 2020)*, paper STh4M.1. Optica Publishing Group. https://doi.org/10.1364/CLEO_SI.2020.STh4M.1

Morais, R. M., & Pedro, J. (2018). Machine learning models for estimating quality of transmission in DWDM networks. *Journal of Optical Communications and Networking*, *10*(10), D84-D99. https://doi.org/10.1364/JOCN.10.000D84

Mosayebi, R., & Lampe, L. (2024). Anomaly detection in optical fiber: A change-point detection perspective. In *Advanced Photonics Congress 2024, Technical Digest Series* (paper SpTh2G.4). Optica Publishing Group.

Mukherjee, B. (2006). *Optical WDM networks*. Springer

Murphy, K., Lavignotte, A., & Lepers, C. (2023). Fault prediction for optical access network equipment using decision tree methods. In *2023 Asia Communications and Photonics Conference/2023 International Photonics and Optoelectronics Meetings (ACP/POEM)* (pp. 1-5). https://doi.org/10.1109/ACP/POEM59049.2023.10369987

Murphy, K., Lavignotte, A., & Lepers, C. (2024). Fault prediction for heterogeneous telecommunication networks using machine learning: A survey. *IEEE Transactions on Network and Service Management*, 21(2), 2515–2538. doi: 10.1109/TNSM.2023.3340351

Musumeci, F., Rottondi, C., Nag, A., Macaluso, I., Zibar, D., Ruffini, M., & Tornatore, M. (2019a). An overview on application of machine learning techniques in optical networks. *IEEE Communications Surveys and Tutorials*, *21*(2), 1383-1408. https://doi.org/10.1109/COMST.2018.2880039

Musumeci, F., Rottondi, C., Corani, G., Shahkarami, S., Cugini, F., & Tornatore, M. (2019b). A tutorial on machine learning for failure management in optical networks. *Journal of Lightwave Technology*, *37*(16), 4125-4139. https://doi.org/10.1109/JLT.2019.2922586

Musumeci, F., Venkata, V. G., Hirota, Y., Awaji, Y., Xu, S., Shiraiwa, M., Mukherjee, B., & Tornatore, M. (2022). Domain adaptation and transfer learning for failure detection and failure-cause identification in optical networks across different lightpaths [Invited]. *Journal of Optical Communications and Networking*, *14*(2), A91-A100. https://doi.org/10.1364/JOCN.438269

Musumeci, F., & Tornatore, M. (2025). Failure management in optical networks with ML: A tutorial on applications, challenges, and pitfalls [Invited]. *Journal of Optical Communications and Networking*, *17*(8), C144-C155. https://doi.org/10.1364/JOCN.551910

Naik, M. K., Jeyachitra, R. K., & Doss M, P. (2023). Machine learning based multi-parameter light path quality of transmission estimation and prediction in optical networks. In *2023 Fifth International Conference on Electrical, Computer and Communication Technologies (ICECCT)* (pp. 01-06). IEEE. https://doi.org/10.1109/ICECCT56650.2023.10179761

Natalino, C., Mohammadiha, N., & Panahi, A. (2023). Machine-learning-as-a-service for optical network automation. In *2023 Optical Fiber Communications Conference and Exhibition (OFC)* (pp. 1-3). Optica Publishing Group. https://doi.org/10.1364/OFC.2023.W4G.3

Natalino, C., Panahi, A., Mohammadiha, N., & Monti, P. (2024). AI/ML-as-a-service for optical network automation: Use cases and challenges [Invited]. *Journal of Optical Communications and Networking*, *16*(2), A169-A179. https://doi.org/10.1364/JOCN.500706

Ouyang, Z., Chen, X., Liu, Z., Chen, X., & Zhu, Z. (2024). Overview of ML-aided QoT estimation in optical networks: A perspective of model generalization. In *2024 IEEE 24th International Conference on Communication Technology (ICCT)* (pp. 130-136). IEEE. https://doi.org/10.1109/ICCT62411.2024.10946528

Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, *22*(10), 1345-1359. https://doi.org/10.1109/TKDE.2009.191

Panayiotou, T., Savva, G., Shariati, B., Tomkos, I., & Ellinas, G. (2019). Machine learning for QoT estimation of unseen optical network states. In *2019 Optical Fiber Communications Conference and Exhibition (OFC)* (pp. 1-3). Optica Publishing Group.

