

ML-Based Spectral Power Profiles Prediction in Presence of ISRS for Ultra-Wideband Transmission

Original

ML-Based Spectral Power Profiles Prediction in Presence of ISRS for Ultra-Wideband Transmission / Rosa Brusin, Ann Margareth; Nespola, Antonino; Zefreh, Mahdi Ranjbar; Piciaccia, Stefano; Poggiolini, Pierluigi; Forghieri, Fabrizio; Carena, Andrea. - In: JOURNAL OF LIGHTWAVE TECHNOLOGY. - ISSN 0733-8724. - ELETTRONICO. - 42:1(2024), pp. 37-47. [10.1109/JLT.2023.3301897]

Availability:

This version is available at: 11583/2984860 since: 2024-01-05T11:21:58Z

Publisher:

IEEE

Published

DOI:10.1109/JLT.2023.3301897

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

IEEE postprint/Author's Accepted Manuscript

©2024 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

ML-based Spectral Power Profiles Prediction in Presence of ISRS for Ultra-wideband Transmission

A. Margareth Rosa Brusin, Antonino Nespola, Mahdi Ranjbar Zefreh, Stefano Piciaccia, Pierluigi Poggiolini, Fabrizio Forghieri, and Andrea Carena

Abstract—A generalized method based on machine learning (ML) and artificial neural networks (ANNs) is proposed for a fast and accurate prediction of spectral and spatial evolution of power profiles in support of performance and quality-of-transmission (QoT) real-time assessment of ultra-wideband links. These systems, operating on bandwidths larger than the standard C-band, are affected by inter-channel stimulated Raman scattering (ISRS), whose impact on power profiles evolution along the fiber is generally estimated by solving numerically a set of nonlinear ordinary differential equations (ODEs). However, the computational effort, in terms of complexity and convergence time to the solution, increases with the bandwidth and the number of transmitted wavelength division multiplexing (WDM) channels, which makes the usual approach no longer particularly suitable to operate in real time.

To meet the speed requirements, three different ANNs are introduced to make fast predictions of power profiles over frequency and distance considering a wide range of scenarios: different power per channel values, different fiber types and different span lengths. Two ANNs are used on synthetic data to estimate the impact of linear and nonlinear fiber impairments in support of system modeling. Specifically, one to directly predict the evolution of spectral power profiles along the fiber and the other to estimate the coefficients to insert in a closed-form version of the EGN model. A third ANN operates on experimental data and it is used to predict power profiles at the end of the fiber for fast estimations of system performance.

The obtained results show highly accurate predictions with values of maximum absolute error, computed between predicted and actual power profiles, not exceeding 0.2 dB for ~97% of cases for synthetic data and always below 0.5 dB for experimental data. Such results prove the potential of the proposed approach making it suitable for real time application of QoT estimation.

Index Terms—Optical fiber communication, multi-band systems, ultra-wideband transmission, inter-channel stimulated Raman scattering (ISRS), machine learning, artificial neural networks.

I. INTRODUCTION

THE disparity in the growth rate between the internet data traffic (~60%/year) and the capacity actually provided by

commercial optical fiber systems (~20%/year), will lead in the near future to the so-called *capacity crunch* [1]. Indeed, in the last few years, the emergence of new Internet applications, 5G technologies, cloud computing, video streaming, Internet-of-Things (IoT) and machine-to-machine (M2M) communications, together with the increase of the number of users and devices connected, has caused an exponential increase of capacity demand [2], [3]. To cope with this persistent growth, improvements of current optical communication systems must be implemented.

For this purpose, different technologies are currently being investigated by researchers, each characterized by noticeable advantages in terms of bandwidth increase. An interesting cost-effective solution consists in ultra-wideband transmission, trying to fully exploit all optical bands (C-, L-, S-, O- and E-band) available in the fiber. In this case, the capacity increases linearly up to 10 times with respect to current standard single mode fiber (SSMF) systems [4], but most importantly this solution allows to exploit the broad spectrum available in the already deployed fibers, avoiding the installation of new cables (and additional further costs).

Even more capacity can be achieved if space division multiplexing (SDM) technologies are considered. Implemented through multiple parallel fibers (MPF), multi-mode fibers (MMF) or multi-core fibers (MCF), SDM is able to provide up to 2-3 orders of magnitude the capacity of the current SSMF [4], [5], scaling with the number of fibers, modes or cores, respectively. Unfortunately, such a larger capacity comes at the cost of deploying new dedicated cables, thus ultra-wideband transmissions represent the most suitable candidate as a short-term solution.

When the transmission bandwidth is extended beyond the usual C-band, other nonlinear impairments, besides the Kerr effect, become stronger and no longer negligible, causing a further degradation of transmission quality. Among them, inter-channel stimulated Raman scattering (ISRS) is particularly noteworthy. Consisting in a power transfer from higher to lower frequency carriers, ISRS depends on the spectral load provided at the input of the fiber and produces a tilt on the power spectral profile [6]. Furthermore, its effect is stronger in the first kilometers of the fiber, where indeed the power is higher.

In optical communication networks design, optimization and monitoring, the accurate estimation of physical layer propagation effects is fundamental. Computation speed is essential as well. Several physical layer models, accounting for both linear and nonlinear effects in the fiber, have been proposed over the

A. M. Rosa Brusin, P. Poggiolini and A. Carena are with the Dipartimento di Elettronica e Telecomunicazioni, Politecnico di Torino, Torino, Italy e-mail: ann.rosabrusin@polito.it.

A. Nespola is with LINKS Foundation.

M. Ranjbar Zefreh, S. Piciaccia and F. Forghieri are with Cisco Photonics Italy.

This work was partially supported by the CISCO Sponsored-Research Agreements (SRA) "OptSys 2022" and "IT-ROCS 2023", by the PhotoNext Center of Politecnico di Torino and by the European Union under the Italian National Recovery and Resilience Plan (NRRP) of NextGenerationEU, partnership on "Telecommunications of the Future" (PE00000001 - program "RESTART").

Manuscript received Month Day, Year; revised Month Day, Year.

last decade [7]–[14]. Among them, the GN and EGN models [10], [11] have achieved wide adoption by both the industry and academia.

While the original version of these models required time-consuming numerical integration, approximate closed-form model (CFM) versions have been worked out. Initially these CFMs were limited to C-band systems, but more recently they have been extended to ultra-wideband (UWB) by making them capable of accounting for frequency-dependent dispersion, loss, non-linearity coefficient and, most importantly, ISRS [15]–[18].

These GN/EGN approximate CFMs can provide a full UWB multi-span system performance estimation in milliseconds. However, as input, they require the spectrally resolved power evolution along the span. These spectral power profiles are shaped by ISRS. Unfortunately, their estimation requires the numerical integration of a set of nonlinear ordinary differential equations (ODEs) with a complexity that scales up with the number of WDM channels.

While approximate closed-form solutions have been proposed for the ISRS ODEs, they are accurate only within rather stringent assumptions on the WDM signal spectrum and general system features. In fact, the numerical integration of the ODEs is often preferred, because power profile accuracy is in turn key for the GN/EGN CFM to provide reliable results. The ISRS ODE integration time turns out to be the limiting factor in the speed of the overall GN/EGN CFM system performance assessment.

To solve this issue, particularly significant in UWB transmission scenario, already in [19], two artificial neural network (ANN) models were presented to predict the evolution of power profiles along the fiber span. One model directly predicts the spectral and spatial evolution of power profiles, while the other predicts the parameters required in the closed form-model formula demonstrated in [20] for NLI estimation, where instead these parameters are determined through fitting. This machine learning (ML) based solution finds support in the promising results obtained when ML and ANNs were applied to other optical communication systems problems, such as in the analysis and design of Raman amplifiers [21]–[24].

The determination of the NLI CFM coefficients through best fitting of the true power profile in presence of ISRS, computed by numerically integrating Eq. (1) as in [20], is also investigated in [25]–[27]. In particular, in [26] and [27] also Raman amplification is taken into account. Nevertheless, in all these works no clear indications are provided neither on the computational time required by the fitting procedure to determine the coefficients nor on the spatial resolution considered in the evaluation of the power profile evolution along frequency and distance.

As a matter of fact, the spatial resolution (as well as the spectral resolution) has a significant impact on the computational time required by the numerical RS to determine the power profile evolution. A higher resolution is preferred to have a more accurate power profile description, but at the same time it is more computationally demanding. Moreover, the fitting optimization needs to be performed again from scratch every time a new link configuration is considered.

In this context, the method presented in [19] was proved to be ultra-fast and highly accurate in power profiles prediction, achieving maximum absolute errors within 0.1 dB. However, the analysis was limited to a single type of fiber (SSMF) and to a single value of launched power per channel ($P_{ch}=0$ dBm). For this reason, in this paper we propose an upgrade of both ANN models to support a generalized scenario. We test the new designed ANNs over different types of fibers and multiple power per channel values. To have a full control of the data-set generation performed for different scenarios of power per channel and fiber type, the analysis is carried out considering synthetic data. For simplicity in the generation of the data-sets, we assume uniform launch power profiles, namely the channels turned on are assumed to have equal power.

