Recalibration Learning: Enabling Universal Transfer of ML Model of Gain and NF for Remote Optically Pumped Amplifiers

Arthur Minakhmetov, Benjamin Prieur, Maël Le Monnier, Delphine Rouvillain and Bruno Lavigne
ASN/Nokia, 1 Avenue du Canada, 91940 Les Ulis, France
Author e-mail address: arthur.minakhmetov@asn.com

Abstract: We demonstrate a novel, physical assumptions-based method – recalibration learning, that transfers Gain and Noise Figure ML models across remote optically pumped amplifiers. Spectral measurements over just two configurations on a target device ensure reliable transfer. © 2024 Nokia

1. Introduction

Optical amplification is a vital part of contemporary optical networks. It comes in many types and forms, with the Erbium-Doped Fiber Amplifier (EDFA) being the most deployed type of amplifier. Its working principle is based on stimulated emission, which requires a light source, a pump, that transfers its energy to the optical signal. The pump is often located in the immediate proximity to EDFA along with its electric power supply, however, it is not always the case, namely in unrepeatered lines. In that case one can use Remotely Pumped Optical Amplifier (ROPA) – an EDFA pumped by remote optical pump delivered over the same fiber as signal, contributing also to Raman amplification [1].

The exhaustive and precise modeling of EDFA is required to reduce transmission margins and to correctly design optical networks [2]. To address these needs, analytical and Machine Learning (ML) based models of EDFA are proposed in the literature [3,4]. ML models are based on exhaustive EDFA characterization with further training on obtained data. The best model to have is a model valid across devices, which is a challenging objective. A model valid for 3 EDFAs was demonstrated [5], but implies a common characterization over all devices, which is time-consuming, and it is unclear how many EDFAs one need characterize to obtain a general model. Furthermore, it comes with general model's precision degradation for a particular device when compared to device-specific model. There are proposals to overcome these problems: a) transfer learning with retraining pretrained ML models using small amount of measurements over new devices, with one ML model per EDFA [6], b) use of auxiliary Artificial Neural Networks (ANN) that are EDFA-specific and paired with a main ANN, trained on a reference EDFA [7], c) an efficient ML model enabling its transfer to a target device via 1 spectral measurement on it, also with an ML model per EDFA [8].

In this study we propose a method of ROPA ML model transfer across other ROPAs: recalibration learning. This method aims at finding a device-specific set of parameters, that helps to adapt the reference ML model to another device, without model retraining. We exploit physics-based assumptions: we suppose that differences between ROPAs originate from efficiencies to receive the input signal, transfer the pump power to the signal and deliver the output signal. Thus, the performance of ROPA may be modelled and adapted to another via 3 recalibration factors/corrections: on the total input power, total output power and power of the optical pump. In particular, we apply the recalibration factors to ROPA ML model's inputs and outputs, so to replicate the differences between ROPAs in reality. We summarize the functioning principle of the model transfer in Fig.1 and develop it in further sections.

We show that characterizing a target device at a flat input spectrum over two pump powers, comprising just 0.21% of original dataset, is enough to determine recalibration factors and transfer the reference ML model to a target device.

2. Characterization of ROPAs and design of reference Gain and NF ML models

We assemble a setup depicted on Fig.2 a) to characterize different ROPAs of the same make: ROPA 1, 2 and 3. ROPA 1 plays the role of a reference device on which we train ML models of wavelength depended Gain and NF, while ROPA 2 and 3 are target devices to which we aim to transfer trained ML models. We control the optical signal sent to ROPA via Wavelength Selective Switch (WSS), which carves from Amplified Spontaneous Emission (ASE) source a spectrum of desired shape, sampled over 84 channels of 50GHz width in C-band. The signal is then attenuated by Variable Optical Attenuator (VOA) to emulate multiple total input powers P_{in}^{tot}. A ROPA is pumped by an optical

