Lightpath QoT Computation in Optical Networks Assisted by Transfer Learning

IHTESHAM KHAN[1*], MUHAMMAD BILAL[1], M UMAR MASOOD[1], ANDREA D'AMICO[1], AND VITTORIO CURRI[1]

[1] DET, Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129, Torino, Italy (planet.polito.it)
[*] Corresponding author: ihtesham.khan@polito.it

Compiled January 9, 2021

Precise computation of the Quality of Transmission (QoT) of Lightpaths (LPs) in transparent optical networks has techno-economic importance for any network operator. The QoT metric of LPs is defined by the Generalized Signal-to-Noise Ratio (GSNR), which includes the effect of both Amplified Spontaneous Emission (ASE) noise and Nonlinear Interference (NLI) accumulation. Generally, the physical layer of a network is characterized by nominal values provided by vendors for the operational parameters of each Network Element (NE). Typically, NEs suffer a variation on the working point that implies an uncertainty from the nominal value, which creates uncertainty in GSNR computation and requires the deployment of a system margin. We propose the use of a Machine Learning (ML) agent trained on a dataset from an *in-service* network to reduce the uncertainty on the GSNR computation on an *un-used* sister network, based on the same optical transport equipment, so following the transfer learning paradigm. We synthetically generate datasets for both networks using the open-source library GNPy and show how the proposed Deep Neural Network (DNN) based on TensorFlow© may substantially reduce the GSNR uncertainty and consequently, the needed margin. We also present a statistical analysis of the observed GSNR fluctuations, showing that the per-wavelength GSNR distribution is always well-approximated as Gaussian, enabling a statistical closed-form approach to the margin setting.

http://dx.doi.org/10.1364/ao.XX.XXXXXX

1. INTRODUCTION

Recently, the remarkable increase in the global IP traffic, compelled by the introduction of 5G technology, the internet of things, and cloud services, has marked-up high demands and new requirements for the capacity enhancement and reliability of optical networks [1]. To serve the rapidly increasing number of internet service users, the technologies of optical networks are continuously evolving. To cater to this dramatic upsurging number of internet service users, the key operator demands the full utilization of network resources of existing infrastructure to maximize returns on CAPEX investments. To achieve this, the data transport layer needs to be driven to reach the maximum available capacity. The center key factor for optimal exploitation of data transport is maximizing the capacity of Wavelength Division Multiplexed (WDM) transmission technique along with network disaggregation. These attributes evolved to technologies: Elastic Optical Networks (EONs) and Software-Defined Networking (SDN) paradigm in an optical network. The EONs provide flexibility to the network controller to scale up or down resources according to the traffic requests in order to efficiently utilize the available spectrum [2, 3]. In addition to this, the SDN implementation enables the management of each NE within a virtualized environment, so permitting a disaggregated approach to the network exploitation, enabling openness and virtual network slicing. These features pave a roadway for a partly or fully disaggregation of optical networks: a disaggregated network includes subsystems that are controlled independently by relying on common data structures and Application Program Interface (API).

The foundation step towards flexible and disaggregated networks is to abstract the WDM optical transport as a topology graph weighted by the Generalized Signal-to-Noise Ratio (GSNR) degradation on transparent Lightpaths (LPs) introduced by each crossed NE and subsystem, mainly by Optical

Line Systems (OLSs) including fibers and amplifiers [4, 5]. Each OLS is controlled by the OLS controller running in the control plane [6] that sets the amplifier working point and consequently determines the GSNR degradation. The more accurate the nominal working point, the better the capability to rely on the overall LP GSNR. So, a lower system margin is requested in LP deployment and, consequently, a larger traffic can be deployed, enabling a better exploitation of the installed equipment. In addition to this, a reliable softwarization of optical transport also helps in the automatic recovery of network failures by reducing downtime.

The state-of-the-art for optical data transport relies on WDM spectral use and dual-polarization coherent optical technologies for which a transparent LP can be well approximated as an Additive White Gaussian Noise (AWGN) channel affected by the Amplified Spontaneous Emission (ASE) noise introduced by the amplifiers and the Nonlinear Interference (NLI) disturbance due to self- and cross-channel nonlinear crosstalk. So, the QoT is well assessed by the GSNR, which includes the effect of ASE noise and NLI accumulation [7]. Given the flex-transceiver characteristics as threshold signal-to-noise ratio per modulation format, the available GSNR on a given LP compared to such a threshold defines the path feasibility and the deployable rate [5]. So, the key element in the SDN application down to the WDM optical transport is a QoT Estimator (QoT-E) that, given the network status, computes the GSNR over a specified transparent LP. Within the consortium Telecom Infra project [7, 8], it has been extensively demonstrated that in case of accurate knowledge of the physical layer, a QoT-E can give extremely accurate GSNR computation. Installed equipment is affected by a variation on the operational point with respect to the nominal one due to hardware aging, change in OLS spectral load and side-effects of in-field operations. The variations induce changes on the actual GSNR with respect to the nominal value computed by the QoT-E [9, 10]. The main causes of GSNR uncertainties are ripples on amplifiers' gain and noise figure (NF) and connector losses. In particular, among connector losses the ones at the fiber input are the most significant as they set the actual power levels triggering most of the nonlinear effects in first kilometers of fiber spans. So, on top of the computed nominal available GSNR on a given LP, a system margin must be conservatively deployed to avoid Out-of-Service (OOS) [11].

The ML paradigm has already been expertly applied to optical networking: consider [12–15] for performance monitoring applications. A comprehensive survey of ML applied in optical networks can be found in [16]. Explicitly, coming towards a particular interest of this study, i.e., QoT-E of LP before its deployment, some useful ML methods such as cognitive Case-Based Reasoning (CBR) technique are proposed [17]. Experimental results corresponding to [17] achieved with real field data are discussed in [18]. In [10], the authors proposed an ML approach to estimate the Optical Signal-to-Noise ratio (OSNR) response over distinct spectral load configurations. The dataset used in [10] has been obtained experimentally from an OLS containing a cascade of 11 pairs of EDFAs and, in place of fibers, variable optical attenuators in order to avoid any NLI generation and to focus the investigation only on the OSNR prediction. A learning process based on a gradient descent algorithm is proposed to exploit the stored database in [19] to reduce uncertainties on network parameters and design margins. In the context of multicast transmission in an optical network, a neural network is trained to predict the Q-factor in [20–22]. Several ML techniques for QoT-E of LP before its deployment are also presented in [23, 24]. In [25], the authors proposed and exploited the ability of several ML models such as wide DNN, multi-layer perceptron, boasted tree, decision tree and random forest regressors for QoT-E. The developed model is trained and tested on the synthetic data generated by the open-source GNPy library. A random forest based binary classifier is presented in [26] to predict the bit-error-rate of un-deployed LPs. In [27], a random forest classifier is proposed, along with the potentiality of two other techniques, i.e., k-nearest neighbor and support vector machine. The authors made a detailed comparison of the proposed ML techniques. This analysis, showed that the support vector machine is more refined in performance but worst in computational time. In [28], the authors used a neural network for the characterization of integrated circuits consequently used for their full and accurate softwarization. In [29], the authors evaluated the performance of two Domain Adaptation (DA) approaches for ML assisted QoT-E of an optical LP, for a fixed/variable number of available training samples from the source/target domain. The authors considered two networks characterized by different topologies, but adopting the same fiber type and transmission equipment and assessed the performance of two DA techniques depending on the number of available training instances from the target domain. The results in [29] reported that DA based approach performed better as compared to standard ML techniques. The authors in [30] presented the use of ML for QoT estimation in case of span length uncertainty. This work described several methods along with comparison for QoT estimation such as Gaussian Noise (GN)-model based approach, ML-based approach or the hybrid (ML+GN) approach. Finally, in [31], the authors compared the QoT-E accuracy achieved by a few Active Learning (AL) and DA methods on two different network topologies. The results presented in this work reported significant improvements using an AL approach with some additional samples acquired from the target domain.

