Auxiliary Neural Network Assisted Machine Learning EDFA Gain Model

Jiachuan Lin, Xiang Lin, Zhiping Jiang
Ottawa Research Centre, Huawei Technologies Canada Co. Ltd., 303 Terry Fox Drive, K2K3J1, Kanata, ON, Canada
zhiping.jiang@huawei.com

Abstract: An enhanced EDFA model employing auxiliary neural networks is proposed. Adaptive to different devices, the model reduces the root mean square error from 0.04 to 0.02 dB with significantly less amount of training data. © 2022 The Author(s)

1. Introduction

Optical networks are evolving towards a more dynamic and flexible paradigm [1,2]. Fast add/drop of the service wavelengths in the network requires estimation on the performance of unestablished lightpath (LP). An accurate gain modelling for erbium doped fiber amplifiers (EDFAs) helps a better prediction on the performance of to-be-established LP, as well as better evaluation of the impact on the existing/remaining LPs during channel add/drop operation. Accuracy is one of the important metrics for the gain modelling: long-haul systems consist of tens of EDFAs, and a small prediction error for single EDFA might result in huge signal power deviation from the actual power. In addition to strict requirement for accuracy, a good EDFA gain model is expected to work under various channel loadings. EDFA gain profile changes drastically under different channel loadings due to the spectrum hole burning (SHB) effect. SHB is very complicated in nature and there lacks a tractable analytical tool so far. Developing an accurate EDFA gain model with SHB effect considered and suitable for a variety of channel loading conditions is desired.

In recent years, machine learning (ML) techniques have been employed to model the EDFA behaviors [3-7]. A pioneering work on ML based EDFA gain modelling was reported in [3], and excellent accuracy with a root mean square error (RMSE) smaller than 0.02 dB is achieved. It is worth mentioning that in that work, the channelized differential gain (i.e. gain profile change from full-fill channel loading) rather than gain profile itself, is modelled. Directly applying the gain model obtained from one EDFA to the others could experience a decrease in accuracy. In [5], a generalizable model is proposed, and its performance is evaluated across different EDFAs from the same make. The generalization leads to a reduced mean square error from about 0.06 to 0.02 dB^2 (equivalent RMSE from 0.25 to 0.14 dB), at the price of 3 times larger training dataset. In [6], a hybrid approach combining ML and physical model is introduced to model the EDFA gain for non-flat input spectrum. The model works well at different gains with a RMSE of 0.05 dB. However, as pointed out in the paper, the SHB effect is ignored, therefore the model is more suitable for submarine application with channel loading close to full-fill. The modelling reported in [3] yields the highest accuracy reported so far, and the model is suitable for terrestrial scenarios where the channel loading varies widely. By further incorporating the in-band tilt parameters, the same methodology has been proved to be capable for EDFA in-band gain ripple prediction with a good prediction accuracy [4]. However, its generalization performance to different gains of the same device, or different devices has yet to be investigated.

In this paper, we first show significant difference in the amplifier's gain behavior between different devices or different gains of the same device. Prediction accuracy reduces when apply the model trained on one device to the others. We then propose to use an auxiliary neural network (Aux-NN) to account for the difference. Results show a significantly improved model accuracy with greatly reduced retraining efforts.

2. Evaluation of model generalization

The ML model used in [3] is based on the multi-layer perceptron (MLP) structure, with 1 input layer, 2 hidden layers and 1 output layer. The input layer takes the channelized input powers of a given channel loading. The output layer gives the model results of channelized gain difference. The gain difference, $\Delta G_k(\lambda_i)$, is defined as

$$\Delta G_k(\lambda_i) = G^k_{\text{partial}}(\lambda_i) - G_{\text{full}}(\lambda_i), \tag{1}$$

where $G^k_{\text{partial}}(\lambda_i)$ is the gain of the i-th wavelength channel when the k-th partial loading pattern is configured, $G_{\text{full}}(\lambda_i)$ is the gain of channel λ_i when a full-fill loading pattern is configured.

The experiment setup we used for data collection is shown in Fig. 1(a). A flat C-band amplified spontaneous emission (ASE) source filtered by a wavelength selective switch (WSS) is used for channel loading generation. The loading spectrum is configured based on 50 GHz grid, where the odd channels are used to load signal tones and probe tones, and the even channels are blocked for noise floor measurement. The probe tone power is much lower than the signal one, ensuring a negligible impact to signal gain performance. As proved in [3], the signal + probe configuration attains accurate full-band gain measurements. The generated loading spectrum is then given as the input of the EDFA under test (EUT). Three EUTs of the same model are tested. An optical spectrum analyser (OSA) is used to measure the input and output spectrums of the EUTs. Then the differential gain spectrum are calculated.

