Application of Machine Learning in Fiber Nonlinearity Modeling and Monitoring for Elastic Optical Networks

Qunbi Zhuge, Xiaobo Zeng, Huazhi Lun, Meng Cai, Xiaomin Liu, Lilin Yi, and Weisheng Hu

***Abstract*—Fiber nonlinear interference (NLI) modeling and monitoring are the key building blocks to support elastic optical networks. In the past, they were normally developed and investigated separately. Moreover, the accuracy of the previously proposed methods still needs to be improved for heterogenous dynamic optical networks. In this paper, we present the application of machine learning (ML) in NLI modeling and monitoring. In particular, we first propose to use ML approaches to calibrate the errors of current fiber nonlinearity models. The Gaussian-noise model is used as an illustrative example, and significant improvement is demonstrated with the aid of an artificial neural network. Further, we propose to use ML to combine the modeling and monitoring schemes for a better estimation of NLI variance. Extensive simulations with 2411 links are conducted to evaluate and analyze the performance of various schemes, and the superior performance of the ML-aided combination of modeling and monitoring is demonstrated.**

***Index Terms*—Coherent optical communication, elastic optical network, fiber nonlinearity, machine learning, optical performance monitoring.**

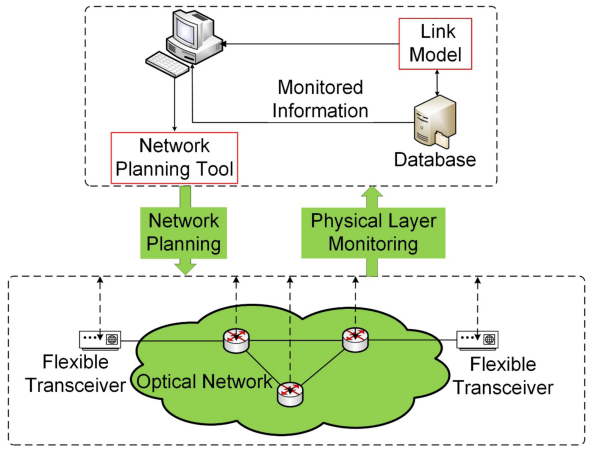

Fig. 1. Optical network architecture of the physical layer.

I. Introduction

NETWORK capacity demand is ever-increasing due to emerging internet applications such as high-definition video streaming, cloud, 5G, internet of things, virtual/augmented reality, and so forth. However, serving as the underlying infrastructure of global communication networks, the capacity of current optical networks based on single mode fiber is approaching the theoretical limit [1]. Commercial products with constellation shaping and strong forward error correction (FEC) techniques are being built to close the gap to the Shannon limit [2]. In the research community, space-division multiplexing (SDM) has been actively investigated to scale fiber channel capacity [3], but its commercial deployment for long-haul optical networks is infeasible in the near future. In the meantime, with a higher commercial viability in the short term, elastic optical networks (EONs) have been proposed to intelligently and efficiently utilize physical layer resources in order to increase network capacity and enable dynamic services [4].

In EONs, channel configurations including modulation format, symbol rate, spacing, and launch power can be remarkably diverse, and they need to be fully optimized in order to extract more capacity from the physical layer infrastructure [5], [6]. As illustrated in the EON architecture schematic in Fig. 1, it is essential to have a planning tool capable of providing accurate prediction of link performance (or quality of transmission), and the required computational time of it can be at the order of milliseconds. A key element of such a planning tool is the estimator of fiber nonlinear interference (NLI) accumulated over signal propagation. On the one hand, modeling of fiber NLI has been widely and deeply investigated. Starting from the nonlinear Schrödinger (NLS) equation, the split-step Fourier method (SSFM) was developed to obtain a highly accurate numerical solution [7]. However, the computation is too intensive for its use in a real-time planning tool. Later, after coherent transmission systems became the mainstream for backbone optical networks, the Gaussian-noise (GN) model was proposed to calculate nonlinear signal-to-noise ratio (SNR) [8] (denoted as

