Experimental Investigation of Physics- and ML-based QoT Estimation for WDM Systems

Md Saifuddin Faruk (1,*), Mariane Mansour(1), Charles Laperle(2), Maurice O'Sullivan (2), Seb J. Savory(1)

(1) Electrical Engineering Division, Department of Engineering, University of Cambridge, Cambridge CB3 0FA, UK, *msf35@cam.ac.uk
(2) Ciena Corporation, Ottawa, Ontario K2K 0L1, Canada

Abstract *With a seven-channel WDM transmission over 1000 km, we experimentally study the data-driven physics- and machine learning (ML)-based SNR estimation techniques. While the ML-based approach provides good estimation accuracy, the physics-based method performs close to it with more explainability and less training data requirements.*

Introduction

In a WDM system, Quality of Transmission (QoT) estimation is an important and challenging task as the performance of a channel depends on the power of other neighbouring channels. Recently, accurate QoT estimation, where signal-to-noise ratio (SNR) is estimated as a QoT metric, is under intense investigation to enable a low-margin network.

Both the analytical approach as well as machine learning (ML) based approach has been investigated for QoT estimation. For analytical approaches, usually, the Gaussian noise (GN) model is used [1] whereas various ML approaches such as neural network (NN), support vector machine (SVM), K-nearest neighbour (KNN), random forest (RF) and so on [2-4] has been demonstrated. The analytical model suffers from parameter uncertainty in a real system whereas data-driven ML techniques rely on large datasets and provides less explainability.

In this paper, two key approaches are considered for SNR estimation. The first one is the physics-based model, which can be referred to as the 'white box' model, where the unknown parameters are extracted from experimental measurement data. The second approach is the 'black box' ML-based approach which completely relies on the data without reference to the underpinning physics [5]. For the ML-based approach, an NN as well as a Gaussian process regression (GPR)-based techniques are used in this work.

To validate the estimation methods, a 7-channel WDM experiment is conducted where 200 Gb/s DP-16-QAM signal is transmitted over 1000 km of standard SMF. A comparable estimation performance in terms of root mean square error (RMSE) and maximum estimation error is obtained for both physics- and ML-based methods with the later performs better to some extend. The maximum error across all the channel for physics-based method is 0.22 dB while that for ML-based approach is 0.17 dB. However, about four-fold less measurement data is required for the physics-based method for the same RMSE.

SNR estimation methods

Herein we describe the SNR estimation technique used in this work.

Physics-based method:

For the physics-based model the SNR of *i*-th channel is given as [6] -

$$SNR_i = \left(\frac{1}{SNR_{0,i}} + \frac{P_{ASE}}{P_i} + \frac{P_i \sum_{j=1}^{N_{ch}} \eta_{ij} P_j^2}{P_i} \right)^{-1}, \quad (1)$$

where, SNR_0 is the back-to-back SNR, P_{ASE} is the ASE noise power from the inline amplifiers, P_i is the *i*-th channel launch power into the fibre span, N_{ch} is the number of WDM channels and η_{ij} is the nonlinear interference coefficient between channels *i* and *j* accounting for both SPM and XPM. Therefore, by measuring the launch power of each channel and corresponding SNR, it is possible to solve for the unknown variables SNR_0, P_{ASE} and η_{ij}. Once these variables are estimated, SNR can be calculated for any other launch power distribution.

Neural network-based method:

NN is a powerful ML tool to model the relationship between channel power and SNR. For the NN-based model, a feed-forward network with one hidden layer is found to be sufficient for this purpose. The launch powers of all WDM channels are used as the input features of the NN while the corresponding SNRs are set as outputs. The known measured data is used for the training of NN while in the testing phase, unknown SNRs are estimated for a particular launch power profile.

Gaussian process regression-based method:

As an alternative to NN, GPR is investigated

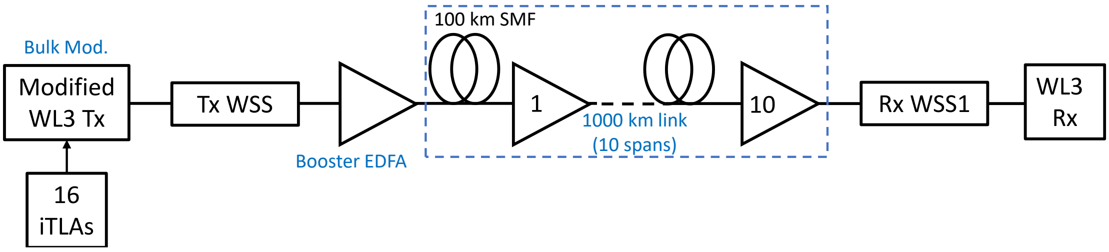

Fig. 1: Experimental setup used to verify the SNR estimation techniques.

which is a probabilistic supervised ML method. GPs use Bayesian statistics to find the most likely function describing the relationship between a set of inputs and a set of outputs, given the data and a set of priors [7]. Unlike NN, GPR is a non-parametric ML method and thus instead of assuming a given parametric form and finding a set of parameters that describe the mapping, the space of functions is searched directly in a probabilistic way. Similar to NN, for QoT estimation, the input and the output of GPR are chosen as the launch power of all WDM channels and corresponding SNRs, respectively.

