This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JLT.2022.3160379, Journal of Lightwave Technology

Machine-Learning-Based Lightpath QoT Estimation and Forecasting

Stéphanie Allogba, Sandra Aladin, and Christine Tremblay, *Senior Member, IEEE*

Abstract— **Machine learning (ML) is more and more used to address the challenges of managing the physical layer of increasingly heterogeneous and complex optical networks. In this tutorial, we illustrate how simple and more sophisticated machine learning methods can be used in lightpath quality of transmission (QoT) estimation and forecast tasks. We also discuss data processing strategies with the aim to determine relevant features to feed the ML classifiers and predictors. We then introduce a preliminary study on the application of transfer learning to try to overcome the scarcity of field data.**

Index Terms— **Optical fiber communication, machine learning (ML), artificial neural networks, quality of transmission, support vector machine, performance prediction, forecasting, recurrent neural networks, long short-term memory, gated recurrent unit, transfer learning**

I. Introduction

THE continuous increase in Internet traffic in recent years has been greatly due to an influx of new technologies and applications (5G, video, cloud computing, etc.). Telecom operators are coping with this by deploying wavelength-division multiplexing (WDM) optical transmission systems that have ever-greater speeds, capacity, and flexibility. However, this increase in traffic accentuates the impact of performance degradation and network failures and with it, the need for flexible, autonomous network management [1]. To this effect, DSP-based coherent transceivers provide performance monitoring data and leverage machine learning (ML) for quality of transmission (QoT) estimation of lightpaths in complex and heterogeneous optical networks.

ML is a branch of artificial intelligence that allows a system to learn by itself by using monitored historical data, how it can work to solve a specific problem. ML has recently been explored for applications at both the physical and network layer of optical networks. In this paper, we present an overview on ML applications in optical communications, with a focus on lightpath QoT estimation and lightpath QoT forecast. The main question here is to be able to identify among all the ML techniques the one that is most suitable for estimating and forecasting lightpath QoT.

Manuscript submitted on November 30, 2021.

This work was supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada under grant RGPIN-2019-03972.

S. Allogba, S. Aladin and C. Tremblay are with the Network Technology Lab, Department of Electrical Engineering, École de technologie supérieure, Montréal, Qc, Canada H3C 1K3 (e-mail: christine.tremblay@etsmtl.ca).

The paper is organized as follows. Section II provides the main motivations behind the use of ML in optical communications. Section III presents an overview of some preliminary ML concepts, mainly those used in the following sections. It also summarizes the applications of ML in optical communications. Sections IV and V describe the ML applications for lightpath QoT estimation and lightpath QoT forecasting, respectively. Section VI summarizes the conclusions of this tutorial and highlights possible future work using ML for QoT estimation and prediction in optical communications.

II. Cognitive Optical Networking

With the emergence of bandwidth-hungry applications, such as 5G and Internet of Things (IoT), optical networks need to be more dynamic and autonomous. Current core networks are composed of high-capacity optical links (up to 80 x 100 Gb/s) between nodes equipped with a wavelength selective switch (WSS)-based reconfigurable optical add-drop multiplexers (ROADMs) for channel add-drop and routing. This wavelength-division multiplexing technique brings about a flexible structure offering multiple lightpath configurations to meet the ever-increasing demand for capacity. In addition, these lightpath configurations include different modulation formats, symbol rates, coding schemes, etc. However, despite this, network agility is still not fully exploited due to the complexity of tasks such as lightpath provisioning and rerouting, network reconfiguration and failure management, in the absence of fast and accurate tools in the network management system.

In this context, cognition has been introduced in optical networks as a way to efficiently control resources while achieving signal quality requirements [2]. As shown in Fig. 1, a cognitive optical network can be seen as a software-defined network (SDN) in which the network nodes are programmed by an SDN controller working in conjunction with a cognitive decision process. The cognitive processes use models fed by monitored performance on the network to make future decisions in an automated fashion and to act on the network through software adaptable elements. Coherent transceivers, can be considered as powerful performance monitors. Their integrated performance monitoring capability allows several system and link parameters to be dynamically monitored and leveraged by the control plane, in complement to the conventional power monitors. Coherent transceivers are also

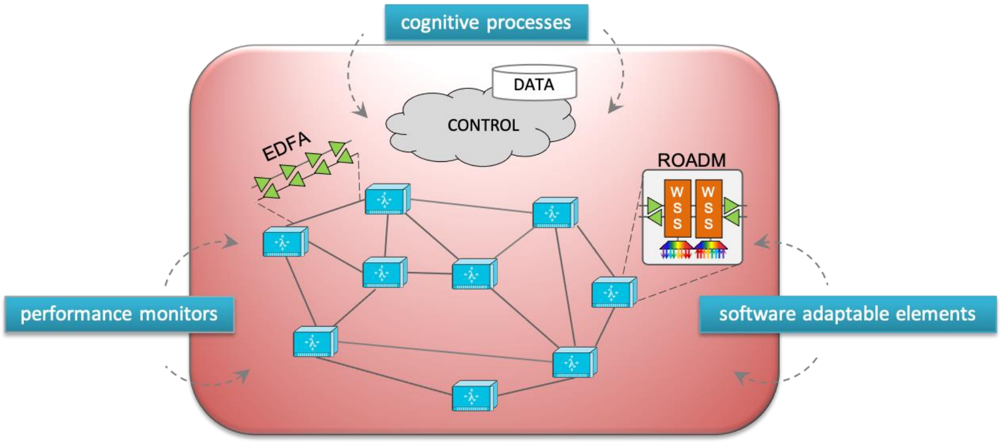


Fig. 1 Cognitive optical network.
EDFA: Erbium-Doped Fiber Amplifier; ROADM: Reconfigurable Optical Add-Drop Multiplexer; WSS: Wavelength Selective Switch

powerful software-defined optical devices. With their digital signal processing (DSP) capability, they can compensate for propagation effects electronically. The most recent flexible transceivers based on multidimensional formats and constellation shaping can also modify or adapt channel bandwidth and bit rate, making them efficient software adaptable elements, together with ROADMs, for enabling cognitive optical networking.

III. Machine Learning in Optical Communications

ML is resulting from contributions in several research areas: computer sciences, mathematics, economics, neurosciences, control theory, etc. It has recently become an enabler to build cognitive optical networks, thanks to increased data availability and computing capabilities. ML is grouped into three families of ML techniques:

- Supervised learning: all database information, including labels and features, is known. These techniques are used for classification problems, such as detection tasks, or for regression problems such as risk assessment or forecasting tasks. Common supervised learning techniques are K-nearest neighbors (KNN), random forest (RF), support vector machine (SVM), neural networks (NN) and case-based reasoning (CBR).
- Unsupervised learning: in this group, the labels are not defined. It is used primarily for clustering, density estimation and dimensionality reduction. The areas of application include medical tasks, pattern recognition, facial recognition, text mining etc. The most widely used algorithms are k-mean partitioning, expectation maximization (EM), principal component analysis (PCA).
- Reinforcement learning (RL): it is a reward-based approach. In this group, an agent learns to make decisions in a complex environment through rewards. One common algorithm used in RL is Q-learning.

Machine learning in optical communications has become a hot research topic in the last years. A good overview of the applications of ML algorithms in the physical and network layers can be found in [3]. In the physical layer, the ML use cases fall into 5 categories [3]:

- Lightpath QoT estimation: this consists of classifying the QoT of a lightpath before its establishment in the network, using supervised ML trained with synthetic or field data, as a potentially faster method (compared to analytical models) to assess the bit error rate (BER), signal-to-noise ratio (SNR) and optical signal-to-noise ratio (OSNR). This also consists of forecasting the QoT of deployed lightpaths using supervised learning methods trained with historical QoT data, for proactive maintenance and channel margin optimization purposes.
- Optical amplifier control or the control of channel power excursions in optical line systems using supervised, unsupervised and reinforcement learning methods.
- Modulation format recognition in coherent digital receivers using supervised and unsupervised learning.
- Nonlinearity mitigation in coherent optical systems using supervised and unsupervised methods.
- Optical performance monitoring which aims at estimating transmission system parameters from a mapping between the lightpath parameters and the properties of the received signal using supervised ML methods.