The primary motivation of this study is to decrease the uncertainty in the GSNR computation of an LP and, consequently, to enable reliable path computation to deploy the candidate LP at the minimum margin. We suppose a disaggregated network scenario, in which the network controller may rely on a QoT-E API. The QoT-E accurately evaluates the fiber span losses together with the amount of introduced NLI and computes the gain and ASE noise from amplifiers. Filtering penalties and crosstalk introduced by Reconfigurable Optical Add-Drop Multiplexers (ROADMs) can be considered as well, but are out of the scope of this work. If the controller can get a reliable network status, i.e., an *exact snap-shot* of operational parameters for each NE, the QoT-E can estimate the GSNR with excellent accuracy, as shown, for instance, in [7, 8]. In the absence of an *exact* description of system parameters, a network operator relies on the *nominal* description of system parameters. Typically, nominal values are the average parameters provided by the vendors. The estimator engine utilizes these values and calculates a nominal GSNR with some degree of uncertainty. In the present study, we suppose to rely on a dataset coming from an *in-service* network; after a statistical analysis on GSNR data, according to the transfer learning paradigm, we use it to train an ML agent to assist the QoT-E in managing an *un-used* sister network. As a sister network, we suppose a different topology based on the same hardware: specifically, fiber type and Erbium-doped fiber amplifiers (EDFAs). Along with this, we focused on random and spectrally flat connector losses and performed a Monte-Carlo analysis of EDFA gain and ripple. GSNR data are then first analyzed to search for a common statistical characterization. We show how

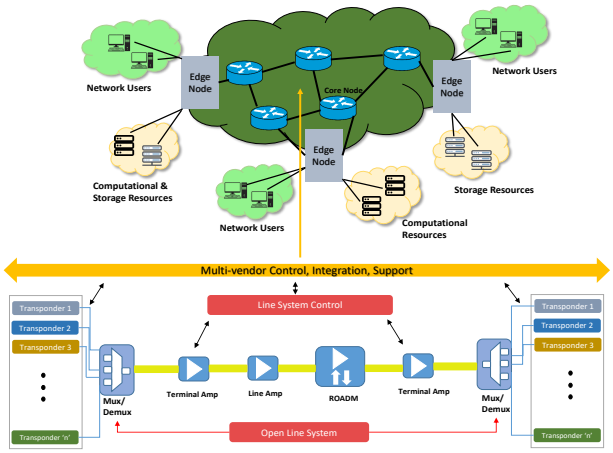

Fig. 1. Schematic description of an optical network as a topology of ROADM nodes connected by OLSs. The inset shows a general setup for an OLS that, in this case, is supposed to be open.

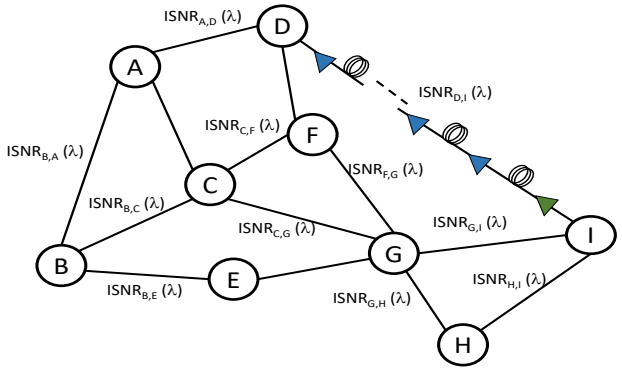

Fig. 2. Abstraction of Optical Network

the per-wavelength GSNR distribution can always be well approximated as Gaussian-distributed. This observation leads to propose a method to set the needed margin on a network, given the standard-deviation of GSNR, with a fixed maximum tolerable OOS percentage. Then, we propose to use the dataset from the *in-service* network to train an ML agent running in the controller on the sister network. The ML agent's scope is to correct the GSNR computation for LP QoT on an *un-used* sister network whose nominal NE parameters have been perturbed to include a realistic degree of uncertainty to be reduced by an ML. We tested the method on several paths on the *un-used* network. We show that using the trained ML agent may substantially reduce the uncertainty in path GSNR computation, consequently reducing the needed margin.

The originality of this work is in proposing the use of a transfer learning method to effectively train an ML agent operating together with a reliable QoT-E in the network controller to correct the GSNR uncertainties due to EDFA ripples and spectral load dependence, also including uncertainties in connector losses. By the statistical characterization of GSNR distribution that is always well approximated as Gaussian, we also propose an original method to set the margin by a maximum tolerable OOS percentage, given the GSNR standard deviation.

The remainder of the paper is organized as follows. In Sec. 2, we briefly describe the physical layer's abstraction to effectively perform a multi-layer optimization, along with the argument that an accurate QoT-E has a key role in minimizing the system margin. In Sec. 3, we describe the analysis to obtain a synthetic dataset, supposing to be able to read from the network status, for each deployed LP the received power and LP OSNR and GSNR, so the accumulated ASE noise and NLI disturbance. The synthetic datasets are generated for both networks: the *in-service* network dataset for ML training, whereas, the *un-used* network dataset for statistical characterizations and testing. The synthetic datasets have been generated using the GNPy opensource software by the Telecom Infra Project [32]. Nominal NE elements parameters are statistically perturbated: we specifically target EDFA NF, ripple gain and insertion losses. In Sec. 4, first, we perform the statistical analysis of GSNR measurements within a dataset, to comment on margin to be considered with respect to the QoT-E. Along with this, we also list different possible approaches to define the margin, with each allowing a distinct reduction in the GSNR uncertainty. Then we also observe that GSNR measurements distribution within a dataset is always well approximated by a Gaussian and the margin can be set according to a maximum tolerable percentage of OOS. Then, in Sec. 5, we describe the basic schematic of the proposed ML agent trained on the dataset of the already *in-services* network and implemented within the QoT-E of the *un-used* network to correctly estimate the GSNR of a specific LP before its actual deployment. We also describe the orchestration of trained ML module specifying features, labels and the ML algorithm's additional configuration parameters. In this work, we do not aim to develop a specific ML model from scratch and focus on showing the effectiveness of ML in this scenario. So, we exploit an extensively tested opensource project: The TensorFlow© library [33]. In Sec. 6, we describe the results in detail. We show for some paths on the *un-used* network that the synergistic use of the trained ML agent with the QoT-E enables a substantial increase in the GSNR prediction, so effectively reducing the needed margin. Finally, the conclusion and future research direction are drawn in Sec. 7.

2. PHYSICAL LAYER ABSTRACTION OF TRANSPARENT OPTICAL NETWORK

A transparent optical transport network is a structure of connected ROADM nodes, where traffic request is added/dropped or routed, as shown in Fig. 1. Topology links are bidirectional OLSs made of fiber pairs and In-Line Amplifiers (ILAs) – EDFA [34]. As links are virtually symmetrical, in this analysis we consider unidirectional links. Following the disaggregated approach, each OLS is independently and autonomously controlled by an OLS controller that sets the operational point of ILAs to minimize QoT degradation. [6, 35]. The OLS controller's settings are based on the available parameters describing the physical layer, among which the most delicate are the fiber connector losses and the ripples of ILAs' gains and NFs. In particular, these ripples vary with the spectral load. So, OLS controllers may guarantee to operate at the nominal working point with a certain degree of uncertainty on the actual GSNR degradation.