Experimental setup

To verify the performance of SNR estimation, we conducted an experiment as shown in Fig.1 with 7 WDM channels transmission over a 1000 km link on a 50 GHz grid. Each channel had a net data rate of 200 Gb/s, using a 34.5 GBd PM-16QAM signal. The 16 iTlAs were bulk modulated by a modified Ciena WaveLogic 3 line card to provide sufficient power to the modulator and later a transmitter WSS was used to select 7 channels centered around 1550 nm. A booster EDFA was used before launching the signal into the first span of fibre.

The signal was transmitted over 10 spans of link each consisting of 100 km standard fibre. A Polatis 32×32 fibre switch was used to connect the individual 100 km spans. This switch was also used to control the launch power into the spans, which in this experiment was uniform in each span. In each span, an EDFA with a fixed gain of 26 dB was used, along with a variable optical attenuator (VOA) to compensate for the extra gain.

At the receiver, a WSS was used to select the channel of interest (COI) and Ciena WaveLogic 3 line card receiver was used to demodulate the signal. The SNR of the received signal was calculated from the receiver pre-FEC BER using the relation $BER = \frac{3}{8}\mathrm{erfc}\sqrt{\frac{SNR}{10}}$.

The launch power of each channel was randomly varied between -3 dBm to 4 dBm and corresponding SNRs were measured to generate a dataset. We found that the optimum launch power was ~0.75 dBm and the chosen data points cover the linear, quasi-linear and nonlinear regimes. We captured a total of 400 datasets each having seven channel powers and corresponding SNRs. The number of training data was varied which is taken from the first 350 datasets whereas the rest 50 datasets were used for testing purposes.

For physics-based estimation, the model of Eq. (1) was used. A Moore-Penrose pseudoinverse matrix-based method was used to solve the equation using the measured data [8]. For the NN, one hidden layer with 64 nodes was used. Also, we used the *elu* activation function, Adam optimization algorithm and MSE loss function. For the GPR a kernel function is used to model the covariance of the data and in this work, we used the ard squared exponential kernel function.

Results and discussion

First, we check the number of training data requirements for each method. For this purpose, we measure estimation error (difference between measured and estimated SNR) from the test data set and then calculate the RMSE. Fig.2 shows the RMSE against the number of training data for all three methods for the central channel. It is found that the physics-based method converged very quickly and reached a steady state with a small number of datasets. On the other hand, the ML-based approach requires relatively larger training data convergence and it continues to improve marginally with more data. As an example, for the same RMSE, the required number of the training dataset is 50 for the physics-based method whereas it requires around 130 datasets to get the same performance for the ML-based technique. For a larger dataset of >250 ML-based approach performs better than the physics-based method. An almost similar requirement is found for NN- and GPR- based methods.

Next, we estimate the RMSE for all seven channels for a training data of 50 for the physics-based method and 250 for ML-based methods. As shown in Fig.3, comparable performance is found for both approaches with the ML-based approach performing marginally better in general. The maximum RMSE over all the channels is 0.065 dB for the physics-based method and 0.055 for the ML-based method.

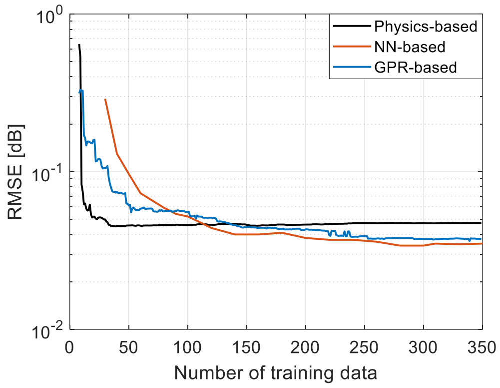

Fig. 2: RMSE versus number of training data for all three methods for the middle channel.

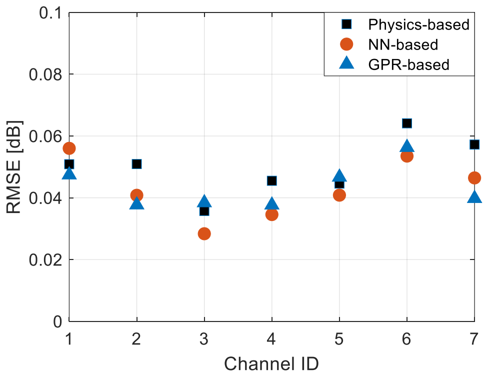

Fig. 3: RMSE of all seven channels for the three SNR estimation methods.

