



HAL
open science

Survey on the use of machine learning for quality of transmission estimation in optical transport networks

Reda Ayassi, Ahmed Triki, Noel Crespi, Roberto Minerva, Maxime Laye

► To cite this version:

Reda Ayassi, Ahmed Triki, Noel Crespi, Roberto Minerva, Maxime Laye. Survey on the use of machine learning for quality of transmission estimation in optical transport networks. *Journal of Lightwave Technology*, 2022, 40 (17), pp.5803-5815. 10.1109/JLT.2022.3184178 . hal-03762481

HAL Id: hal-03762481

<https://hal.science/hal-03762481>

Submitted on 7 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Survey on the Use of Machine Learning for Quality of Transmission Estimation in Optical Transport Networks

| | |
|-------------------------------|--|
| Journal: | <i>Journal of Lightwave Technology</i> |
| Manuscript ID | JLT-29941-2022 |
| Manuscript Type: | Original Paper |
| Date Submitted by the Author: | 14-Mar-2022 |
| Complete List of Authors: | Ayassi, Reda; Orange Labs Lannion, OLN Triki, Ahmed; Orange Labs Lannion, OLN Crespi, Noel; Institut Polytechnique de Paris, Telecom SudParis Minerva, Roberto; Institut Polytechnique de Paris, Telecom SudParis Laye, Maxime; Orange Labs Lannion, OLN |
| Key Words: | Machine learning, Optical networks, WDM |
| Category: | Systems, subsystems and networks for fiber optics communications |
| | |

Survey on the Use of Machine Learning for Quality of Transmission Estimation in Optical Transport Networks

Reda Ayassi, Ahmed Triki, Noel Crespi, Roberto Minerva and Maxime Laye

Abstract— Estimating the Quality of Transmission (QoT) of the optical signal from source to destination nodes is the cornerstone of design engineering and service provisioning in optical transport network. Recent studies have turned to Machine Learning (ML) techniques to improve the accuracy of QoT estimation. In this paper, we survey the literature on this topic and classify the studies into categories based on their scope. Accordingly, we distinguish four categories of ML-based solutions: i) check lightpath feasibility, ii) estimate a lightpath's QoT, iii) enhance existing analytical models and iv) improve model generalization. We describe the proposed solutions in each category in terms of ML algorithms, inputs/outputs of the models, source of data and performance evaluation. Deploying a ML-based solution in the real field is not straightforward and presents several challenges. Therefore, we also discuss from an operator's perspective the potential of these solutions for real-field deployment.

Index Terms—Machine Learning, WDM, optical networks, QoT.

I. INTRODUCTION

In a constant attempt to meet increasing capacity demands, Optical transport networks have steadily evolved through a number of technological advances. Technologies such as coherent transmission, flexible modulation and tunable transceivers have led to a plethora of new parameters and configurations that complicate network design and operation. At the same time, revolutionary initiatives are emerging with the introduction of Software Defined Networks (SDN) [1] that could open up new opportunities to deal with these complexities in optical networks. Among these initiatives are those that push towards the promotion of the openness of Application Programming Interfaces APIs and the definition of common data models (e.g., T-API [2] and OpenROADM initiative [3][4]). Other initiatives propose new control and monitoring protocols such as Netconf and gRPC [5][6]. These solutions will lead to providing large amounts of data in standardized format that could be harnessed to solve network issues using

This paragraph of the first footnote will contain the date on which you submitted your paper for review. It will also contain support information, including sponsor and financial support acknowledgment. For example, "This work was supported in part by the U.S. Department of Commerce under Grant BS123456."

The next few paragraphs should contain the authors' current affiliations, including current address and e-mail. For example, F. A. Author is with the

new paradigms in optical networks like Artificial Intelligence (AI).

AI is the introduction of cognition to machines in order to perform intelligent tasks in a similar manner to humans. The most popular subfield of AI is machine learning (ML). ML consists of algorithms that capture patterns and behaviors in the data in order to produce models for a variety of tasks such as estimating a value based on inputs (i.e., regression techniques) and classifying data into groups (i.e., classification techniques). ML has seen an increase in popularity in research in recent years in multiple fields, particularly computer vision, natural language processing and speech recognition [7].

ML is subject to the same scrutiny in optical networks. A large amount of papers have been published on the application of ML techniques to multiple use cases: routing and wavelength/spectrum assignment (RWA/RSA) [8][9], Quality of Transmission (QoT) estimation [10] and fault management [11].

QoT estimation is of particular interest for optical networks. It consists of ascertaining the performance of an existing or candidate lightpath based on its characteristics and the network configuration. The QoT is used to monitor the health of an existing lightpath or check the feasibility of a candidate one by comparing its predicted QoT to the receiver's threshold. Estimating the QoT of a lightpath is crucial in network design and service provisioning. In fact, an underestimated QoT value can lead to significant loss in capacity and increase the network deployment cost (e.g., unnecessary equipment expenditure). On the other hand, an overestimation of QoT can lead to unstable lightpath. QoT estimation is also the basis of network optimization, as an accurate QoT is required for optimal RWA/RSA and capacity maximization.

The difficulty of QoT estimation stems mainly from the various impairments in the fiber that optical transmission is subject to [12]. Linear impairments are due to the signal attenuation, chromatic/polarization dispersion and the noise generated by the equipment. Nonlinear impairments include

National Institute of Standards and Technology, Boulder, CO 80305 USA (e-mail: author@boulder.nist.gov).

S. B. Author, Jr., was with Rice University, Houston, TX 77005 USA. He is now with the Department of Physics, Colorado State University, Fort Collins, CO 80523 USA (e-mail: author@lamar.colostate.edu).

T. C. Author is with the Electrical Engineering Department, University of Colorado, Boulder, CO 80309 USA, on leave from the National Research Institute for Metals, Tsukuba, Japan (e-mail: author@nrim.go.jp).

effects such as Kerr and scattering effects. QoT estimation must also take into consideration the behavior of various transmission equipment that vary widely in their performance according to their models, types and vendors. QoT estimation is generally performed using analytical models. An analytical model features a model of the fiber transmission also called a Physical Layer Model (PLM) that estimates the linear and nonlinear impairments. Analytical models vary in their accuracy, execution time and considered assumptions which always leaves room for improvement. ML is being used to improve QoT estimation in several creative ways such as building ML-based QoT estimators, modeling the response of optical equipment or estimating nonlinear effects.

Several papers in recent literature have proposed surveys of studies applying ML in optical networks [13][14][15]. In this paper, we focus specifically on papers that apply ML for the QoT estimation use case. Our goal is to understand and discuss the current state of the art and give a perspective for future work. To the best of our knowledge, only two surveys [10], [16] focused on the same scope. They both list ML based solutions that have been proposed in literature. While [10] focuses of the sources of uncertainty in QoT estimation, [16] focuses more on describing the proposed ML algorithms. Although our goal is similar, in this survey we also offer a network operator's point of view on the surveyed papers and we discuss the feasibility of the solutions in an operational context.

This paper is structured in the following manner. Section II gives an in-depth look into the QoT estimation problem from an analytical view. Section III gives an overview of papers that apply ML to estimate QoT in terms of ML models, used data and the obtained performances. Section IV discusses the challenges of the proposed solutions and their applicability in the real field.

II. ANALYTICAL QOT ESTIMATION

QoT estimation is traditionally performed using analytical models. An analytical model is built upon four elements: i) the transmission impairments taken into consideration, ii) the QoT indicator to be estimated, iii) the physical layer model (PLM) used to model the transmission and iv) the information about the transmission required as input. In this section, we provide an overview of each of these factors in order to show the current challenges of analytical QoT estimation.