On the optical infrastructure, LPs are deployed as transparent optical circuits on the WDM flexible grid [36] connecting transceivers, hence, they support dual-polarization multiple level modulation formats. In this framework, a higher operative GSNR enables an higher deployable modulation cardinality, which provides a larger deployable rate. Thus, the key operation in LP deployment is to selecting the feasible modulation format,

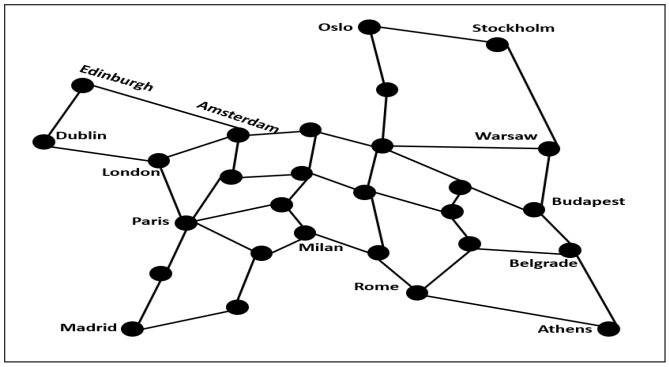

Fig. 3. EU network topology

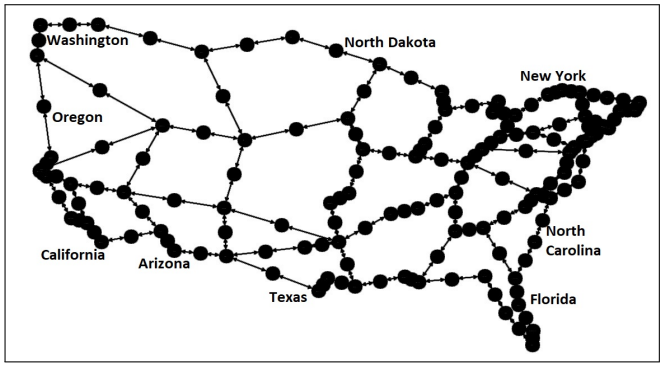

Fig. 4. USA network topology

so maximizing the rate according to the available LP QoT. Such an operation performed by the network controller requires an accurate QoT-E, as uncertainties on the actual LP QoT require system margins and a consequent reduction in the deployed traffic. LPs suffer from several propagation impairments: ASE noise from ILAs, linear and nonlinear fiber propagation effects and ROADM filtering effects. Thanks to DSP techniques in coherent transceivers, fiber propagation linear effects can be managed as well as some amount of the additive Gaussian disturbance (NLI) and phase noise, both nonlinear and generated by self- and cross-channel nonlinear crosstalk. In general, the DSP module at the receiver effectively compensates for the phase noise using the carrier phase estimator algorithm and must be considered in QoT-E only for transceivers operated by modulation formats with large-constellation cardinality designed for short-reach [37]. Finally, the ROADMs filtering effects and crosstalk also apply some level of degradation to QoT, generally considered as an extra loss and additional Gaussian noise source. ROADM impairments are out of the scope of this work.

So, within core networks operated by transceivers using QPSK, 8-QAM and 16-QAM constellations, or hybrid formats [38], a transparent LP can be effectively modeled as an AWGN channel, affected by ASE noise and NLI Gaussian disturbances. QoT over an AWGN channel is summarized by the Signal-to-Noise ratio (SNR), that for LPs is typically identified as GSNR:

$$\mathrm{GSNR} = \frac{P_{\mathrm{Rx}}}{P_{\mathrm{ASE}} + P_{\mathrm{NLI}}} = \left(\mathrm{OSNR}^{-1} + \mathrm{SNR}_{\mathrm{NL}}^{-1}\right)^{-1}, \quad \textbf{(1)}$$

where $\mathrm{OSNR} = P_{\mathrm{Rx}}/P_{\mathrm{ASE}}$ is the optical signal to noise ratio detectable by optical channel monitors, $\mathrm{SNR}_{\mathrm{NL}} = P_{\mathrm{Rx}}/P_{\mathrm{NLI}}$ is the nonlinear SNR due to NLI only and so observable only on the DSP recovered constellation. P_{Rx} is the power of the channel at the receiver, P_{ASE} is the power of the accumulated ASE noise and P_{NLI} is the power of the accumulated NLI.

Following a disaggregated approach, the NLI introduced by each fiber span can be separated into two main contributions: the Self-channel NLI (SC-NLI) and the Cross-channel NLI (XC-NLI) depending on the input spectrum [39]. Several mathematical models have been proposed for the NLI evaluation with a disaggregated approach, with different accuracy levels, e.g., [40, 41]. For an accurate spectrally-resolved evaluation, NLI calculation must consider the simultaneous effect of Stimulated Raman Scattering (SRS) [42]. So, the NLI per span can be evaluated by a proper algorithm following a spectral disaggregated approach superimposing the effects of each channel. This is the approach we exploit by using the GNPy library. For the

Table 1

Simulation Parameters	
Launch Power/ Channel	0 dBm
Dispersion (D)	16.0 $ps/nm/km$
Attenuation coefficient (α)	0.3 dB/km
Channel Spacing	50 GHz
Span Length	80 km
WDM Comb (C-Band)	76
Baud Rate	32 $Gbaud$
Amplifier Noise Figure	[6 - 11] dB [8]
Nominal Amplifier Noise Figure	8.75 dB
Amplifier Gain Ripple	Variation of 1 dB
Nominal Amplifier Gain Ripple	Flat
Fiber Type	Standard SMF
Insertion losses	Exponential Distribution ($\lambda = 4$)[43]
Nominal Insertion losses	0.3 dB [44]

ASE noise accumulation, we consider the gain and NF for a typical amplifier set. So, a data structure effectively abstract WDM optical transport network as a topology weighted graph G(V,E), where graph-vertices (V) are switches and ROADMs, while graph-edges (E) are OLSs. Weights on edges are the GSNR characterizing each OLS as exposed by the OLS controller. Fig. 2 shows an example, where Inverse-SNR (ISNR)=1/GSNR. Over such a data structure, LP GSNR can be evaluated by *navigating* the route by accumulating metrics for the wavelength under test:

$$GSNR = \frac{1}{\sum_i^{OLS} \frac{1}{GSNR_i}}. \quad \textbf{(2)}$$

The accuracy in GSNR computation depends on the mathematical models for fiber propagation and amplifiers and on the accuracy in the knowledge of physical layer parameters. The perturbation from nominal values used for OLS controlling and for QoT-E yields an uncertainty on the LP GSNR, which requires system margins and reduces the deployable traffic, with respect to the nominal amount. The main uncertainties are the fiber connector losses defining the fiber input power and consequently the NLI and the ILAs' gain and NF determining the ASE noise, this depending on the spectral load.

3. SYNTHETIC DATASET GENERATION

In this section, we describe the simulation performed to model a particular network and its transmission components. Along with this modeling, we also report the mimicked datasets, which are generated synthetically for two different networks; *in-service*

and *un-used* network by using the GNPy library. red In particular, a dataset is generated for the *in-service* network for ML training, whereas, the *un-used* network dataset is used for statistical characterization of GSNR and to test the effectiveness of the ML agent.

Table 2

EU: Training					
Source	**Destination**	**Spans**	**Source**	**Destination**	**Spans**
Vienna	Warsaw	7	Brussels	Bucharest	30
Amsterdam	Berlin	8	Frankfurt	Istanbul	34

Table 3

USA : Testing					
Source	**Destination**	**Spans**	**Source**	**Destination**	**Spans**
Louisville	Memphis	7	Charleston	Charlotte	8
Memphis	Miami	24	Los Angeles	Louisville	46
Little Rock	Long Island	25	Kansas City	Las-Vegas	30

We suppose each OLS is controlled targeting the maximum GSNR [45] according to the Local-Optimization Global-Optimization (LOGO) strategy [6]. We assume that both the OLS controllers and QoT-E might rely on the nominal values for NE parameters. red Although both datasets are synthetically generated, we consider the *in-service* network dataset as obtained by reading data from transceivers and monitors for each deployed LP, specifically the channel power and OSNR from monitors, GSNR from transceivers and the number of crossed spans from the controller. Therefore, in this investigated hypothetical scenario, the *in-service* network dataset contains the exact channel powers, OSNRs and GSNRs values, without any uncertainty. On the other hand, the *un-used* network dataset includes inaccuracies due to network parameter uncertainties. To mimic reality, a synthetic dataset is generated considering random connector losses and ILAs' NF and gain ripples. In particular, we first randomly set connector losses, we performed a Monte-Carlo analysis on different OLS spectral load and consequent ripples [46], and collect the dataset. Insertion losses are determined by an exponential distribution with $\lambda = 4$ as described in the study [43]. To generate synthetic data, we used the GNPy library [8, 32]. The GSNR computation of the GNPy library is spectrally resolved and is based on the Generalized Gaussian Noise (GGN) model for NLI. The GGN model always considers worst effect for the NLI by supposing Gaussian-modulated interfering channels [8, 42]. Because of the computational effort needed to generate a proper dataset, the considered OLSs carry only 76 channels over the standard 50 GHz grid on the C-band, having total bandwidth close to 4 THz. We do not expect a substantial difference in results when considering standard 96 channels on the entire C-band. We supposed to rely on transceivers at 32 GBaud, shaped with a root-raised-cosine filter. OLS launch power is defined by the booster at the ROADM output set to define as 0 dBm per channel, while ILAs are set at transparency, i.e., at the nominal gain completely recovering fiber losses. All OLSs are supposed to operate on ITU-T G.652 Standard Single-Mode Fiber (SSMF) with a span length of 80 km. These hypotheses are used for both the *in-service* and *un-used* network, as we suppose the sister network exploiting the same optical transport technologies. Network parameters are summarized in Tab. 1.