The plot of maximum absolute estimation error for the three methods is also depicted in Fig.4. It is found that, like RMSE performance, the ML-based method performs better than the physics-based method. The worst-case maximum absolute error for the physics-based method is 0.21 dB (ch-6) whereas that for NN- and GPR-based method is 0.17 dB and 0.16 dB, respectively (ch-1).

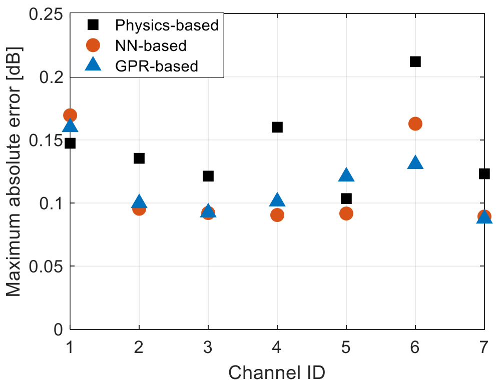

Fig. 4: Maximum absolute error of all the channels for the three methods. The number of training data is 50 for physics and 250 for ML-based methods.

Unlike the ML-based method, the physics-based approach is more explainable as it allows the estimation of unknown parameters as in Eq. (1). As an example, we illustrated the back-to-back SNR estimation results as in Fig. 5. For the back-to-back measurement, we exclude the 1000 km link and connect the booster amplifier directly to the receiver WSS of Fig.1. On the other hand, for the physics-based method, the estimation value, $\widehat{SNR_0}$ is calculated from the 1000 km transmission data as

$$\widehat{SNR_0} = \underset{\mathrm{SNR}_0}{\mathrm{argmin}}\, E\left\{\left|SNR_{mes} - SNR_{\mathrm{phy}}\right|^2\right\},$$

where, SNR_{mes} is the measured SNR and SNR_{phy} is the estimated SNR by the physics-based method. Provided that back-to-back SNR is transmission distance independent, we found a good agreement between the measured value and physics-based estimated value with a maximum estimation error of below 0.2 dB.

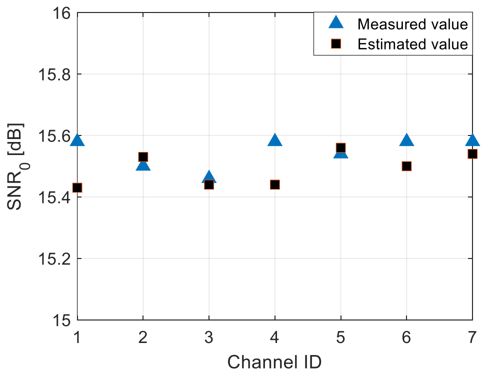

Fig. 5: Comparison of measured back-to-back SNR with that from the physics-based estimation method.

Conclusion

An SNR estimation technique is experimentally demonstrated with 7-channel WDM systems from its measured launch power per channel using physics-based and two ML-based approaches, namely NN and GPR. While the physics-based method is more interpretable and requires fewer measurement data for training, the ML-based approach provides marginally better accuracy.

Acknowledgments

MSF, MM and SJS thank Ciena for the donation of equipment and support. MM also thanks Ciena for funding her PhD studentship. MSF and SJS thank UK EPSRC for funding via the Programme Grant TRANSNET (EP/R035342/1). This research was performed under the auspices of a Ciena University collaborative research grant.

Data underlying the results presented here are at https://doi.org/10.17863/CAM.96555.

References

[1] J. Lu *et al*., "Performance comparisons between machine learning and analytical models for quality of transmission estimation in wavelength-division-multiplexed systems [Invited]," *J Opt. Commun. Net.*, vol. 13, no. 4, pp. B35-B44, 2021, doi: 10.1364/JOCN.410876.

[2] Y. Pointurier, "Machine learning techniques for quality of transmission estimation in optical networks," *J. Opt. Commun. Netw.*, vol. 13, pp. B60–B71, 2021.

[3] S. Allogba, S. Aladin and C. Tremblay, "Machine-learning-based lightpath QoT estimation and forecasting," *J. Lightw. Technol.,* vol. 40, no. 10, pp. 3115-3127, 15 May15, 2022, doi: 10.1109/JLT.2022.3160379.

[4] J. W. Nevin *et al*. "Machine learning for optical fiber communication systems: An introduction and overview," *APL Photonics*, vol.6, no. 12, art. 121101, 2021. doi: 10.1063/5.0070838.

[5] Seb J. Savory, "Digital twins and measurement informed physical layer models [Invited Tutorial]," in Proc. *Opt. fibre commun. Conf. (OFC)*, paper M2F.1, 2023.

[6] H. -M. Chin et al., "Probabilistic Design of Optical Transmission Systems," *J. Lightw. Technol.*, vol. 35, no. 4, pp. 931-940, 15 Feb.15, 2017, doi: 10.1109/JLT.2016.2640451.

[7] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. MIT press, Cambridge, MA, 2006.

[8] G. H. Golub and C. F. Van Loan, *Matrix Computations*, The Johns Hopkins Univ. Press, Baltimore, MD, USA, 2012.