Impairments in the transmission come either from the propagation of the signal through the fiber, or from the behavior of optical equipment. Fiber optical impairments can be split into linear and nonlinear effects. Linear effects include signal attenuation, Chromatic Dispersion (CD), Polarization Dependent Loss (PDL) and Polarization Mode Dispersion (PMD). Signal attenuation is rectified using optical amplifiers, however, this degrades the signal by adding Amplified Spontaneous Emission (ASE) noise. CD, PDL and PMD, on the other hand, are compensated using modern digital signal processing (DSP) techniques in the receiver. Nonlinear impairments are either due to the Kerr effects that include Self-Phase Modulation (SPM), Cross-Phase Modulation (CPM) and Four-Wave Mixing (FWM), or the inelastic scattering

phenomenon that includes Stimulated Brillouin-Scattering (SBS) and Stimulated Raman-Scattering (SRS). The scattering effects manifest themselves as a tilt in the spectrum which can be corrected using power equalizers. Kerr effects are usually modeled using equations that calculate the Power Spectral Density (PSD) of the signal such as the Schrödinger or Manakov equations [12]. Significant research effort is dedicated to the mitigation of nonlinear effects, and ML-based solutions have recently been proposed for this purpose [17].

In addition to propagation impairments, equipment generate impairments that contribute to signal degradation. The ASE noise generated by amplifiers significantly degrades the signal [18]. Additionally, wavelength and polarization dependent gain in the amplifiers introduces a tilt and ripple effect on signal spectrum. The impairments in a Reconfigurable Optical Add-Drop Multiplexer (ROADM) include PMD and PDL effects, insertion losses, ASE noise from internal amplifiers, filtering effects from imperfect filters and crosstalk effects between the channels. The impairments from the equipment are estimated by doing laboratory characterization, or by modeling each equipment analytically.

QoT is generally measured using either the Bit Error Rate (BER) or the Signal to Noise Ratio (SNR). SNR represents the ratio of the power of the optical signal to the noise contribution of all the optical impairments mentioned above. Linear SNR, referred to as Optical SNR (OSNR), is defined as the ratio of optical power of the signal P_{SIG} to optical noise added to the signal by optical amplifiers P_{ASE} ($OSNR = \frac{P_{SIG}}{P_{ASE}}$). The OSNR can be measured using an optical spectrum analyzer (OSA) [19], which is not possible for the SNR. BER, on the other hand, is a measure of the amount of errors in the received bits. A lightpath is considered healthy if its BER is above a certain threshold and the receiver's Forward Error Correction (FEC) module is able to correct the error in the bits.

SNR is computed before the deployment of a lightpath to check its feasibility taking into account optical impairments as well as various margins such as end of life margins and equipment aging [20]. BER is measured at the transceiver in real time, so it can be used to monitor the health of a lightpath. Before service deployment, BER cannot be estimated but deduced from the SNR once the modulation format, and the transceiver's back to back penalty are provided. This characterization is done by mapping the back-to-back OSNR to the BER response of the transceiver [21].

Analytical models are based on PLMs that attempt to model the propagation of the signal through the fiber medium. This generally comes down to estimating the PSD that is defined in the Schrödinger equation. Each PLM takes into consideration a number of impairments based on the assumptions taken into consideration. A large number of models have been proposed in literature [22].

We distinguish two families of QoT estimation analytical models. The first family consists of exact models that use comprehensive and extremely accurate methods. These models are heavy to execute and require a large number of parameters to model the transmission line. Therefore, they are more

suitable for laboratory simulations as their execution time and parameter requirements make them inconvenient to be used in the field. Among the models of this family, we find the Split-Step Fourier method [23] (SSFM) which is a numerical method of solving Schrödinger's equation by splitting the transmission into a succession of small linear and nonlinear steps. It is highly flexible and can be used to simulate network scenarios that have not yet been deployed. Its high computational requirements make this method unsuited for online QoT estimation.

The second family consists of approximate models that are able to estimate the QoT accurately provided a set of assumptions are satisfied. The most popular class of these models are the ones that consider nonlinear interference as a small perturbation of the signal. Among perturbation models, we find models based on truncated Volterra Series [24], logarithmic perturbation models [25] as well as Gaussian Noise (GN) models [26]. The light computational load makes these models more likely to be used in an operational context. The GN-model [26] for instance considers that the nonlinear interference in the fiber can be modeled as white Gaussian noise. It is based on three main assumptions [26]: i) nonlinear noise is a perturbation of the signal, ii) the transmitted signal statistically behaves as stationary Gaussian noise and iii) interference in the fiber is an additive Gaussian noise. These assumptions simplify the expression of the PSD defined by the Manakov equation (which itself is a simplification of the Schrödinger equation [27]). In GN-model, the SNR is redefined as the generalized SNR (GSNR) with $GSNR^{-1} = \frac{P_{sig}}{P_{NL} + P_{ASE}}$.

Several versions of the GN-model have been proposed such as the enhanced GN-model (EGN) [28] which removes the third assumption (iii), or the generalized GN-model (GGN) [29] which includes the SRS noise contribution. Analytical models based on the GN-model are fast to compute and have been experimentally demonstrated to have satisfying results [30]. However, their performance drops when the aforementioned assumptions fail (e.g. in highly nonlinear regimes).

III. ML-BASED QOT SOLUTIONS

In this section, we provide a survey of papers dealing with QoT estimation using machine learning. We distinguish four categories of solutions according to the scope of using ML: i) check lightpath feasibility, ii) estimate lightpath's QoT, iii) improve model generalization and iv) enhance analytical models.

Tables I, II, III and IV give an overview of the particularities of each proposed algorithm based on a set of criteria:

- *algorithm*: the ML techniques used by the solutions.
- *input*: the selected features to feed the machine learning model.
- *output*: the output of the ML model. This is generally the QoT indicator that the study seeks to compute.
- *source of data*: information about how data was collected (i.e., synthetic, simulated or experimental or real data).

- *results*: the key results from the study. As most of the papers provide results from multiple experiments, we choose to only mention the most relevant results.

A. ML based models to check lightpath feasibility

The goal of a ML classification model is to attribute a class to each data entry composed of a combination of features. In the case of QoT estimation, ML classification models are used in literature to decide if a candidate lightpath is feasible or not based on a set of optical parameters. The classes in this case are generally binary: the lightpath's QoT indicator is beyond a predefined threshold or not. Thus, ML is used as a simple decision tool for lightpath deployment. Table I lists the characteristics of surveyed papers that fall into this category. The performance of a classification model is usually represented by the accuracy score; the ratio of correctly classified lightpaths to the total number of lightpaths.

A case-based reasoning (CBR) approach is proposed in [31] to classify lightpaths based on a Q-factor threshold, proving that only a simple classification model is required to achieve high accuracy scores up to 98.7%. More conventional ML models like Support Vector Machines (SVM) and random forest, are used in [32][33][34][35][36] and [37]. These two models usually have the best performance. Authors in [38] uses a more complex ML model based on a deep graph Convolutional Neural Network (CNN) to model inter-channel interferences in multi-core fibers which can classify lightpaths with up to 97% accuracy.

Most studies use end-to-end line features, such as total lightpath length and number of spans. However, authors in [33] prove that using additional features from neighboring channels improve the classification results. In [36], the statistical representations of the features are calculated and used in the classification in order to reduce the number of features without losing information. Almost all studies use the BER as a QoT indicator and choose the FEC limit as the threshold to separate the two classes (i.e., feasible/unfeasible lightpath). Authors in [33] additionally provide a degree of certainty to the classification, which can be used to choose between multiple feasible lightpaths.

B. ML based models to estimate a lightpath's QoT

In this category, the scope of the proposed solutions is to estimate the precise value of a QoT indicator. Therefore, ML-regression models are used. This scope is more challenging than the previous one since the model output space is continuous. Having the exact QoT value allows to compare two feasible potential lightpaths. Regression models are usually scored using an error operator, such as Root Mean Squared Error (RMSE) or Mean Absolute Error (MAE). MAX and MIN errors are also used, as they allow to set network margins. Table II lists the surveyed papers that belong to this category.

The features used in regression models are more diverse than those used in classification. Basic line features, such as the number of spans or length of the link, are always used, noticeably in [39] and [40]. Additionally, authors in [41]

1 consider per-channel features for instance input power and
2 amplifier gain for each channel in the link. The per-channels
3 features are generally flattened into a single vector, except in
4 the case of [42], where a two dimensional CNN is used to obtain
5 a one-vector representation of all the features. Furthermore,
6 [41] [43] use features from multiple points of the line, such as
7 input power at each amplifier.