Afterwards, we propose an additional ANN model to be used for in-field applications to predict power profiles at the end of the fiber span. Indeed, this ANN represent an interesting tool to obtain fast estimations of system performance. In this case, to consider a more reliable and realistic scenario, the ANN model is trained, validated and tested using experimental data.

The paper is organized as follows. Section II first presents the system setup used for the synthetic data-sets generation. After that, the two ANN models, referred to as Model 1 and Model 2, introduced to support modeling of linear and nonlinear propagation effects, are described together with an extensive discussion on their prediction performance. Then, in Section III the third ANN model, referred to as Model 3, is presented for the prediction of power profiles at the end of the fiber span for in-field system performance evaluation. Since the ANN Model 3 operates on experimental data, the experimental setup and the generation of the training and testing data-sets are illustrated. The ANN prediction performance is discussed at the end of Section III. Finally, the conclusions are drawn in Section IV.

II. ANNS TO SUPPORT MODELING OF LINEAR AND NONLINEAR PROPAGATION EFFECTS: MODEL 1 AND MODEL 2

In this Section, we first describe the scenario under analysis and the system setup considered for the generation of the synthetic data-sets used in the ML-framework. Then, we present the considered ANN models, with a brief description of the training and validation process, followed by the testing results.

Similarly to the study presented in [19], the effect of different input spectral loads causing inter-channel stimulated Raman scattering (ISRS) is analyzed over a single fiber span, but compared to [19], here the study is extended to a more general scenario including multiple types of fibers and different values of channel power.

A. System Setup and Synthetic Data-sets Generation

To generate the training and testing data-sets required in the machine learning framework, we follow the same procedure presented in [19], where the testing data-set was generated on a *channel-basis*, while the training data-set was generated on

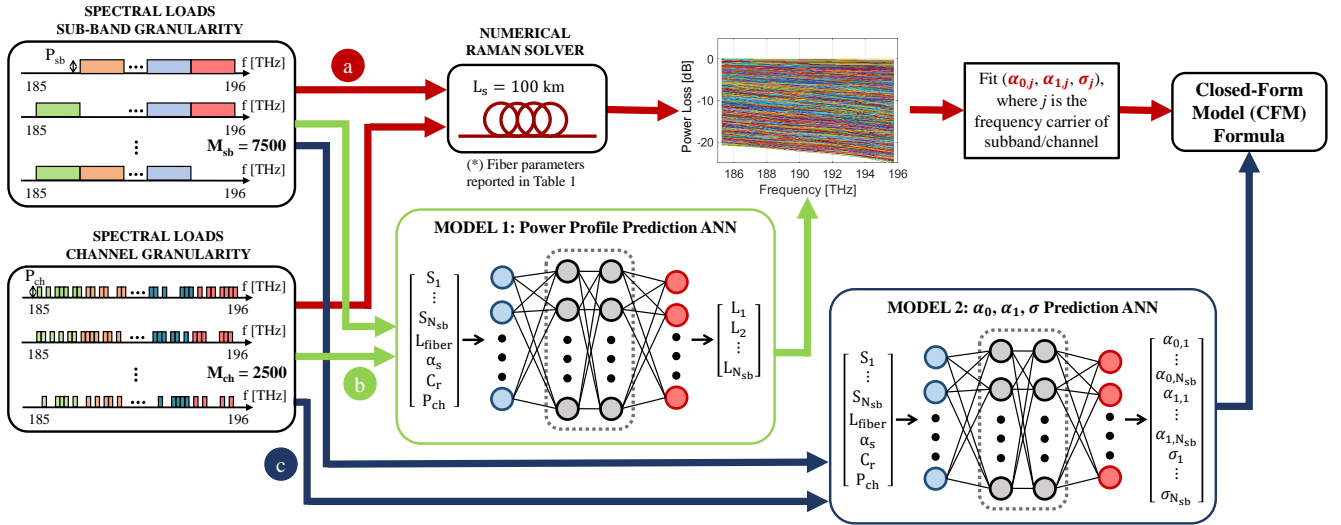


Fig. 1: Paths: (a) Simulation setup considered for data-sets generation, (b) ANN Model 1 used for power profiles prediction and (c) ANN Model 2 used for prediction of parameters α_0 , α_1 and σ used in the closed-form model (CFM) formula to evaluate power profile evolution. M_{sb} represents the number of different spectral loads extracted for the sub-band granularity data-set for each discrete value of sub-band power P_{sb} , while M_{ch} corresponds to the number of different spectral loads extracted for the channel granularity data-set for each discrete value of channel power P_{ch} .

a *subband-basis*, with each subband represented by a group of 10 adjacent channels. This assumption is based on the fact that power loss profiles do not change significantly along the frequency of adjacent channels. In [19], this choice was motivated by the need of reducing the space to explore in the generation of different partial spectral load conditions for the training data-set over the 220 channels, as 2^{220} combinations of channels with on/off states were possible. Moreover, operating on a *subband-basis* was beneficial also to reduce the complexity at the neural network since the number of inputs and outputs is reduced. In particular, the choice of 10 adjacent channels represented a good trade-off between accuracy and complexity.

The system setup is shown in Fig. 1, path (a). Similarly to [19], also here we assume to operate both on channels and on subbands. More precisely, in case of channels the system under study consists of an ideal Nyquist wavelength division multiplexing (WDM) comb, composed of 220 channels (N_{ch}) in the ITU-T grid assuming 50 GHz frequency spacing over the C+L-band operating between 185 THz and 196 THz, for a total of 11 THz bandwidth. The power carried by each channel can assume the following discrete values $P_{ch} = [-10, -7.5, -5, -2.5, 0, 2.5]$ dBm. Instead, in case of subbands, each subband is identified by 10 adjacent channels with same frequency spacing and symbol rate as before, corresponding to a total of 22 subbands with 500 GHz frequency spacing and with power given by the contribution in power of these 10 channels. Therefore, the power levels that can be assumed by each subband are $P_{sb} = [0, 2.5, 5, 7.5, 10, 12.5]$ dBm.

Once generated, the WDM comb is given at the input of a single span of fiber of $L_s = 100$ km length and then it is propagated along the span. Its power evolution along

frequency and distance can be modelled by the set of ordinary differential equations (ODEs) reported in Eq. (1), one for each frequency channel (subband) j [20], [28]:

$$\begin{cases} \frac{dP_1(z)}{dz} = P_1(z) \left[-2\alpha_1 + \sum_{i=2}^N C_r(f_i - f_1) P_i(z) \right] \\ \vdots \\ \frac{dP_j(z)}{dz} = P_j(z) \left[-2\alpha_j + \sum_{i=j+1}^N C_r(f_i - f_j) P_i(z) + \right. \\ \quad \left. - \sum_{i=1}^{j-1} \frac{f_j}{f_i} C_r(f_j - f_i) P_i(z) \right] \\ \vdots \\ \frac{dP_N(z)}{dz} = P_N(z) \left[-2\alpha_N - \sum_{i=1}^{N-1} \frac{f_N}{f_i} C_r(f_N - f_i) P_i(z) \right] \end{cases} \quad (1)$$

Specifically, N is number of channels (subbands) considered, z is the distance from the fiber span origin, $P_j(z)$ is the power of channel j at distance z , α_j is the fiber attenuation coefficient, which can be different channel by channel (subband by subband) as in general it depends on frequency, $C_r(f_i - f_j)$ is the true Raman gain profile, or Raman gain efficiency coefficient, expressed as a function of $(f_i - f_j)$, which is a characteristic of the fiber related to the effective area, and finally $P_j(z)$ is the power transferred from/to the other WDM channels due to ISRS. Indeed, since the system operates over the C+L-band, the effects caused by ISRS are no longer negligible.

Since no analytical solutions are available to the expressions in Eq. (1), in general they are solved numerically by means of Raman solvers (RS). In our case, we use the RS available within the open source library GNPpy [29]. Thus, given both

input and output power profiles, it is possible to determine the power loss L for any channel/subband j along the fiber span at any desired distance z as:

$$L_j(z) = 10 \log_{10} \left[\frac{P_j(z)}{P_j(0)} \right] \quad (2)$$

A qualitative representation is shown in Fig. 1, where the power loss profiles of a subband granularity data-set are plotted with respect to the frequency at different distances.

Depending on whether the training or the testing data-set is generated, the parameters of the fiber and the generation procedure can be different, for this reason they will be explained separately in the following.

1) *Training Data-set Generation*: Real deployed optical communication networks are not made up of only one type of fiber, but instead different ones can be found, which are characterized by different attenuation coefficient α and Raman gain efficiency coefficient C_r (related to the fiber effective area). Therefore, since the ANN should handle different scenarios, the training data-set is not generated for just a specific type of fiber. On the contrary, both attenuation coefficient and Raman gain efficiency coefficient are extracted from uniform distributions ($\alpha \sim \mathcal{U}[0.14, 0.22]$ dB/km and $C_r \sim \mathcal{U}[0.19, 0.74]$ W⁻¹km⁻¹, as reported in Table I), whose lower and upper limits have been selected to cover the range of values of commercial fibers. To simplify the generation of the data-sets, the fiber attenuation coefficient is assumed to be flat along frequency. Although this assumption might be too simplistic, it would not impact the validity of the methodology, because the ANN would just learn the relation between the input spectral load and the power profile at the desired distance, regardless the assumptions considered for the power profiles generation. To consider a more realistic scenario where the fiber loss profile is a function of the frequency, the ANN just need to get at its input the fiber loss profile vector instead of a single scalar.