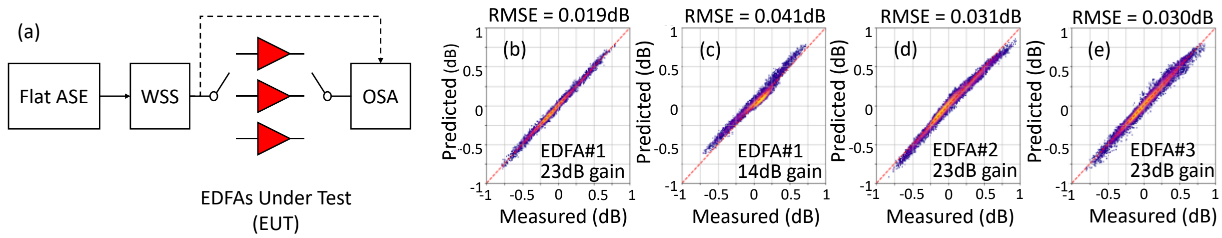

Fig. 1. (a) Experiment setup for data collection. (b)-(e) Predicted vs. measured gain difference results: (b) the baseline model of EDFA#1 with 23dB gain (RMSE = 0.019dB); and adapt the baseline model to different conditions of (c) EDFA#1 at 14dB gain (RMSE = 0.041dB), (d) EDFA#2 with 23dB gain (RMSE = 0.031dB), and (e) EDFA#3 with 23dB gain (RMSE = 0.030dB)

We first collected 4500 measurements from EDFA#1 at 23 dB gain and use 3500 of them to train the ML model, as the baseline model. The rest 1000 sets are used to verify the baseline performance. Fig. 1(b) shows the baseline model result of predicted $\Delta G_k(\lambda_i)$ vs. measured $\Delta G_k(\lambda_i)$. In this result, the training and prediction are under the same condition. The warmer colour (yellow) represents higher probability while the colder ones (blue) are low probability. In the baseline case, the scattered points fall on the straight diagonal line with low noise, and the model RMSE is 0.019 dB.

For different conditions, our data collection covers two scenarios: the same EDFA with different gains and different EDFAs with same gain. In total we tested 3 different conditions, including 1) EDFA #1 of 14, 18, 21 dB gains, 2) EDFA #2 of 23 dB gain and 3) EDFA #3 of 23 dB gain. Note that, for condition-1, only 14 dB data is discussed in the rest of the paper, as the most severe degradation happens under this gain setting. For each condition, we collected 1000 dataset. Fig. 1(c)-(e) show the corresponding results of applying the baseline model to EDFA#1 14 dB gain and EDFA#2/EDFA#3 with 23 dB gain. The predicted vs. measured results become noisy (larger errors) and show loading dependent bias or rotation. The model RMSE degrades to 0.03~0.04 dB. To be noted, the max error can be greater than 0.2 dB, which may result in 2.8 dB~8 dB power prediction error for a long link scenario with 40 EDFAs deployed (depending on the error is more biased towards a same direction or random directions).

3. Aux-NN assisted EDFA gain model and performance verification

To adapt the baseline model to new conditions without losing performance, the straightforward way is retrain the model using additional large amount data collections covering all new conditions, or apply transfer learning. However, a large number of data measurement is time consuming, and should be avoided if possible.

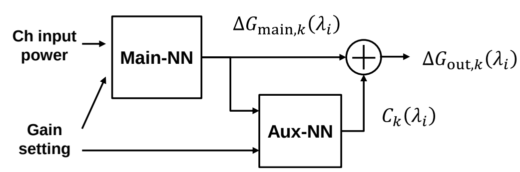

Fig. 2. Aux-NN assisted ML EDFA gain model.

To reduce the required retaining efforts, we propose to use an auxiliary NN to more efficiently adapt the baseline mode to different conditions. The proposed model is shown in Fig. 2, which contains a main-NN and an Aux-NN. The main NN is the same as the baseline one. The Aux-NN also uses MLP structure with two hidden layers, taking the main-NN outputs and pre-set gain as its inputs, outputting a channelized correction factor $C_k(\lambda_i)$ for the given k-th channel loading. The final model's outputs $\Delta G_{out,k}(\lambda_i)$ is

$$\Delta G_{out,k}(\lambda_i) = \Delta G_{main,k}(\lambda_i) + C_k(\lambda_i). \quad (2)$$

The main-NN adopts the NN weights of the baseline model, which are trained by 3500 measurements from EDFA#1 of 23dB gain. We individually train the Aux-NN for each condition, using 600 dataset for training and the rest 400 for verification. Fig. 3(a)-(c) show the predicted vs. measured results of the final model outputs, corresponding to the corrected versions of Fig. 1(c)-(e). After Aux-NN correction, the model results are much improved: the overall model prediction error has been reduced and the loading dependent bias has been removed.

The model error probability distributions are illustrated in Fig. 3(d)-(f), where the light blue colored bars are the error probability of the main-NN outputs, and the light red bars are the Aux-NN corrected ones. The errors are counted based on the bin size (step size of x-axis) of 0.005dB. As can be seen, without applying Aux-NN, the error probability distributions are more spread and in some cases they are no longer symmetric. The main-NN model RMSEs are 0.041 dB (EDFA #1 with 14 dB gain), 0.031 dB (EDFA #2 with 23 dB) and 0.030 dB (EDFA #3 with 23 dB gain). After applying the Aux-NN, the error probability distributions are more confined to the center, and the final model RMSEs are improved to 0.015 dB, 0.018 dB, and 0.021 dB respectively.