Manuscript received January 29, 2019; revised April 1, 2019; accepted April 6, 2019. Date of publication April 11, 2019; date of current version May 24, 2019. This work was supported by the National Natural Science Foundation of China under Grant 61801291. *(Qunbi Zhuge and Xiaobo Zeng contributed equally to this work.) (Corresponding author: Qunbi Zhuge.)*

The authors are with the State Key Laboratory of Advanced Optical Communication Systems and Networks, Shanghai Institute for Advanced Communication and Data Science, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: qunbi.zhuge@sjtu.edu.cn; xiaobozeng@sjtu.edu.cn; nigulasikaochuan@sjtu.edu.cn; caimeng0922@sjtu.edu.cn; liuxiaomin1997@sjtu.edu.cn; lilinyi@sjtu.edu.cn; wshu@sjtu.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JLT.2019.2910143

SNR_{nl}), which is defined as the ratio of signal power to nonlinear noise power on received symbols after the nonlinear propagation over fiber. The simplicity of the GN model (and its variants) makes it a promising candidate for network planning, even though its accuracy still needs to be improved for heterogenous dynamic links [9]. Another shortcoming of the model-based performance prediction is that the uncertainty of link parameters, e.g., inaccurate measurement of optical power, might lead to further deviations.

On the other hand, real-time optical performance monitoring (OPM) is a key building block for future software defined elastic optical networks (EON), which is also depicted in Fig. 1 [10]. For example, the monitored link performance such as SNR can be used to iteratively learn link parameters for accurate quality of transmission (QoT) estimation [11]–[13]. In [13], it further demonstrates a supervised machine learning (ML) QoT estimation model which is trained based on the monitored data. Therefore, it is essential to monitor various useful information in the transmission system. In terms of fiber nonlinearity monitoring, a few techniques have been proposed to report nonlinear noise variance based on receiver signal processing. For example, methods aided by ML were proposed in [14], [15] to isolate nonlinear noise variance in the presence of linear noise such as amplified spontaneous emission (ASE).

In this paper, we present the application of ML techniques in the NLI modeling and monitoring for EONs. In particular, we first propose the use of ML to perform the calibration of NLI models to reduce deviations in heterogenous dynamic optical networks. Afterwards, the ML is further applied to combine the NLI modeling and monitoring. To investigate and demonstrate their performances, 2411 links with diverse configurations and link parameters are simulated and an artificial neural network (ANN) is adopted in these schemes. With the proposed ML-aided combination of NLI modeling and monitoring, we show that the estimation accuracy is significantly improved. Finally, the sensitivity of these schemes against the uncertainties in launch powers is evaluated.

The remainder of the paper is organized as follows. In Section II, we briefly introduce the GN models and the noise covariance-based nonlinearity monitoring, which are used in the following studies. In Section III, the application of ML in the NLI modeling is first described, followed by the principle of applying ML to combine the NLI modeling and monitoring. Section IV presents the simulation setup, results and analysis, showing a significantly improved accuracy of the proposed approach. The paper is finally concluded in Section V.

II. Technical Background

In modern coherent systems, the Kerr nonlinearity effects of optical fiber have become the major limiting factors to further improve the capacity of wavelength-division multiplexing (WDM) systems [16]. The nonlinear propagation of optical signals in a single mode fiber can be modeled by the well-known NLS equation [7]. Considering the averaged random evolution of polarization effects over long distance fiber, the Manakov equation was derived to describe dual-polarization (DP) signal propagations as [17]

$$\frac{\partial \boldsymbol{E}}{\partial z} = j\frac{8}{9}\gamma\left|\boldsymbol{E}^2\right|\boldsymbol{E} - j\frac{\beta_2}{2}\frac{\partial^2 \boldsymbol{E}}{\partial t^2} - \alpha\boldsymbol{E} \tag{1}$$

where $\boldsymbol{E} = [\boldsymbol{E}_x,\ \boldsymbol{E}_y]^T$ is the optical field of the DP signal. γ is the fiber nonlinear coefficient, β_2 is the second order dispersion coefficient, and α is the fiber loss of the optical field. Note that higher-order dispersion coefficients are neglected in Eq. (1).