The generated dataset is obtained by perturbating the above frame of reference by varying the parameters of EDFA: NFs, amplifier gain ripples and fiber connector losses. The selection of NF is made by a uniform distribution varying between 6 dB to 11 dB , while the amplifier gain ripples are varied uniformly within a 1 dB interval. (Note that such a wide range of NF is typical in commercial devices at spectral loading far from full load.)

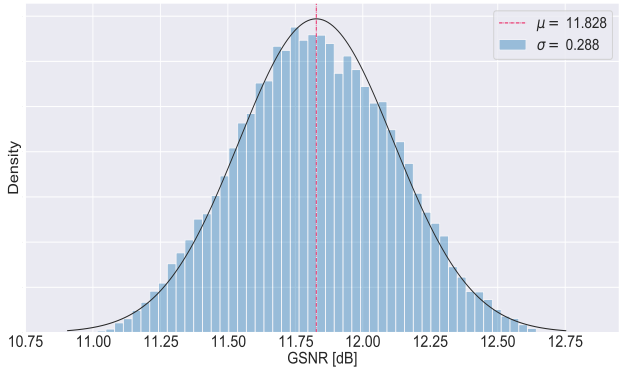

Fig. 5. Overall GSNR Distribution for single path Louisville-Memphis of Acquired dataset

The considered spectral loads in dataset generation is a subset of the entire 2^{76} possible loads, where 76 is the total number of WDM channels. In the considered subset of spectral loads, each source-to-destination, (s –> d), pair has 1024 realizations of random load ranging from 34% to 100% of total bandwidth utilization. Given the (s –> d), the add/drop variation of the spectral load has been reproduced by randomly selecting amplifier ripples. Nevertheless, we always assume a full spectral load when the NLI impairment is calculated, as it represents the worst-case scenario. The first dataset is generated for the EU network topology which consists of 4096 data realizations for four (s –> d) pairs (1024 combinations for each (s –> d) pair) and is used as an *in-service* network as shown in Tab. 2, while for the *un-used* network, the dataset is generated for the USA network topology that includes six (s –> d) pairs having 6144 data realizations as shown in Tab. 3. The considered networks are characterized by distinct topologies shown in Fig. 3 and Fig. 4, adopting the same fiber type and transmission equipment. Still, they are different in terms of random amplifiers' parameters, i.e., ripples on NF and gain and fiber insertion losses. The basic parameters for both network topologies are shown in Tab. 4.

The distribution of *GSNR* for a single particular path *Louisville- Memphis* for all the considered realizations of varying spectral load of WDM channels is depicted in Fig. 5, which includes the resulting GSNR values obtained varying spectral loads, WDM channels, gain ripples, NFs and insertion losses. We analyze only one path as we did not observe a substantial difference in the statistical characteristics of GSNR for other paths.

4. GSNR MARGIN & ANALYSIS OF GSNR DATASET

In this section, we analyze the GSNR dataset of an *un-used* network and its statistics to comment on the margin to be considered with respect to the QoT-E. Along with this, we also comment on different possible approaches to define the margin: considering only variation ranges or relying on a statistical characterization of GSNR.

In a scenario of *un-used* network, to estimate the GSNR, the network controller can rely only on the nominal description of system parameters reported in Tab. 1. Using only this nominal description of system parameters, the network controller estimates a *nominal* GSNR value. This estimated nominal GSNR has some degree of uncertainty due to the variation in NEs' working

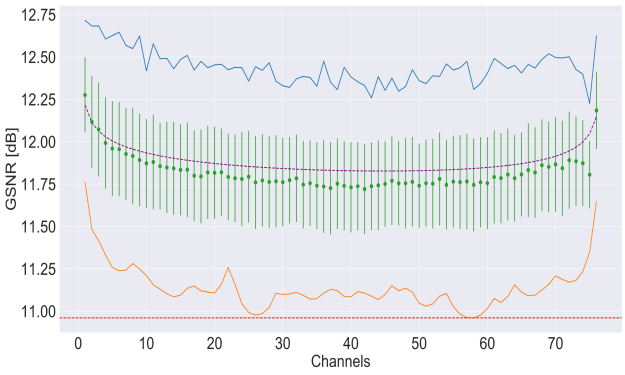

Fig. 6. Overall, GSNR measurements for a single path Louisville- Memphis of an *un-used* network in the frequency domain. The green dots are the mean values over the entire sample for each channel; the error bars are equal to the standard deviations. The purple line shows the nominal values for this path in the frequency domain. In blue and orange, the maximum and the minimum for each channel are outlined, respectively. The dashed red line indicates the overall GSNR minimum of 10.81 dB.

Table 4

Topology Details		
Parameters	**EU: Training [19]**	**USA : Testing[47]**
Number of Nodes	28	100
Number of Links	41	171
Average path distance (km)	2014.06	2541.75
Maximum path distance (km)	3051.10	5481.07
Minimum path distance (km)	669.30	568.33
Average number of spans per Link	19.75	27.49

point. Fig. 5 shows the GSNR distribution for all WDM channels and all realizations for the path *Lousiville-Menphis*: it can be observed that it is distributed around the average nominal value according to a probability density function that is well approximated as Gaussian. Fig. 6 shows the same results for all wavelengths on the same path. In this figure, the exact variation ranges are shown. In general, it can be observed that the system uncertainties induce a variation of the actual GSNR ($\text{GSNR}^{\text{actual}}$) with respect to the nominal value ($\text{GSNR}^{\text{nominal}}$) that we suppose is used for OLS control and computed by the QoT-E for LP deployment. So, in general, we get an uncertainty

$$\Delta\text{GSNR} = \text{GSNR}^{\text{nominal}} - \text{GSNR}^{\text{actual}} \quad \textbf{(3)}$$

in the GSNR computation that must be taken into account. In particular, all cases when $\Delta\text{GSNR} > 0$ are critical because the actual GSNR is smaller than the estimated one and so relying on the QoT-E computation in these cases leads to unwanted out-of-services. To overcome such an issue, the operative GSNR ($\text{GSNR}^{\text{operative}}$) to be used for a reliable LP deployment is obtained by reducing the nominal one of a given GSNR margin δ: $GSNR^{\text{operative}} = GSNR^{\text{nominal}} - \delta$. Note that in this work we focus on the margin needed by ripples in gain NF and variations in connector losses. Other uncertainties and hardware aging may require to further enlarge the deployed GSNR margin. In the following, we'll discuss how to set the margin, first considering only the worst cases, then statistically approaching the problem. According to our analyses, we focus our comments on a specific path, but results can be generalized.