8 In terms of QoT indicator, we notice that noise based QoT
9 indicators (SNR and OSNR) are more frequently used than
10 BER. Authors in [40] propose to estimate the distribution of the
11 QoT indicator rather than just a single value. Authors in [41]
12 compare the estimation of a ML model to that of an analytical
13 model and show that ML models perform better than analytical
14 ones for lightpaths in the edge of the spectrum.

15 While the majority of the surveyed papers are based on
16 synthetic and experimental data, a couple of papers have also
17 used real data. Noticeably, authors in [44] showcase a model
18 trained on synthetic data then tested on real network data.

20 C. ML models to enhance analytical models

21 The goal of the first and second categories was to build a
22 standalone ML estimator, while the goal of this category is to
23 use ML models and analytical models in tandem. This means
24 that ML is used to improve the accuracy of analytical models
25 instead of replacing them. This can be achieved by, either
26 improving the accuracy of input parameters of an analytical
27 model (i.e., reducing the uncertainty on input parameters) or
28 assessing hard to compute impairments or physical coefficients.
29 Studies focusing on “*Reducing uncertainty on parameters*”
30 justify the usefulness of their approach by the fact that some
31 parameters values are not up to date in the operators’ databases
32 although they undergo changes due to multiple factors (e.g.,
33 temperature and equipment aging). This could be due to the
34 inability of equipment to measure these parameters (e.g. the
35 fiber nonlinear coefficient) or the inability of the monitoring
36 protocols deployed between the equipment and network
37 management system (NMS) to communicate parameters values
38 in real-time. As an alternative, fixed values like
39 design/beginning-of-life values are used to compute QoT [45].
40 In literature, optimization algorithms such as gradient descent,
41 are generally used to reduce uncertainty in parameters. The
42 objective function of the optimization is generally set to the
43 difference between measured and estimated QoT indicator (also
44 called QoT error). Then, the values of a set of uncertain
45 parameters are iteratively changed until the QoT error is
46 minimal. The amplifier’s noise figure is the most commonly
47 considered uncertain parameter [46]–[48]. The performance of
48 the solution is assessed according to the ability of the algorithm
49 to improve the QoT estimation of the analytical model and
50 reduce the QoT error. This performance closely depends on the
51 number of uncertain parameters assumed in the experiments.

52 Authors in [46] propose two different approaches for QoT
53 estimation. The first is a purely ML-based estimator that
54 assesses SNR value. The second approach iteratively reduces
55 the uncertainty of the parameters required as input of an
56 analytical model in order to improve the model’s QoT
57 estimation. An elaborate closed-loop controller architecture for
58 optical networks is proposed [49] in order to use feedback from

network measurements to improve the accuracy of the ML-
model by reducing parameter uncertainty.

To improve QoT estimation models, other studies focused on
estimating optical parameters that represent hard to compute
impairments. This means to assess a certain coefficient required
by the analytical model such as the nonlinear coefficient in GN
models [26] or by modeling the behavior of an equipment for
instance amplifier’s ripple [65]. In Table IV, we survey the
papers that follow one of these two approaches. Authors in [50]
and [51] attempt to estimate the nonlinear SNR. Thus, a neural
network model is proposed in [50] to directly estimate the
nonlinear noise from line features and fiber characteristics.
Authors in [51] propose to mix line features with covariance
coefficient calculated from DSP constellations, as well as the
output of an analytical model. They show that the nonlinear
SNR estimation can be enhanced by feeding all this information
as an input of an artificial neural network (ANN). Papers [52]
and [53] are focused on modeling the effect of different spectral
loads on amplifiers. The aim is to estimate the SNR and OSNR
taking into consideration only the impairments generated by the
amplifier. This estimation is used alongside an analytical model
in order to estimate the overall QoT. In a similar vein, authors
in [54] propose a ML model to estimate the SNR of a lightpath.
The ML model is used when the total length of the lightpath is
inferior to 200 km, otherwise, the GNLI is used.

D. ML techniques to improve model generalization

Model generalization refers to the ability of the model to adapt
to data with different distributions. A ML model trained on data
extracted from a specific network would not necessarily
perform similarly using another network dataset. In order to
improve model generalization, more diverse datasets are
needed to train the model which is not always possible due to
the lack of datasets extracted from different networks.
Techniques such as transfer learning and active learning are
generally used to resolve this issue. In Table III, we have
surveyed papers that try to improve ML model performance
using these techniques. It is worth noting that each of these
papers propose their own QoT estimation model which means
that they can also fit into the category A or B. However, we
choose to survey them in this category as we consider their main
contribution is to solve the model generalization problem. In the
characteristics, we provide both the ML algorithm used for QoT
estimation and the ML algorithm used for model generalization.
The performance is assessed according to the ability to reduce
the dataset requirements or to improve error rates.

The goal of Transfer Learning (TL) algorithms is domain
adaptation, which means ensuring that a ML model is
generalizable to multiple datasets with different feature
distributions. Pesic et al. study in [55] the impact of using
networks with different span lengths on the performance of
ML-based solution. They show the importance of pre-training
a model on unbiased data. This study is extended in [56] to
include more network features. Authors also study the effect of
domain adaptation on the structure of an ANN model. In [57],
[58] and [55], authors use a ML model initially trained on data

from a network A in another network B. To adapt the model, a dataset from network B, up to 50 times smaller than that of network A is used for retraining.

Active Learning (AL) approaches are proposed in [59] and [60]. AL algorithms seek to reduce the dataset size by selecting data that best improve model performance. The method typically starts with an initial small dataset, then progressively adds more data entries using an algorithm to compute and rank the importance of each data point. Only the selected data points are measured which avoid the need of a large dataset from the

beginning. Both papers use a Monte Carlo based process for data selection. Authors in [60] also propose three domain adaptation techniques, namely Bayesian updating, feature augmentation, correlation alignment, then provide an extensive benchmark of each method.

TABLE I
COMPARISON OF PAPERS PROPOSING A ML MODEL TO CHECK LIGHTPATH FEASIBILITY

| Ref | Algorithm | Inputs | Outputs | Data Source | Performance |
|------|--|---|---|--------------|---|
| [31] | Case based reasoning | Channel wavelength, the launch power, loss per span, number of spans, active lightpaths, total input power, total power of the adjacent channels | Q-factor | Experimental | Accuracy between 79% and 98.7% |
| [37] | SVM | Number of ROADMs, of links, of fiber spans; length of fiber span, launch channel power | OSNR | Synthetic | Accuracy >95% |
| [61] | ANN | Total length, max link length, central frequency, number of allocated slots, modulation format, number of amplifiers, number of links | BER based on slice requirement | Simulated | Accuracy >90%, depends on slices and class per experiment |
| [36] | Logistic regression, Decision trees, SVM, random forest, xgboost | Hop lengths, number of channels, hop losses, number of hops, modulation format, bitrate, aggregation-based feature engineering | 2 classes: bad and good configuration | Simulated | Best performance at 0.9 AUC |
| [32] | KNN, logistic regression, SVM, ANN | Number of hops, number of spans, total length, average link length, maximum link length, average attenuation, average dispersion, modulation format | SNR above threshold | Simulated | Accuracy >95% |
| [33] | KNN, Random Forrest | Lightpath length, longest link length, number of links, traffic volume, modulation format, left/right guardband, left/right traffic volume, left/right modulation format. | P_{pos} probability that the BER of the lightpath exceeds a predefined threshold | Synthetic | Accuracy up to 96% on certain topologies |
| [34] | SVM, ANN | Total link length, span length, launch power, modulation format, data rate | SNR | Synthetic | Accuracy >99% |
| [35] | SVM, logistic regression, CART, random forest | Lightpath length, link lengths, wavelength, statistics on co-propagating light paths | BER | Simulated | Best performance at 99.9% accuracy |
| [38] | Deep graph convolutional neural network | Channel adjacency matrix, lightpath length, max link length, central frequency, number of slots, core identifier, modulation format, number of amplifiers, number of links, BER of the deployed lightpaths. | BER classification based on threshold | Synthetic | Accuracy rates between 92% and 97% |