In general, Eq. (1) does not have analytic solutions. Alternatively, it can be numerically solved by the SSFM, which has been widely used in simulating long-haul optical transmissions. The SSFM approach requires a large number of fast Fourier transform (FFT) and inverse FFT (IFFT), making it extremely computation-intensive for full-filled WDM simulations. Therefore, it is not suitable for network planning applications.

A. GN Model

The GN model and its variants were proposed and extensively investigated in recent years to estimate the accumulated nonlinear noise variance for coherent optical transmission systems [8], [18]. The GN models are derived based on Eq. (1) with a perturbation assumption. In addition, the nonlinear noise is considered as additive Gaussian noise. According to the GN model, the power spectral density (PSD) of the NLI per polarization at frequency *f* after transmission can be expressed as [8]

$$\begin{aligned} &G_{NLI}(f) = \\ &\frac{16}{27}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} G_{WDM}(f_1)\, G_{WDM}(f_2) \\ &\quad \cdot G_{WDM}(f_1+f_2-f)\cdot \\ &\left|\sum_{n=1}^{N_s}\gamma_n \cdot \left[\prod_{k=1}^{n-1} exp\left(-3\alpha_k L_k\right)\cdot \Gamma_k^{3/2}\right]\right. \\ &\quad \left.\cdot \left[\prod_{k=n}^{N_s} exp\left(-\alpha_k L_k\right)\cdot \Gamma_k^{1/2}\right]\cdot\right. \\ &exp\left(j4\pi^2\left(f_1-f\right)\left(f_2-f\right)\cdot\sum_{k=1}^{n-1}\beta_{2,k}L_k\right) \\ &\quad \cdot \int_0^{L_n}\left[exp\left(-2\alpha_n z\right)\cdot\right. \\ &\left.exp\left(j4\pi^2\left(f_1-f\right)\left(f_2-f\right)\cdot\beta_{2,n}z\right]dz\right|^2 df_1 df_2 \end{aligned} \tag{2}$$

where Γ_k is the lumped power-gain placed at the end of the *k*-th span, L_k is the length of the *k*-th span, α_k is the field loss of the *k*-th span, and N_s is the total number of spans. In this study, we only consider dispersion-uncompensated systems with "link transparency", which means that the loss of each span is exactly compensated by the optical amplifier at the end of each span. Note that the transmitted signals are polarization-division multiplexed, and each span only consists of a single fiber type.

Based on Eq. (2), we can obtain the variance of the NLI coherently accumulated over different spans. This type of the GN

model is referred to as "coherent GN (CGN) model". Another type of the GN model neglects the coherence among the NLI generated in different spans, and it is referred to as "incoherent GN (IGN) model". In the IGN model, the total NLI PSD can be obtained by simply summing up the PSD from each span:

$$G_{IGN}(f) = \sum_{n=1}^{N_s} G_{NLI,n}^{1span}(f) \tag{3}$$

where $G_{NLI}^{1span}(f)$ is the NLI PSD based on Eq. (2) over a single span, which can be expressed as

$$\begin{aligned} G_{NLI}^{1span}(f) = \frac{16}{27}\gamma^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} & G_{WDM}(f_1) G_{WDM} \\ & \times (f_2) G_{WDM}(f_1+f_2-f) \\ & \cdot \left| \frac{1-exp(-2\alpha L)\cdot exp\left(j4\pi^2\beta_2 L(f_1-f)(f_2-f)\right)}{2\alpha - j4\pi^2\beta_2(f_1-f)(f_2-f)} \right| \\ & \times df_1 df_2 \end{aligned} \tag{4}$$

Although the IGN model is an approximation of the CGN model, it is reported that the results of the IGN model are also relatively accurate, especially when there exists a large number of channels [8]. In these scenarios, the IGN model is preferable because of its simple analytical form and low computational complexity. For a small number of channels, however, the IGN model may be inaccurate [8]. In addition, when the WDM configurations and link conditions are highly heterogenous and dynamic, the accuracy of the GN models are found to be significantly degraded [9].