Analyzing results displayed in Fig. 6, it can be observed that the actual GSNR varies around the nominal values with different worst-case $GSNR$ per wavelength and if we set $GSNR^{\text{operative}}$

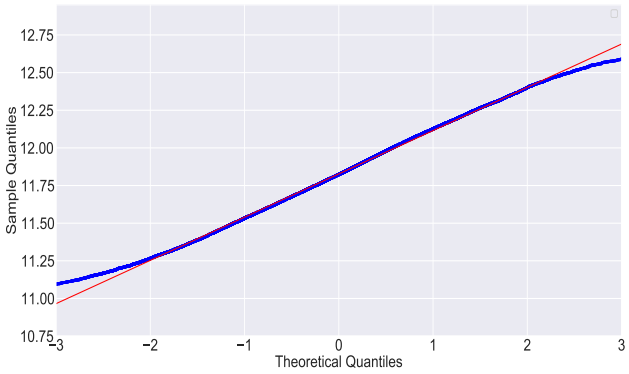

Fig. 7. Normal Q-Q plot of overall GSNR measurements for the path Louisville-Memphis

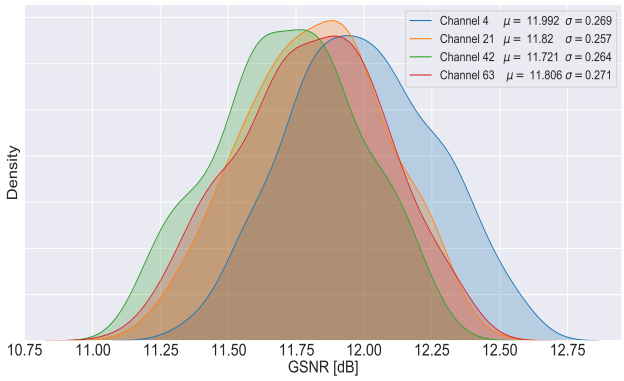

Fig. 8. GSNR measurements of randomly selected channels for path Louisville-Memphis

equal to the worst-case $GSNR$, the margin is given by $\delta = GSNR^{\text{nominal}} - GSNR^{\text{operative}}$. According to our dataset analysis, on the path under test, we observe a variation of required δ ranging from 0.8 dB up to 1.8 dB. Considering to set a unique value for the margin over the entire exploited band, it means a request for 1.8 dB of margin, that will correspond to a large waste in potential capacity, mostly in the case of flexible transceivers [38]. Deploying a per-wavelength margin partially reduces the issues, but keeps an open problem. Being the GSNR fluctuation a random process around the nominal value, it is convenient to approach the problem statistically. If we suppose to know – analytically or numerically – the Probability density function (PDF) $f_{GSNR,\lambda}(x)$ for the GSNR fluctuations for every wavelength λ on a given path, the per-wavelength margin δ_λ can be set by a maximum tolerable percentage of OOS cases p_{oos}, according to the following expression:

$$\int_{-\infty}^{GSNR^{\text{nominal}}-\delta_\lambda} f_{GSNR,\lambda}(x)\,\mathrm{d}x \le p_{\text{oos}}/100\,. \quad \textbf{(4)}$$

This approach can be further strengthened if we are able to find an effective analytical approximation for the GSNR *PDF*. In the following, we show how the per-wavelength GSNR *PDF* is always well approximated by a Gaussian distribution, so Eq. 4 becomes closed form just depending on the per-wavelength mean and variance of GSNR distributions.

By observing the GSNR statistics, the aggregated statistics of the overall GSNR is Gaussian. Even more, also the per-wavelength statistics suggest that the GSNR varies according to

a distribution well approximated as Gaussian. This Gaussian-*ity* of aggregated statistics of overall GSNR is verified in Fig. 7 by Normal Q–Q (quantile-quantile) plot, which is one of the methods frequently used for the dataset Gaussian-*ity* test [48] because of its graphical representation. In parallel to this, the per-wavelength GSNR statistics are shown in Fig. 8, where the relative GSNR distribution for several randomly selected channels (4, 21, 42, 63) are plotted along with their mean (μ) and standard deviation (σ) values. Also in this case Q-Q analyses lead to good results, confirming that that for practical purposes the GSNR statistics can be assumed as Gaussian. In general, the mean GSNR is not exactly equal to the nominal GSNR, but practically we can consider such an equivalence. It further simplifies the approach by just requiring per-wavelength standard deviation for a full characterization and the margin setting Eq. 4 assumes the following expression:

$$\frac{1}{2}\mathrm{erfc}\left(\frac{\delta}{\sigma_{GSNR}}\right) \leq p_{\mathrm{oos}}/100\,, \tag{5}$$

that yields to a closed-form expression for margin setting:

$$\delta \geq \sigma_{GSNR}\ \mathrm{inverfc}\left(2\ p_{\mathrm{oos}}/100\right)\ \ [\mathrm{dB}], \tag{6}$$

where inverfc is the inverse of the complementary error function erfc. A practical engineering rule could be setting $\delta = 3\ \sigma_{GSNR}$ that corresponds to $p_{\mathrm{oos}} = 0.13\%$. This results opens-up future investigation on statistical regression on GSNR variations given by different line uncertainties, as the margin setting just needs the GSNR standard deviation.

As an application example, we consider the usual path under test and apply the statistical margin setting by considering a maximum tolerable percentage of OOS of $p_{\mathrm{oos}} = 1\%$. Results are margin requests ranging from 0.3 to 1.1 dB, depending on the channel. Margin setting is pictorially described in Fig. 9 for channel-58 of the *Louisville-Memphis* path. The blue shaded part shows the maximum allowable oos while the red line represents the nominal GSNR value for the considered channel. The black line represents the minimum value of GSNR. The GSNR margin is defined by the difference between the nominal GSNR (red line) and the operative GSNR (blue line).

The out-of-services penalty in the statistical approach of margin setting creates a road path for proposing a more flexible architecture. As we have the dataset of the already *in-service* network, we can exploit this data to assist an *un-used* network controller in the estimation of GSNR. A data-driven technique such an ML architecture can be ideal in this kind of scenario where the operator has a dataset of an already *in-service* network. In this work, we propose a trained ML module running over the controller of an *un-used* network. The proposed ML module is trained on the data provided by the already *in-service* network and specifically used to assist the core QoT-E engine of *un-used* network, to provide a correcting mechanism for the estimation of the QoT shown in Fig. 10.

In Fig. 10, the controller of an *un-used* network (USA network) fed its core QoT-E engine (GNPy) with nominal values of system parameters, which estimates nominal GSNR with some degree of uncertainty due to variation in the working point of NE. This error (uncertainty) in the estimation of GSNR is corrected by using trained ML module, which is trained on the stored telemetry data of already *in-service* (EU network).

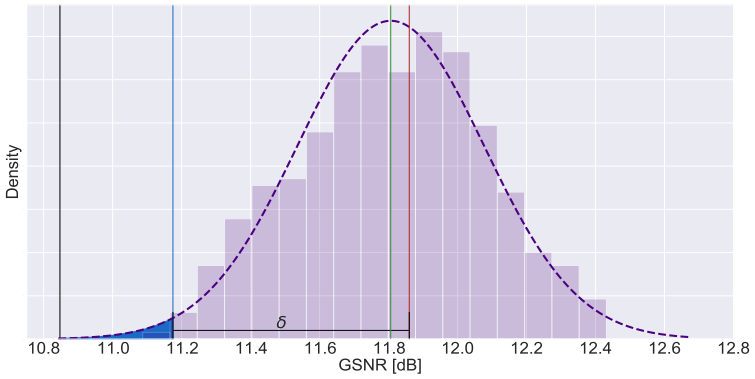

Fig. 9. Analytical GSNR margin estimation for channel-58 path *Louisville-Memphis*: green, red, blue and black vertical lines represent the values of the distribution mean and the nominal, operative and worst-case GSNR, respectively. The dashed line sketches a Gaussian shape with the same mean and standard deviation of the considered distribution.

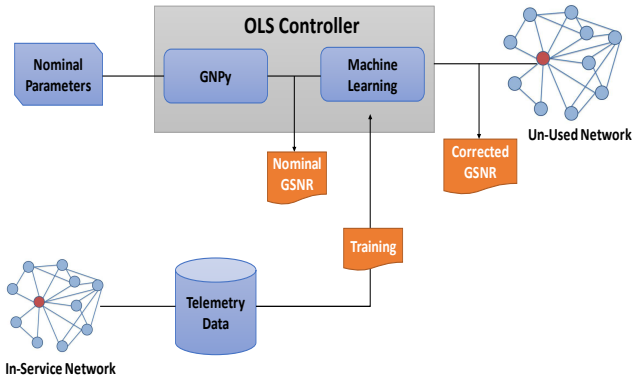

Fig. 10. Model Schematic: The Machine Learning module assists the QoT engine (GNPy)

5. TRANSFER LEARNING AGENT FOR QOT-E CORRECTION

This section describes the ML module, which is trained on the dataset of the already *in-services* network. Besides, we also describe the orchestration of trained ML module specifying features, labels and the additional configuration parameters of the ML algorithm.