TABLE II
COMPARISON OF PAPERS PROPOSING A ML MODEL TO ESTIMATE A LIGHTPATH'S QOT METRIC

| Ref | Algorithm | Inputs | Outputs | Data Source | Performance |
|--------------|---|---|-------------------|--|---|
| [41] | ANN | Per channel launch power, per amplifier [gains, NFs, gain tilts], per span input power | SNR | Experimental | SNR standard deviation < 0.14 dB |
| [43] | ANN | Launch power, laser bias, per amplifier [input power, output power] | Q-factor | Experimental | Q RMSE < 0.02 dB |
| [40] | Random forest | Number of links, total length, max link length, traffic volume, modulation format | GNSR distribution | Synthetic | Best performance: 0.02 RMSE |
| [62] | Decision tree, random forest, MLP | Received signal power, NLI, ASE, channel frequency, total length | GSNR | Synthetic | Best performance: 0.16 dB in average prediction error |
| [63] | ANN | Per channel [power, frequency], number of spans, analytical model output [ASE, nonlinear noise] | SNR | Simulated | Max error < 0.5 dB |
| [64] [65] | Gaussian process | Wavelength, measured OSNR, OSNR noise | OSNR | Field trial testbed | MSE < 0.7 dB |
| [44] | ANN (ReLU, SeLU) | P_{tx} CUT, R_{sym} CUT, $4x$ [fPtx, R_{sym} , Df], total used bandwidth, number of WDM channels, number of spans, span length parameters, average power level, N_{ch} around CUT | SNR | Trained on synthetic and applied real data | Max error < 0.5 dB |
| [66] | ANN | Channel, noise power on each link | OSNR | Experimental | Average error < 0.5 dB* |
| [67] | ANN | | OSNR | Experimental | RMSE < 0.2 dB |
| [42] | CNN for feature extraction + ANN for prediction | Per channel [Power, ASE, NLI, number of spans, total length] | GSNR | Synthetic | Maximum error = 0.37 dB * |

| | | | | | |
|------|-------------------------|---|---------------|----------------------------|------------------------|
| [39] | GPR | Input power, number of spans, baud rate, interchannel spacing | BER, Q-factor | Simulated and experimental | Average error < 0.3 dB |
| [68] | Random Forest, ANN, KNN | Distance, number of spans, ASE noise, nonlinear interference, power | GSNR | Synthetic | MAE score < 0.007 |

TABLE III
COMPARISON OF PAPERS ATTEMPTING TO ENHANCE ANALYTICAL MODELS THROUGH MACHINE LEARNING

| Ref | Algorithm | Inputs | Target parameter | Data Source | Performance |
|---|---|---|---|--------------|--|
| <i>Reducing uncertainty on parameters</i> | | | | | |
| [69] | Gradient descent | SNR | Noise figure, power profile | Synthetic | Design margins reduced from few dBs to 0.1 dB |
| [70] | 3 step optimizations based on Gradient descent | Per OMS power estimation, end to end SNR estimation | Lumped loss before/after fiber spans, amplifier power ripple | Experimental | SNR estimation improved from 2dB to 0.2dB |
| [46] | Custom linear regression | Target Q-factor value, initial noise figure value, initial nonlinear coefficient value. | Noise figure, nonlinear coefficient | Experimental | Error reduced from 1.4 dB to 0.6 dB |
| [47] | Gradient descent | Target SNR value, initial noise figure value, initial input power value. | Noise figure, input power | Synthetic | Error reduced by up to 1.9 dB and 4.18 dB compared with SAMBA and EGN respectively |
| [49] | Nonlinear curve fitting | Target SNR value, initial values of attenuation, dispersion and non-linear coefficients. | Attenuation, dispersion and non-linear coefficients | Synthetic | Design margin reduction is up to 1.95 dB |
| [48] | Metropolis algorithm | OSNR estimation | Nonlinear distortion coefficient, filter wavelength detuning, amplifier gain, amplifier noise figure NF | Experimental | Error reduction from 3.7dB to 0.5dB. |
| <i>Impairment modeling</i> | | | | | |
| [52] | ANN | Spectral load | OSNR | Experimental | Average error < 0.2dB |
| [50] | ANN | Noise covariance, number of spans, max span length, average power, launch power, link length, chromatic dispersion, average fiber gamma, average fiber alpha, number of channels. | Nonlinear SNR | Synthetic | 0.33 dB of SNR _{nl} deviation using combination of all features |
| [51] | ANN | Fiber attenuation, dispersion coefficient, effective area, non-linear refractive index, analytical model [output span length, number of active channels, launch power, channel bandwidth and frequency] | Nonlinear SNR | Simulated | Error below 0.5 dB for 99% of cases |
| [53]* | SVM | Per channel power, optical spectrum, pre-FEC BER | SNR | Synthetic | Average error < 0.2 dB |
| [54] | Linear Regression, Multivariate polynomial Regression, DT, Random Forest, SVM, KNN, ANN | Number of fiber spans, span length, channel bandwidth, guard band, number of channels, channel power | SNR (when lightpath length < 200km) | Synthetic | Decreased the cases where the absolute error was higher than 2 dB from 2.30% to 0.47%. |

TABLE IV
COMPARISON OF PAPERS FEATURING MACHINE LEARNING TECHNIQUES FOR DATASET MANAGEMENT

| Ref | Algorithm | Inputs | Outputs | Data Source | Performance |
|------|---|---|--------------------------------------|--------------|--|
| [59] | ANN, Active learning through MC dropping for uncertainty sampling | Signal bandwidth, modulation format, Vpp, ROP | Generalized mutual information (GMI) | Experimental | Requires 25% less data than random sampling, while maintaining blow 0.055 in MSE |
| [71] | DNN, Evolutionary transfer learning | Power profile at each WSS | Q-Factor | Experimental | Only 10% data size required for retraining |
| [72] | DNN | Post ADC constellation | SNR | Experimental | Average Error < 0.2dB |
| [60] | Gaussian process, Active learning using a MC method, Domain adaptation using Bayesian updating, feature augmentation, Correlation alignment | total lightpath length, longest link length, number of traversed links, traffic volume, and modulation format | SNR | Synthetic | Depends on method and dataset size |

| | | | | | |
|------|-----|--|--------------------|--------------|--|
| [55] | ANN | Span lengths | SNR | Simulated | RMSE improved by 2 after retraining |
| [56] | DNN | Q-factor of different lines | Q-factor | Experimental | 50% less dataset size for retraining |
| [58] | SVM | Total length, number of links, maximum link length, demand capacity, modulation format | BER classification | Synthetic | 20 times less data required for retraining |
| [57] | DNN | Amplitude histogram of received samples | OSNR | Experimental | 20 times less data required for retraining |

IV. DISCUSSION

In this section, we discuss the ML-based solutions described in the previous section. Through this discussion, we present the challenges of these solutions and explore their feasibility in the context of operational networks. We also propose some perspectives for future work. Our discussion focuses on four main points: i) the relevance of the tackled scopes, ii) the dataset used for training, iii) the feature and model choice and iv) the evaluation metrics.

A. Adopted approaches

The four categories that we have identified to classify the solutions proposed in the literature tackle the use of ML to estimate QoT in different manners. However, each solution has its challenges that make it more or less useable in an operational context.

By proposing an ML-based QoT estimator, whether it is a classification or a regression model (Categories A and B), the objective is to provide an alternative to analytical models. In order to justify the substantial efforts needed to train the model correctly and the extra-costs to deploy such solution, especially when it comes to data collection, ML-based solution must outperform the analytical models in terms of QoT-estimation precision, number of required input parameters or/and computational load. Guaranteeing, at least, the performance achieved by analytical model seems to be feasible because the capacity of ML-solutions to model complex physical phenomenon has been extensively proven not only in optics but also in various areas of research [7]. Reducing the execution time is also not a concern, because outside of the training phase, model inference is generally not computationally extensive. In [44], ANN prediction takes microseconds, compared to minutes using a full EGN model. Additionally, ML models may use feature representations with fewer input parameters, for instance by averaging the list of span lengths as in [32]. ML models can also be more robust to uncertain parameters as featured in [73]. However, the challenge facing solutions involved in these categories (i.e., A and B) is to build a model that is effective, easy to maintain and works on all the scenarios where it is applied with acceptable deployment extra-cost.

The choice between classification and regression depends on the operational requirement. If only checking the feasibility of an optical path is needed, classification is better. However, if the value of the QoT is needed (e.g., to compare the performance of different models of equipment), the regression model will be the best option. For both methods, it is useful, from an operational point of view, to additionally assess the

estimation error in order to facilitate the computation of operational margins, which might be easier to compute with a regression approach.