As mentioned above and similarly to [19], the training data-set generation is performed on a *subband-basis*. Assuming that each subband has 50% probability of being turned on or off to emulate partial loads, for each power level value, we extract 7500 different configurations of input spectral load (M_{sb}), attenuation coefficients α and Raman gain efficiency coefficient C_r . Then, the corresponding power profile and loss profile are determined by means of the numerical RS, with a resolution of 1 km. The frequency position of the subbands that are turned on is randomly selected. The value of M_{sb} is selected by increasing by 50% the number of partial spectral loads generated in [19], i.e. 5000, since in this analysis also α and C_r are randomly extracted. Then, the overall training data-set is obtained by merging and shuffling all the obtained profiles, for a total of 7500 (*partial spectral loads*) \times 100 (*distance points*) \times 6 (*channel power values*) = $4.5 \cdot 10^6$ samples.

2) *Testing Data-set Generation*: Since the trained ANN needs to be used to make predictions on existing fibers, the testing data-sets are generated considering real commercial fibers: standard single mode fiber (SSMF), two non-zero dispersion shifted-fibers (NZDSF) Truewave-RSTM and LEAFTM

TABLE I: Parameters for different types of fiber.

Fiber Type	α @1550 nm [dB/km]	C_r [W ⁻¹ km ⁻¹]	A_{eff} [μm^2]
Random	[0.14,0.22]	[0.19,0.74]	-
SSMF	0.21	0.3841	80
Truewave-RS TM [30]	0.21	0.73	55
LEAF TM [30]	0.19	0.56	72
TeraWave [®] SCUBA 125 [31]	0.158	0.2458	125
TeraWave [®] SCUBA 150 [32]	0.155	0.2008	153

α : attenuation coefficient, C_r : Raman gain efficiency coefficient, A_{eff} : effective area.

[30], and two ultra-low loss fibers TeraWave[®] SCUBA 125 [31] and TeraWave[®] SCUBA 150 [32]. Their attenuation coefficient, Raman gain efficiency coefficient and effective area are reported in Table I. In case of TeraWave[®] SCUBA 125 and TeraWave[®] SCUBA 150, the Raman gain coefficients C_r are not provided in the datasheets, therefore we analytically computed them using Eq. (3):

$$C_r = \frac{g_R}{A_{eff}} \quad (3)$$

where A_{eff} is the effective area given in the datasheets and for the Raman gain coefficient g_R we assume the same value as for SSMF.

Also in this case, channels are assumed to be on or off with 50% of probability. For each fiber type and power per channel level, we extract 2500 configurations of load (M_{ch}), such that the position of channels turned on is randomly selected. The evolution of power and loss profiles is numerically determined at every kilometer along the fiber span using the RS. The resulting testing data-set is composed of 2500 (*partial spectral loads*) \times 100 (*distance points*) \times 6 (*channel power values*) \times 5 (*fiber types*) = $7.5 \cdot 10^6$ samples.

B. Machine Learning Framework and Artificial Neural Network Models

Once generated, the two data-sets, which are actually independent, are used to train and then test the ANNs. As discussed above, already in [19] the use of ANNs to make accurate spatial and spectral predictions of loss profiles was proposed and successfully demonstrated, but there it was limited to a single type of fiber (SSMF) and to 0 dBm power per channel. In particular, in [19] two ANN models were presented. A first ANN model, Model 1, used to predict the loss and the output power profiles at any desired distance. Then, a second model was introduced, Model 2, predicting the triplets of coefficients α_0 , α_1 and σ . These coefficients are the contributing terms to the equivalent channel loss in a span, expressed as [33]:

$$\alpha_j(z) = \alpha_{0,j} + \alpha_{1,j} \cdot \exp(-\sigma_j \cdot z) \quad (4)$$

where j represents the channel (sub-band) index. $\alpha_{0,j}$ corresponds to the fiber loss without accounting for ISRS, $\alpha_{1,j}$ represents the loss variation induced by ISRS at the beginning of the span, while σ_j is an indication of the speed of ISRS effect in vanishing with the optical power along the fiber span. Once $\alpha_{0,j}$, $\alpha_{1,j}$ and σ_j have been assigned to each channel (sub-band), the nonlinear interference power can be fully calculated in closed-form.

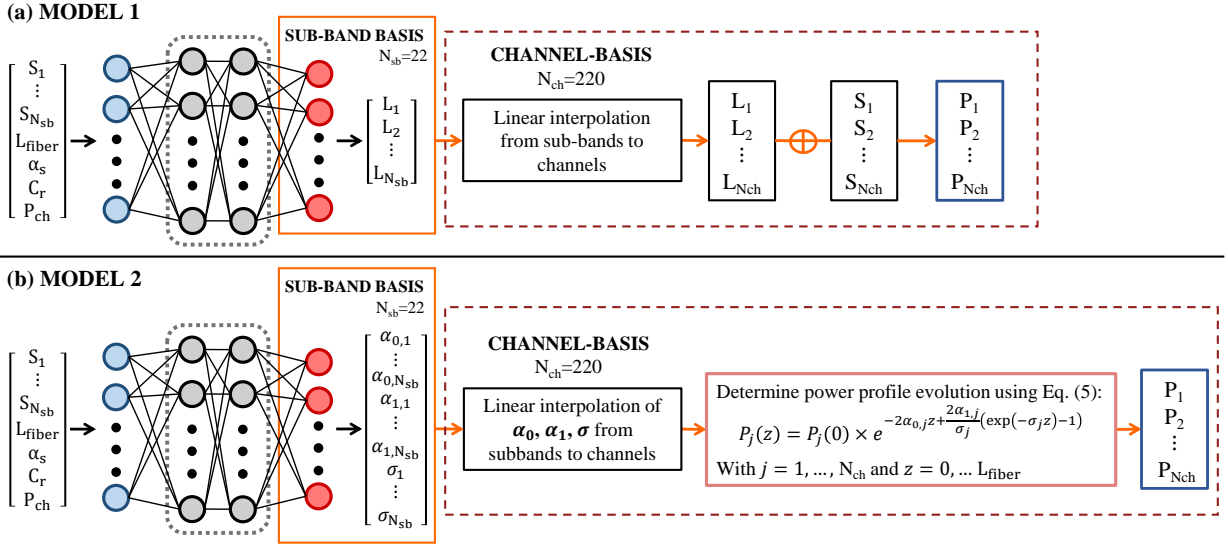


Fig. 2: Testing process from subbands to channels for (a) ANN Model 1 and (b) ANN Model 2. For Model 1, the loss profiles \mathbf{L} , the input power loads \mathbf{S} and the output power profiles \mathbf{P} are in logarithmic units.

Nevertheless, these coefficients, that can assume different values channel by channel (sub-band by sub-band), can not be computed in closed-form. Instead, they are determined by fitting the actual power profile with Eq. (5):

$$P_j(z) = P_j(0) \times \exp \left(-2\alpha_{0,j}z + \frac{2\alpha_{1,j}}{\sigma_j} (e^{-\sigma_j z} - 1) \right) \quad (5)$$

such that a cost function looking at the errors along the fiber span is minimized [20]. Additionally, since stronger nonlinear effects mostly occur in the first kilometers of the fiber span, where indeed the power is stronger, higher weights are assigned to the cost function in this section.

In extending these two approaches to handle multiple fiber types and channel power levels, it is fundamental to provide the ANN with useful information regarding the scenario for which it is making predictions. Fig. 1, paths (b) and (c), shows the ANN models considered in this paper. Compared to the ANNs in [19], together with the vector of spectral load $\mathbf{S} = [S_1, S_2, \dots, S_{N_{sb}}]^T$, with N_{sb} being the number of subband, and the position along the fiber L_{fiber} , also the fiber attenuation coefficient α , the Raman gain efficiency coefficient C_r and the power per channel P_{ch} are given as *features* at the input of both ANN models. The vector of spectral load \mathbf{S} is in logarithmic units (dBm). Instead, the *labels* at the output of the ANNs are the same as those in [19], which indeed are: the vector of loss profile $\mathbf{L} = [L_1, L_2, \dots, L_{N_{sb}}]^T$ in dB for Model 1, and the vector of coefficients $[\alpha_0, \alpha_1, \sigma] = [\alpha_{0,1}, \dots, \alpha_{0,N_{sb}}, \alpha_{1,1}, \dots, \alpha_{1,N_{sb}}, \sigma_1, \dots, \sigma_{N_{sb}}]^T$ expressed in dB/km for Model 2. We can notice that, although we have added further features, the size of the ANNs remains limited, with 26 nodes at the input and 22 (66) for Model 1 (Model 2) at the output, confirming the advantage of operating on a subband-basis.