In the network layer, ML use cases can be divided into 4 categories [3]:

- Traffic prediction and virtual topology redesign tasks to allow proactive traffic rerouting and network configuration using supervised and unsupervised learning algorithms.
- Failure management: this includes failure detection, failure localization and failure root cause methods based on supervised learning to determine appropriate restoration, traffic reconfiguration or field interventions.
- Traffic flow classification for efficient resource allocation and traffic priority management using supervised and unsupervised learning methods.
- Path computation using supervised and unsupervised learning fed with service demand and network status information for selecting appropriate paths and available network resources that meet the desired quality of service (QoS) without affecting existing services provisioned in the network.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JLT.2022.3160379, Journal of Lightwave Technology

In this paper, we focus on signal quality estimation tasks which deal with QoT estimation of unestablished lightpaths and short-term QoT forecast of established lightpaths. These two specific tasks - which aim at fast connection provisioning and proactive network management, respectively - can be handled by supervised learning approaches.

First, let's consider QoT estimation before lightpath establishment. Lightpath QoT can be defined by different metrics, such as the BER, SNR and OSNR). The accurate estimation of lightpath QoT involves a precise assessment of the linear and nonlinear interference (NLI) noise contributions. When using analytical models, these metrics are calculated as a function of system and link parameters. Analytical models achieved this with methods such as the split-step Fourier method or the Gaussian noise with the assumption that the parameters used are accurate.

In the additive white Gaussian noise (AWGN) model proposed in [4-6], the nonlinear OSNR is calculated from the channel power, the amplified spontaneous emission (ASE) noise power and the nonlinear interference (NLI) noise power. These ASE and NLI noise power depend on variables that describe the status of particular lightpaths and require computing steps before acquiring the final QoT metric of interest. Equations (1), (2) and (3), as detailed in [4-6], show how these parameters are used to estimate the nonlinear OSNR.

$$OSNR_{NL} = P_{TX}/(P_{ASE} + P_{NLI}) \tag{1}$$

$$P_{ASE} = N_S F(A_S - 1)h\mu B_N \tag{2}$$

$$P_{NLI} = 8\, N_S \gamma^2 L_{eff} P_{TX}^3 B_N \log\left(\pi^2 \beta L_{eff} N_{ch}^2 R_S^2\right)/(27\pi\beta R_S^3) \tag{3}$$

where P_{TX} is the average channel power, N_{ch} is the number of channels, R_S is the channel baud rate and μ is the channel frequency. These are the WDM channel's parameters. The noise figure F and the gain A_S are the Erbium-doped fiber (EDFA) amplifier parameters. The noise bandwidth B_N is generally assumed to be 0.1 nm (12.5 GHz at 1550 nm). The link parameters are the effective length L_{eff}, the number of spans N_S, as well as the dispersion coefficient ß and nonlinearity coefficient γ of the optical fiber.

The basic AWGN analytical model, or the generalized OSNR (GOSNR) model proposed later on, can be used to estimate the QoT of candidate lightpaths [4-9]. These analytical formulations, well known for their applications in the GNPy tool proposed in [10] and in the data generation algorithm implemented in the cognitive QoT tool described in [11], can provide accurate QoT estimations in simple deployment scenarios where all the channel, system and link parameters are known. The problem becomes more complex in production networks which are very heterogenous (fiber types, transmission systems, outside plants, etc.). Equipment inventory and network topology data are often incomplete or not up-to-date. Elastic networks carrying channels of different formats and bandwidth, and open line (or disaggregated) systems composed of components from different vendors make things even more complex. In other words, heterogeneity and flexibility in the physical layer of optical networks make it a challenge to perform the analytical computation of the QoT of lightpaths. In such a context, can ML be used in lightpath QoT estimation and provide guidance for complex routing and resource allocation tasks in heterogeneous optical networks? Section IV deals with this problem.

Second, let's consider QoT forecast of deployed lightpaths. Predicting lightpath QoT can be a simple task in stable environments where system and link parameters remain constant over time. However, system performance can fluctuate over time due to temperature effects, aging or malfunction. Optical fiber is not only an excellent transmission medium but also a powerful distributed sensing device whose transmission characteristics can be affected by environmental conditions (temperature and wind effects, cable handling and bending effects, vibrations, etc.). In such dynamic and complex environments that cannot be handled by analytical models, predicting lightpath QoT can become a much more challenging task where ML can potentially be useful. In this case, the objective of QoT forecasting is to detect performance degradations early enough to trigger changes in the parameters (modulation format, symbol rate, optical power, etc.) of programmable transceivers, proactive maintenance or network reconfiguration before transmission errors occur. Section V deals with this problem.

IV. Machine Learning for Lightpath QoT Estimation

ML-based QoT estimation of candidate lightpaths prior to their establishment can help overcome the heavy computation time and parameter uncertainty in analytical models. ML-based QoT estimation models can learn from labeled data and predict the QoT of unestablished lightpaths.

Fig. 2 presents a cognitive lightpath QoT estimation tool using four classifiers based on K-nearest neighbors (KNN), support vector machine (SVM), random forest (RF) and neural networks (NN) techniques, as proposed in [11-12]. KNN and RF classifiers use a majority vote concept to determine the class labels of new records. SVM aims at maximizing the distance between hyperplane and support vectors while NN use interconnected processing units (neurons) for classification.

Fig. 2 shows the lightpath classification process. Classifiers based on supervised learning require labeled data (i.e., records of both "good QoT" and "poor QoT" types) for training. It is very difficult to get poor QoT data in a production network. Therefore, synthetic or lab data must be used to train the QoT estimators. In [11-12], synthetic data were generated with a MATLAB tool using the AWGN model, which estimates the BER for combinations of parameters describing different lightpaths of different optical link configurations.

The constructed knowledge base (KB) was then split into training and test datasets. Then four classifiers were trained with the training data:

- KNN classifier determines class labels through majority vote among K nearest neighbors. It is considered a "lazy

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JLT.2022.3160379, Journal of Lightwave Technology

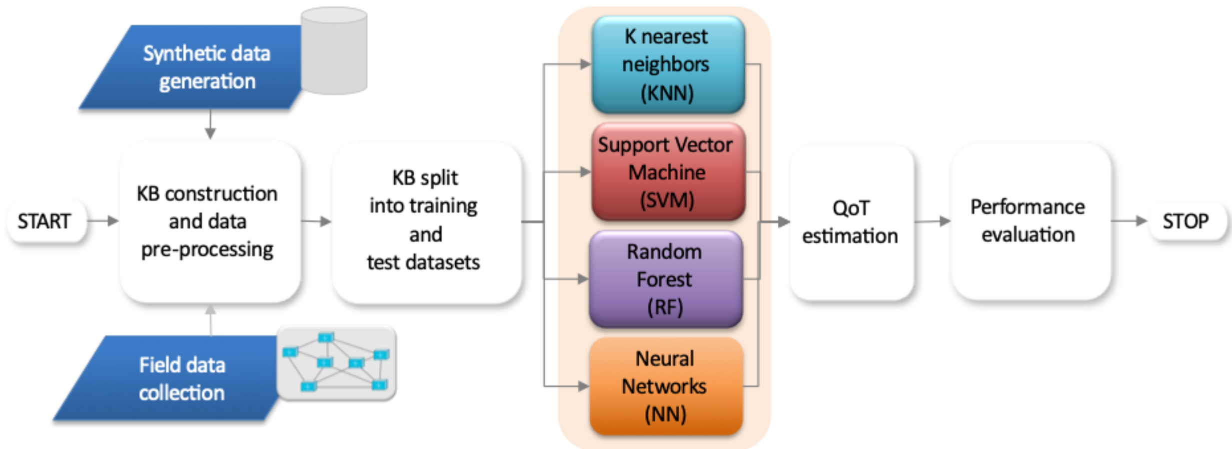


Fig. 2 Cognitive lightpath QoT estimation tool.

learner" because it requires that most computation be done during the testing to determine the class of the new record. This can be expensive for large datasets. The training dataset is split into 5 subsets (folds) for training and validation (5-fold cross validation). The model is iteratively fitted 5 times, each time training the data on 4 of the folds and evaluating on the 5th fold (called the validation data).

- SVM aims at maximizing the distance between hyperplane and support vectors. In the SVM model, kernel functions are used as parameters and help to determine the shape of the hyperplane and decision boundary. In nonlinear SVM, a Gaussian radial basis function kernel is used, the parameters C and gamma are computed to optimize the classification model:
  - C makes it possible to control the influence of each support vector;
  - Gamma is linked to the Gaussian kernel whereby a large value leads to high bias and low variance, and vice-versa.
- RF uses majority voting amongst random decision forests to classify new records. Among others, the hyperparameters that can be tuned are the number of decision trees in the forest (estimators), which affects the model variance for more generalized results, but also the training time. The maximum number of leaf nodes which helps restricts the growth of the trees and the maximum depth of individual trees which defines the longest path between the root node and the leaf node.
- NN is a sophisticated classifier which uses interconnected processing units (neurons) for classification. The hyperparameters considered in this technique are the optimizer, the activation and the loss functions, the learning rate, the number and size of hidden layers, the number of neurons, etc.