Panayiotou, T., Michalopoulou, M., & Ellinas, G. (2023). Survey on machine learning for traffic-driven service provisioning in optical networks. *IEEE Communications Surveys & Tutorials, 25*(2), 1412-1443. https://doi.org/10.1109/COMST.2023.3247842

Pang, G., Shen, C., Cao, L., & Van Den Hengel, A. (2021). Deep learning for anomaly detection: A review. *ACM Computing Surveys*, *54*(2), Article 38. https://doi.org/10.1145/3439950

Paudyal, P., Shen, S., Yan, S., & Simeonidou, D. (2021). Toward deployments of ML applications in optical networks. *IEEE Photonics Technology Letters*, *33*(11), 537-540. https://doi.org/10.1109/LPT.2021.3074586

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825-2830.

Pei, Y., Côté, D., Sigue, P. A. N., Mahmoudialami, A., Tremblay, C., & Desrosiers, C. (2022, October 18). *Identifying root causes of network service degradation* (U.S. Patent No. 11,477,070). U.S. Patent and Trademark Office.

Pesic, J., Lonardi, M., Seve, E., Rossi, N., & Zami, T. (2020). Transfer learning from unbiased training data sets for QoT estimation in WDM networks. In *Proceedings of the 46th European Conference on Optical Communications (ECOC 2020)* (pp. 1-4). IEEE. https://doi.org/10.1109/ECOC48923.2020.9333399

Pilori, D., Cantono, M., Carena, A., & Curri, V. (2017). FFSS: The fast fiber simulator software. In *2017 19th International Conference on Transparent Optical Networks (ICTON)* (pp. 1-4). IEEE. https://doi.org/10.1109/ICTON.2017.8025002

Poggiolini, P. (2012). The GN model of non-linear propagation in uncompensated coherent optical systems. *Journal of Lightwave Technology*, *30*(24), 3857-3879. https://doi.org/10.1109/JLT.2012.2217729

Poggiolini, P., Bosco, G., Carena, A., Curri, V., Jiang, Y., & Forghieri, F. (2014). The GN-model of fiber non-linear propagation and its applications. *Journal of Lightwave Technology*, *32*(4), 694-721. https://doi.org/10.1109/JLT.2013.2295208

Ramaswami, R., Sivarajan, K. N., & Sasaki, G. H. (2010). *Optical networks: A practical perspective* (3rd ed.). Morgan Kaufmann

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1135–1144). Association for Computing Machinery. https://doi.org/10.1145/2939672.2939778

Rottondi, C., Barletta, L., Giusti, A., & Tornatore, M. (2018). Machine-learning method for quality of transmission prediction of unestablished lightpaths. *Journal of Optical Communications and Networking*, *10*(2), A286-A297. https://doi.org/10.1364/JOCN.10.00A286

Rottondi, C., di Marino, R., Nava, M., Giusti, A., & Bianco, A. (2021). On the benefits of domain adaptation techniques for quality of transmission estimation in optical networks. *Journal of Optical Communications and Networking*, *13*(1), A34-A43. https://doi.org/10.1364/JOCN.401915

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533-536. https://doi.org/10.1038/323533a0

Safari, P., Shariati, B., Bergk, G., & Fischer, J. K. (2021a). Deep convolutional neural network for network-wide QoT estimation. In *2021 Optical Fiber Communications Conference and Exhibition (OFC)* (pp. 1-3). Optica Publishing Group.