Compared to retraining using full data for each new condition, the required training data of Aux-NN approach is much reduced, from several thousands to hundreds. One of the reason could be that, the condition deviations from the baseline are relative small, as those deviations usually come from the slight different performance of used physical components such as Erbium doped fibers (EDFs), gain flattening filters, etc. Thus, the Aux-NN can be trained to predict the corrections using reduced dataset.

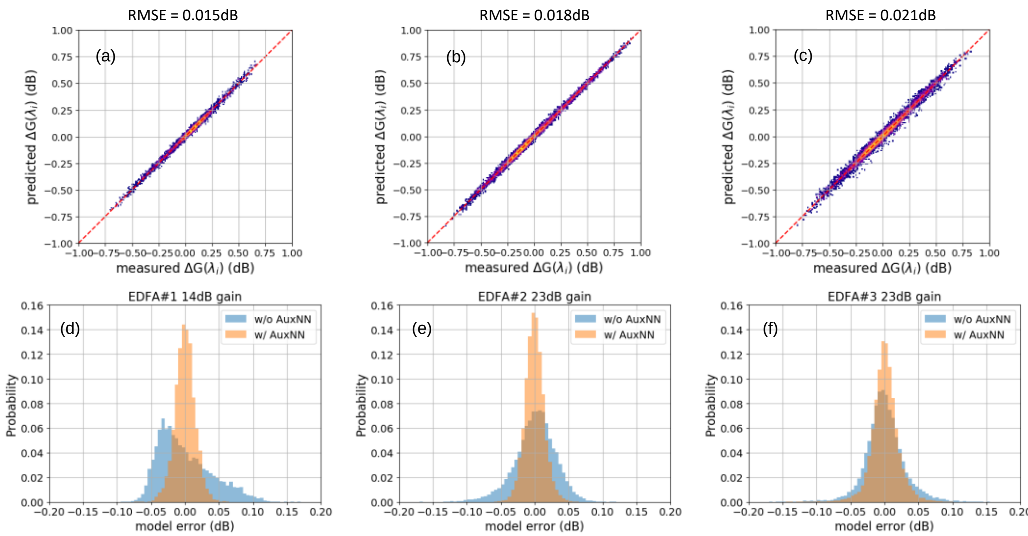

Fig. 3. Predicted vs. measured gain difference results after applying Aux-NN corrections for (a) EDFA#1 with 14dB gain, (b) EDFA#2 with 23dB gain and (c) EDFA#3 with 23dB gain; (d-f) are the error probability distributions corresponding to three conditions with and without applying Aux-NN corrections.

4. Conclusion

We propose an improved ML EDFA gain model, whose performance when adapting to new conditions is enhanced by an Aux-NN. In the three tested conditions, the model RMSEs are reduced from 0.04dB to around 0.02dB. The required retraining data for each condition is reduced from few thousands to few hundreds. This reduced training efforts could potentially benefit the whole network modelling with a much improved efficiency. Moreover, the Aux-NN approach could also be potentially applied to other device ML modeling to improve the training efficiency when adapting to new conditions.

5. References

[1] M. Freire-Hermelo, D. Sengupta, A. Lavignotte, C. Tremblay, and C. Lepers, “Reinforcement learning for compensating power excursions in amplified WDM systems”, J. Lightwave Technol., **39**(21), 6805-6813 (2021).

[2] Z. Jiang, S. Wang, “Fast Optical Performance Monitoring for Diagnosing Transient Behavior during Channel Add/Drop”, in Proceedings IEEE European Conference on Optical Communication (ECOC), Bordeaux, France, 2021, pp.1-4.

[3] Y. You, Z. Jiang, and C. Janz, “Machine learning-based EDFA gain model”, in Proceedings IEEE European Conference on Optical Communication (ECOC), Rome, Italy, 2018, pp. 1-3.

[4] Z. Jiang, J. Lin, and H. Hu, “Machine Learning Based EDFA Channel In-band Gain Ripple Modeling”, in Proceedings Optical Fiber Communications Conference and Exhibition (OFC), San Diego, USA, 2022, pp. W4I.2.

[5] F. Da Ros, U. C. De Moura, and M. P. Yankov, “Machine learning-based EDFA gain model generalizable to multiple physical devices”, in Proceedings IEEE European Conference on Optical Communications (ECOC), Brussels, Belgium, 2020, pp. 1-4.

[6] A. C. Meseguer, J. C. Antona, A. Bononi, J. Cho, S. Grubb, P. Pecci, O. Coutois and V. Letellier, “Highly accurate measurement-based gain model for constant-pump EDFA for non-flat WDM inputs”, in Proceedings Optical Fiber Communications Conference and Exhibition (OFC), San Francisco, USA, 2021, pp. M5C.4.

[7] A. Mahajan, K. Christodoulopoulos, R. Martinez, S. Spadaro, and R. Munoz, “Modeling EDFA gain ripple and filter penalties with machine learning for accurate QoT estimation”, J. Lightwave Technol., **38**(9), 2616–2629 (2020)