The methodology used to build the models with the four aforementioned techniques consisted in a training and optimization step. For the KNN model, a 5-fold cross-validation method was used with the Euclidean distance to determine the best K nearest neighbors necessary for the classification of new lightpaths. The parameters C and gamma were computed using 5-fold cross-validation, as in the KNN training; and a Gaussian radial basis function kernel was used to consider the nonlinearity aspect of the SVM classifier when no linear classification of the data is possible. For the RF algorithm, the number of estimators was the parameter used to optimize the built model. For the NN, the number of epochs, the learning rate as well as the number of hidden layers and neurons were the hyperparameters tuned to optimize the model.

The classification accuracy was used to assess the classifiers' performance. Also, confusion matrices provided a good view of classification errors and made it possible to compute other metrics useful when working with imbalanced datasets.

A. QoT Estimation: Do We Need Machine Learning?

Analytical QoT estimators are computationally heavy or necessitate a margin to account for parameter uncertainty when used in production networks.

ML predictive models are used to make predictions on unseen data by learning the mapping function between the training data and the analytically estimated labels. In other words, ML QoT classifiers use specific features to predict classes of candidate lightpaths before their establishment, while providing better estimation performance than analytical QoT estimators, according to experimental results [11].

B. QoT Estimation: Do We Need Neural Networks?

Several ML-based QoT estimators have been proposed in the literature [11-19]. Good overviews of the most recent ML models and tools developed for lightpath QoT estimation can be found in [20-22]. The methodology consists in predicting whether the QoT of the candidate lightpath is above or below a predefined threshold. Artificial neural networks (ANN) are powerful information processing models which provide very efficient classification and regression models once they are properly configured and trained. However, building NN

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JLT.2022.3160379, Journal of Lightwave Technology

TABLE I. PERFORMANCE EVALUATION OF SIMPLE LIGHTPATH QOT CLASSIFIERS BASED ON KNN, RF AND SVM [11]

Performance metric	**KNN**	**RF**	**SVM**
Classification accuracy (%)	81.8	96.3	99.2
Error rate (%)	18.3	3.7	0.9
False positives (%)	16.1	2.1	0.4
Computation time (s)	0.40	0.83	1.26

models can be challenging. The objective in this section is to determine whether it is worth using complex NN for QoT estimation tasks and whether simpler models can perform better than NN.

Simple QoT estimators based on KNN, RF and SVM were proposed in [11]. These models were built using a synthetic KB of 25,600 instances generated by varying 6 parameters in the Gaussian Noise model: the total link length, the span length, the number of spans, the channel input power, the modulation format, and the data rate [11].

A comparative analysis of the best performing QoT estimator in [11], the SVM model, and a NN-based model was presented in [12]. The SVM and NN models were trained by using a synthetic KB of 38,400 instances generated by considering a wider range of parameter values in the Gaussian model.

The performance results obtained in [11-12] are summarized in TABLE I and TABLE II. The QoT estimators were evaluated using classification accuracy, error rate and the computation time. The classification accuracy is used to define the proportion of the total number of predictions that were correct while the error rate represents the proportion of the total number of incorrect predictions. Additionally, the false positive rate was computed to assess the performance of the classifiers by determining the incorrectly identified positive instances. In TABLE I, we can see that for the simple classifiers the best performance was achieved with SVM.

TABLE II shows that the NN model performed slightly better than SVM in terms of classification accuracy. The accuracy of the QoT estimator of the ANN model is 99.6% with a recall of 98.9% and an F1 score of 99.1%. Note that these last two metrics determine the correctly predicted classes in the case of unequal class distribution. The NN classifier performed better than the SVM with a false positive rate of 0.2%, compared to 0.4% for SVM, and the shortest computing time (12 times faster than SVM). The models were executed on a system with an Intel Core i5-8600K 3.6 GHz CPU, 16 GB RAM and a GTX 970 GPU.

Feature engineering aims at evaluating the impact of different features on the classification accuracy in the hope of reducing the feature set to only those that are the most important. In the study presented in [12], this concept was used for building NN and SVM QoT estimators with two different feature sets. Interestingly, it was found that while the two models performed comparably with a set of 6 features, the SVM classifier performed better than the NN classifier with a reduced feature set.

TABLE II. PERFORMANCE COMPARISON OF SVM AND NN QOT ESTIMATORS [12]

Performance metric	**SVM**	**NN**
Classification accuracy (%)	99.4	99.6
Error rate (%)	0.6	0.4
False positives (%)	0.4	0.2
Computation time (ms)	3.45	0.28

These results justify the need to deepen the experiments on the application of NN in QoT estimation to possibly benefit from their reputed ability as universal function approximators.

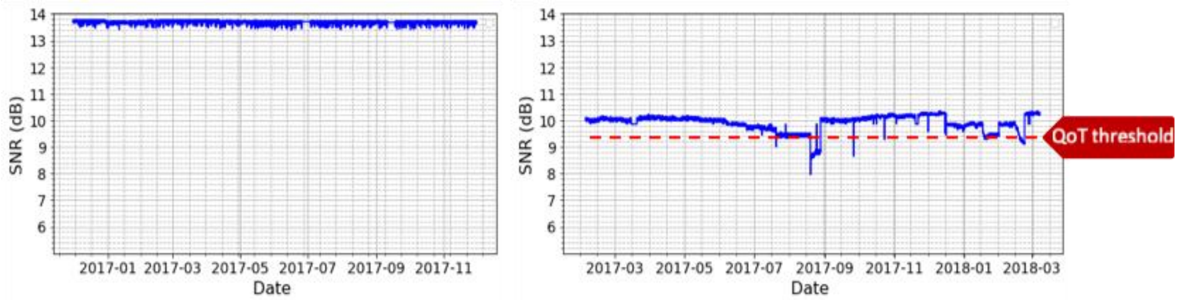
C. Neural Networks For Lightpath QoT Estimation

Several recent studies show a tendency towards favorable results for ANN based estimators [13-19].

In [13], the authors considered four ML models in both classification and regression by using a knowledge base built based on four reference network topologies and different features: KNN, logistic regression, SVM and ANN. Their ANN-based QoT estimator presented good generalization and a prediction accuracy of almost 99.9% with 0.04% false positives for residual channel margin prediction. In [14], an ANN model using five features was implemented for lightpath QoT prediction. The OSNR prediction model achieved a mean relative error (MRE) close to 1% with the full set of features. The authors in [15] used a transfer learning approach with an ANN-based QoT predictor. The results showed good performance for a database size as small as 20 records. The authors in [16] presented true OSNR vs. predicted OSNR, with points appearing along the diagonal of the graph. Such a behavior describes the good accuracy of the ANN-based QoT estimator. Moreover, in [17] a deep graph convolutional neural network (DGCNN) QoT estimator could predict the bit error rate with accuracies ranging between 92% and 97%. In [18], it was possible for an ANN-based model to predict the SNR with a standard deviation of the SNR estimation error of 0.13 dB as opposed to 0.2 dB for an analytical model. The authors in [19] used a synthetic knowledge base (KB) built from the GNPy simulation tool, with three different ML techniques (RF, NN and KNN) to predict the generalized signal-to-noise ratio (GSNR) of unestablished lightpaths. The ANN-based QoT estimator was proven to produce the best results for the replicated European (EU) and USA network topologies with a mean absolute error (MAE) of 0.001 dB and 0.005 dB respectively.

These results suggest that NN techniques are suitable for QoT estimation and offer the advantage of classification and prediction accuracies. However, we have observed in [9] that with reduced number of features, the SVM classifier performed better than its ANN counterpart. It achieved accuracies of 93.3% and 88.5% with 4 and 3 features respectively, as opposed to 92.3% and 85.0% for ANN. Moreover, with half the number of features, we noticed a reduction in the computation time by a factor of 3 for the SVM models while it remained approximately the same for

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JLT.2022.3160379, Journal of Lightwave Technology



(a) (b)

Fig. 3. Evolution of SNR over time for 2 lightpaths. QoT threshold level set arbitrarily for illustration purposes.

the ANN models.