To create our ANN models and to train, validate and test them, we rely on TensorFlow™ and Keras libraries. For Model 1, the ANN architecture is composed of 2 hidden layers

(HLs) and 1000 hidden nodes (HNs). The ANN weights are initialized based on Glorot initialization with uniform distribution, the default option in Keras. The training is based on the back-propagation training algorithm considering Rectifier Linear Unit (ReLU) activation function over 1000 epochs and assuming learning rate $\lambda=0.001$. The loss function used in the training is the mean square error (MSE) computed between the target and the predicted outputs of the ANN, i.e. the true power profile from the RS and the power profile predicted by the ANN, respectively. During the training phase, 10% of the training data-set is held out to validate the accuracy of predictions and to keep under control any possibility of over-fitting and under-fitting of the ANN. A similar ANN architecture is considered for Model 2: 2 HLs, 500 HNs, ReLU activation function, MSE loss function, learning rate $\lambda=0.001$ and 1000 epochs, but here the MSE is computed between the fitted and the predicted coefficients. Also in this case, 10% of the training data-set is used for ANN validation.

Once trained, the ANNs are used to make predictions on unseen data, in our case on the testing datasets for different fiber types and different levels of channel power. The entire testing process is schematized in Fig. 2 for both ANN models. More precisely, the power per subband, given as the contribution of the channels turned on in each subband, is provided at the input of both ANNs. Since the output layer of Model 1 provides 22 points for the loss profile, linear interpolation is performed over the 220 channels. Then, given the input power profile along channels, it is possible to compute the power profile at the desired distance.

Similarly for Model 2, for which also the coefficients α_0 , α_1 and σ , predicted with a subband granularity just for $L_{fiber}=[50-100]$ km range of distances (assuming 1 km step), are linearly interpolated over the 220 channels and afterwards inserted in Eq. 5 to compute the overall power profile evolution from the fiber start to any desired distance L .

To assess the prediction performance of ANNs, we compare

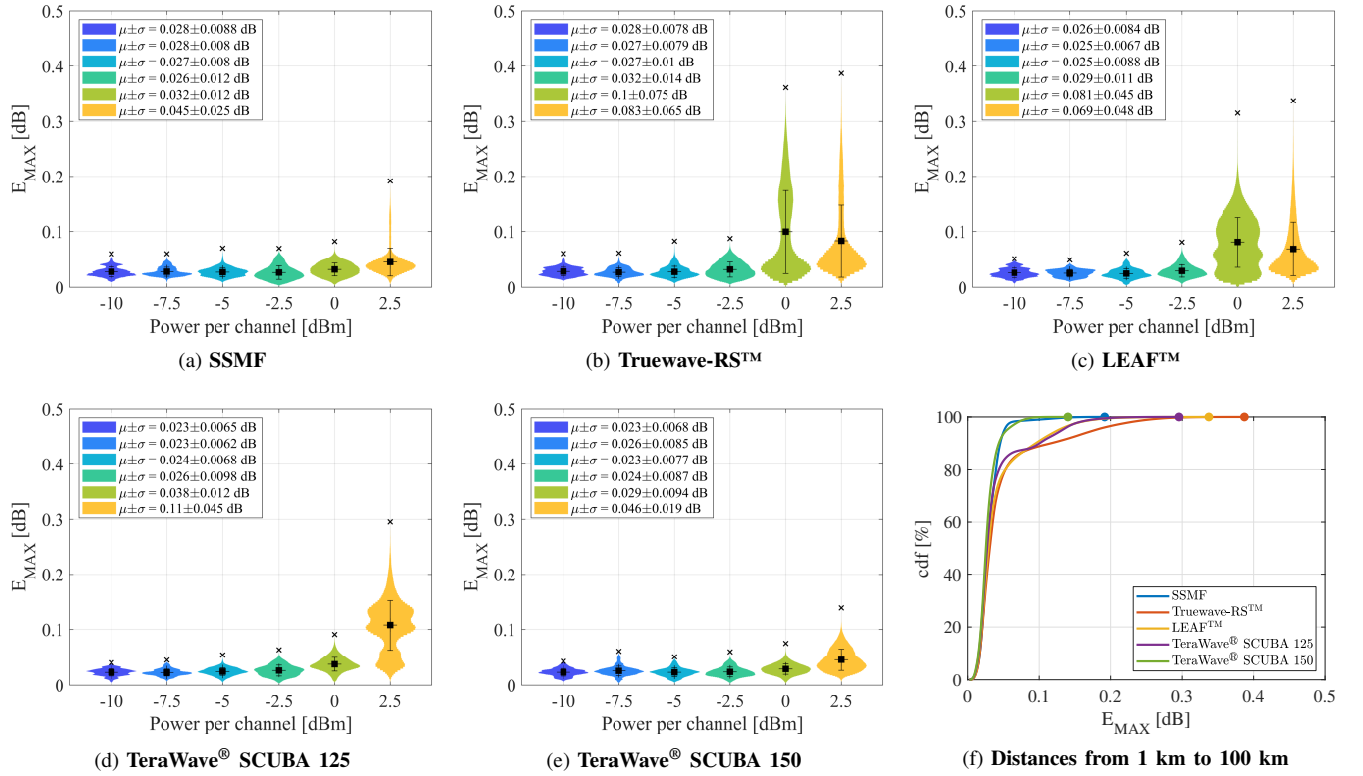


Fig. 3: Violin plots versus power per-channel of the *maximum absolute error* (E_{MAX}) computed between the power profiles predicted using *Model 1* and the actual power profiles at every kilometer for: (a) SSMF, (b) Truewave-RSTM, (c) LEAFTM, (d) TeraWave[®] SCUBA 125 and (e) TeraWave[®] SCUBA 150, at different spectral uniform launch power profiles. For each violin plot the mean value and the standard deviation are also reported, together with the maximum value, plotted with cross marker. (f) Cumulative distribution functions of E_{MAX} for the different fiber types including all power per channel values.

each predicted power profile P^{pred} and the corresponding target power profile P^{targ} in terms of maximum absolute error (E_{MAX}) defined as follows:

$$E_{MAX} = \max_{n \in \{1, \dots, N_f\}} |P_n^{pred} - P_n^{targ}| \quad (6)$$

where P^{pred} , P^{targ} and E_{MAX} are in logarithmic units. Instead, N_f is the number of points in frequency, which actually corresponds to the number of channels. In case of ANN Model 1, P^{pred} and P^{targ} are represented by the power profile directly predicted by the ANN and the power profile generated by the numerical RS, respectively. Instead, in case of ANN Model 2, P^{pred} and P^{targ} are the power profiles obtained inserting respectively the predicted and the fitted coefficients α_0 , α_1 and σ in Eq. (5).

C. Testing Results of ANN Model 1

Fig. 3 shows the violin plots for the E_{MAX} of the power profiles predicted at every kilometer using ANN Model 1 plotted with respect to the different power per-channel levels for the five considered types of fiber. For each violin plot we also highlight the mean value and the standard deviation in black with their values reported in the legend. Additionally, the respective maximum values are plotted as single cross marker. From Fig. 3, we can observe that up to $P_{ch} = -2.5$ dBm the E_{MAX} is always below 0.1 dB with average around 0.025 dB

and reduced standard deviation for all fiber types. A similar behavior is observed when $P_{ch} = 0$ dBm in case of SSMF, TeraWave[®] SCUBA 125 and TeraWave[®] SCUBA 150, while for the Truewave-RSTM and the LEAFTM this is no longer true. Indeed, as can be seen in Figs. 3(b) and 3(c), for these two last types of fiber, prediction errors can be larger than 0.1 dB at 0 dBm of power per channel, with values up to 0.36 dB and 0.31 dB, respectively. Consequently, also their mean and standard deviation values increase, corresponding to 0.1 ± 0.075 dB and 0.081 ± 0.045 dB, respectively for Truewave-RSTM and LEAFTM.

To explain this behavior, we should recall that, due to a smaller effective area, the Truewave-RSTM and the LEAFTM have higher Raman gain efficiency coefficient C_r than the other three types of fiber, as reported in Table I. As a consequence, the ISRS effect is stronger, which produces a stronger tilt on the power profile along the fiber span.