The proposed work presents a trained ML module to assist the core QoT-E engine in order to correct the QoT-E estimated $GSNR$ of a specific LP before its actual deployment in an *un-used* network. As ML algorithm we select a DNN [49], which is a powerful tool that has shown significant results in numerous frameworks as the one under investigation. Like all other supervised ML-based learning methods, in order to perform the training and prediction processes, the proposed model requires the definition of the features and labels, which represent the system inputs and outputs, respectively. The manipulated features include the received signal powers, NLIs, ASEs, channel frequencies and the number of spans between source to the destination node. Values that we suppose to retrieve on the *in-service* network from transceivers, optical channel monitors and controller. We fixed as labels the $\Delta GSNR$, the difference between the nominal and the actual GSNR values expressed in Eq. 3, obtained for each channel: Fig. 11 shows the final ML module structure.

The total number of input features consists of *305* entries, the number of LP spans plus the received signal power, the NLI, the ASE and frequency for each channel ($1 + 4 \times 76 = 305$).

In order to enhance the performance of the DNN algorithm, we trained it on a normalized dataset [50]. In general, each feature and label in a normalized dataset has a distribution with zero mean and unit variance. In the present work, the normalization of the dataset is done using z-score normalization, expressed in Eq. 7, where μ and σ are, respectively, the mean and standard deviation of each particular feature or label.

$$Z = \frac{X - \mu}{\sigma} \tag{7}$$

The proposed DNN is developed by using higher-level APIs of the TensorFlow© platform [51], which provides a variety of learning algorithms as well as appropriate functions to refine the dataset before using it as ML model input. The considered DNN is configured by several parametric values that have been optimized, such as *training steps* = 1000, loaded with default *Adaptive Gradient Algorithm (ADAGRAD)* keras optimizer with default *learning rate* = 0.01 and default L_1 regularization = 0.001 [52]. Moreover, during the model building, several non-linear activation functions such as *Relu, tanh, sigmoid* have been tested. After testing, *Relu* as been selected to empowered DNN as it outperforms the others in terms of prediction and computational load [53]. Finally, another important DNN configuration is the number of *hidden-layers*, the model has been tuned on several numbers of *hidden-layers* and neurons to achieve the best trade-off between precision and computational time. These two parameters are linked to the complexity of the DNN, which is tied to the complexity of the problem. Although an increase in the number of layers and neurons improves the accuracy of the DNN up to a certain extent, a further increase in these values has an adverse effect that causes over-fitting and increases in the computational time. After this trade-off analysis, we decided upon a DNN with *3 hidden-layers* containing *20 neurons* each. This choice results in approximately 6 minutes training time using a workstation having specifications, 32 GB of 2133 MHz RAM and an Intel® Core™ i7 6700 3.4 GHz CPU. Given these configurations, first we performed the training, validation and testing on three separate subset of the *in-service* EU network dataset; we choose the conventional rule of 70%,15% and 15%, respectively, as subset proportions. The train set in this scenario consists of data realizations for four (s –> d) pairs of the *in-service* EU network already described in Sec. 3. We set the *training steps* as the stopping factor in-order to avoid over fitting of the model and *Mean square error* (MSE) as loss function given by:

$$\mathrm{MSE} = \frac{\sum_{i=0}^{n} \left(\Delta\mathrm{GSNR}_i^{\mathrm{P}} - \Delta\mathrm{GSNR}_i\right)^2}{n}, \tag{8}$$

where n is the number of tested realizations and, for each tested case i, $\Delta\mathrm{GSNR}_i^{\mathrm{P}}$ and $\Delta\mathrm{GSNR}_i$ are, respectively, the predicted and calculated errors given by the nominal QoT estimation. Once the accuracy level of the model predictions has been reached, the trained ML module can be used together with the core QoT-E engine GNPy to enhance the GSNR estimation of the LP before its actual deployment in an *un-used* USA network.

6. RESULTS AND DISCUSSION

In this Section, we compare the GSNR predictions and the consequent required margins in both the QoT estimation with and without the ML model support.

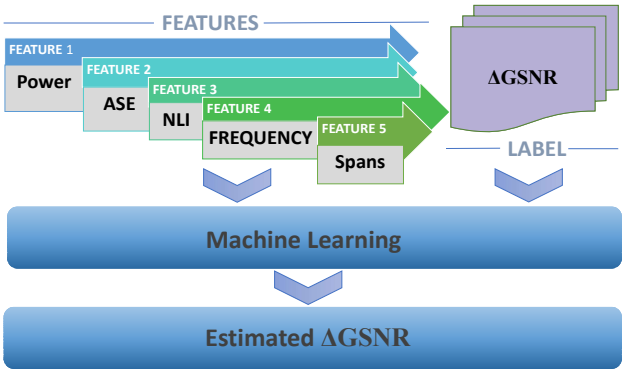

Fig. 11. Description of the Machine Learning Module.

In both these scenarios, we assume that the Gaussian statistical characterization of the GSNR distribution holds and that the required margin evaluation can be performed imposing a fixed percentage of OOS as described in Sec. 4. Summarizing, due to the degree of uncertainty of any GSNR prediction, system margins are required in order to avoid unwanted OOS events. These margins lead to a resizing of the actual LP capacity. To have a uniform comparison metric that quantifies the performance of different GSNR evaluation methods, we choose the following expression:

$$\Delta = \sum_{i}^{n} \left| \frac{\mathrm{GSNR}_i^{\mathrm{operative}} - \mathrm{GSNR}_i^{\mathrm{actual}}}{n} \right| \tag{9}$$

where n is the sample dimension and the operative $\mathrm{GSNR}^{\mathrm{operative}}$ take into account both the GSNR predicted by the specific method used for the GSNR estimation and the GSNR margin obtained by means of Eq. 6. In this investigation, the test set includes six (s –> d) pairs of the *un-used* USA network already described in Sec. 3. Moreover, we further reduce our analysis to those realizations that have a lower actual GSNR with respect to the nominal one, as these cases are more critical, since they result in an OOS event, if no margin deployment is considered.

In order to provide a clear description of the obtained results, we first focus the results presentation to a single path, *i.e. Louisville- Memphis* of the *un-used* USA network. Given the actual GSNR synthetic dataset, we set as ground reference the Δ evaluation obtained by considering the minimum GSNR value. This solution represents a rough approach and it is presented to provide a reference scale. This method produce an average margin of $\Delta = 0.98$ dB on the worst-case scenario channel. In this analysis, we focus on the worst-case scenario channel as it sets the required margin to guarantee all channels to be in-service. The QoT-E engine significantly refine the previous approach, providing a nominal GSNR value. In this case, without the ML support, the average margin is $\Delta = 0.69$ dB on the single path *Louisville- Memphis* for the worst-case scenario channel requiring a $p_{\mathrm{oos}} = 1\%$.

Finally, when the QoT-E engine is supported by the ML trained module, the proposed implementation provides a final QoT-E that enables to reach an average margin of $\Delta = 0.27$ dB on the same worst-case scenario channel requiring a $p_{\mathrm{oos}} = 1\%$.