In summary, we can say that the jury is still out as to determining whether neural networks are the best classifiers to use for lightpath QoT estimation. More importantly, field data collected in real-world scenarios are required to fully assess the potential and applicability of ML-based and more specifically NN-based QoT estimators.

V. Machine Learning for Lightpath QoT Forecast

Unlike the QoT estimation statement, the problem of ML-based QoT prediction is considered once the lightpath is deployed within the network. Lightpath QoT is monitored at the receiver level through various performance metrics, such as pre-FEC BER, SNR, Q-factor, and received optical power (P_{RX}). It is therefore a prediction problem whose objective is to predict the future QoT lightpath over a certain forecast horizon.

A. *Lightpath QoT Forecast: Is Machine Learning Needed?*

QoT forecast has been studied extensively in complex dynamic wireless environments. ML-based forecasters based on recurrent neural networks have been proposed for link quality prediction in wireless community networks to improve the performance of routing protocols [23,24]. Models based on deep learning have also been proposed for signal quality prediction in radio networks [25,26].

QoT forecast in optical networks has started to be explored only very recently [28,29,34,35], thanks to the availability of field data showing that optical network environments and lightpath performance behavior can be quite dynamic.

Fig. 3 shows the SNR evolution over a period of several months for two optical lightpaths carried in different optical links of production networks. Note that the performance of the lightpath can be very stable over time, as shown in Fig. 3(a). In such a case, performance prediction is a simple task that can be handled without the need for complex methods. However, as shown in Fig. 3(b), it may happen that some lightpaths exhibit a more dynamic performance behavior over time. This notion of performance dynamism is visible in Fig. 3(b) from the performance degradations observed in the spring as well as the decreases and increases in the performance in winter. Note that the variations in performance in this particular example, can drop by a few dBs for periods and can last from a few days to several months and cannot be explained by channel add-drop and switching. This behavior can generate errors affecting data transmission and ultimately the customers and triggering the need for the network operator to apply corrective actions (if the SNR drops below a critical performance threshold, for example). Such a dynamic behavior cannot be explained or predicted by theoretical models (such as the GN model). These variations in SNR can last from a few days to several months and cannot be explained by channel add-drop and switching.

For this reason, lightpath QoT prediction could help network operators to detect performance degradation early enough to proactively reroute traffic or trigger maintenance actions. It could also be used for margin optimization, by adapting the bandwidth occupation and bit rate of flexible transceivers.

Thus, the objective of ML-based QoT forecasting is to predict the future QoT lightpath over a certain forecast horizon. Fig. 4 illustrates the implementation of QoT forecast models. It is divided into three steps. The first step is the construction of the KB. This includes preprocessing of the database, namely the management of missing or outlier data, and the statistical analysis (seasonality, stationarity, etc.) performed on the data. The second step is the construction of the forecast models. The last step is the performance evaluation phase for assessing the robustness and the precision of the predictions made by the models.

B. *Performance Monitoring Data*

ML-based QoT forecasters use historical data from established lightpaths to predict performance degradations and equipment failure or to perform channel margin optimization. This historical data come from performance metrics (PMs) collected in three different production networks:

- CANARIE production network: the KB includes PMs collected on 100 Gb/s polarization multiplexed quadrature phase shift keying (PM-QPSK) channels transported over different routes of the CANARIE production network. The channels are deployed on different routes, both buried and aerial fiber types, up to 1,500 km. Moreover, PMs are composed of time series of P_{RX} and SNR collected at 15-minute sampling intervals over an observation period of up

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JLT.2022.3160379, Journal of Lightwave Technology

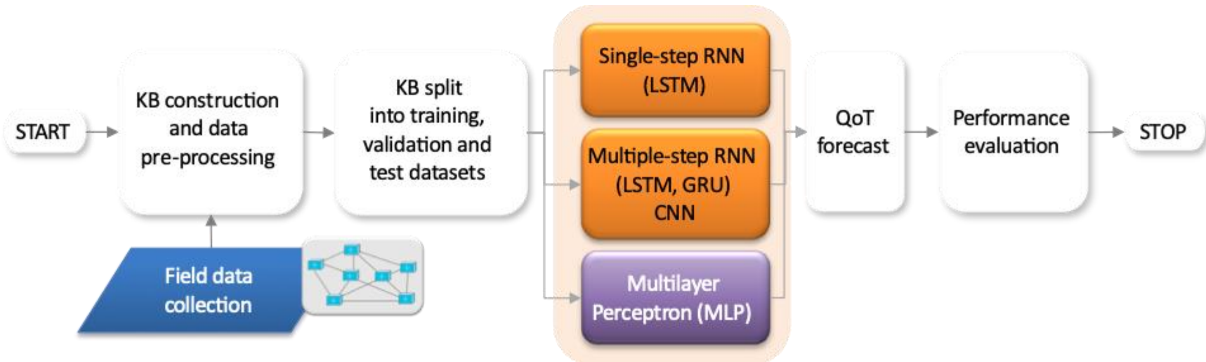


Fig. 4. Lightpath QoT forecasters based on recurrent neural networks.

to more than 12 months.

- North American service provider (NASP) network: the KB includes PMs on 140 PM-QPSK channels at 100 Gb/s transported over 10 routes in the network. These channels are deployed on routes, both buried and aerial fiber types, ranging from 100 to 1,400 km. PMs are composed of time series of P_{RX}, Q-factor and BER, collected at 15-minute sampling intervals over an observation period ranging from 5 to 12 months. Note that the Q-factor is a figure of merit that is directly related to the BER.
- Microsoft (MS) network: The KB includes PMs across 4,000 channels transported over 15 routes in the Microsoft optical backbone. The PM considered here is the Q-factor, collected at 15-minute sampling intervals over a 14-month observation period.

Note that additional link parameters were also monitored using coherent transceivers for impairment compensation, such as chromatic dispersion (CD), polarization mode dispersion (PMD), polarization dependent loss (PDL), etc.

The resulting KBs contain gaps (missing records) and outliers (i.e., sudden drops such as the ones observed in Fig. 3 (b)) in the time series. These gaps and outliers, which vary in duration from a few minutes to several hours, are typically the result of software upgrades or maintenance activities. Additionally, studies of field data collected on buried and aerial fiber links in the NASP network revealed daily and seasonal patterns in SNR time series and daily variations in the PDL time series [27].

The data preprocessing step in Fig. 4 mainly involves filling the gaps and removing the outliers in the time series as both gaps and outliers can impact the performance of the models. The missing data can be replaced by a moving average, as in [28]. In the case of the NASP and MS datasets, outliers were observed. The outliers can be defined by data instances whose distance from the mean QoT value is greater than 3 times the standard deviation and replaced by the corresponding median value in the time series, as in [29].

Data preprocessing also includes stationary tests applying transformation such as differencing to make the time series stationary. However, recurrent neural networks such as long-short-term memory can learn the non-stationary nature of time series, making such transformations optional [30].

C. Recurrent Neural Networks

Recurrent neural networks (RNN) are a class of neural networks that are used in speech recognition and natural language processing tasks. They can recognize patterns in time series and use them to make predictions. Therefore, they seem to be well suited for ML problems that involve sequential data.

As shown in Fig. 4, four different variants of neural networks were used in the development of lightpath QoT forecasters.

1) Long Short-Term Memory

Long-term memory (LSTM) is a type of RNN that learns long-term dependencies between time steps of sequence data. Thus, it can memorize information for long periods of time.

Fig. 5(a) shows the topology of the LSTM network that groups the N time steps units of the time series. As such, it contains information in gated cells and uses structures called gates to control cell states and input information to determine the outputs. Thus, the updated output and cell states are calculated using the current network state and the next time step.

Several parameters come into play in the optimization of LSTM models, which make the optimization process complex:

- the number of hidden layers, i.e., the number of units in an LSTM cell,
- the number of epochs, i.e., the number of iterations to train the model,
- the function to be used as a solver to determine the states of the cells,
- the function to calculate the gates in the LSTM unit.

Note that the number of hidden layers and number of epochs can be determined by selecting the parameters minimizing the prediction errors. On the other hand, the function Adam can be used as the solver, as in [31, 32].