Safari, P., Shariati, B., Bergk, G., & Fischer, J. K. (2021b). Secure multi-party computation and statistics sharing for ML model training in multi-domain multi-vendor networks. In *2021 European Conference on Optical Communication (ECOC)* (pp. 1-4). IEEE. https://doi.org/10.1109/ECOC52684.2021.9606082

Schölkopf, B., Platt, J. C., Shawe-Taylor, J., Smola, A. J., & Williamson, R. C. (2001). Estimating the support of a high-dimensional distribution. *Neural Computation*, *13*(7), 1443-1471. https://doi.org/10.1162/089976601750264965

Semrau, D., Killey, R. I., & Bayvel, P. (2019a). A closed-form approximation of the Gaussian noise model in the presence of inter-channel stimulated Raman scattering. *Journal of Lightwave Technology*, *37*(9), 1924–1936. https://doi.org/10.1109/JLT.2019.2895237

Semrau, D., Sillekens, E., Killey, R. I., & Bayvel, P. (2019b). A modulation format correction formula for the Gaussian noise model in the presence of inter-channel stimulated Raman scattering. *Journal of Lightwave Technology*, *37*(19), 5122–5131. https://doi.org/10.1109/JLT.2019.2929461

Silva, M. F., Sgambelluri, A., Pacini, A., Paolucci, F., Green, A., Mascarenas, D., & Valcarenghi, L. (2023). Confidentiality-preserving machine learning algorithms for soft-failure detection in optical communication networks. *Journal of Optical Communications and Networking*, *15*(8), C212-C222. https://doi.org/10.1364/JOCN.481690

Singhal, P., Walambe, R., Ramanna, S., & Kotecha, K. (2023). Domain adaptation: Challenges, methods, datasets, and applications. *IEEE Access*, *11*, 6973-7020. https://doi.org/10.1109/ACCESS.2023.3237025

Sinkin, O. V., Holzlöhner, R., Zweck, J., & Menyuk, C. R. (2003). Optimization of the split-step Fourier method in modeling optical-fiber communications systems. *Journal of Lightwave Technology*, *21*(1), 61-68. https://doi.org/10.1109/JLT.2003.8086283

Skoog, R. A., Banwell, T. C., Gannett, J. W., Habiby, S. F., Pang, M., Rauch, M. E., & Toliver, P. (2006). Automatic identification of impairments using support vector machine pattern classification on eye diagrams. *IEEE Photonics Technology Letters*, *18*(22), 2398-2400. https://doi.org/10.1109/LPT.2006.886146

Souza, A., Costa, N., Pedro, J., & Pires, J. (2023). Comparison of fast quality of transmission estimation methods for C + L + S optical systems. *Journal of Optical Communications and Networking*, *15*(11), F1-F12. https://doi.org/10.1364/JOCN.486898

Sun, K., Yu, Z., Shu, L., Wan, Z., Huang, H., Lei, Y., & Xu, K. (2023). Digital residual spectrum-based generalized soft failure detection and identification in optical networks. *IEEE Transactions on Communications*, *71*(1), 324-338. https://doi.org/10.1109/TCOMM.2022.3222519

Sun, Z., Zhang, C., Zhang, M., Ye, B., & Wang, D. (2024). Semi-supervised learning model synergistically utilizing labeled and unlabeled data for failure detection in optical networks. *Journal of Optical Communications and Networking*, *16*(5), 541-552. https://doi.org/10.1364/JOCN.516128

Tariq, H., Usmani, F., Khan, I., Masood, M. U., Ahmad, A., & Curri, V. (2023). Iterative transfer learning approach for QoT prediction of lightpath in optical networks. In *2023 23rd International Conference on Transparent Optical Networks (ICTON)* (pp. 1-4). IEEE. https://doi.org/10.1109/ICTON59386.2023.10207173

Tremblay, C., & Aladin, S. (2018). Machine learning techniques for estimating the quality of transmission of lightpaths. In *2018 IEEE Photonics Society Summer Topical Meeting Series (SUM)* (pp. 237-238). IEEE. https://doi.org/10.1109/PHOSST.2018.8456791

Tremblay, C., Allogba, S., & Aladin, S. (2019). Quality of transmission estimation and performance prediction of lightpaths using machine learning. In *Proceedings of the 45th European Conference on Optical Communication (ECOC 2019)* (pp. 1-3). IET. https://doi.org/10.1049/cp.2019.0956