The reliability of the joint estimation of the QoT-E engine with the trained ML module has been further verified on five

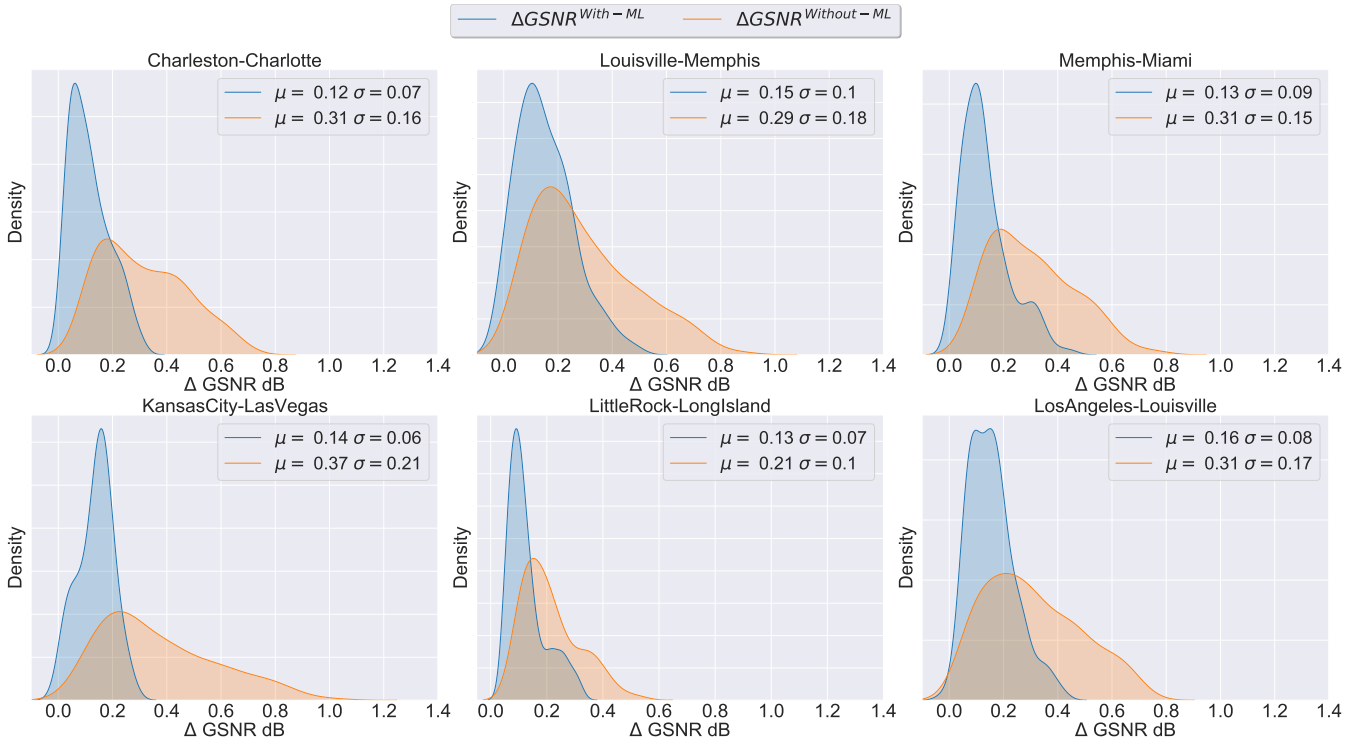

Fig. 12. ΔGSNR distributions obtained with and without ML module for all the investigated path of the *un-used* USA network

more paths of *un-used* USA network shown in Tab. 3. For all the considered paths, in Fig. 12 we report the distribution of the prediction error in both with and without ML support scenarios. In the figure, it can be observed that the ML agent significantly enhances the QoT-E; in all the cases, both the error mean, μ, and standard deviation, σ, values decrease considerably. Therefore, the required margin calculated with Eq. 6 can be reduced maintaining the same p_{oos}. For example, on the *Louisville- Memphis* path, the ML module application reduces the margin required to reach p_{oos} from 0.76 dB to 0.58 dB. This leads, in conclusion, to a smaller average margin and, therefore, to a more extensive exploitation of the LP capacity.

7. CONCLUSION

In this work, we proposed the use of a ML agent trained by a dataset of *in-service* network to correct the GSNR computation by the QoT-E in a sister *un-used* network from uncertainties caused by variations in the operational point of NEs. We analyzed a EU and a USA topology as *in-service* and *un-used* network, respectively. We synthetically generated dataset for both networks, for training and testing purposes, using the opensource GNPy library, specifically addressing the GSNR uncertainties due to connector losses and to ripples of the gains and NFs depending on the spectral load. We first analyzed the GSNR statistics in datasets showing that the GSNR statistics is always well approximated as Gaussian, so enabling a statistical approach to margin setting based on the maximum tolerable percentage of OOS.

Then, the main cognitive unit of a transfer learning agent is developed by using higher-level APIs of the *TensorFlow*© library. We show that the synergistic use of a transfer learning unit with the QoT-E may substantially reduce the inaccuracy in QoT-E, so enabling a reduction in the needed margin.

Finally, further studies can be performed enlarging the number of LPs under investigation. Moreover, in future analyses the considered set of configurations can be further enriched by including different fiber types and lengths and, even more, adding Raman amplification alongside EDFAs.

DISCLOSURES

The authors declare no conflicts of interest.

REFERENCES

1. Cisco, "Cisco Visual Networking Index: Forecast and Trends, 2017–2022," Tech. rep., Cisco (2017).
2. O. Gerstel, M. Jinno, A. Lord, and S. B. Yoo, "Elastic optical networking: A new dawn for the optical layer?" IEEE Commun. Mag. **50**, s12–s20 (2012).
3. G. Zhang, M. De Leenheer, A. Morea, and B. Mukherjee, "A survey on ofdm-based elastic core optical networking," IEEE Commun. Surv. & Tutorials **15**, 65–87 (2012).
4. V. Curri, A. Carena, A. Arduino, G. Bosco, P. Poggiolini, A. Nespola, and F. Forghieri, "Design strategies and merit of system parameters for uniform uncompensated links supporting nyquist-WDM transmission," JLT **33**, 3921–3932 (2015).
5. V. Curri, "Software-defined wdm optical transport in disaggregated open optical networks," in *ICTON 2020,* (2020), p. We.C2.1.
6. R. Pastorelli, "Network optimization strategies and control plane impacts," in *OFC,* (OSA, 2015).
7. M. Filer, M. Cantono, A. Ferrari, G. Grammel, G. Galimberti, and V. Curri, "Multi-Vendor Experimental Validation of an Open Source QoT Estimator for Optical Networks," JLT **36**, 3073–3082 (2018).
8. A. Ferrari, M. Filer, K. Balasubramanian, Y. Yin, E. Le Rouzic, J. Kundrát, G. Grammel, G. Galimberti, and V. Curri, "Gnpy: an open source application for physical layer aware open optical networks," JOCN **12**, C31–C40 (2020).
9. A. Ferrari, G. Borraccini, and V. Curri, "Observing the generalized snr statistics induced by gain/loss uncertainties," in *45th European Conference on Optical Communication (ECOC 2019),* (2019), pp. 1–4.