B. Fiber Nonlinearity Monitoring

Coherent optical receiver provides a resourceful platform through digital signal processing (DSP) to build OPM functionalities [19]. For instance, polarization mode dispersion can be monitored based on the coefficients of the adaptive butterfly filter. Meanwhile, ML-aided OPM approaches have been widely investigated. For direct detection systems, ML-based OSNR monitoring schemes based on delay-tap asynchronous sampling [20] and single channel sampling [21] were proposed. For coherent systems, deep artificial neural network (ANN) was adopted to perform OSNR monitoring and format identification simultaneously [22].

Regarding the fiber nonlinearity monitoring, a few techniques have been proposed. In [23], the nonlinearity-induced amplitude noise correlation across neighboring symbols was characterized and employed to aid optical signal-to-noise ratio (OSNR) monitoring in the presence of strong fiber nonlinearities. The fiber nonlinear effects also cause phase noise correlations over neighboring symbols [24]. In [14], ML-based nonlinear noise monitoring using amplitude noise covariance (ANC) and phase noise covariance (PNC) was proposed as illustrated in Fig. 2, and a significant improvement in monitoring accuracy was numerically demonstrated and experimentally verified. Note that the Machine Learning Engine in Fig. 2 is an ML based scheme/algorithm that can take the nonlinearity-related features as the input to obtain the desired output, i.e., SNR_{nl} in this work. The dimensions of the inputs depend on the specific implementation.

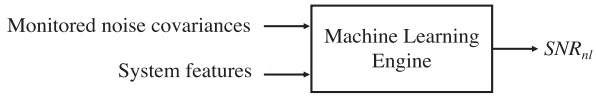

Fig. 2. Block diagram of ML-aided fiber nonlinearity monitoring.

The ANC across neighboring symbols can be calculated as

$$\begin{aligned} ANC_{ij}(m) &= cov\left(\Delta A_i(k), \Delta A_j(k-m)\right) \quad i,j \in \{x,y\} \\ &= \frac{1}{N-1}\sum_{k=1}^{N}\left(\Delta A_i(k) - \frac{1}{N}\sum_{k=1}^{N}\Delta A_i(k)\right)^* \\ &\quad \cdot \left(\Delta A_j(k-m) - \frac{1}{N}\sum_{k=1}^{N}\Delta A_j(k-m)\right) \end{aligned} \tag{5}$$

where $\Delta A_{i/j}(k)$ represents the amplitude noise of the *k*-th received symbol for the *x*- and *y*-polarization expressed as

$$\Delta A_{i/j}(k) = \left|A_{i/j}^R(k)\right| - \left|A_{i/j}(k)\right| \quad i,j \in \{x,y\} \tag{6}$$

where $A_{i/j}^R(k)$ is the received symbol and $A_{i/j}(k)$ is the transmitted symbol for the *x*- and *y*-polarization. The subscript *i/j* means either *i* or *j*.

The PNC across neighboring symbols can be calculated as

$$\begin{aligned} PNC_{i\,j}(n) &= corr\left(\theta_i(k), \theta_j(k-n)\right) \\ &= \frac{\sum_{k=1}^{N}\theta_i(k)\cdot\theta_j(k-n)}{N\cdot(N-n)} \quad i,j \in \{x,y\} \end{aligned} \tag{7}$$

where *N* is the total number of symbols involved in the calculation, and $\theta_{i/j}(k)$ is the phase noise of the *k*-th received symbol for the *x*- and *y*-polarization, which can be expressed as

$$\theta_{i/j}(k) = arg\left(\frac{A_{i/j}^*(k)\cdot A_{i/j}^R(k)}{\left|A_{i/j}^*(k)\cdot A_{i/j}^R(k)\right|}\right) \quad i,j \in \{x,y\} \tag{8}$$

where A^* denotes the complex conjugation of transmitted symbol *A*.

Following [14], the input features to the ML algorithm include R_{ij} and P_{ij}, which are obtained by

$$R_{ij} = 10\cdot log_{10}\left(1\Big/\sum_{m=1}^{6} ANC_{ij}(m)\right) \tag{9}$$

$$P_{ij} = 10\cdot log_{10}\left(1\Big/\sum_{n=1}^{30} PNC_{ij}(n)\right) \tag{10}$$

In addition, $ANC_{xy}(0)$ and $PNC_{xy}(0)$ are also used as the input.