The idea behind the category C is to use ML in tandem with analytical models. Analytical models are already used extensively in the field, so proposing solutions to improve their usability is better than replacing them. In fact, missing or inaccurate parameter values is a recurring issue in data extracted from operational networks as outlined in [74]. Studies that focus on reducing the inaccuracy of parameters values have mostly focused on optimization approaches using the error in QoT estimation between measurements and the analytical model as an objective function [69]. The solutions proposed within this scope risk not being able to converge towards the real values because some impairments that are not taken into consideration by the analytical model may contribute to the error in QoT estimation. In this case, alternative values of the input parameters are provided by the ML algorithm which improves the accuracy of the QoT, although the values of input parameters do not correspond to the real values [75].

As analytical models are showing promising results with decent rapidity [45], we believe that solutions that aim to assist analytical models (i.e., Category C) are pragmatic and could provide promising results in short/medium terms. For solutions of categories A and B, we consider that the main benefit of using a ML estimator rather than an analytical one could be in one of these cases: i) the ML estimator can achieve accurate performance when the analytical model is not appropriate because its assumptions are not satisfied [54], ii) some input parameters of analytical model could not be provided whereas, they are not needed by ML-based solution (e.g. a better feature representation in [46] leads to reduce the number of input parameters), or iii) the analytical model is unusable due to execution time constraints.

The interest of improving the generalization of ML models (Category D) is justified by the issues around data collection in optical networks. Models proposed within this category do not focus only on computing the QoT itself but to make models proposed under the other categories technically feasible in terms of data and able to be generalized in many scenarios. Concepts presented in this category could be applied to any ML-model that requires training. We believe that this approach will be the key to making ML models a viable solution in the field, especially when it comes to domain adaptation between heterogeneous networks as in [60]. Another point of interest that was rarely addressed in papers is continuous learning. Operational networks are susceptible to change, which might make ML models inefficient if they are not updated regularly. Therefore, proposing a closed-loop process, based for example on active learning as in [76] would help to solve this problem.

B. Dataset collection

Training ML models requires large datasets. Some approaches, do not have a proper training phase, as in the case of solutions based on optimization algorithms such as [46] and [47]. However, even in this case QoT indicator measurements are still required. In general, the performance of ML methods is tightly linked to the quality of the dataset used in the training phase. The different steps to build the dataset: data collection, data annotation, feature engineering, data augmentation and splitting datasets for training and validation, must be carefully performed in order to ensure a successful training and avoid biased results.

Data collection remains the first bottleneck in optical networks due to several factors: equipment lock-in (i.e., inability to access equipment data), lack of standardized data models and monitoring protocols, lack of data collection and monitoring tools and the cost of deploying optical signal probes in the network. Datasets used in the surveyed papers are either synthetic, experimental or operational. Synthetic datasets are generated by simulating network scenarios using an analytical model, for instance, authors in [40] and [68] use the GNPpy tool [77]. This method allows greater control over data entry points, feature variation, flexibility in the definition of the scenario and setting the dataset size. However, models trained with synthetic data learn the behavior of the simulation platform and the analytical model behind which might not faithfully represent a real optical network behavior. In fact, the performance of the ML in this case is tightly related to the accuracy of the simulation platform. Transfer learning approaches proposed in Section III.D could solve this problem by retraining the model using field data in order to increase its accuracy and remove the synthetic data bias, but, to our knowledge, this has not yet been demonstrated in the literature.

Experimental setups on the other hand better reproduce the conditions of an operational network, while keeping the flexibility of simulation approaches. For example, in [52], a setup consisting of cascaded amplifiers is used to model amplifier response to spectral load, while in [41], an experimental setup of a full transmission line is used. Through experimental data, models can learn the behavior of a real transmission using physical fibers and equipment. Generating this kind of data is costly and time-consuming given the large amount of measurements required to train an ML model. Therefore, a full automation of the experimental setup, as proposed in [78], is highly recommended. Experimental setup usually includes a small-scale network. Thus, applying the ML model in a large-scale operational network requires the use of adaption and generalization techniques.

Training ML algorithms with operational data confronts the model to the real condition of the field. However, the data collection process in this case is complex due to the lack of monitoring and data extraction tools in the optical layer and the inability to define on-demand data extraction scenario. For instance, data cannot be extracted from unfeasible or low QoT lightpath and equipment configuration settings cannot be changed for training purposes. Moreover, operational datasets are less diverse in terms of features availability and variation. In fact, these datasets are tightly linked to the network from which they were collected. Since feature distributions could

change from one network to another, transfer learning could be applied to generalize the model between heterogeneous networks (or domains) as shown in [56].

We notice that most papers use synthetic or experimental data to train ML models. Only a few studies have used data from an operational network such as [44]. [64] and [65] use data extracted from field trial testbed with a total 436.4 km optical path over the national dark fiber facility in UK. The choice of a data source requires a balance between simulation flexibility and representation. The ideal scenario is to have enough variation in field data to train the models correctly. But since this is far from being immediately achievable, we consider that it is more convenient to train models with synthetic or experimental data mixed with samples of operational data to generalize the model's performance.

Data collection concerns hamper research focused on the application of ML in optical networks. While it is justified to adopt an optimistic outlook and assume that the data scarcity will be resolved in the future, we believe that it will be more beneficial to actively tackle the problem by proposing detailed data collection schemes alongside the ML solutions as in [64], or by only taking into consideration parameters that are available in the field as features as in [75].

C. ML models and input features

The choice of input features and ML model is motivated primarily by the scope of the solution. If we assume that the goal is to estimate a QoT indicator, the input feature of the model must fully describe the factors that impact this indicator. Similarly, the ML model must be sufficiently complex to model the impact of said features on the indicator. In the surveyed papers, we find different levels of features. The first level concerns the end-to-end lightpath features, such as total length of the lightpath, and number of hops/spans [37]. These features are often related to the lightpath under study rather than considering the entire network. The next level include data related to specific equipment/fiber through the path such as the attenuation coefficient of the fiber spans, or the amplifier gains in [41]. The third level concern information about the co-propagating lightpaths (i.e., neighboring wavelengths). These features can range from a simple number of wavelengths to a detailed description of the spectral load [52]. The fourth level is to use a feature representation of the whole network and its lightpaths. This can be modeled using matrixes or graphs as in [38]. Feature representation of the network can also be provided with other information depending on the use case: information about slices [61], calculated features from DSP constellations [44], or analytical model output [51]. Setting the level of details in feature depends on the impairments to be considered and the level of precision to be achieved by the ML-model. For instance, study [38] proposes a graph based feature representation of all the lightpaths in the networks (i.e., fourth level of feature representation) because they aim to take into consideration inter-core crosstalk effects between all the lightpaths.

D. Performance and evaluation metrics

Several metrics exist to assess the performance of ML models. The choice of metric depends primarily on the type of model (regression or classification), and on the performance to evaluate. In the case of classification, some studies such as [61] use the accuracy metric, which gives the rate of successful classifications but does not give information on false positives or false negatives. In the case of regression, a Root Mean Squared (RMSE) or/and Mean Absolute Error (MAE) operator are used [40][68]. Using a varied list of evaluation metrics assesses better the model's performance, for instance, providing AUC (Area Under the Curve) scores for classification as in [36], and error distribution for regression as in [40]. These evaluation metrics allow to precisely know the model capability which could be helpful for some operational settings like the specification of network margins.

Furthermore, careful attention must be given to dataset biases, such as a dataset with higher percentage of a class over another. Multiple papers have used cross validation to detect overfitting problems like in [68] and [33]. But the best approach is to validate the performance on completely different dataset such as in [44].

In addition to the accuracy of QoT estimation, other evaluation metrics such as network capacity gain [49] or potential resource saving [33] are relevant to show the added-value of the proposed solution. Nevertheless, the assessment of resource saving should also take into consideration the extra cost of deploying the ML solution such as the probes for data monitoring. Finally, a comparison must be established with existing solutions, especially with analytical models in the case of ML-based QoT estimators as in [54].