2) Encoder-Decoder LSTM

Fig. 5(b) shows the topology of the encoder-decoder LSTM. First, it compresses the input SNR sequence into a set size context vector. Then, the internal memory cell is updated at each time step until it reaches the end of the sequence. Finally, the decoder LSTM produces an output sequence from the context vector with its input taken from the previous forecast.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JLT.2022.3160379, Journal of Lightwave Technology

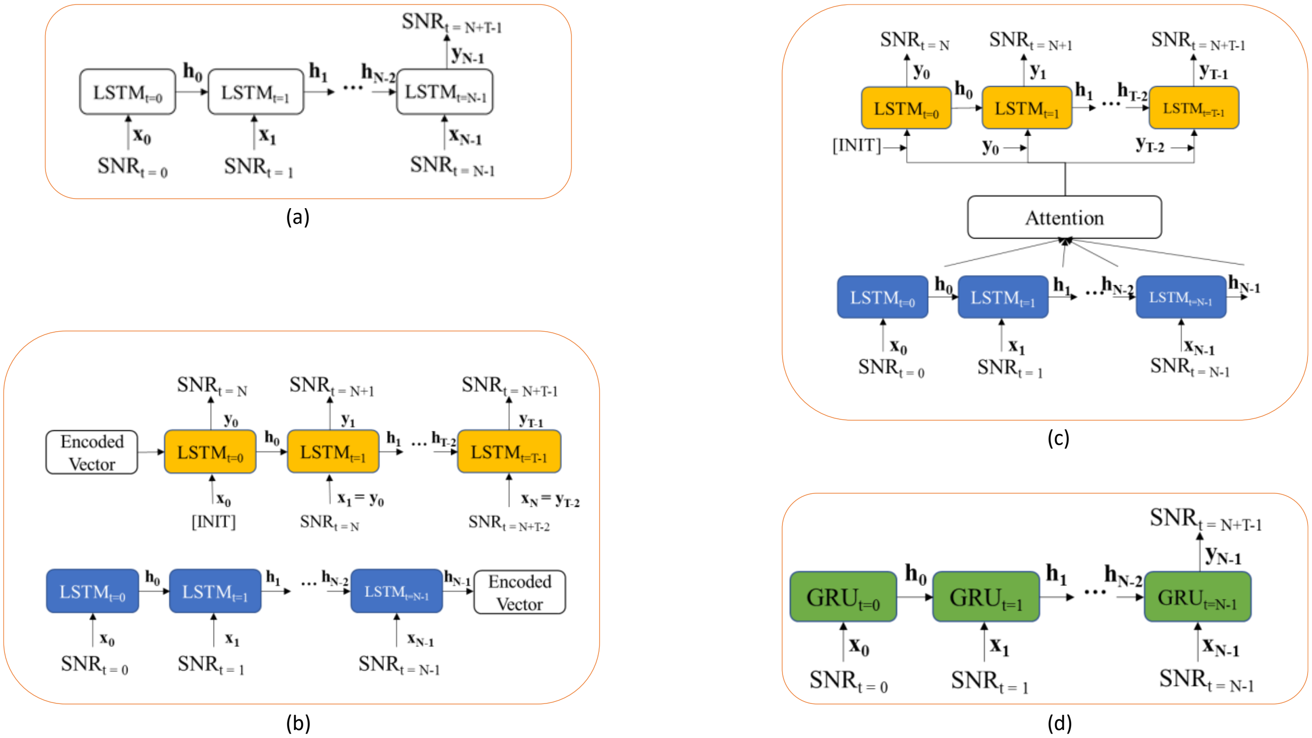


Fig. 5. Multi-step prediction: (a) LSTM topology; (b) Encoder-Decoder LSTM; (c) Encoder-Decoder LSTM with attention; (d) GRU

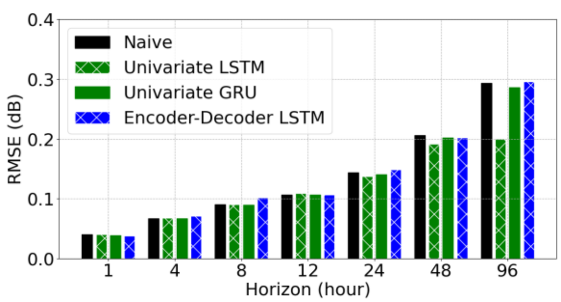


Fig. 6. RMSE of the RNN models for forecast horizons ranging from 1 to 96 hours. From [28].

Note that the attention mechanism, as shown in Fig. 5(c), adds focus capability on different parts of the input sequence by producing weighted encoder outputs as inputs for the decoder LSTM.

Both these mechanisms are well described in [32,33].

3) Gated Recurrent Unit

Fig. 5(d) shows the topology of the gated recurrent unit (GRU). The GRU is a lighter version of LSTM and it contains less gates for controlling the cell states. It is therefore less complex and faster to run than the LSTM [31].

4) Multilayer Perceptron

An MLP is a simple class of feedforward neural networks, which maps an input (historical window) to an output (future target). It consists of at least three layers of nodes: an input layer, a hidden layer, and an output layer. Except for the input nodes, each node is a neuron that uses a nonlinear activation function. An MLP is characterized by several layers of input nodes connected as a directed graph between the input and output layers. Note that, in a feedforward network, there are no feedback connections in which outputs of the model are fed back into itself, contrary to the LSTM and GRU models [31].

D. Performance Analysis

1) Univariate RNN Models

The first studies on lightpath QoT forecast involved RNN models trained with single-lightpath field data [12, 28].

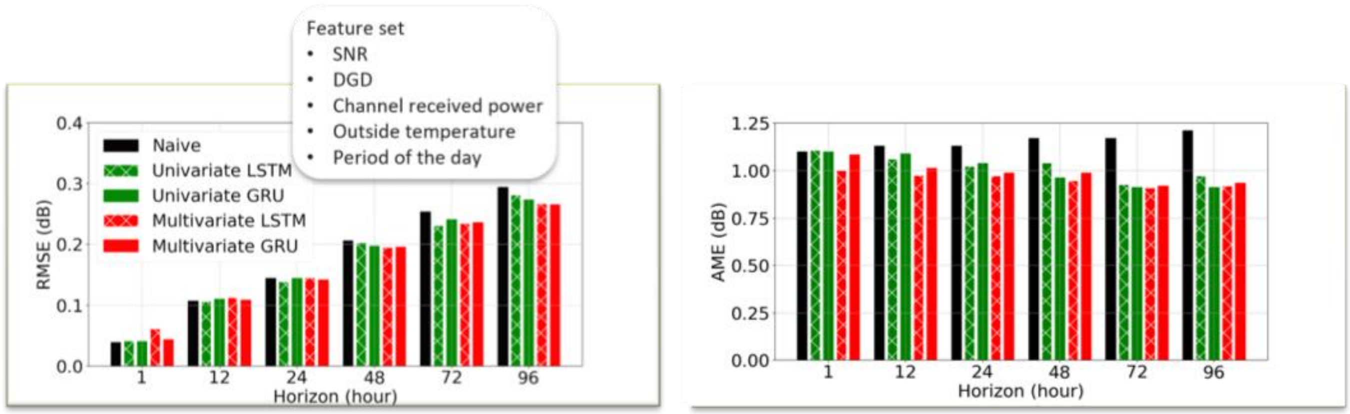
The main reason for that was the very limited field data available for such studies. These RNN models, using historical single-lightpath SNR data as the only input, included LSTM, encoder-decoder LSTM and GRU. A persistence model (also called naïve model) was also implemented to evaluate the performance of the models as in [28]. The naive method consists in assigning to the SNR value predicted at the horizon T, the last value of the observation window N.

The KB used in these initial studies was built from the CANARIE dataset. To build the models, first the KB was split into validation, training, and test datasets. The validation and training datasets are used to determine the appropriate hyperparameters and to train the models.

The univariate models are evaluated using the RMSE as the performance metric. This metric is used to determine the accuracy of models. It indicates how close the observed values are to the predicted values. Thus, low RMSE values indicate better prediction. The obtained RMSE is shown in Fig. 6, presented in [28]. The performance of SNR prediction models was evaluated for forecast horizons ranging from 1 to 96 hours. The RMSE increased with the forecast horizon for all models, as expected, with a maximum improvement of 0.02 dB at 96 hours for the LSTM compared to the naive method.