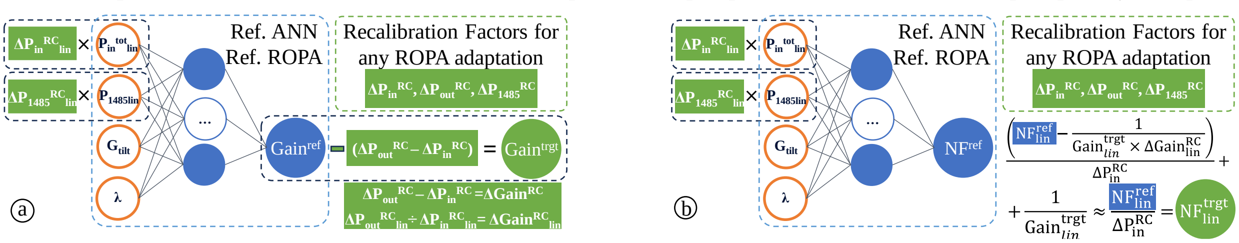

Fig. 1. Functional scheme of adaptation of a reference ML model of ROPA to any other ROPA (i.e., recalibration model) for: a) Gain, b) NF

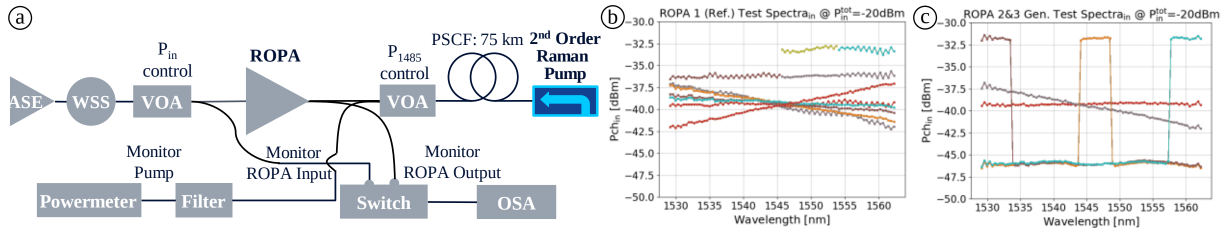

Fig. 2 a) Setup for a ROPA characterization, b) spectral loadings for reference ROPA test, c) generic spectral loadings for target ROPA 2&3 test.

pump at λ=1485nm, brought over 75km of fiber. We emulate multiple powers of the pump P_{1485} entering ROPA via another VOA. A ROPA is characterized by measuring Gain and NF under different forms of input spectrum, under different values of P_{in}^{tot} and P_{1485}. To characterize ROPA 1 we set P_{in}^{tot} and P_{1485} as next: $P_{in}^{tot}\in[-30,-5]$ dBm at 5 dB step and $P_{1485}\in[5,12]$ dBm at 1 dB step, totaling 48 configurations. We test these configurations under next spectral loadings: a) for training dataset we have 15 special spectral profiles, b) for cross-validation dataset we have 5 spectral profiles with -20, -15, -10, 10, 15 dB of tilt (over 84 channels), c) for test dataset we have 5 flat profiles (Fig.2b): one full, 1st and 2nd halves, 3rd and 4th quarters of C-band, plus 5 spectral profiles with -5, -4, -2, -1, 5 dB of tilt. Considering training and cross-validation dataset, we conclude that we need 48×(15+5)=960 configurations for ML model training.

The measurements on ROPA 2 can be summarized as next: we set P_{in}^{tot} and P_{1485} to cover $P_{in}^{tot}\in[-30,-5]$ dBm at 5 dB step and nominal $P_{1485}\in[6,13]$ dBm at 1 dB step, totaling 48 configurations. We measure each configuration under 5 profiles: flat C-band profile, one profile at -5 dB tilt, and 3 spectral profiles with a group of 12 channels elevated by 15 dB over others, placed on the left, center, or right part of the spectrum (Fig.2c). The latter profiles are fabricated to test ROPA under extreme conditions. Thus, Gain and NF of the ROPA 2 are measured over 48×5=240 configurations. Measurements on ROPA 3 are similar to those of ROPA 2 with next differences: $P_{1485}\in[6,12]$ dBm at 1 dB step, so 42 configurations per spectral profile, 210 in total, and for the tilted profile we have -2dB of tilt, instead of -5dB.