Tremblay, C., Mahmoudialami, A., Ngani Sigue, P. A., Pei, Y., Côté, D., Boertjes, D., Demeter, D., & Desrosiers, C. (2023). Detection and root cause analysis of performance degradation in optical networks using machine learning. In *Proceedings of the 49th European Conference on Optical Communications (ECOC 2023)* (pp. 1278-1281). IET. https://doi.org/10.1049/icp.2023.2525

Tremblay, C., Aladin, S., & Wosinska, L. (2026). Towards automated and interpretable ML for quality of transmission estimation in optical networks. In *Advanced Photonics Congress (APC) 2026 (Optica Publishing Group),* Long Beach, California, July 26–30. [Invited]

Truong, C., Oudre, L., & Vayatis, N. (2018). *ruptures*: change point detection in Python. arXiv preprint arXiv:1801.00826. https://doi.org/10.48550/arXiv.1801.00826

Truong, C., Oudre, L., & Vayatis, N. (2020). Selective review of offline change point detection methods. *Signal Processing*, *167*, 107299. https://doi.org/10.1016/j.sigpro.2019.107299

Usmani, F., Khan, I., Siddiqui, M., Khan, M., Bilal, M., Masood, M. U., Ahmad, A., Shahzad, M., & Curri, V. (2021). Evaluating cross-feature trained machine learning models for estimating QoT of unestablished lightpaths. *2021 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)*, 1-6. https://doi.org/10.1109/ICECCE52056.2021.9514154

Usmani, F., Khan, I., Masood, M. U., Ahmad, A., Shahzad, M., & Curri, V. (2022a). Transfer learning aided QoT computation in network operating with the 400ZR standard. In *2022 International Conference on Optical Network Design and Modeling (ONDM)* (pp. 1-6). IEEE. https://doi.org/10.23919/ONDM54585.2022.9782856

Usmani, F., Khan, I., Tariq, H., Masood, M. U., Shahzad, M., Ahmad, A., & Curri, V. (2022b). Performance analysis of transfer-learning approaches for QoT estimation of network operating with 400ZR. In *2022 Asia Communications and Photonics Conference (ACP)* (pp. 1238-1242). IEEE. https://doi.org/10.1109/ACP55869.2022.10088462

Usmani, F., Khan, I., Masood, M. U., Ahmad, A., Shahzad, M., & Curri, V. (2022c). QoT-estimation assisted by transfer learning in extended C-band network operating on 400ZR. In *2022 IEEE Photonics Society Summer Topicals Meeting Series (SUM)* (pp. 1-2). IEEE. https://doi.org/10.1109/SUM53465.2022.9858280

Usmani, F., Khan, I., Mehran, Ahmad, A., & Curri, V. (2024). Integrating knowledge distillation and transfer learning for enhanced QoT-estimation in optical networks. *IEEE Access*, *12*, 156785-156802. https://doi.org/10.1109/ACCESS.2024.3485999

Vela, A. P., Ruiz, M., Fresi, F., Sambo, N., Cugini, F., Meloni, G., Potì, L., Velasco, L., & Castoldi, P. (2017). BER degradation detection and failure identification in elastic optical networks. *Journal of Lightwave Technology*, *35*(21), 4595-4604. https://doi.org/10.1109/JLT.2017.2747223

Wan, Y., Yang, H., Yao, Q., Bao, B., Li, C., & Zhang, J. (2021). Anomaly detection based on correlative prediction for elastic optical network. *2021 Opto-Electronics and Communications Conference (OECC)*, 1-3. https://doi.org/10.1364/OECC.2021.T2A.7

Wang, D., Zhang, C., Chen, W., Yang, H., Zhang, M., & Lau, A. P. T. (2022b). A review of machine learning-based failure management in optical networks. *Science China Information Sciences*, *65(11)*, 211302. https://doi.org/10.1007/s11432-022-3557-9

Wang, Q., Cai, Z., Lau, A. P. T., Li, Y., & Nadeem Khan, F. (2023). Invariant convolutional neural network for robust and generalizable QoT estimation in fiber-optic networks. *Journal of Optical Communications and Networking*, *15*(7), 431-441. https://doi.org/10.1364/JOCN.488689

Wass, J., Thrane, J., Piels, M., Jones, R., & Zibar, D. (2017). Gaussian process regression for WDM system performance prediction. In *2017 Optical Fiber Communications Conference and Exhibition (OFC)* (pp. 1-3). Optica Publishing Group.