Moreover, the GRU RMSE curve closely followed the naive RMSE curve, and the LSTM encoder-decoder obtained the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JLT.2022.3160379, Journal of Lightwave Technology



(a) (b)

Fig. 7. Performance evaluation of the multivariate forecasting models: (a) RMSE; (b) AME. From [35]

TABLE III. LIGHTPATH QOT FORECASTERS: SUMMARY OF PERFORMANCE RESULTS

Metric	**Forecast horizon (hours)**	**Baseline**	**Univariate model**		**Multivariate model**	
		Naive model	**LSTM**	**GRU**	**LSTM**	**GRU**
		Features				
		SNR			Channel SNR, DGD, P_{RX} Outside T, Period of the day	
RMSE (dB)	1	0.04	0.04	0.04	0.06	0.04
	24	0.14	0.14	0.14	0.14	0.14
	96	0.29	0.28	**0.27**	0.27	0.27
AME (dB)	1	1.10	1.11	1.10	**1.00**	1.08
	24	1.13	1.02	1.04	**0.97**	0.99
	96	1.21	0.97	**0.91**	0.92	0.93

best performance in the very short-term forecast horizon (1-hour). Hence, the LSTM was positioned as the best model among the forecast models tested at higher forecast horizons with the lowest RMSE values. However, it could not beat the naive method at lower forecast horizons (1 to 12 hours).

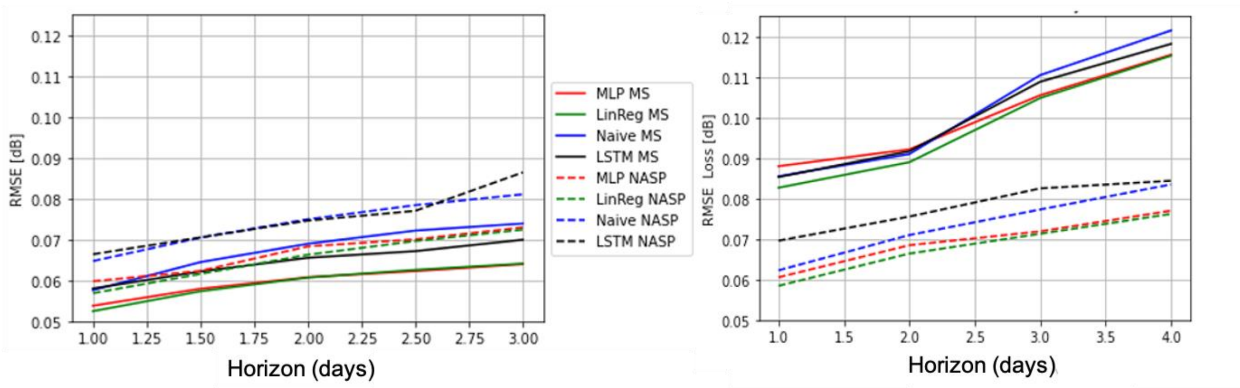
Moreover, in [34] the authors developed a one-dimensional convolution neural network (1D-CNN) model, using the NASP field data for short-term multi-step performance prediction. The model performed well in capturing and predicting the temporal change in SNR 24 hour ahead.

2) Multivariate Models

The next question is whether the prediction performance of the univariate QoT forecasters can be improved by considering multivariate models. Multivariate models aim to improve the performance of the forecasters by adding supplementary inputs (or features) to the models. This was the objective of the study realized in [35], using a KB of data collected for one lightpath in the CANARIE production network including the channel BER and 5 additional features: the channel received power and DGD, as well the outside temperature and time of day.

In Fig. 7, multivariate LSTM and GRU models (in red) using the 5 available features are compared to their univariate counterparts (in green), using a naive model (in black) as a baseline as in the original univariate case, and the RMSE and the absolute max error (AME) as performance metrics. The AME, which is the maximum error of the forecast model, was used to evaluate the impact of bad predictions on model performance. The results show a slightly better performance for multivariate models at certain forecast horizons. This performance advantage (especially in AME) could potentially increase as more features become available for such studies.

We can see that the performance differential between the ML models and the naive model (namely the AME advantage) increases at longer forecast horizons, showing the advantages of ML and its potential applicability to longer forecast horizons.



(a) (b)

Fig. 8. Performance of the lightpath QoT forecasters, as obtained by executing the models on two different subsets of the MS and NASP datasets: (a) subset 1; (b) subset 2. From [29].

TABLE III presents a summary of the performance results obtained with the multivariate and univariate lightpath QoT forecasters trained field data from a single lightpath [35]. The RMSE and AME values are displayed for 1-hour, 24-hour and 96-hour forecast horizons, with the best performers (lowest values) shown in bold. The first observation is that complex RNN models did not show an RMSE advantage over a simple naive method at very short forecast horizons (24 hours or less). However, according to the AME performance metric, the multivariate models outperformed their univariate counterparts as well as the naive models on all forecast horizons, except at the 96-hour horizon where the univariate GRU model shows a slight performance advantage.

Moreover, the longer the forecast horizon, the greater the advantage of AME over univariate and naive models. Thus, to conclude, the best overall performance over all forecast horizons, in terms of RMSE and AME, was obtained using the multivariate LSTM model.

3) Simpler Forecast Models and Larger Field Datasets

RNN models, although a good fit for forecasting tasks, are rather complex to build and optimize. Furthermore, the univariate and multivariate models presented in the previous section were built using single-lightpath data due to the limited availability of field data. The next question is whether a comparable or better performance can be achieved by using larger datasets of field data for training the RNN models, and by using simpler models such as multilayer perceptron (MLP) and linear regression for lightpath QoT forecasting.

This was the objective of the study presented in [29]. In this work, univariate lightpath QoT forecasters based on LSTM, MLP and linear regression were built using two KBs of field data collected in two production networks. The first dataset includes 5-month PM data from 140 channels deployed on 11 routes of a Tier-1 North American Service Provider (NASP). The second dataset includes 15-month PM data from 4000 channels deployed on 115 routes of the Microsoft (MS) backbone network in North America. The models were trained and executed multiple times by randomly selecting subsets of the NASP and MS datasets, using the estimated Q-factor as the target to be predicted and the RMSE as the performance metric.

Fig. 8 (a) and Fig. 8 (b) show the RMSE values versus the forecast horizon, as obtained by executing the models on two subsets of the original MS and NASP datasets. Interestingly, the best performance was achieved by the MLP and linear regression models, which seems to show the potential of simple models in lightpath QoT forecasting tasks. Second, we can see that the performance of the models varied from one run to another. This results from the presence of outliers in the time series. Prediction errors increased in the presence of outliers. This also shows that RNNs perform differently on different scenarios and datasets, which remains one of their challenges and limitations.

In summary, an interesting conclusion of this study is that simple QoT forecasters based on linear regression and MLP can outperform more complex RNN models. Another interesting conclusion is that the performance of the forecasters can be impacted significantly by the presence of outliers in the time series. These outliers (or sudden SNR drops) can be caused by cable manipulations during maintenance activities and last 15 minutes to 2 hours, typically. Outliers can be removed during the data preprocessing phase before training the models, but they cannot be avoided in normal network operations and therefore can impact the performance of the forecasters.

E. Transfer Learning

We explored supervised ML-based approach for QoT estimation of unestablished lightpaths and QoT forecast of deployed lightpaths in optical networks. To achieve accurate predictions of lightpath QoT, supervised ML models need large amounts of training data. However, this may be difficult to achieve due to their acquisition cost.

Transfer learning (TL) has been proposed to enable ML models to share relevant structures, so as to effectively reduce the size of the training dataset and the time required for prediction or classification tasks. A preliminary study of a domain adaptation approach that used the complete set of SNR data from one lightpath (in red in Fig. 9) to train a univariate

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JLT.2022.3160379, Journal of Lightwave Technology

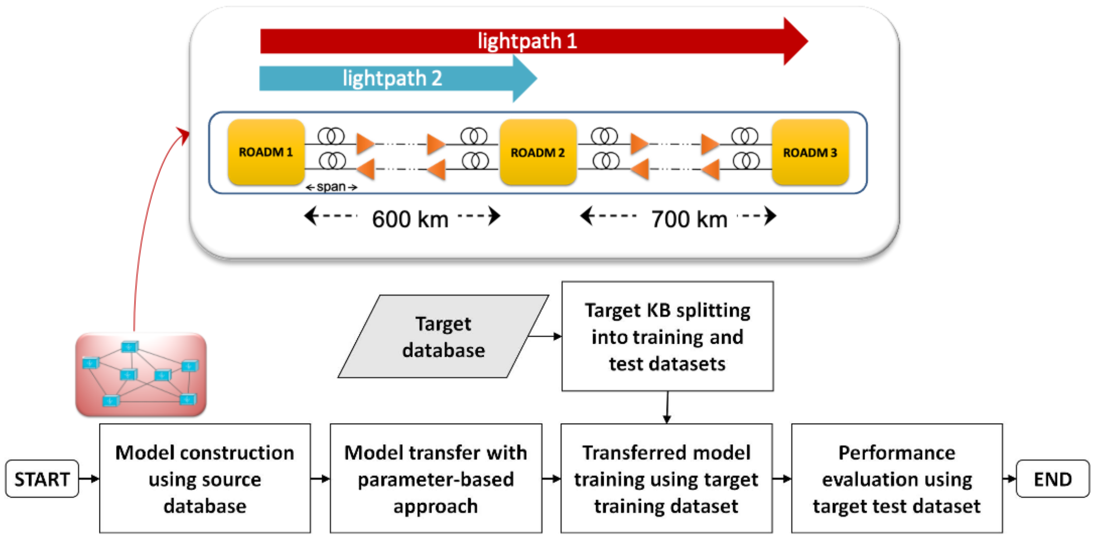


Fig. 9. Transfer learning process to forecast QoT of lightpath-2 using model pre-trained with lightpath-1 data. From [35].