With $P_{ch} = 2.5$ dBm, the highest value of power considered, both ISRS effect and tilt of power profiles become particularly significant also for SSMF, TeraWave[®] SCUBA 125 and TeraWave[®] SCUBA 150. Indeed, the more power is transferred from higher to lower frequency carriers, the more different power profiles can be at a given distance, also depending on the input spectral load. This affects also the prediction performance of the ANN, as it is required to handle a larger space of input-output combinations. This is clear in Fig. 3,

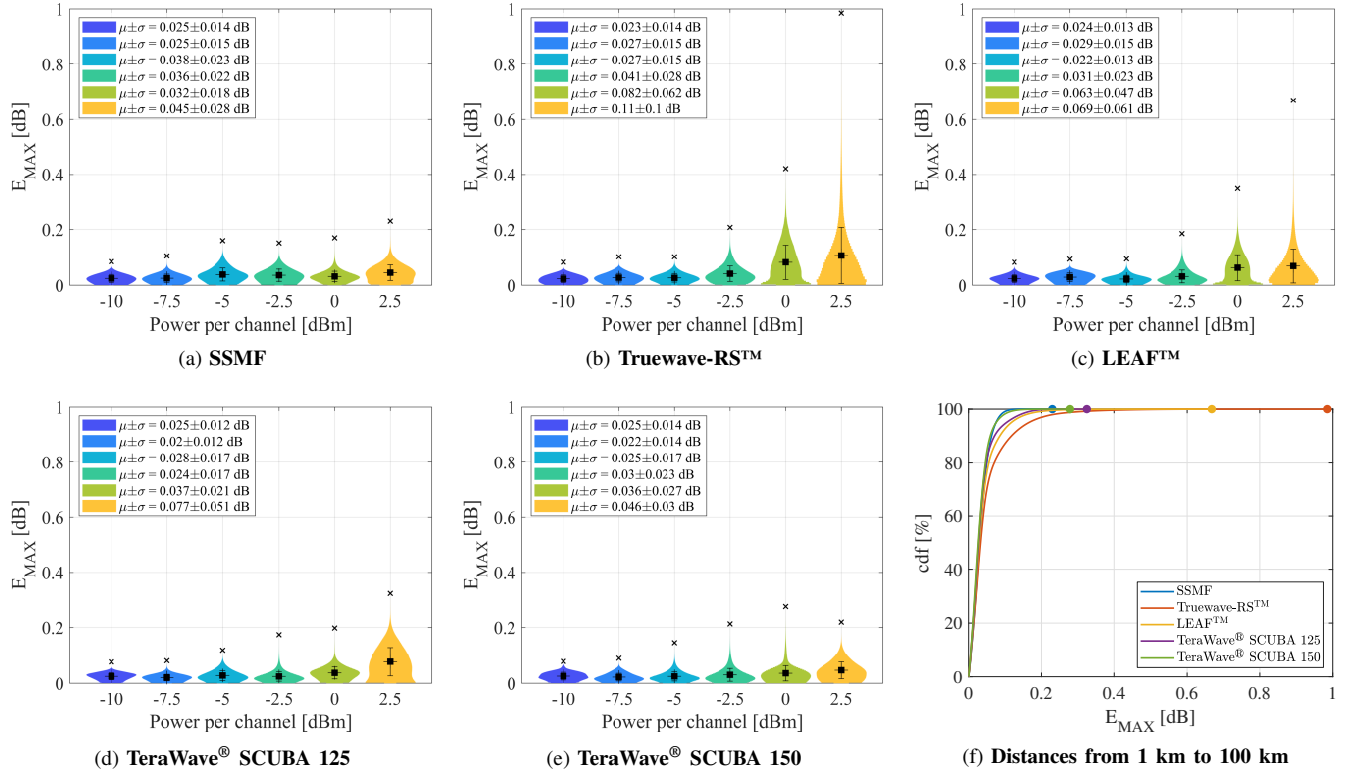


Fig. 4: Violin plots versus power per-channel of the *maximum absolute error* (E_{MAX}) computed between the power profiles predicted using *Model 2* and the actual power profiles for distances from 1 km to $L_{fiber} = [50-100]$ km in case of: (a) SSMF, (b) Truewave-RSTM, (c) LEAFTM, (d) TeraWave[®] SCUBA 125 and (e) TeraWave[®] SCUBA 150, at different spectral uniform launch power profiles. For each violin plot the mean value and the standard deviation are also reported, together with the maximum value, plotted with cross marker. (f) Cumulative distribution functions of E_{MAX} for the different fiber types including all power per channel values.

where for all fiber types E_{MAX} can assume values larger than 0.1 dB, with larger values obtained in case of Truewave-RSTM and LEAFTM. However, the mean values stay within 0.1 dB, except for the TeraWave[®] SCUBA 125, which indeed on average is characterized by the worst ANN prediction performance for $P_{ch}=2.5$ dBm.

Interestingly, we can notice that for Truewave-RSTM and LEAFTM, although the maximum values of E_{MAX} are the largest with respect to the power per channel, the mean value is slightly lower when $P_{ch}=2.5$ dBm than when $P_{ch}=0$ dBm.

Nevertheless, the ANN Model 1 analyzed here shows great abilities in making accurate predictions, performing even better than the ANN models considered in [19] under the same scenario, i.e. SSMF and $P_{ch}=0$ dBm. As a matter of fact, here with just a single ANN, trained to handle different scenarios, we are able to obtain predictions with E_{MAX} way lower than 0.1 dB without performing span slicing into sub-spans and without considering data-sets with sub-bands partially turned on. In fact in [19], sub-bands were not only considered completely turned on or off (2 states), but they could also assume intermediate states.

The ability of ANN Model 1 in making highly accurate predictions is also confirmed by the cumulative distribution functions (cdf) of E_{MAX} showed in Fig. 3(f) for each fiber type and considering all together the different power per

channel values. Indeed, there we can see that $E_{MAX} \leq 0.2$ dB for more than 96.6% of cases when Truewave-RSTM is considered, while this percentage is $\sim 100\%$ for all other types of fibers. The best performance is observed for SSMF and TeraWave[®] SCUBA 150, for which E_{MAX} is always below 0.2 dB and assumes values ≤ 0.1 dB for more than 99% of cases.

Apart from the prediction accuracy of the trained ANN, it is fundamental to compare its computational effort with respect to the numerical RS. For this purpose, we run a speed test on the same server (Intel Xeon CPU E5-2690 v4 @ 2.60 GHz) at same conditions. The RS requires ~ 3 minutes to provide the complete evolution of the power profile for a single input partial load along frequency channels (220 channels) and distance up to the span length (1 km spatial granularity is assumed). On the contrary, the ANN Model 1 is capable of predicting the overall power profile evolution in only 0.24 s, proving the potential of the ANN to be used in real-time in-field applications.

D. Testing Results of ANN Model 2

Once trained, the ANN Model 2 is used to make predictions for fiber spans of length $L_{fiber} = [50-100]$ km with a resolution of 1 km. This means that the coefficients predicted for a desired length L_{fiber} are inserted in Eq. (5) to compute the

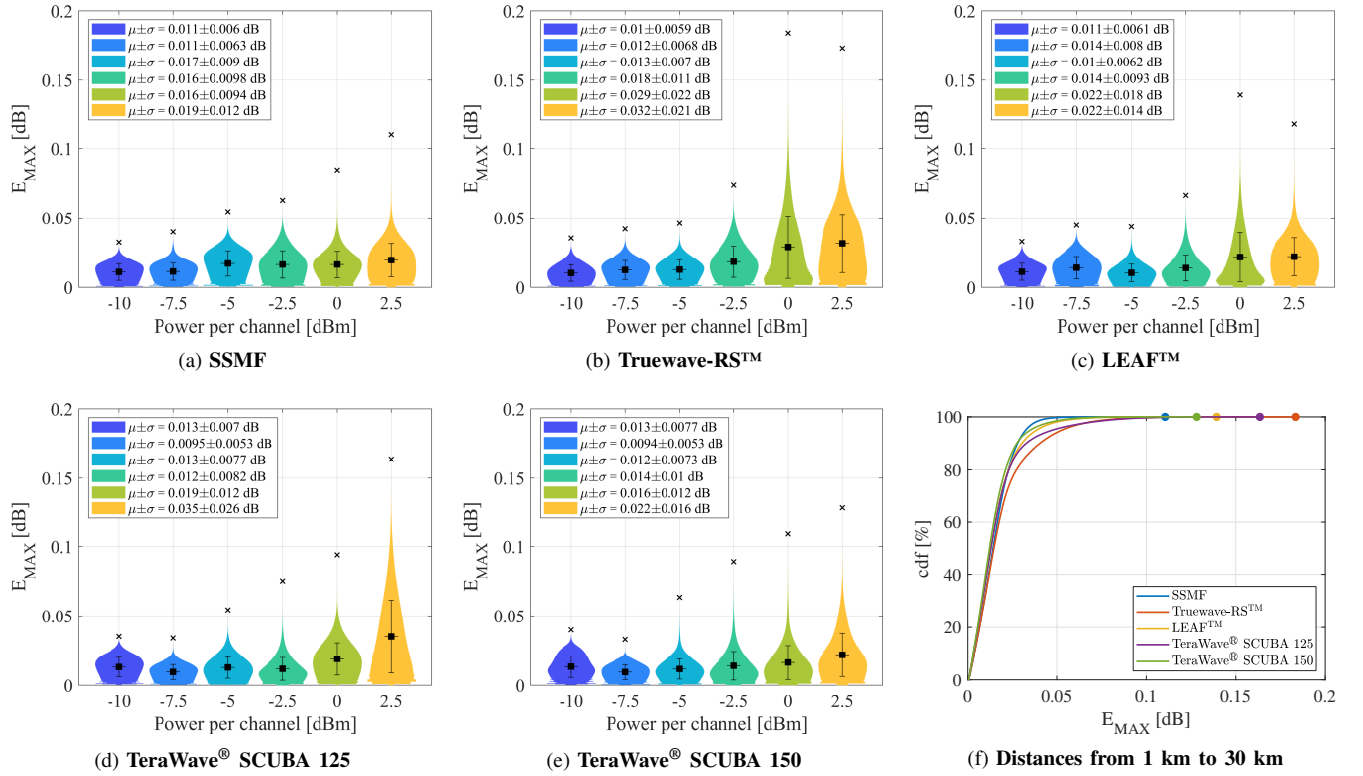


Fig. 5: Violin plots versus power per-channel of the *maximum absolute error* (E_{MAX}) computed between the power profiles predicted using *Model 2* and the actual power profiles for distances *from 1 km to 30 km* in case of: (a) SSMF, (b) TrueWave-RSTM, (c) LEAFTM, (d) TeraWave[®] SCUBA 125 and (e) TeraWave[®] SCUBA 150, at different spectral uniform launch power profiles. For each violin plot the mean value and the standard deviation are also reported, together with the maximum value, plotted with cross marker. (f) Cumulative distribution functions of E_{MAX} for the different fiber types including all power per channel values.