Note that Eq. (5), (7), (9) and (10) cover the four polarization combination cases: (*x*, *x*), (*x*, *y*), (*y*, *x*), and (*y*, *y*).

Reasonable accuracy of the ML-based fiber nonlinearity monitoring has been demonstrated for the system and link cases in [14]. However, due to the limited information in the coherent

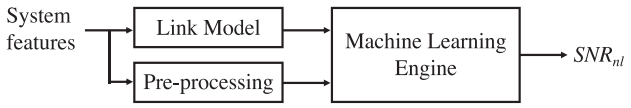

Fig. 3. Block diagram of ML-aided fiber nonlinearity modeling.

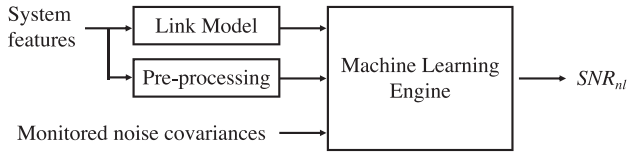

Fig. 4. Block diagram of ML-aided combination of fiber nonlinearity modeling and monitoring.

receiver, it is quite difficult to guarantee the accuracy in more complicated link conditions.

In this paper, we propose a novel framework leveraging ML to combine the NLI modeling and monitoring, and it will be detailed in the next section.

III. Principle

A. ML-Aided Fiber Nonlinearity Modeling

As mentioned earlier, fast analytical models such as the GN model and its variants provide relatively accurate results for certain links. However, in practical systems with more complex link conditions, their accuracy still needs to be improved. On the other hand, the deviations with respect to the true SNR_{nl} are found to be related to system configurations and parameters, which implies that it is possible to calibrate the deviations using system parameters [9]. Nevertheless, the deviations are generally caused by high order nonlinearity, and thus could be very difficult, if not impossible, to obtain through an analytical approach. Meanwhile, ML-based approaches have shown the capacity to deliver exceptional performance in scenarios where the underlying physics and mathematics of the problem are too difficult to be described explicitly, and the numerical procedures involved require computational time/resources [25]. Therefore, we propose to use the ML algorithm with system parameters as the input features to perform the calibration function as depicted in Fig. 3. Training the ML algorithm is necessary and critical. Fortunately, large size data sets can be generated using SSFM simulations, whose computation time is reasonable for the offline training process especially when graphics processing unit (GPU) is adopted.

The GN model and ANN are used as an illustrative example in this paper, and the same framework can be extended to other NLI models and ML algorithms. All system parameters are fed into the GN model including the WDM configurations and the parameters of each fiber span. These parameters are pre-processed to obtain the input features of the ANN, in order to avoid overfitting problems. Specifically, the input features include span number, maximum span length, average span length, launch power, link length, net chromatic dispersion, average gamma and average alpha of the fiber spans, and number of WDM channels. Note that the set of input features is similar to the one used in [26] to predict residual margin. The output of the ANN is SNR_{nl}. In the training process, the SSFM results are used as the desired output.

B. ML-Aided Combination of Modeling and Monitoring

So far, fiber nonlinearity modeling and monitoring are developed separately, which is apparently not the optimal solution, considering that both of them still suffer from inaccuracies in practical systems to some extent. First, for link models, the input parameters might contain errors caused by incorrect link information such as fiber type, which can happen due to historical reasons, and inaccurate measurement such as optical power. Moreover, some information related to the received waveforms, such as ANC and PNC, might not be available in the models. Then, for link monitoring, the nonlinear noise estimation purely relying on the information and processing in receiver has limited accuracy, and its measurement can be slow in noisy environment.