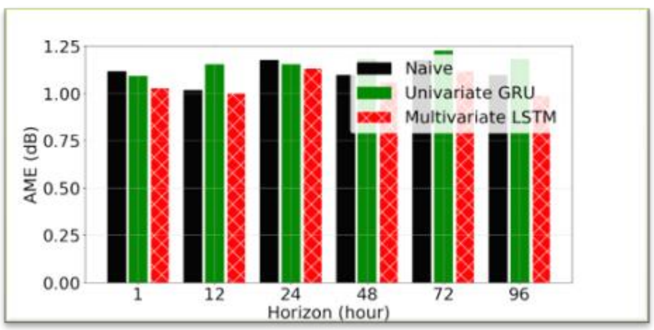


Fig. 10. Performance evaluation (AME vs. forecast horizon) using KB-2 (lightpath carried in the same optical fiber on a portion of same route). From [35].

GRU and a multivariate LSTM model, and to further transfer the structures to forecast the SNR of another lightpath carried in the same fiber but on a shorter portion of the route, was presented in [35]. An additional scenario, exploiting data from another lightpath carried on the same route but in the opposite direction, was also considered. Domain adaptation is a category of transfer learning in which the feature spaces between the source and target domains are similar, but the marginal probability distributions of the input data are different.

A parameter-based approach was adopted for the transfer structure. More specifically, we implemented a weight-sharing tactic between the source and the target domains. This strategy enables the knowledge transfer from the pre-trained models (using 24,445 instances of the source domain) to the target domain through a partial retraining with up to 1.9X less data. These results are certainly far from those found in [36], for the similar use case where only the path length is different between the source and target domains. This is probably because no optimization was performed in this preliminary study. However, Fig. 10 shows that transfer learning could be suitable for the short-term performance forecast of lightpaths carried in the same optical fiber of the same route.

VI. Concluding Remarks

In this tutorial, we explored how ML, and more specifically neural networks, can be used to solve optical communications and networking problems, with a special focus on two specific use cases: lightpath QoT estimation and forecast tasks.

In the first use case, we showed how a cognitive QoT tool based on supervised learning can be used for QoT estimation before lightpath provisioning. The objective here was to provide a fast QoT estimation method in real-time network operations. Among the considered ML methods, we established that an SVM estimator can perform as well as an NN estimator and even better with reduced feature sets. Note that these models were trained with synthetic data, given that field data from production networks is difficult to obtain, especially for poor QoT lightpaths. Therefore, it still remains to be demonstrated that cognitive QoT estimation tools can be integrated beneficially in the control system of a production network or used jointly with analytical tools based on the GN model in lightpath provisioning tasks. Such tools would be particularly interesting for network operators dealing with incomplete or outdated inventory system and fiber plant data.

In the second use case, we showed how machine learning, and more specifically RNNs, can be used in lightpath QoT forecast tasks. The objective was to detect degradations in lightpath performance over forecast horizons of up to 4 days for proactive maintenance purposes. Such models would be particularly useful to track dynamic lightpaths over time (for margin optimization purposes, for example). Univariate and multivariate models are based on LSTM and GRU and trained with single lightpath data. These studies reveal some very interesting findings. A better performance was observed for single-step univariate LSTM over multi-step encoder-decoder LSTM and GRU. Multivariate LSTM using 4 extra features exhibited a better performance than its univariate counterpart. Recent studies have shown that MLP and linear regression can

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JLT.2022.3160379, Journal of Lightwave Technology

outperform RNNs but not always. The models were shown to be very sensitive to outliers (i.e., sudden drops) in the time series, which can impact model performance and increase the prediction errors. Recent studies performed with field data from three lightpaths suggest that QoT forecasters trained with single-lightpath data could be used to test lightpaths carried in the same fiber by using the concept of transfer learning. These interesting findings still need to be validated with richer datasets (greater number of lightpaths) and in quality (number of features).

Field data remains a critical issue for demonstrating the full potential of ML in tasks such as lightpath QoT estimation and lightpath QoT forecast in real-world production networks. If successful, this research will pave the way to applications aiming at network automation, such as fast lightpath provisioning, margin optimization as well as proactive maintenance.

Acknowledgments

This work was supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada under grants RGPIN-2014-05898 and RGPIN-2019-03972.

References

[1] R. Gu, Z. Yang and Y. Ji. “Machine Learning for Intelligent Optical Networks: A Comprehensive Survey.” J. Netw. Comput. Appl. 157 (2020): 102576.

[2] I. T. Monroy, D. Zibar, N. G. Gonzalez and R. Borkowski, "Cognitive Heterogeneous Reconfigurable Optical Networks (CHRON): Enabling technologies and techniques," 2011 13th International Conference on Transparent Optical Networks, 2011, pp. 1-4, doi: 10.1109/ICTON.2011.5970833

[3] F. Musumeci et al., "An Overview on Application of Machine Learning Techniques in Optical Networks," in IEEE Communications Surveys & Tutorials, vol. 21, no. 2, pp. 1383-1408, Secondquarter 2019, doi: 10.1109/COMST.2018.2880039.

[4] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang and F. Forghieri, "The GN-Model of Fiber Non-Linear Propagation and its Applications," in Journal of Lightwave Technology, vol. 32, no. 4, pp. 694-721, Feb.15, 2014, doi: 10.1109/JLT.2013.2295208.

[5] P. Poggiolini, A. Carena, V. Curri, G. Bosco and F. Forghieri, "Analytical Modeling of Nonlinear Propagation in Uncompensated Optical Transmission Links," in IEEE Photonics Technology Letters, vol. 23, no. 11, pp. 742-744, June1, 2011, doi: 10.1109/LPT.2011.2131125.

[6] P. Poggiolini and Y. Jiang, "Recent Advances in the Modeling of the Impact of Nonlinear Fiber Propagation Effects on Uncompensated Coherent Transmission Systems," in Journal of Lightwave Technology, vol. 35, no. 3, pp. 458-480, 1 Feb.1, 2017, doi: 10.1109/JLT.2016.2613893.

[7] J. Lin, J. Li, and Y. Wang, “Study on the applicable distance scope of GN model to open cable”, in <i>Seventh Symposium on Novel Photoelectronic Detection Technology and Applications</i>, 2021, vol. 11763. doi:10.1117/12.2587430.

[8] L. Berg, "Demystifying Transceiver and Line Characterization Metrics: A Tutorial," 2019 Optical Fiber Communications Conference and Exhibition (OFC), 2019, pp. 1-14.

[9] 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," in Journal of Lightwave Technology, vol. 36, no. 15, pp. 3073-3082, 1 Aug.1, 2018, doi: 10.1109/JLT.2018.2818406.

[10] A. Ferrari et al., "GNPy: an open source planning tool for open optical networks," 2020 International Conference on Optical Network Design and Modeling (ONDM), 2020, pp. 1-6, doi: 10.23919/ONDM48393.2020.9133027.

[11] S. Aladin and C. Tremblay, "Cognitive Tool for Estimating the QoT of New Lightpaths," 2018 Optical Fiber Communications Conference and Exposition (OFC), 2018, pp. 1-3.

[12] S. Aladin, A. V. S. Tran, S. Allogba and C. Tremblay, "Quality of Transmission Estimation and Short-Term Performance Forecast of Lightpaths," in Journal of Lightwave Technology, vol. 38, no. 10, pp. 2807-2814, 15 May15, 2020, doi: 10.1109/JLT.2020.2975179.

[13] R. M. Morais and J. Pedro, "Machine learning models for estimating quality of transmission in DWDM networks," in Journal of Optical Communications and Networking, vol. 10, no. 10, pp. D84-D99, Oct. 2018, doi: 10.1364/JOCN.10.000D84.