complete power profile evolution from 1 km up to L_{fiber} . However, unlike the ANN Model 1, the prediction performance of the ANN Model 2 is analyzed over two ranges of distances, since the NLI generation is different along the fiber span. In particular, the prediction errors are evaluated from 1 km to L_{fiber} and from 1 km to 30 km, a value close to the effective length of the considered fibers spans, where indeed most of the NLI is generated.

Figs. 4(a)-(e) shows the resulting violin plots of the E_{MAX} for distances from 1 km to L_{fiber} for the five different types of fibers and different power per channel values. Similar observations can be made with respect to the results obtained considering ANN Model 1, but slightly larger errors are observed here. Also in this case, the prediction errors are larger for higher values of power per channel, especially for TrueWave-RSTM and LEAFTM. Indeed, for $P_{ch}=2.5$ dBm, the average E_{MAX} are 0.11 dB and 0.069 dB, respectively for TrueWave-RSTM and LEAFTM, and for some few cases E_{MAX} assumes values beyond 0.5 dB.

Slightly lower errors are obtained when $P_{ch}=0$ dBm, as the mean E_{MAX} reduces to 0.082 dB and 0.063 dB, respectively for TrueWave-RSTM and LEAFTM, and all values are within 0.42 dB. Also the TeraWave[®] SCUBA 125 shows larger errors when $P_{ch}=2.5$ dBm, with average and maximum E_{MAX} values corresponding to 0.077 dB and 0.32 dB, respectively. On

the other hand, highly accurate predictions are achieved for the other fiber types and power per channel values. This is demonstrated by their average E_{MAX} which is always below 0.05 dB.

Like for ANN Model 1, to have an immediate idea of the ANN Model 2 prediction performance, also here we evaluate the cdfs for the different types of fibers including all power per channel values. The results, reported in Fig. 4(f), show that for all fibers, regardless of the value of power per channel, $E_{MAX} \leq 0.2$ dB for more than 97% of cases, confirming again the capability of the ANN model to provide accurate predictions.

However, if we analyze the prediction performance in the first 30 km of the fiber span L_{fiber} , represented by the violin plots in Fig. 5, we can observe that the trained ANN model is able to provide highly accurate predictions also for TrueWave-RSTM and LEAFTM with higher P_{ch} , as all E_{MAX} values are below 0.2 dB and their average never exceeds 0.035 dB. Interestingly, results obtained for SSMF and $P_{ch}=0$ dBm are comparable to those presented in [19], which correspond to the same scenario. This means that, although the ANN Model 2 has been trained considering a more extended data-set, with different fiber types and power levels, its prediction accuracy was not affected. On the contrary, the ANN has been enhanced to handle a more general scenario without losing in accuracy.

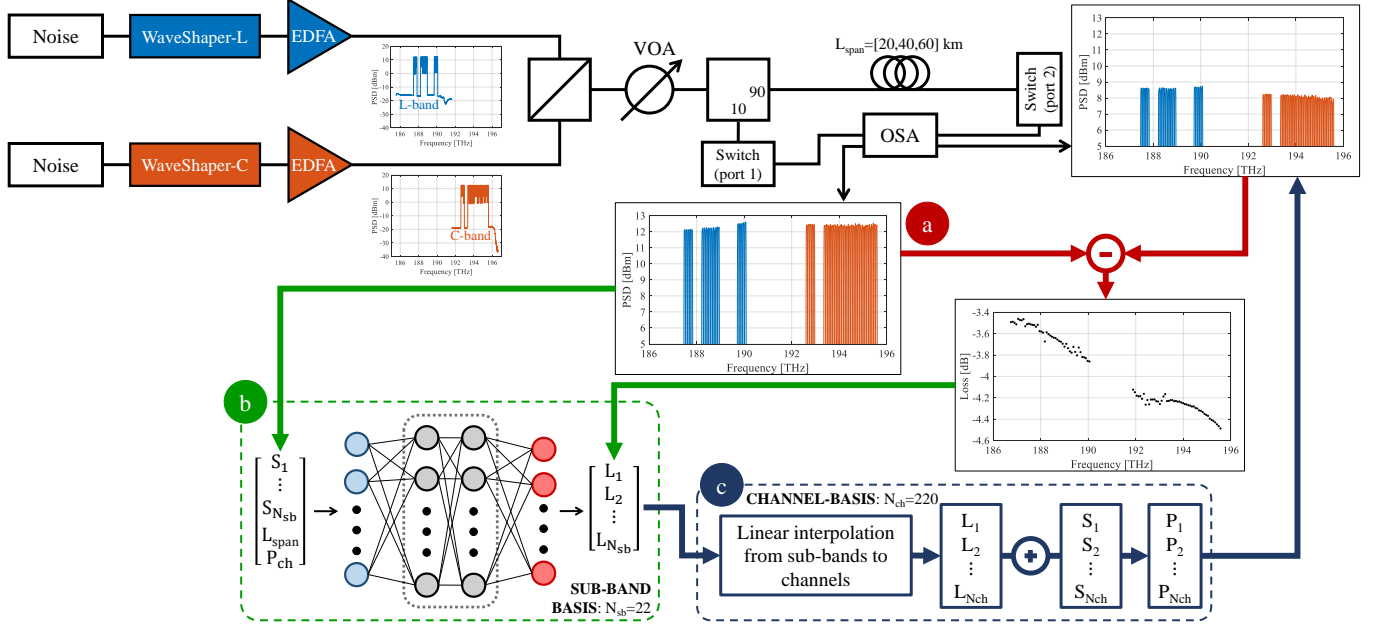


Fig. 6: Paths: (a) Experimental setup used to generate training and testing data-sets for different span length (L_{span}) nominal values. EDFA: erbium doped fiber amplifier, WXC: wavelength cross-connect, VOA: variable optical attenuator, OSA: optical spectrum analyzer; (b) ANN Model 3 considered for the ML-framework; (c) ANN testing process over channels.

In addition, the overall prediction performance of ANN Model 2 in the first 30 km of fiber span is summarized by the cdfs in Fig. 5(f) for the different fiber types analyzing together different values of uniform channel power, where we can see that $E_{MAX} \leq 0.1$ dB for $\sim 100\%$ of cases.

Like for the ANN Model 1, also for the ANN Model 2 we observe a substantial reduction of the computational time required to determine the spatial and spectral evolution of the power profiles. In particular, when using the trained ANN, the determination of α_0 , α_1 and σ for a single power profile is obtained in only 0.24 s. This is a very low value compared to the 3.48 s taken by the approach described in [20], where 3 s were required by the numerical RS and 0.48 s by the fitting.

III. ANN FOR IN-FIELD SYSTEM PERFORMANCE EVALUATION: MODEL 3

Whether in the previous section we extensively analyze ANNs Model 1 and Model 2 and the potentiality of using them in support of linear and nonlinear modeling for QoT estimation, in this section we present the ANN Model 3 for an in-field system performance evaluation. For this purpose the data-sets used for ANN training, validation and testing are generated experimentally.

A. Experimental Setup and Data-sets Generation

The experimental setup considered for the training and testing data-sets generation is shown in Fig. 6, path (a). Like for the synthetic data, also here the system considers a WDM comb over the C+L-band, which consists of $N_{ch} = 95$ channels with 75 GHz frequency spacing and 64 GBd symbol rate. Two different wide-spectrum noise sources are exploited

to mimic 45 channels in the L-band and 50 channels in the C-band. In particular, two programmable optical filters (Finisar Waveshaper) shape the noise signals to generate an equivalent flat WDM spectrum at the output of the following erbium-doped fiber amplifiers (EDFA). Afterwards, C and L-band channels are combined in a single WDM comb whose power is controlled by a variable optical attenuator (VOA) so that the power per channel assumes the following discrete values $P_{ch} = [-10, -7.5, -5, -2.5, 0]$ dBm, where the highest value is 0 dBm due to hardware limitations. An optical spectrum analyzer (OSA) measures the spectrum of the transmitted signal after it goes through a 10% splitter and the first port of a switch. The remaining 90% of the signal propagates over an ITU-T G.652 standard compliant Corning® SMF-28® fiber span and the switch before reaching the OSA again. Finally, the measured spectra are re-scaled to recover losses introduced by the splitter and the switch. For our analysis, three fiber spools with different nominal span lengths $L_{span} = [20, 40, 60]$ km and approximatively the same loss of 0.19 dB/km are considered.