The reference Gain and NF ML models are trained on ROPA 1 measurements, as in [1]. We employ 3-layer ANN [9] with 30, 20, 10 neurons/layer with *tanh(x)* activation function. ANN has 4 standardly scaled [9] inputs: ${P_{in}^{tot}}_{lin}$ [mW], $P_{1485lin}$ [mW], G_{tilt}, λ [nm] and Gain [dB] or NF [dB] as output. The G_{tilt} parameter represents a generalized tilt over spectrum, and ordinary tilt over 84 channels is proportional to it; λ is a channel's wavelength. We train ANNs using LBFGS optimizer [9] with an objective to minimize Mean Squared Error (MSE) over the train dataset, we track MSE evolution over validation dataset to decide when to stop training (early stopping method). We test ANNs over the test dataset for ROPA 1 and present the results on Fig.3a) for Gain and Fig.3d) for NF. We consider only realistic Gain>10 dB and NF<10 dB: and we get Gain Root MSE (RMSE)=0.17 dB, and NF RMSE=0.04 dB.

3. Recalibration of ML models of Gain and Noise Figure and their adaptation to an arbitrary amplifier.

The objective of recalibration learning is to find adjustments to ML models of Gain and NF, so they are applicable to other devices. We suppose that the performance of an arbitrary ROPA could be scaled towards the reference ROPA if additional controls (like VOA or ideal amplifiers) placed on input/output ports and pump of ROPA. We assume that by adjusting such controls we could mimic the performance of a reference ROPA, as if reference ROPA was in place, but with additional recalibrations. This would mean a ReCalibration (RC) of powers measured at input/output of ROPA and its pump by ΔP_{in}^{RC}, ΔP_{out}^{RC}, ΔP_{1485}^{RC} respectively as: $P_{1485}^{ref}=P_{1485}^{trgt}+\Delta P_{1485}^{RC}$, $P_{in}^{tot\text{-}ref}=P_{in}^{tot\text{-}trgt}+\Delta P_{in}^{RC}$, $P_{out}^{tot\text{-}ref}=P_{out}^{tot\text{-}trgt}+\Delta P_{out}^{RC}$ [dB], with suffix "ref" for reference device, and "trgt" for target device. Thus, recalibration learning means finding recalibration factors and applying them on ML model inputs/outputs accordingly.

Let us consider recalibrated model of Gain, described in Fig.1a). The reference ANN takes as inputs powers read by reference device, so we shall provide as inputs the powers at a target device scaled to the reference (in linear domain): ${P_{in}^{tot\text{-}ref}}_{lin}={P_{in}^{tot\text{-}trgt}}_{lin}\times{\Delta P_{in}^{RC}}_{lin}$, ${P_{1485}^{ref}}_{lin}={P_{1485}^{trgt}}_{lin}\times{\Delta P_{1485}^{RC}}_{lin}$. Similarly, as reference ANN predicts Gain valid for reference device, ANN will predict $Gain^{ref}=Gain^{trgt}+\Delta Gain^{RC}$, where $\Delta Gain^{RC}=\Delta P_{out}^{RC}-\Delta P_{in}^{RC}$. Thus, the sought Gain of a target device will be defined as $Gain^{trgt}=Gain^{ref}-\Delta Gain^{RC}$. Now, let us consider NF ML model recalibration, with mechanics described in Fig.1b). As the inputs for NF ANN are the same as for Gain ANN, we replicate the same transformation for inputs. When it comes to the output, we conclude that in linear domain ${NF^{trgt}}_{lin}$ shall be calculated by the formula provided in Fig.1b) and then converted to dB. Such transformation requires knowledge of $Gain^{trgt}$, which is available when finding recalibration parameters, but is not available for prediction. So, we use an estimation of $Gain^{trgt}$ by Gain ML model. Otherwise, one can approximate $NF^{trgt}=NF^{ref}-\Delta P_{in}^{RC}$, without knowledge of $Gain^{trgt}$.