The fusion of the link modeling and monitoring schemes is expected to provide more accurate results. However, in terms of fiber nonlinearities, it is not straightforward to develop an analytical solution with desired performance to achieve this target. In this work, we propose a ML-aided fusion of the fiber nonlinearity modeling and monitoring to provide more consistent and accurate results for optical network planning. The conceptual block diagram is illustrated in Fig. 4. A single ML engine can be used to combine the input features of the ML-aided monitoring in Section II-B and the ML-aided modeling in Section III-A.

The ML-aided scheme can be used in any modeling and monitoring methods, and in the following study we employ the GN model, noise covariance-based monitoring, and a feedforward ANN.

IV. Simulation Results and Discussions

A. Simulation Setup

The performance evaluations of the ML-based combination of fiber nonlinearity modeling and monitoring were numerically carried out over a wide range of system configurations.

The simulation setup is depicted in Fig. 5. At the transmitter (Tx) side, root-raised-cosine (RRC) pulse shaping with a roll-off factor of 0.02 was applied. To emulate the flex-grid environment in EONs, three different symbol rates (R_s) of 35, 70 and 90 Gbaud were evaluated, located in 50, 75, and 100 GHz channel spacing (Δf_{ch}), respectively. The length of the symbol sequence was 2^{16}. For the fiber link, lumped Erbium-doped fiber amplifiers (EDFA) were employed but the ASE noise was ignored. At the receiver (Rx) side, the center channel was filtered out and processed. Particularly, after chromatic dispersion (CD) compensation, a matched filter was applied, followed by down-sampling to obtain received symbols. Since the Tx laser and local oscillator phase noise was ignored, carrier recovery was not included except that the overall phase rotation caused by fiber nonlinearities was removed to align the received symbols with the transmitted symbols. Afterwards, the ANC and PNC calculations were conducted based on the received symbols. The SNR of the received symbols were considered as the

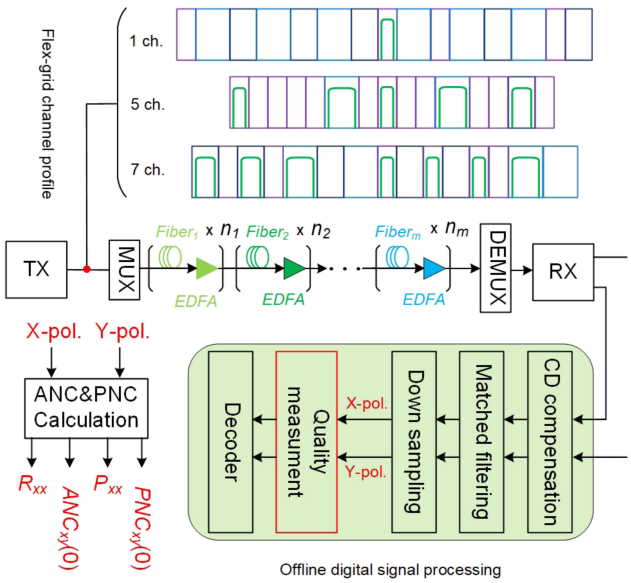

Fig. 5. Simulation setup.

TABLE I
SUMMARY OF LINK CONFIGURATIONS AND PARAMETERS

Fiber type	SSMF, TWC, ELEAF, PSCF
Span length (km)	10, 20, 30, 40, 50, 60, 70, 80, 90, 100
Span number	1:1:20
Launch power (dBm)	-4, -3, -2, -1, 0, 1, 2, 3, 4
Modulation format	DP-QPSK, DP-16QAM
Total WDM slots	15
Channel number	1:1:15
Symbol rate (Gbaud)	35, 70, 90

true SNR_{nl} since the system only contains fiber nonlinear noise. The interactions of ASE noise and laser phase noise with fiber nonlinearities are typically small, and the study of their impacts is left for future works.