Similarly to Section II, also here training and testing data-sets are generated on *subband-basis* and *channel-basis*, respectively, where each subband is formed by five adjacent channels, for a total of 19 subbands. Thus, channels and subbands are assumed to have 50% probability of being on or off to enable partial spectral loads. In case of subbands the selection of the on/off state is on the subband, therefore when the subband is on, this means that all channels of that subband are on. Similarly when the subband is off, all channels of that subband are off.

Specifically on the data-sets generation, for each value of span length L_{span} and power per channel P_{ch} , we extract

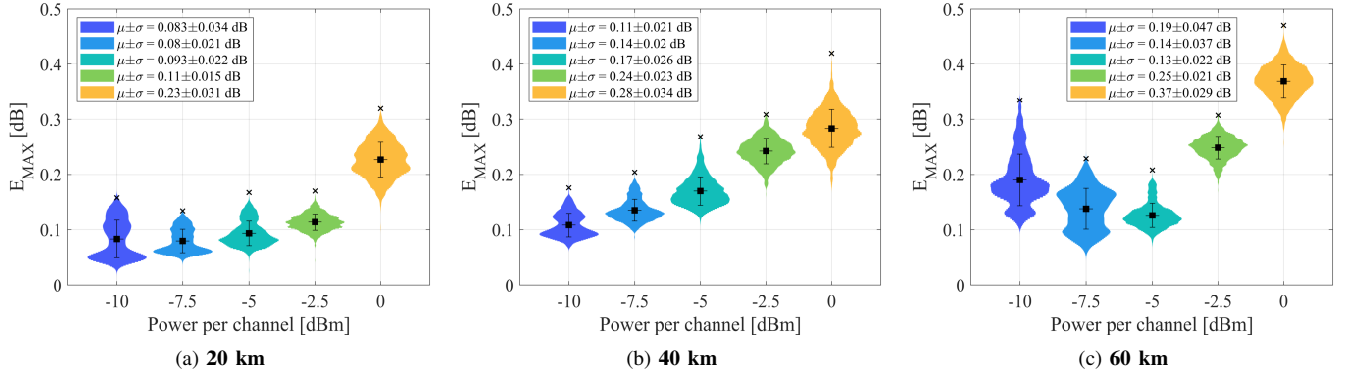


Fig. 7: Violin plots of the *maximum absolute error* (E_{MAX}) computed between predicted and actual power profiles of testing data-set for different power per channel values and for different span lengths: (a) 20 km, (b) 40 km and (c) 60 km. For each violin plot the mean value and the standard deviation are also reported, together with the maximum value, plotted with cross marker.

$M_{sb}=3000$ and $M_{ch}=1500$ different partial spectral loads, for a total of 45000 and 22500 samples, respectively for training and testing data-sets. Then, once we have the power spectral density (PSD) of the received spectrum PSD_{dBm}^{RX} at the fiber output, we compute the corresponding loss profile L , in logarithmic units, as:

$$L = PSD_{dBm}^{TX} - PSD_{dBm}^{RX} \quad (7)$$

where PSD_{dBm}^{TX} is the power spectral density of the input spectral load.

The data have been collected during different measurement sessions, thus, despite the measurement process is repeatable, the experimental setup conditions might slightly change from one session to another one.

B. Machine Learning Framework and Artificial Neural Network Model

The considered ANN model, reported in Fig. 6 path (b), is a feed-forward neural network based on TensorFlowTM and Keras libraries and characterized by 2 HLs and 500 HNs. Also for Model 3, the weights of the ANN nodes are initialized using Glorot initialization with uniform distribution. The *features* at the input are represented by the spectral load information $\mathbf{S} = [S_1, S_2, \dots, S_{N_{sb}}]$, the length of the span L_{span} and the power per channel P_{ch} . The loss profile at the fiber output $\mathbf{L} = [L_1, L_2, \dots, L_{N_{sb}}]$ corresponds to the *labels* at the output of the ANN. The ANN training is based on back-propagation algorithm assuming ReLU activation function, MSE loss function, $\lambda=0.001$ learning rate and it is performed over 1000 epochs considering 90% of the training data-set. The remaining 10% is held out for ANN validation.

It is worth noticing that, like in case of synthetic data, here as well we operate on subbands, which advantageously allows to reduce the complexity of the ANN without significantly affecting the prediction accuracy. Therefore, when we test the trained ANN, also the spectral load profiles of the testing data-set at the input of the neural network are considered with subband granularity. To determine the power profiles at the end of the fiber span, for simplicity referred to as *predicted* power

profiles, a linear interpolation from subbands to channels is performed over the predicted loss profiles, to which we add the original input spectral load at channel granularity (Fig. 6 path (c)).

C. Testing Results

Like in case of ANNs Model 1 and Model 2, the prediction accuracy of the ANN Model 3 is assessed by computing the E_{MAX} between the PSDs of the power profile measured at the end of the span and the profile predicted by the ANN. The results for the three different fiber spools at different power per channel are reported in Fig. 7 in the form of violin plots. The average and the maximum values of E_{MAX} are plotted as single square and cross markers, respectively. Similarly to the results obtained for the synthetic data, in general the errors are smaller for lower values of P_{ch} , with an increasing trend for the mean E_{MAX} . Specifically, it increases from ~ 0.08 dB to 0.23 dB for $L_{span}=20$ km, and from 0.11 dB to 0.28 dB for $L_{span}=40$ km.

This is no longer observed for $L_{span}=60$ km, as shown in Fig. 7(c), where the average E_{MAX} is in the range from 0.13 dB up to 0.37 dB, but the increasing trend from lower to higher power per channel is no longer respected. This might be explained by recalling that the experimental data have been collected during different sessions, thus some samples of training and testing data-sets may be affected by uncertainties affecting losses. Nevertheless, with values of E_{MAX} all lower than 0.5 dB, the trained ANN Model 3 is able to provide predictions with particularly high accuracy, showing great potentialities to be used for in-field system performance evaluation in an ultra-wideband transmission scenario.

IV. CONCLUSIONS

In this paper, three different ANN models have been presented to operate in an UWB system scenario affected by inter-channel stimulated Raman scattering (ISRS).

Two ANN models, Model 1 and Model 2, have been proposed to predict the evolution of power profiles along the fiber span in support of linear and nonlinear modeling of

propagation effects, respectively. In particular, this analysis was performed considering synthetic data over a wide range of system scenarios, including different types of fibers and a set of discrete power per channel values. These assumptions significantly extend the study presented in [19], where only a single type of fiber and a single power per channel value were considered. From the obtained results, with $E_{\text{MAX}} \leq 0.2$ dB for more than 96.6% of cases using Model 1, and $E_{\text{MAX}} \leq 0.2$ dB for 96.9% of cases using Model 2, both ANNs proved to be highly accurate in predicting power profiles at any desired distance for real fibers. Noticeably, in case of ANN Model 1, E_{MAX} is always below 0.1 dB for $P_{\text{ch}} \leq -2.5$ dBm regardless the type of tested fiber, while in case of Model 2 $E_{\text{MAX}} < 0.1$ dB for all fibers and power per channel values when the prediction errors are evaluated in the first 30 km of the fiber span, where indeed most of the NLI generation takes place, as the power is higher.

The third ANN model (Model 3) has been introduced as a proof of principle to determine the power profiles at the end of the fiber span for a fast in-field system performance evaluation based on actual measurements. The study was carried out over experimental data generated considering three fiber spools of SMF-28® with different span lengths and for a set of discrete values of power per channel. Assuming a measurement uncertainty of ~ 0.1 dB, the results showed pretty accurate power profiles predictions for all considered scenarios, with E_{MAX} always below 0.5 dB.

The obtained results demonstrate the feasibility of using machine learning and artificial neural networks to obtain fast and accurate power profile prediction both in support of system modeling for the evaluation of QoT and in in-field applications.

In addition, the ANN-based solution becomes even more beneficial with respect to standard approaches when more than one power profile needs to be evaluated at a time. As a matter of fact, the computational time for the numerical solver and for the fitting process increases linearly with the number of profiles to predict, as they are invoked iteratively on each profile. Instead, the ANN is capable of handling large sets of data, since it operates on matrices, providing the prediction of power profiles and of coefficients α_0 , α_1 and σ in few seconds for thousands of different spectral loads.