V. CONCLUSION

Using machine learning to improve the QoT estimation has seen a surge in popularity over the last years. Therefore, we provide in this paper a survey on studies that tackle this research topic from different angles. We distinguish four categories of models using ML for QoT estimation. The first category consists of building ML model to check the feasibility of a path. The second category aims to make the ML-based model as full alternative of analytical models. The third category uses ML to improve analytical models by either reducing the uncertainty on input parameters, modeling equipment or assessing hard to compute impairments or coefficients to supplement the analytical models. The last category consists of improving the performance and generalization ability of ML-based solution by enhancing the samples of the dataset in the training phase through transfer learning or active learning techniques.

While the results of the proposed algorithms are generally satisfactory, some concerns remain regarding their ability to be generalized in order to support complex optical transport network topologies and various equipment configurations. Moreover, the data scarcity and additional cost related to monitoring data and implementing these solutions are among the challenges that hinder the deployment of ML-based QoT estimator in the operational networks.

REFERENCES

- [1] M. Channegowda, R. Nejabati, and D. Simeonidou, "Software-Defined Optical Networks Technology and Infrastructure: Enabling Software-Defined Optical Network Operations [Invited]," *J. Opt. Commun. Netw.*, vol. 5, no. 10, p. A274, Oct. 2013, doi: 10.1364/JOCN.5.00A274.
- [2] V. Lopez *et al.*, "Transport API: A Solution for SDN in Carriers Networks," p. 3.
- [3] M. Birk *et al.*, "The OpenROADM initiative [Invited]," *J. Opt. Commun. Netw.*, vol. 12, no. 6, p. C58, Jun. 2020, doi: 10.1364/JOCN.380723.
- [4] S. Oda *et al.*, "A Learning Living Network With Open ROADMs," *J. Lightwave Technol.*, vol. 35, no. 8, pp. 1350–1356, Apr. 2017, doi: 10.1109/JLT.2017.2660540.
- [5] M. Dallaglio, N. Sambo, F. Cugini, and P. Castoldi, "Control and Management of Transponders With NETCONF and YANG," *J. Opt. Commun. Netw.*, vol. 9, no. 3, p. B43, Mar. 2017, doi: 10.1364/JOCN.9.000B43.
- [6] F. Paolucci, A. Sgambelluri, F. Cugini, and P. Castoldi, "Network Telemetry Streaming Services in SDN-Based Disaggregated Optical Networks," *J. Lightwave Technol.*, vol. 36, no. 15, pp. 3142–3149, Aug. 2018, doi: 10.1109/JLT.2018.2795345.
- [7] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi, "A survey of deep neural network architectures and their applications," *Neurocomputing*, vol. 234, pp. 11–26, Apr. 2017, doi: 10.1016/j.neucom.2016.12.038.
- [8] I. Martin *et al.*, "Machine Learning-Based Routing and Wavelength Assignment in Software-Defined Optical Networks," *IEEE Trans. Netw. Serv. Manage.*, vol. 16, no. 3, pp. 871–883, Sep. 2019, doi: 10.1109/TNSM.2019.2927867.
- [9] M. Salani, C. Rottondi, and M. Tornatore, "Routing and Spectrum Assignment Integrating Machine-Learning-Based QoT Estimation in Elastic Optical Networks," in *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications*, Paris, France, Apr. 2019, pp. 1738–1746. doi: 10.1109/INFOCOM.2019.8737413.
- [10] Y. Pointurier, "Machine learning techniques for quality of transmission estimation in optical networks," *J. Opt. Commun. Netw.*, vol. 13, no. 4, p. B60, Apr. 2021, doi: 10.1364/JOCN.417434.
- [11] T. Ahmed and M. Coates, "Machine Learning Algorithms for Anomaly Detection in Optical Networks," p. 1.
- [12] G. P. Agrawal, *Fiber-optic Communication Systems (Wiley series in microwave and optical engineering)*. John Wiley & Sons Incorporated.
- [13] F. N. Khan, Q. Fan, C. Lu, and A. P. T. Lau, "An Optical Communication's Perspective on Machine Learning and Its Applications," *J. Lightwave Technol.*, vol. 37, no. 2, pp. 493–516, Jan. 2019, doi: 10.1109/JLT.2019.2897313.
- [14] F. Musumeci *et al.*, "An Overview on Application of Machine Learning Techniques in Optical Networks," *arXiv:1803.07976 [cs, stat]*, Dec. 2018, Accessed: Dec. 05, 2019. [Online]. Available: <http://arxiv.org/abs/1803.07976>
- [15] R. Boutaba *et al.*, "A comprehensive survey on machine learning for networking: evolution, applications and