10. A. D'Amico, S. Straullu, A. Nespola, I. Khan, E. London, E. Virgillito, S. Piciaccia, A. Tanzi, G. Galimberti, and V. Curri, "Using machine learning in an open optical line system controller," JOCN **12**, C1–C11 (2020).
11. Y. Pointurier, "Design of low-margin optical networks," IEEE/OSA J. Opt. Commun. Netw. **9**, A9–A17 (2017).
12. M. Freire, S. Mansfeld, D. Amar, F. Gillet, A. Lavignotte, and C. Lepers, "Predicting optical power excursions in erbium doped fiber amplifiers using neural networks," in *2018 (ACP),* (IEEE, 2018), pp. 1–3.
13. J. Thrane, J. Wass, M. Piels, J. C. Diniz, R. Jones, and D. Zibar, "Machine learning techniques for optical performance monitoring from directly detected pdm-qam signals," JLT. **35**, 868–875 (2017).
14. F. N. Khan, C. Lu, and A. P. T. Lau, "Optical performance monitoring in fiber-optic networks enabled by machine learning techniques," in *2018 (OFC),* (IEEE, 2018), pp. 1–3.
15. L. Barletta, A. Giusti, C. Rottondi, and M. Tornatore, "Qot estimation for unestablished lighpaths using machine learning," in *OFC,* (OSA, 2017).
16. J. Mata, I. de Miguel, R. J. Duran, N. Merayo, S. K. Singh, A. Jukan, and M. Chamania, "Artificial intelligence (ai) methods in optical networks: A comprehensive survey," OSN. **28**, 43–57 (2018).
17. T. Jiménez, J. C. Aguado, I. de Miguel, R. J. Durán, M. Angelou, N. Merayo, P. Fernández, R. M. Lorenzo, I. Tomkos, and E. J. Abril, "A cognitive quality of transmission estimator for core optical networks," JLT. **31**, 942–951 (2013).
18. A. Caballero, J. C. Aguado, R. Borkowski, S. Saldaña, T. Jiménez, I. de Miguel, V. Arlunno, R. J. Durán, D. Zibar, J. B. Jensen *et al.*, "Experimental demonstration of a cognitive quality of transmission estimator for optical communication systems," Opt. express **20**, B64–B70 (2012).
19. E. Seve, J. Pesic, C. Delezoide, S. Bigo, and Y. Pointurier, "Learning process for reducing uncertainties on network parameters and design margins," J. Opt. Commun. Netw. **10**, A298–A306 (2018).
20. T. Panayiotou, S. P. Chatzis, and G. Ellinas, "Performance analysis of a data-driven quality-of-transmission decision approach on a dynamic multicast-capable metro optical network," J. Opt. Commun. Netw. **9**, 98–108 (2017).
21. W. Mo, Y.-K. Huang, S. Zhang, E. Ip, D. C. Kilper, Y. Aono, and T. Tajima, "Ann-based transfer learning for qot prediction in real-time mixed line-rate systems," in *2018 (OFC),* (IEEE, 2018), pp. 1–3.
22. R. Proietti, X. Chen, A. Castro, G. Liu, H. Lu, K. Zhang, J. Guo, Z. Zhu, L. Velasco, and S. B. Yoo, "Experimental demonstration of cognitive provisioning and alien wavelength monitoring in multi-domain eon," in *Optical Fiber Communication Conference,* (Optical Society of America, 2018), pp. W4F–7.
23. I. Khan, M. Bilal, M. Siddiqui, M. Khan, A. Ahmad, M. Shahzad, and V. Curri, "Qot estimation for light-path provisioning in un-seen optical networks using machine learning," in *ICTON,* (IEEE, 2020).
24. I. Khan, M. Bilal, and V. Curri, "Advanced formulation of qot-estimation for un-established lightpaths using cross-train machine learning methods," in *ICTON,* (IEEE, 2020).
25. I. Khan, M. Bilal, and V. Curri, "Assessment of cross-train machine learning techniques for qot-estimation in agnostic optical networks," OSA Continuum **3**, 2690–2706 (2020).
26. C. Rottondi, L. Barletta, A. Giusti, and M. Tornatore, "Machine-learning method for quality of transmission prediction of unestablished lightpaths," J. Opt. Commun. Netw. **10**, A286–A297 (2018).
27. S. Aladin and C. Tremblay, "Cognitive tool for estimating the qot of new lightpaths," in *OFC,* (OSA, 2018).
28. I. Khan, M. Chalony, E. Ghillino, M. U. Masood, J. Patel, D. Richards, P. Mena, P. Bardella *et al.*, "Machine learning assisted abstraction of photonic integrated circuits in fully disaggregated transparent optical networks," in *ICTON,* (IEEE, 2020).
29. R. Di Marino, C. Rottondi, A. Giusti, and A. Bianco, "Assessment of domain adaptation approaches for qot estimation in optical networks," in *2020 Optical Fiber Communications Conference and Exhibition (OFC),* (IEEE, 2020), pp. 1–3.
30. M. Lonardi, J. Pesic, T. Zami, and N. Rossi, "The perks of using machine learning for qot estimation with uncertain network parameters," in *Photonic Networks and Devices,* (Optical Society of America, 2020), pp. NeM3B–2.
31. D. Azzimonti, C. Rottondi, A. Giusti, M. Tornatore, and A. Bianco, "Active vs transfer learning approaches for qot estimation with small training datasets," in *Optical Fiber Communication Conference,* (Optical Society of America, 2020), pp. M4E–1.
32. Telecominfraproject, "Telecominfraproject/oopt-gnpy," (2019).
33. "Tensorflow core v2.1.0," https://www.tensorflow.org/.
34. V. Curri and A. Carena, "Merit of raman pumping in uniform and uncompensated links supporting nywdm transmission," J. Light. Technol. **34**, 554–565 (2016).
35. R. Pastorelli, S. Piciaccia, G. Galimberti, E. Self, M. Brunella, G. Calabretta, F. Forghieri, D. Siracusa, A. Zanardi, E. Salvadori, G. Bosco, A. Carena, V. Curri, and P. Poggiolini, "Optical control plane based on an analytical model of non-linear transmission effects in a self-optimized network," in *39th European Conference and Exhibition on Optical Communication (ECOC 2013),* (2013), pp. 1–3.
36. https://www.itu.int/rec/T-REC-G.694.1/en.
37. D. Pilori, F. Forghieri, and G. Bosco, "Residual non-linear phase noise in probabilistically shaped 64-qam optical links," in *OFC,* (2018).
38. M. Cantono, R. Gaudino, and V. Curri, "Potentialities and criticalities of flexible-rate transponders in dwdm networks: A statistical approach," IEEE/OSA J. Opt. Commun. Netw. **8**, A76–A85 (2016).
39. E. Virgillito, A. D'Amico, A. Ferrari, and V. Curri, "Observing and modeling wideband generation of non-linear interference," in *2019 21st International Conference on Transparent Optical Networks (ICTON),* (IEEE, 2019), pp. 1–4.
40. A. Carena, V. Curri, G. Bosco, P. Poggiolini, and F. Forghieri, "Modeling of the impact of nonlinear propagation effects in uncompensated optical coherent transmission links," JLT. **30**, 1524–1539 (2012).
41. A. Mecozzi and R.-J. Essiambre, "Nonlinear shannon limit in pseudo-linear coherent systems," J. Light. Technol. **30**, 2011–2024 (2012).
42. M. Cantono, D. Pilori, A. Ferrari, C. Catanese, J. Thouras, J.-L. Augé, and V. Curri, "On the Interplay of Nonlinear Interference Generation with Stimulated Raman Scattering for QoT Estimation," JLT. **PP**, 1–1 (2018).
43. Y. Ando, "Statistical analysis of insertion-loss improvement for optical connectors using the orientation method for fiber-core offset," IEEE Photonics Technol. Lett. **3**, 939–941 (1991).
44. M. Filer, J. Gaudette, Y. Yin, D. Billor, Z. Bakhtiari, and J. L. Cox, "Low-margin optical networking at cloud scale [invited]," IEEE/OSA J. Opt. Commun. Netw. **11**, C94–C108 (2019).
45. V. Curri, A. Carena, A. Arduino, G. Bosco, P. Poggiolini, A. Nespola, and F. Forghieri, "Design strategies and merit of system parameters for uniform uncompensated links supporting nyquist-wdm transmission," JLT. **33**, 3921–3932 (2015).
46. M. Bolshtyansky, "Spectral hole burning in erbium-doped fiber amplifiers," JLT. **21**, 1032–1038 (2003).
47. S. K. Korotky, "Network global expectation model: a statistical formalism for quickly quantifying network needs and costs," J. Light. Technol. **22**, 703–722 (2004).
48. P. Mishra, C. M. Pandey, U. Singh, A. Gupta, C. Sahu, and A. Keshri, "Descriptive statistics and normality tests for statistical data," Annals cardiac anaesthesia **22**, 67 (2019).
49. C. M. Bishop, *Pattern recognition and machine learning* (springer, 2006).
50. G. Hackeling, *Mastering Machine Learning with scikit-learn* (Packt Publishing Ltd, 2017).
51. M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard *et al.*, "Tensorflow: A system for large-scale machine learning," in *12th {USENIX} ({OSDI} 16),* (2016), pp. 265–283.
52. J. Duchi, E. Hazan, and Y. Singer, "Adaptive subgradient methods for online learning and stochastic optimization," JMLR **12**, 2121–2159 (2011).
53. C. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, "Activation functions: Comparison of trends in practice and research for deep learning," arXiv preprint arXiv:1811.03378 (2018).