REFERENCES

- [1] P. J. Winzer, D. T. Neilson, and A. R. Chraplyvy, "Fiber-optic transmission and networking: the previous 20 and the next 20 years [Invited]," *Opt. Express*, vol. 26, no. 18, pp. 24190–24239, 9 2018. [Online]. Available: <http://opg.optica.org/oe/abstract.cfm?URI=oe-26-18-24190>
- [2] Cisco, "Cisco Annual Internet Report," <https://www.cisco.com/c/en/us/solutions/executive-perspectives/annual-internet-report/index.html>, accessed: 2021-10-29.
- [3] P. J. Winzer and D. T. Neilson, "From Scaling Disparities to Integrated Parallelism: A Decathlon for a Decade," *Journal of Lightwave Technology*, vol. 35, no. 5, pp. 1099–1115, 2017.
- [4] A. Napoli, N. Costa, J. K. Fischer, J. ao Pedro, S. Abrate, N. Calabretta, W. Forysiak, E. Pincemin, J. P.-P. Gimenez, C. Matrakidis, G. Roelkens, and V. Curri, "Towards multiband optical systems," in *Advanced Photonics 2018 (BGPP, IPR, NP, NOMA, Sensors, Networks, SPPCom, SOF)*. Optical Society of America, 2018, p. NeTu3E.1. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=Networks-2018-NeTu3E.1>
- [5] R. Ryf and C. Antonelli, *Springer Handbook of Optical Networks*. Cham: Springer International Publishing, 2020, ch. 10.
- [6] W. S. Pelouch, "Raman Amplification: An Enabling Technology for Long-Haul Coherent Transmission Systems," *Journal of Lightwave Technology*, vol. 34, no. 1, pp. 6–19, 2016.
- [7] M. Secondini and E. Forestieri, "Analytical Fiber-Optic Channel Model in the Presence of Cross-Phase Modulation," *IEEE Photonics Technology Letters*, vol. 24, no. 22, pp. 2016–2019, 2012.
- [8] P. Johannisson and M. Karlsson, "Perturbation Analysis of Nonlinear Propagation in a Strongly Dispersive Optical Communication System," *Journal of Lightwave Technology*, vol. 31, no. 8, pp. 1273–1282, 2013.
- [9] R. Dar, M. Feder, A. Mecozzi, and M. Shtaf, "Properties of nonlinear noise in long, dispersion-uncompensated fiber links," *Opt. Express*, vol. 21, no. 22, pp. 25 685–25 699, Nov 2013.
- [10] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "The GN-Model of Fiber Non-Linear Propagation and its Applications," *Journal of Lightwave Technology*, vol. 32, no. 4, pp. 694–721, 2014.
- [11] A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, "EGN model of non-linear fiber propagation," *Opt. Express*, vol. 22, no. 13, pp. 16 335–16 362, Jun 2014. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-22-13-16335>
- [12] P. Serena and A. Bononi, "A Time-Domain Extended Gaussian Noise Model," *Journal of Lightwave Technology*, vol. 33, no. 7, pp. 1459–1472, 2015.
- [13] M. Cantono, D. Pileri, A. Ferrari, C. Catanese, J. Thouras, J.-L. Augé, and V. Curri, "On the Interplay of Nonlinear Interference Generation With Stimulated Raman Scattering for QoT Estimation," *Journal of Lightwave Technology*, vol. 36, no. 15, pp. 3131–3141, 2018.
- [14] D. Semrau, E. Sillekens, R. I. Killey, and P. Bayvel, "A Modulation Format Correction Formula for the Gaussian Noise Model in the Presence of Inter-Channel Stimulated Raman Scattering," *Journal of Lightwave Technology*, vol. 37, no. 19, pp. 5122–5131, 2019.
- [15] D. N. Christodoulides and R. B. Jander, "Evolution of stimulated Raman crosstalk in wavelength division multiplexed systems," *IEEE Photonics Technology Letters*, vol. 8, no. 12, pp. 1722–1724, 1996.
- [16] M. Ranjbar Zefreh, F. Forghieri, S. Piciaccia, and P. Poggiolini, "Accurate Closed-Form Real-Time EGN Model Formula Leveraging Machine-Learning Over 8500 Thoroughly Randomized Full C-Band Systems," *Journal of Lightwave Technology*, vol. 38, no. 18, pp. 4987–4999, 2020.
- [17] D. Semrau, R. I. Killey, and P. Bayvel, "A Closed-Form Approximation of the Gaussian Noise Model in the Presence of Inter-Channel Stimulated Raman Scattering," *Journal of Lightwave Technology*, vol. 37, no. 9, pp. 1924–1936, 2019.
- [18] P. Poggiolini and M. Ranjbar-Zefreh, "Closed Form Expressions of the Nonlinear Interference for UWB Systems," in *European Conference on Optical Communication (ECOC) 2022*. Optica Publishing Group, 2022, p. Tu1D.1. [Online]. Available: <https://opg.optica.org/abstract.cfm?URI=ECOC-2022-Tu1D.1>
- [19] A. M. R. Brusin, M. R. Zefreh, P. Poggiolini, S. Piciaccia, F. Forghieri, and A. Carena, "Machine Learning for Power Profiles Prediction in Presence of Inter-channel Stimulated Raman Scattering," in *2021 European Conference on Optical Communication (ECOC)*, 2021, pp. 1–4.
- [20] M. Ranjbar Zefreh and P. Poggiolini, "A Real-Time Closed-Form Model for Nonlinearity Modeling in Ultra-Wide-Band Optical Fiber Links Accounting for Inter-channel Stimulated Raman Scattering and Co-Propagating Raman Amplification," <https://arxiv.org/abs/2006.03088>, 2020.
- [21] A. M. Rosa Brusin, U. C. de Moura, V. Curri, D. Zibar, and A. Carena, "Introducing Load Aware Neural Networks for Accurate Predictions of Raman Amplifiers," *Journal of Lightwave Technology*, vol. 38, no. 23, pp. 6481–6491, 2020.
- [22] G. Marcon, A. Galtarossa, L. Palmieri, and M. Santagiustina, "Model-Aware Deep Learning Method for Raman Amplification in Few-Mode Fibers," *Journal of Lightwave Technology*, vol. 39, no. 5, pp. 1371–1380, 2021.
- [23] X. Ye, A. Arnould, A. Ghazisaeidi, D. Le Gac, and J. Renaudier, "Experimental Prediction and Design of Ultra-Wideband Raman Amplifiers using Neural Networks," in *2020 Optical Fiber Communications Conference and Exhibition (OFC)*, 2020, pp. 1–3.
- [24] U. C. de Moura, D. Zibar, A. Margareth Rosa Brusin, A. Carena, and F. Da Ros, "Fiber-Agnostic Machine Learning-Based Raman Amplifier Models," *Journal of Lightwave Technology*, vol. 41, no. 1, pp. 83–95, 2023.
- [25] H. Buglia, M. Jarmolovičius, A. Vasylenkova, E. Sillekens, L. Galdino, R. I. Killey, and P. Bayvel, "A Closed-Form Expression for the Gaussian Noise Model in the Presence of Inter-Channel Stimulated

- Raman Scattering Extended for Arbitrary Loss and Fibre Length,” *Journal of Lightwave Technology*, vol. 41, no. 11, pp. 3577–3586, 2023.
- [26] M. R. Zefreh, F. Forghieri, S. Piciaccia, and P. Poggiolini, “A Closed-Form Nonlinearity Model for Forward-Raman-Amplified WDM Optical Links,” in *Optical Fiber Communication Conference (OFC) 2021*. Optica Publishing Group, 2021, p. M5C.1. [Online]. Available: <https://opg.optica.org/abstract.cfm?URI=OFC-2021-M5C.1>
 - [27] H. Buglia, M. Jarmolovicius, L. Galdino, R. I. Killey, and P. Bayvel, “A Closed-form Expression for the Gaussian Noise Model in the Presence of Raman Amplification,” <https://arxiv.org/abs/2304.11974>, 2023.
 - [28] S. Tariq and J. Palais, “A computer model of non-dispersion-limited stimulated Raman scattering in optical fiber multiple-channel communications,” *Journal of Lightwave Technology*, vol. 11, no. 12, pp. 1914–1924, 1993.
 - [29] “GitHub repository of GNPpy,” DOI: 10.5281/zenodo.3458320, <https://github.com/Telecominfraproject/oopt-gnpy>.
 - [30] E. Pincemin, D. Grot, L. Bathany, S. Gosselin, M. Joindot, S. Bordaïs, Y. Jaouen, and J.-M. Delavaux, “Raman gain efficiencies of modern terrestrial transmission fibers in S-, C- and L-band,” in *Nonlinear Guided Waves and Their Applications*. Optical Society of America, 2002, p. NLTuC2. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=NLGW-2002-NLTuC2>
 - [31] OFS, “TeraWave® SCUBA 125 Ocean Optical Fiber,” <https://stage.fiber-optic-catalog.ofsoptics.com/documents/pdf/TeraWave-SCUBA-125-Single-mode-Fiber-fap-170-web.pdf>, 03 2018.
 - [32] —, “TeraWave® SCUBA 150 Ocean Optical Fiber,” <https://stage.fiber-optic-catalog.ofsoptics.com/documents/pdf/TeraWave-Scuba-150-Ocean-Fibers-fiber-168-web.pdf>, 03 2018.
 - [33] P. Poggiolini, “A generalized GN-model closed-form formula,” <https://arxiv.org/abs/1810.06545>, 2018.