In total, 2411 links were simulated to perform the training and test of the ML-based schemes. The selection of the link configurations based on Table I is described as follows. First of all, all the configurations in Table I were randomly selected with a uniform distribution. The fiber types included standard single mode fiber (SSMF), TrueWave classic (TWC), enhanced large effective area fiber (ELEAF), and pure silica core fiber (PSCF). They represent typical fiber types deployed in the field with a wide variety of dispersion, attenuation and nonlinearity coefficient. The parameters of these fiber types are summarized in Table II. The span length was between 10 and 100 km with a step size of 10 km. The span number was between 1 and 20 with a step size of 1. Note that the span length and fiber type were selected independently for each span in each link, which means a single link could contain multiple fiber types and span lengths. The launch power of the signals was between -4 and 4 dBm with a step size of 1 dB. DP-QPSK and DP-16QAM were both considered. The total number of WDM channel slots was set to 15. The center channel was fixed to be 35 Gbaud and 50 GHz spacing. For the rest 14 channels, we first chose a channel number that was equal to or less than 14. Then this number of channels were randomly distributed in these slots. Afterwards, the symbol rate/channel spacing of each channel slot was independently selected from the three options. In Fig. 5 we depict three flex-grid configurations for illustrative purposes.

TABLE II
SUMMARY OF FIBER PARAMETERS

Fiber type	α [dB/km]	D [ps/nm/km]	γ [1/w/km]
SSMF	0.2	16.7	1.3
TWC	0.21	2.8	2
ELEAF	0.21	4.3	1.47
PSCF	0.18	20.1	0.9

TABLE III
INPUT FEATURES TO THE ANN

Features of modeling	1. SNR_{nl} from GN model 2. Span number 3. Maximum span length 4. Average span length 5. Launch power 6. Link length 7. Net chromatic dispersion 8. Average gamma of fiber spans 9. Average alpha of fiber spans 10. Number of WDM channels
Features of monitoring	11. $ANC_{xy}(0)$ from Eq. (5) 12. $PNC_{xy}(0)$ from Eq. (7) 13. R_{ij} from Eq. (9) 14. P_{ij} from Eq. (10)

As mentioned earlier, the waveform propagation simulations were conducted based on the SSFM to obtain received symbols and the true SNR_{nl}. The step size of the SSFM was 10 m. The adopted ANN has only one hidden layer with 15 neurons, which was carefully chosen to obtain the lowest mean square error (MSE) without introducing overfitting. The input features are described in the previous sections and summarized in Table III for clarity.

1688 links were randomly selected from the total 2411 links to train the ANN and the rest 723 links were used to test the trained ANN. The training was performed using MATLAB R2018a's neural network toolbox. The log-sigmod function $(y(n) = \frac{1}{1+e^{-n}})$ and linear function $(y(n) = n)$ were used as the activation function (or transfer function) for the hidden and output layers, respectively. In the training phase, the Bayesian regularization learning algorithm was employed, which updates the weight and bias values to minimize the linear combination of squared errors and weights for better generalization properties.

B. Results of GN Model

Fig. 6(a) and (b) plot the histogram of the SNR_{nl} deviation for the IGN and CGN models, respectively. We can see that in the simulated test cases the range of the offset is quite large, which

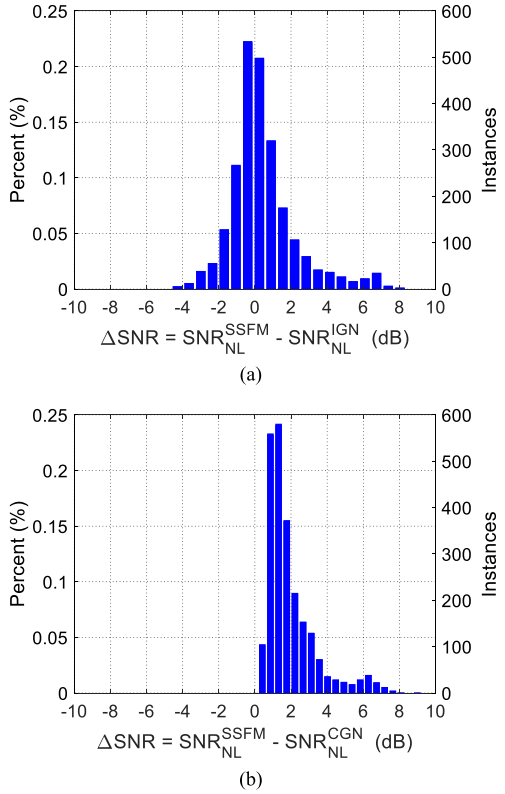

Fig. 6. Histogram of SNR_{nl} deviation for (a) IGN and (b) CGN.