[14] L. Li, G. Shen, X. Fu, S. Jie and S. K. Bose, "Which Features Most Impact: Prediction of ANN-Based Lightpath Quality of Transmission?," 2018 Asia Communications and Photonics Conference (ACP), 2018, pp. 1-3, doi: 10.1109/ACP.2018.8595979.

[15] J. Yu, W. Mo, Y. Huang, E. Ip and D. C. Kilper, "Model transfer of QoT prediction in optical networks based on artificial neural networks," in Journal of Optical Communications and Networking, vol. 11, no. 10, pp. C48-C57, October 2019, doi: 10.1364/JOCN.11.000C48.

[16] F. N. Khan, Q. Fan, C. Lu and A. P. T. Lau, "An Optical Communication's Perspective on Machine Learning and Its Applications," in Journal of Lightwave Technology, vol. 37, no. 2, pp. 493-516, 15 Jan.15, 2019, doi: 10.1109/JLT.2019.2897313.

[17] T. Panayiotou, G. Savva, B. Shariati, I. Tomkos and G. Ellinas, "Machine Learning for QoT Estimation of Unseen Optical Network States," 2019 Optical Fiber Communications Conference and Exhibition (OFC), 2019, pp. 1-3.

[18] Q. Fan et al., "Experimental Comparisons between Machine Learning and Analytical Models for QoT Estimations in WDM Systems," 2020 Optical Fiber Communications Conference and Exhibition (OFC), 2020, pp. 1-3.

[19] F. Usmani et al., "Evaluating Cross-feature Trained Machine Learning Models for Estimating QoT of Unestablished Lightpaths," 2021 International Conference on Electrical, Communication, and Computer Engineering (ICECCE), 2021, pp. 1-6, doi: 10.1109/ICECCE52056.2021.9514154.

[20] L. Zhang, X. Li, Y, Tang, J. Xin, S. Huang, “A survey on QoT prediction using machine learning in optical networks,” Optical Fiber Technology. Volume 68, 2022, 102804, ISSN 1068-5200

[21] R. Ayassi, A. Triki, M. Laye, N. Crespi, R. Minerva et al., “An overview on machine learning-based solutions to improve lightpath QoT estimation.” ICTON 2020: 22nd International Conference on Transparent Optical Networks, Jul 2020, Bari (online), Italy. pp.1-4, ⟨10.1109/ICTON51198.2020.9203755⟩. ⟨hal-03161225⟩

[22] Yvan Pointurier, "Machine learning techniques for quality of transmission estimation in optical networks," J. Opt. Commun. Netw. 13, B60-B71 (2021)

[23] X. Long and B. Sikdar, "A Real-Time Algorithm for Long Range Signal Strength Prediction in Wireless Networks," 2008 IEEE Wireless Communications and Networking Conference, 2008, pp. 1120-1125, doi: 10.1109/WCNC.2008.202.

[24] Abdel-Nasser, Mohamed, et al. "Link quality prediction in wireless community networks using deep recurrent neural networks." Alexandria Engineering Journal 59.5 (2020): 3531-3543.

[25] J. Thrane, B. Sliwa, C. Wietfeld and H. L. Christiansen, "Deep Learning-based Signal Strength Prediction Using Geographical Images and Expert Knowledge," GLOBECOM 2020 - 2020 IEEE Global Communications Conference, 2020, pp. 1-6, doi: 10.1109/GLOBECOM42002.2020.9322089.

[26] C.J. Lowrance, A. P. Lauf, “Link quality estimation in ad hoc and mesh networks: a survey and future directions”, Wireless Pers. Commun. 96 (1) (2017) 475–508.

[27] A. Michel, “Monitoring of performance parameters in a metro ring using a 100G coherent transponder,” M.Eng. Final Rep., Dept. Electr. Eng., École de technologie supérieure, Montreal, 2016.

[28] S. Aladin, S. Allogba, A. V. S. Tran and C. Tremblay, "Recurrent Neural Networks for Short-Term Forecast of Lightpath Performance," 2020 Optical Fiber Communications Conference and Exhibition (OFC), 2020, pp. 1-3.

[29] H. Chouman et al., "Forecasting Lightpath QoT with Deep Neural Networks," Optical Fiber Communication Conf, (OFC) 2021, Paper Th4J.5.

[30] Preeti, R. Bala and R. P. Singh, "Financial and Non-Stationary Time Series Forecasting using LSTM Recurrent Neural Network for Short and Long Horizon," 2019 10th International Conference on Computing,

Communication and Networking Technologies (ICCCNT), 2019, pp. 1-7, doi: 10.1109/ICCCNT45670.2019.8944624.

[31] Francois Chollet. Deep learning with python. 2017. Manning Publications.

[32] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. 2016. Deep Learning. The MIT Press.

[33] Z. Wang, X. Su and Z. Ding, "Long-Term Traffic Prediction Based on LSTM Encoder-Decoder Architecture," in IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 10, pp. 6561-6571, Oct. 2021, doi: 10.1109/TITS.2020.2995546.

[34] A. Mezni, D. W. Charlton, C. Tremblay and C. Desrosiers, "Deep Learning for Multi-Step Performance Prediction in Operational Optical Networks," 2020 Conference on Lasers and Electro-Optics (CLEO), 2020, pp. 1-2.

[35] S. Allogba, S. Aladin and C. Tremblay, "Multivariate Machine Learning Models for Short-Term Forecast of Lightpath Performance," in J. Lightwave Technol., doi: 10.1109/JLT.2021.3110513, 2021.

[36] C. -Y. Liu, X. Chen, R. Proietti and S. J. B. Yoo, "Performance studies of evolutionary transfer learning for end-to-end QoT estimation in multi-domain optical networks [Invited]," in Journal of Optical Communications and Networking, vol. 13, no. 4, pp. B1-B11, April 2021, doi: 10.1364/JOCN.409817.

Stéphanie Allogba received her B.Eng. degree in telecommunications from ESME Sudria, France, in 2011 and her M.A.Sc. degree in telecommunications networks engineering from École de technologie supérieure, Montréal, Canada, in 2015. Her Master's thesis focused on control and management algorithms for femtocell wireless networks. In 2021, she received her Ph.D. degree at École de technologie supérieure, with the Network Technology Lab. Her Ph.D. activities and research interests include the application of machine learning algorithms for lightpath classification and performance prediction in optical networks. She is currently working as a research assistant with the Network Technology Lab at École de technologie supérieure.

Sandra Aladin received her B.Eng. degree in electronics from Université d'État d'Haiti, Port au Prince, Haiti, in 2009 and her M.A.Sc. degree in telecommunications network engineering from École de technologie supérieure, Montréal, QC, Canada, in 2018. Her Master's thesis focused on machine learning-based quality of transmission estimation in optical networks. From 2018 to 2021, she has been working as a Research Assistant on machine learning for optical networking with the Network Technology Lab, Department of Electrical Engineering, École de technologie supérieure, Montréal. She is currently working towards a Ph.D. degree at École de technologie supérieure, with the Network Technology Lab. Her research interests include machine learning for applications in optical communications and networking.

Christine Tremblay received her B.Sc. degree in Engineering Physics from Université Laval, Quebec City, Canada, in 1984, M.Sc. degree (Energy) from INRS-Énergie, Varennes, Canada, in 1985 and Ph.D. degree (Optoelectronics) from École Polytechnique de Montréal, Canada, in 1992. She is a Full Professor with the Department of Electrical Engineering and Associate Director for the Ph.D. Program at École de technologie supérieure. She is the Founding Researcher and Head of the Network Technology Lab. Before joining ÉTS, she was a Research Scientist with the National Optics Institute (INO) where she conducted research on integrated optical devices for communication and sensing applications. She has held senior R&D and technology management positions for several organizations. As Engineering Manager at EXFO and Director of Engineering at Roctest, she was responsible for the development of fiber-optic test equipment. She also served as Product Manager at Nortel for DWDM systems. Her research interests include machine learning for optical networking applications, optical performance monitoring and filterless optical networking. She has been co-instructor for SC314 and SC210 hands-on courses of the Optical Society of America on optical fiber and polarization measurements.

Dr. Tremblay is a member of the Optical Society of America (OSA), the IEEE Photonics Society, as well as STARaCom and COPL Strategic Clusters of FRQNT. She is currently serving as Technical Program Committee (TPC) Member for OFC N3 Subcommittee "Architecture and software-defined control for metro and core networks".