- research opportunities,” *J Internet Serv Appl*, vol. 9, no. 1, p. 16, Dec. 2018, doi: 10.1186/s13174-018-0087-2.
- [16] L. Zhang, X. Li, Y. Tang, J. Xin, and S. Huang, “A survey on QoT prediction using machine learning in optical networks,” *Optical Fiber Technology*, vol. 68, p. 102804, Jan. 2022, doi: 10.1016/j.yofte.2021.102804.
- [17] C. Catanese, A. Triki, E. Pincemin, and Y. Jaouen, “A Survey of Neural Network Applications in Fiber Nonlinearity Mitigation,” in *2019 21st International Conference on Transparent Optical Networks (ICTON)*, Angers, France, Jul. 2019, pp. 1–4. doi: 10.1109/ICTON.2019.8840355.
- [18] S. Donati and G. Giuliani, “Noise in an optical amplifier: formulation of a new semiclassical model,” *IEEE J. Quantum Electron.*, vol. 33, no. 9, pp. 1481–1488, Sep. 1997, doi: 10.1109/3.622626.
- [19] D. Garipey, S. Searcy, G. He, S. Tibuleac, M. Leclerc, and P. Gosselin-Badaroudine, “Novel OSNR Measurement Techniques Based on Optical Spectrum Analysis and Their Application to Coherent-Detection Systems,” *J. Lightwave Technol.*, vol. 37, no. 2, pp. 562–570, Jan. 2019, doi: 10.1109/JLT.2018.2878744.
- [20] Y. Pointurier, “Design of Low-Margin Optical Networks,” *J. Opt. Commun. Netw.*, vol. 9, no. 1, p. A9, Jan. 2017, doi: 10.1364/JOCN.9.0000A9.
- [21] G. Grammel, V. Curri, and J.-L. Auge, “Physical Simulation Environment of The Telecommunications Infrastructure Project (TIP),” in *Optical Fiber Communication Conference*, San Diego, California, 2018, p. M1D.3. doi: 10.1364/OFC.2018.M1D.3.
- [22] P. Poggiolini and Y. Jiang, “Recent Advances in the Modeling of the Impact of Nonlinear Fiber Propagation Effects on Uncompensated Coherent Transmission Systems,” *J. Lightwave Technol.*, vol. 35, no. 3, pp. 458–480, Feb. 2017, doi: 10.1109/JLT.2016.2613893.
- [23] Ya. I. Bogomolov and A. D. Yunakovsky, “Split-step Fourier method for nonlinear Schrodinger equation,” in *DAYS on DIFFRACTION 2006*, St.Petersburg, Russia, 2006, pp. 34–42. doi: 10.1109/DD.2006.348170.
- [24] K. V. Peddaranappagari and M. Brandt-Pearce, “Volterra series transfer function of single-mode fibers,” *J. Lightwave Technol.*, vol. 15, no. 12, pp. 2232–2241, Dec. 1997, doi: 10.1109/50.643545.
- [25] M. Secondini, E. Forestieri, and C. R. Menyuk, “A Combined Regular-Logarithmic Perturbation Method for Signal-Noise Interaction in Amplified Optical Systems,” *J. Lightwave Technol.*, vol. 27, no. 16, pp. 3358–3369, Aug. 2009, doi: 10.1109/JLT.2009.2012873.
- [26] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, “The GN-Model of Fiber Non-Linear Propagation and its Applications,” *J. Lightwave Technol.*, vol. 32, no. 4, pp. 694–721, Feb. 2014, doi: 10.1109/JLT.2013.2295208.
- [27] D. Marcuse, C. R. Menyuk, and P. K. A. Wai, “Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence,” *J. Lightwave Technol.*, vol. 15, no. 9, pp. 1735–1746, Sep. 1997, doi: 10.1109/50.622902.
- [28] A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, “EGN model of non-linear fiber propagation,” *Opt. Express*, vol. 22, no. 13, p. 16335, Jun. 2014, doi: 10.1364/OE.22.016335.
- [29] D. Semrau, R. I. Killey, and P. Bayvel, “The Gaussian Noise Model in the Presence of Inter-Channel Stimulated Raman Scattering,” *J. Lightwave Technol.*, vol. 36, no. 14, pp. 3046–3055, Jul. 2018, doi: 10.1109/JLT.2018.2830973.
- [30] A. Nespola *et al.*, “Extensive Fiber Comparison and GN-model Validation in Uncompensated Links using DAC-generated Nyquist-WDM PM-16QAM Channels,” in *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2013*, Anaheim, California, 2013, p. OTh3G.5. doi: 10.1364/OFC.2013.OTh3G.5.
- [31] A. Caballero *et al.*, “Experimental demonstration of a cognitive quality of transmission estimator for optical communication systems,” p. 3, 2012.
- [32] R. M. Morais and J. Pedro, “Machine Learning Models for Estimating Quality of Transmission in DWDM Networks,” *J. Opt. Commun. Netw.*, vol. 10, no. 10, p. D84, Oct. 2018, doi: 10.1364/JOCN.10.000D84.
- [33] C. Rottondi, L. Barletta, A. Giusti, and M. Tornatore, “Machine-Learning Method for Quality of Transmission Prediction of Unestablished Lightpaths,” *J. Opt. Commun. Netw.*, vol. 10, no. 2, p. A286, Feb. 2018, doi: 10.1364/JOCN.10.00A286.
- [34] S. Aladin, A. V. S. Tran, S. Allogba, and C. Tremblay, “Quality of Transmission Estimation and Short-Term Performance Forecast of Lightpaths,” *J. Lightwave Technol.*, vol. 38, no. 10, pp. 2807–2814, May 2020, doi: 10.1109/JLT.2020.2975179.
- [35] J. Mata *et al.*, “Supervised Machine Learning Techniques for Quality of Transmission Assessment in Optical Networks,” in *2018 20th International Conference on Transparent Optical Networks (ICTON)*, Bucharest, Jul. 2018, pp. 1–4. doi: 10.1109/ICTON.2018.8473819.
- [36] S. Kozdrowski, P. Cichosz, P. Paziewski, and S. Sujecki, “Machine Learning Algorithms for Prediction of the Quality of Transmission in Optical Networks,” *Entropy*, vol. 23, no. 1, p. 7, Dec. 2020, doi: 10.3390/e23010007.
- [37] A. A. Diaz-Montiel, S. Aladin, C. Tremblay, and M. Ruffini, “Active Wavelength Load as a Feature for QoT Estimation Based on Support Vector Machine,” in *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, Shanghai, China, May 2019, pp. 1–6. doi: 10.1109/ICC.2019.8761369.
- [38] T. Panayiotou, G. Savva, B. Shariati, I. Tomkos, and G. Ellinas, “Machine Learning for QoT Estimation of Unseen Optical Network States,” in *Optical Fiber Communication Conference (OFC) 2019*, San Diego, California, 2019, p. Tu2E.2. doi: 10.1364/OFC.2019.Tu2E.2.
- [39] J. Wass, J. Thrane, M. Piels, R. Jones, and D. Zibar, “Gaussian Process Regression for WDM System Performance Prediction,” in *Optical Fiber Communication Conference*, Los Angeles, California, 2017, p. Tu3D.7. doi: 10.1364/OFC.2017.Tu3D.7.
- [40] M. Ibrahim *et al.*, “Machine learning regression for QoT estimation of unestablished lightpaths,” *J. Opt. Commun.*

- 1 *Netw.*, vol. 13, no. 4, p. B92, Apr. 2021, doi:
2 10.1364/JOCN.410694.
- 3 [41] Q. Fan *et al.*, “Experimental Comparisons between
4 Machine Learning and Analytical Models for QoT
5 Estimations in WDM Systems,” in *Optical Fiber
6 Communication Conference (OFC) 2020*, San Diego,
7 California, 2020, p. M2J.2. doi:
8 10.1364/OFC.2020.M2J.2.
- 9 [42] F. Usmani, I. Khan, M. U. Masood, A. Ahmad, M.
10 Shahzad, and V. Curri, “Convolutional neural network for
11 quality of transmission prediction of unestablished
12 lightpaths,” *Microw Opt Technol Lett*, vol. 63, no. 10, pp.
13 2461–2469, Oct. 2021, doi: 10.1002/mop.32996.
- 14 [43] Z. Gao *et al.*, “ANN-Based Multi-Channel QoT-
15 Prediction Over a 563.4-km Field-Trial Testbed,” *J.
16 Lightwave Technol.*, vol. 38, no. 9, pp. 2646–2655, May
17 2020, doi: 10.1109/JLT.2020.2971104.
- 18 [44] J. Muller *et al.*, “Estimating Quality of Transmission in a
19 Live Production Network using Machine Learning,” p. 3,
20 2021.
- 21 [45] A. Ferrari *et al.*, “Experimental Validation of an Open
22 Source Quality of Transmission Estimator for Open
23 Optical Networks,” in *Optical Fiber Communication
24 Conference (OFC) 2020*, San Diego, California, 2020, p.
25 W3C.2. doi: 10.1364/OFC.2020.W3C.2.
- 26 [46] I. Sartzetakis, K. (Kostas) Christodouloupoulos, and E.
27 (Manos) Varvarigos, “Accurate Quality of Transmission
28 Estimation With Machine Learning,” *J. Opt. Commun.
29 Netw.*, vol. 11, no. 3, p. 140, Mar. 2019, doi:
30 10.1364/JOCN.11.000140.
- 31 [47] E. Seve, J. Pesic, C. Delezoide, S. Bigo, and Y.
32 Pointurier, “Learning Process for Reducing Uncertainties
33 on Network Parameters and Design Margins,” *J. Opt.
34 Commun. Netw.*, vol. 10, no. 2, p. A298, Feb. 2018, doi:
35 10.1364/JOCN.10.00A298.
- 36 [48] F. Meng *et al.*, “Robust Self-Learning Physical Layer
37 Abstraction Utilizing Optical Performance Monitoring
38 and Markov Chain Monte Carlo,” in *2017 European
39 Conference on Optical Communication (ECOC)*,
40 Gothenburg, Sep. 2017, pp. 1–3. doi:
41 10.1109/ECOC.2017.8346217.
- 42 [49] M. Bouda *et al.*, “Accurate Prediction of Quality of
43 Transmission Based on a Dynamically Configurable
44 Optical Impairment Model,” *J. Opt. Commun. Netw.*, vol.
45 10, no. 1, p. A102, Jan. 2018, doi:
46 10.1364/JOCN.10.00A102.
- 47 [50] Q. Zhuge *et al.*, “Application of Machine Learning in
48 Fiber Nonlinearity Modeling and Monitoring for Elastic
49 Optical Networks,” *J. Lightwave Technol.*, vol. 37, no.
50 13, pp. 3055–3063, Jul. 2019, doi:
51 10.1109/JLT.2019.2910143.
- 52 [51] R. M. Morais, B. Pereira, and J. Pedro, “Fast and High-
53 Precision Optical Performance Evaluation for Cognitive
54 Optical Networks,” in *Optical Fiber Communication
55 Conference (OFC) 2020*, San Diego, California, 2020, p.
56 Th3D.3. doi: 10.1364/OFC.2020.Th3D.3.
- 57 [52] A. D’Amico *et al.*, “Using Machine-Learning in an Open
58 Optical Line Systems Controller,” p. 9, 2019.
- 59 [53] A. Mahajan, K. Christodouloupoulos, R. Martinez, S.
60 Spadaro, and R. Munoz, “Modeling EDFA Gain Ripple
and Filter Penalties With Machine Learning for Accurate
QoT Estimation,” *J. Lightwave Technol.*, vol. 38, no. 9,
pp. 2616–2629, May 2020, doi:
10.1109/JLT.2020.2975081.
- [54] D. Uzunidis, A. Stavdas, P. Kasnesis, C. Patrikakis, and
A. Lord, “Enhancing Closed-Form Based Physical Layer
Performance Estimations in EONs Via Machine Learning
Techniques,” p. 4.
- [55] J. Pesic, M. Lonardi, E. Seve, N. Rossi, and T. Zami,
“Transfer Learning from Unbiased Training Data Sets for
QoT Estimation in WDM Networks,” in *2020 European
Conference on Optical Communications (ECOC)*,
Brussels, Belgium, Dec. 2020, pp. 1–4. doi:
10.1109/ECOC48923.2020.9333399.
- [56] J. Yu, W. Mo, Y.-K. Huang, E. Ip, and D. C. Kilper,
“Model transfer of QoT prediction in optical networks
based on artificial neural networks,” *J. Opt. Commun.
Netw.*, vol. 11, no. 10, p. C48, Oct. 2019, doi:
10.1364/JOCN.11.000C48.
- [57] L. Xia, J. Zhang, S. Hu, M. Zhu, Y. Song, and K. Qiu,
“Transfer learning assisted deep neural network for
OSNR estimation,” *Opt. Express*, vol. 27, no. 14, p.
19398, Jul. 2019, doi: 10.1364/OE.27.019398.
- [58] C. Rottondi, R. di Marino, M. Nava, A. Giusti, and A.
Bianco, “On the benefits of domain adaptation techniques
for quality of transmission estimation in optical
networks,” *J. Opt. Commun. Netw.*, vol. 13, no. 1, p. A34,
Jan. 2021, doi: 10.1364/JOCN.401915.
- [59] S. Yao *et al.*, “Data Efficient Estimation for Quality of
Transmission Through Active Learning in Fiber-Wireless
Integrated Network,” *J. Lightwave Technol.*, vol. 39, no.
18, pp. 5691–5698, Sep. 2021, doi:
10.1109/JLT.2021.3091377.
- [60] D. Azzimonti, C. Rottondi, A. Giusti, M. Tornatore, and
A. Bianco, “Comparison of domain adaptation and active
learning techniques for quality of transmission estimation
with small-sized training datasets [Invited],” *J. Opt.
Commun. Netw.*, vol. 13, no. 1, p. A56, Jan. 2021, doi:
10.1364/JOCN.401918.
- [61] T. Panayiotou, G. Savva, I. Tomkos, and G. Ellinas,
“Centralized and Distributed Machine Learning-Based
QoT Estimation for Sliceable Optical Networks,”
arXiv:1908.08338 [cs, eess], Sep. 2019, Accessed: Dec.
10, 2019. [Online]. Available:
<http://arxiv.org/abs/1908.08338>
- [62] I. Khan, M. Bilal, and V. Curri, “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)*, Bari, Italy, Jul. 2020, pp. 1–4. doi:
10.1109/ICTON51198.2020.9203334.
- [63] I. Khan *et al.*, “QoT Estimation for Light-path
Provisioning in Un-Seen Optical Networks using
Machine Learning,” in *2020 22nd International
Conference on Transparent Optical Networks (ICTON)*,
Bari, Italy, Jul. 2020, pp. 1–4. doi:
10.1109/ICTON51198.2020.9203364.
- [64] F. Meng *et al.*, “Field Trial of Gaussian Process Learning
of Function-Agnostic Channel Performance Under
Uncertainty,” in *Optical Fiber Communication*