Wei, S., & Xie, Y. (2022). Online kernel CUSUM for change-point detection. In *Advances in Neural Information Processing Systems 35 (NeurIPS 2022)*. https://arxiv.org/abs/2211.15070

Wolpert, D. H., & Macready, W. G. (1997). No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, *1*(1), 67–82. https://doi.org/10.1109/4235.585893

Xu, R., Song, Z., Wu, J., Wang, C., & Zhou, S. (2025). Change-point detection with deep learning: A review. *Frontiers of Engineering Management*, *12*(1), 154-176. https://doi.org/10.1007/s42524-025-4109-z

Yaméogo, B. L. M., Charlton, D. W., Doucet, D., Desrosiers, C., O'Sullivan, M., & Tremblay, C. (2020). Trends in optical span loss detected using the time series decomposition method. *Journal of Lightwave Technology*, *38*(18), 5026-5035. https://doi.org/10.1109/JLT.2020.3000967

Yan, S., Khan, F. N., Mavromatis, A., Gkounis, D., Fan, Q., Ntavou, F., Nikolovgenis, K., Meng, F., Salas, E. H., Guo, C., Lu, C., Lau, A. P. T., Nejabati, R., & Simeonidou, D. (2017). Field trial of machine-learning-assisted and SDN-based optical network planning with network-scale monitoring database. *2017 European Conference on Optical Communication (ECOC)*, 1-3. https://doi.org/10.1109/ECOC.2017.8346091

Yousefi, S., Chouman, H., Djukic, P., Golaghazadeh, F., Tremblay, C., & Desrosiers, C. (2023). Forecasting lightpath quality of transmission and implementing uncertainty in the forecast models. *Journal of Lightwave Technology, 41*(15), 4871-4881. https://doi.org/10.1109/JLT.2023.3252441

Zhang, L., Li, X., Tang, Y., Xin, J., & Huang, S. (2022). A survey on QoT prediction using machine learning in optical networks. *Optical Fiber Technology*, *68*, 102804. https://doi.org/10.1016/j.yofte.2021.102804

Zhang, Y., Song, Y., Pang, Y., Li, S., Jiang, X., Wang, Y., Li, J., Zhang, M., & Wang, D. (2025). Design and evaluation of an LLM-based agent for QoT estimation and performance optimization in optical networks. *IEEE Open Journal of the Communications Society*, *6*, 7470–7484. https://doi.org/10.1109/OJCOMS.2025.3608290

Zhou, Y., Gu, Z., Zhang, J., & Ji, Y. (2023). Neuron-level transfer learning for ANN-based QoT estimation in optical networks. In *2023 Asia Communications and Photonics Conference/2023 International Photonics and Optoelectronics Meetings (ACP/POEM)* (pp. 1-4). IEEE. https://doi.org/10.1109/ACP/POEM59049.2023.10368618

Zhou, Y., Gu, Z., Zhang, J., & Ji, Y. (2024). Evolutionary neuron-level transfer learning for QoT estimation in optical networks. *Journal of Optical Communications and Networking*, *16*(4), 432-448. https://doi.org/10.1364/JOCN.514618

Zuo, X., Gu, Z., Fan, S., Ren, D., Wei, Z., Feng, N., & Zhao, J. (2025). Domain discrepancy feedback optimization method for few-sample QoT estimation in optical networks. *Journal of Optical Communications and Networking, 17*(12), 1161–1175. https://doi.org/10.1364/JOCN.572972