To find recalibration parameters we employ the next framework: we get the reference ANN (Gain or NF) and we make measurements on a target device. We freeze weights and biases of ANN, and then we find ΔPin^{RC}, $\Delta Pout^{RC}$, and $\Delta P1485^{RC}$ values by minimizing MSE error on Gain or NF predictions on a target device, while using the recalibration model described above. It is more practical to find ΔPin^{RC}, $\Delta Gain^{RC}$, and $\Delta P1485^{RC}$ as exactly these parameters are used by recalibration models. We employ LBFGS optimizer to find these recalibration parameters. Normally, they are

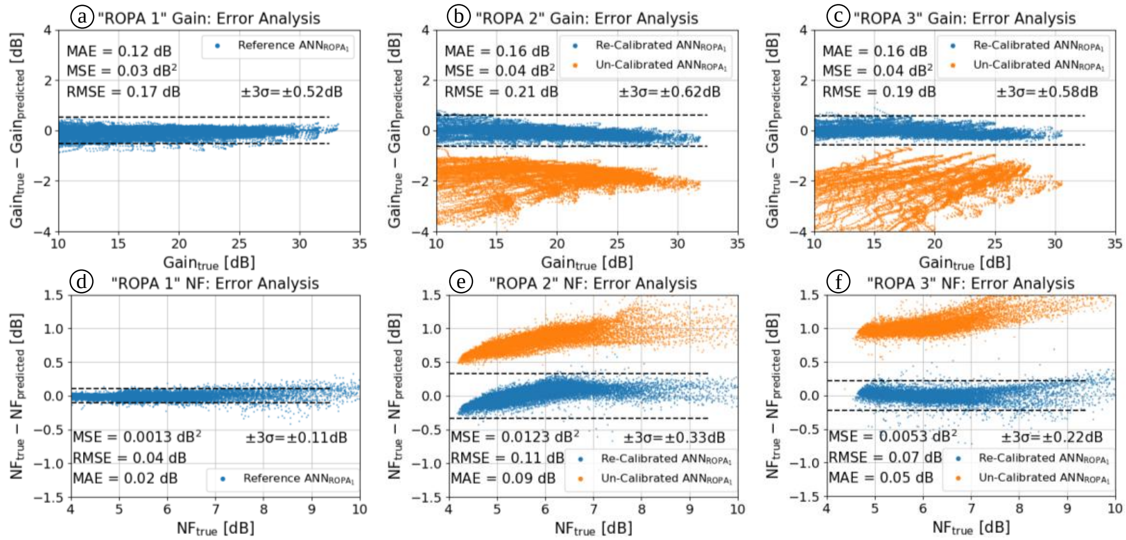

Fig. 3: ML model and its transfer analysis for Gain: a) ROPA 1 (ref), b) ROPA 2, c) ROPA 3; for NF: d) ROPA 1 (ref), e) ROPA 2, f) ROPA 3

Table.1: The summary of recalibration learning application to different ROPAs.