- 1 *Conference*, San Diego, California, 2018, p. W4F.5. doi:
2 10.1364/OFC.2018.W4F.5.
- 3 [65] S. Yan *et al.*, “Field trial of Machine-Learning-assisted
4 and SDN-based Optical Network Planning with Network-
5 Scale Monitoring Database,” in *2017 European
6 Conference on Optical Communication (ECOC)*,
7 Gothenburg, Sep. 2017, pp. 1–3. doi:
8 10.1109/ECOC.2017.8346091.
- 9 [66] R. Proietti *et al.*, “Experimental Demonstration of
10 Machine-Learning-Aided QoT Estimation in Multi-
11 Domain Elastic Optical Networks with Alien
12 Wavelengths,” *J. Opt. Commun. Netw.*, vol. 11, no. 1, p.
13 A1, Jan. 2019, doi: 10.1364/JOCN.11.0000A1.
- 14 [67] P. Samadi, D. Amar, C. Lepers, M. Lourdiane, and K.
15 Bergman, “Quality of Transmission Prediction with
16 Machine Learning for Dynamic Operation of Optical
17 WDM Networks,” in *2017 European Conference on
18 Optical Communication (ECOC)*, Gothenburg, Sep. 2017,
19 pp. 1–3. doi: 10.1109/ECOC.2017.8346216.
- 20 [68] F. Usmani *et al.*, “Evaluating Cross- feature Trained
21 Machine Learning Models for Estimating QoT of
22 Unestablished Lightpaths,” in *2021 International
23 Conference on Electrical, Communication, and Computer
24 Engineering (ICECCE)*, Kuala Lumpur, Malaysia, Jun.
25 2021, pp. 1–6. doi:
26 10.1109/ICECCE52056.2021.9514154.
- 27 [69] E. Seve, J. Pesic, and Y. Pointurier, “Associating
28 machine-learning and analytical models for quality of
29 transmission estimation: combining the best of both
30 worlds,” *J. Opt. Commun. Netw.*, vol. 13, no. 6, p. C21,
31 Jun. 2021, doi: 10.1364/JOCN.411979.
- 32 [70] N. Morette, “Leveraging ML-based QoT Tool Parameter
33 Feeding for Accurate WDM Network Performance
34 Prediction,” p. 3, 2021.
- 35 [71] C.-Y. Liu, X. Chen, R. Proietti, and S. J. B. Yoo,
36 “Performance studies of evolutionary transfer learning for
37 end-to-end QoT estimation in multi-domain optical
38 networks [Invited],” *J. Opt. Commun. Netw.*, vol. 13, no.
39 4, p. B1, Apr. 2021, doi: 10.1364/JOCN.409817.
- 40 [72] I. Khan, M. Bilal, M. Umar Masood, A. D’Amico, and V.
41 Curri, “Lightpath QoT computation in optical networks
42 assisted by transfer learning,” *J. Opt. Commun. Netw.*,
43 vol. 13, no. 4, p. B72, Apr. 2021, doi:
44 10.1364/JOCN.409538.
- 45 [73] M. Lonardi, J. Pesic, T. Zami, and N. Rossi, “The Perks
46 of Using Machine Learning for QoT Estimation with
47 Uncertain Network Parameters,” in *OSA Advanced
48 Photonics Congress (AP) 2020 (IPR, NP, NOMA,
49 Networks, PVLED, PSC, SPPCom, SOF)*, Washington,
50 DC, 2020, p. NeM3B.2. doi:
51 10.1364/NETWORKS.2020.NeM3B.2.
- 52 [74] A. Ferrari *et al.*, “Assessment on the in-field lightpath
53 QoT computation including connector loss uncertainties,”
54 *J. Opt. Commun. Netw.*, vol. 13, no. 2, p. A156, Feb.
55 2021, doi: 10.1364/JOCN.402969.
- 56 [75] R. Ayassi, A. Triki, E. L. Rouzic, N. Crespi, and R.
57 Minerva, “Bayesian optimization-based algorithm to
58 improve the quality of transmission estimation,” p. 3.
- 59 [76] K. Christodoulouopoulos *et al.*, “Toward efficient, reliable,
60 and autonomous optical networks: the ORCHESTRA
61 solution [Invited],” *J. Opt. Commun. Netw.*, vol. 11, no.
62 9, p. C10, Sep. 2019, doi: 10.1364/JOCN.11.000C10.
- 63 [77] A. Ferrari *et al.*, “GNPy: an open source application for
64 physical layer aware open optical networks,” *J. Opt.
65 Commun. Netw.*, vol. 12, no. 6, p. C31, Jun. 2020, doi:
66 10.1364/JOCN.382906.
- 67 [78] J. Lu *et al.*, “Automated training dataset collection system
68 design for machine learning application in optical
69 networks: an example of quality of transmission
70 estimation,” *J. Opt. Commun. Netw.*, vol. 13, no. 11, p.
71 289, Nov. 2021, doi: 10.1364/JOCN.431780.