Gain and NF	Recalibration: Spectral Configurations			Gain: MAE [dB]			NF: MAE [dB]		
Recalibration	Count	P_{in}^{tot}[dBm]	P_{1485}[dBm]	ROPA 1 (ref)	ROPA 2	ROPA 3	ROPA 1 (ref)	ROPA 2	ROPA 3
Individual	1	-15	6	0,12	0,35	0,38	0,02	0,11	0,06
Individual	2	-15	[6, 12]	0,12	0,16	0,16	0,02	0,09	0,05
Individual	4	[-15, -25]	[6, 12]	0,12	0,13	0,14	0,02	0,10	0,04
Common	1	-15	6	0,12	0,18	0,30	0,02	0,13	0,06
Common	2	-15	[6, 12]	0,12	0,12	0,24	0,02	0,10	0,05
Common	4	[-15, -25]	[6, 12]	0,12	0,12	0,21	0,02	0,10	0,05

individual for Gain and NF recalibration models, but they can be made common via next technique: find ΔP_{in}^{RC} using recalibration model for NF, then fix this ΔP_{in}^{RC} for the recalibration of Gain and find $\Delta Gain^{RC}$, and $\Delta P1485^{RC}$.

4. Assessment of the recalibrated ML models of Gain and Noise Figure.

We assess the search for individual and common recalibration parameters for Gain/NF using 1, 2 and 4 configurations taken from flat profile measurements of ROPA 2 and 3 (cf. Table 1). We have 240 and 210 configurations, each with 84 Gain/NF points, for ROPA 2 and 3 respectively to test their recalibration parameters. Fig. 3 shows the outcome of this test for individual recalibration parameters found via 2 configurations (we remove them from the test). We provide Gain/NF prediction error as a function of known Gain/NF: blue points represent recalibration model and orange points represent direct application of reference ANNs; we identify Mean Absolute Error (MAE), MSE, and RMSE on recalibrated models. The divergence observed between orange and blue points attests to differences between devices.

Table 1 sums up tests for individual/common recalibration over 1, 2 and 4 configurations via Gain/NF prediction MAE: just 2 configurations (2/960 ~ 0.21% of data needed to train ML model) with individual calibrations are enough to transfer reference ML models. For Gain predictions we get similar MAE of 0.12, 0.16, 0.16 dB for ROPA 1, 2 and 3 respectively. As for NF we get MAE of 0.02, 0.09, 0.05 dB for ROPA 1, 2 and 3 respectively. These metrics are getting worse for Gain if to consider 1 configuration, but they are acceptable for NF. The common recalibration on 2 configurations also delivers similar MAE both Gain and NF across ROPAs, a little worse than with individual option. All of those results confirm the initial assumption that just 3 recalibration parameters over input/output ROPA ports and pump power are enough to reliably transfer the ML model of Gain and NF from a reference ROPA to any ROPA.

5. Conclusion

In this study we proposed a simple and accurate method of ROPA ML models transfer based on physical assumptions using just 0.21% of data, needed to train ML model. This method is warranted to be studied further in a generic EDFA.

6. References

[1] A. Minakhmetov et al., "Digital Twin of Unrepeatered Line Based on Raman and Remote Optically Pumped Amplifier Machine Learning Models," OFC 2023
[2] Y. Pointurier, "Machine learning techniques for quality of transmission estimation in optical networks," IEEE/OSA JOCN, vol. 13, B60-B71, 2021
[3] A. C. Meseguer et al., "Highly Accurate Measurement-Based Gain Model for Constant-Pump EDFA for non-Flat WDM Inputs," OFC 2021
[4] Y. You et al., "OSNR prediction using machine learning-based EDFA models," ECOC 2018
[5] F. da Ros et al., "Machine learning-based EDFA Gain Model Generalizable to Multiple Physical Devices," ECOC 2020
[6] Z. Wang et al., "Transfer Learning-based ROADM EDFA Wavelength Dependent Gain Prediction Using Minimized Data Collection," OFC 2023
[7] J. Lin et al., "Auxiliary Neural Network Assisted Machine Learning EDFA Gain Model," OFC 2023
[8] A. Raj et al. "Self-Normalizing Neural Network, Enabling One Shot Transfer Learning for Modeling EDFA Wavelength Dependent Gain," ECOC 2023
[9] I. Goodfellow et al, "Deep Learning," MIT Press, 2016.