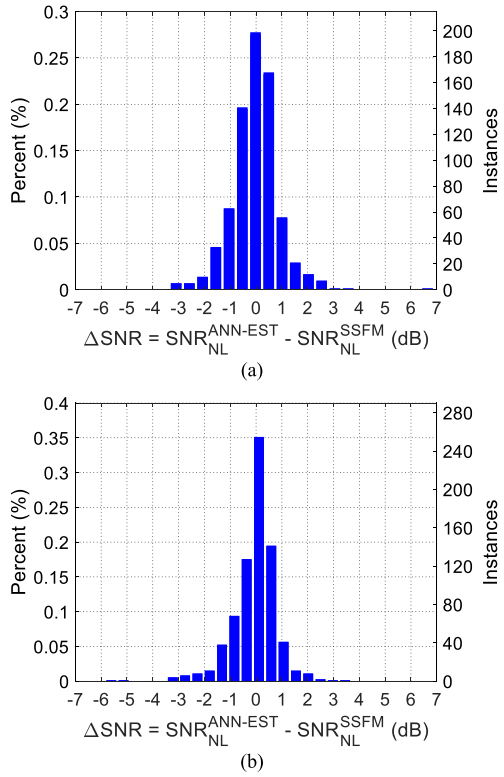

Fig. 7. Histogram of SNR_{nl} deviation for (a) IGN+ANN and (b) CGN+ANN.

can reach >7 dB for a few cases. This result shows that even though the GN models achieve good results in some certain link scenarios, their accuracy needs to be improved for heterogenous dynamic networks. For the CGN model, the deviation is always positive (the SNR_{nl} is overestimated), which is consistent with the observation in previous works [8].

C. Results of GN+ANN Model

Fig. 7 plots the histogram of the SNR_{nl} deviation for the GN+ANN models. With the aid of the ANN, the deviation is significantly reduced. Most of the cases are now within ±4 dB for both models. To quantify the improvement, we plot the cumulative distribution function (CDF) of the absolute SNR_{nl} deviation in Fig. 8. At 95% cumulative probability, the absolute SNR_{nl} deviation is 4.34 dB, 5.41 dB, 1.91 dB and 1.76 dB for the IGN, CGN, IGN+ANN and CGN+ANN schemes, respectively, indicating the effectiveness of the ANN in calibrating the GN models.

To illustrate the process of the ANN calibration, in Fig. 9 we plot the deviation versus the number of spans for the links with homogenous 80 km SSFM spans. The WDM system has 11 channels with 50 GHz spacing and 35 Gbaud symbol rate, which are located together without empty channel slots among them. For the IGN model, it is observed that the deviation decreases from ∼3.8 dB to ∼0.5 dB as the span number increases from 1 to 15. The CGN model manifests a smaller dynamic range since it better approximates the NLE equation as discussed in Section II. Nevertheless, it still changes from ∼3.8 dB to ∼1.8 dB.

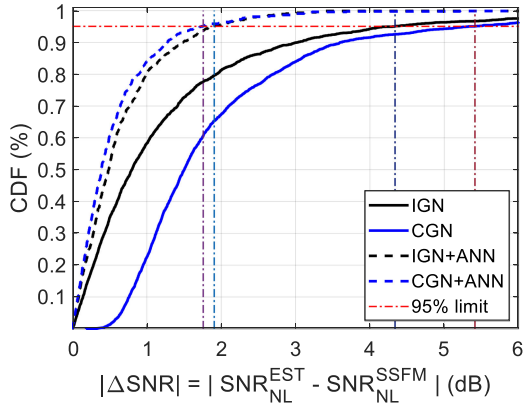

Fig. 8. CDF of absolute SNR_{nl} deviation for the IGN, CGN, IGN+ANN, and CGN+ANN schemes.

Again, we can see that the deviation is correlated to specific link conditions, i.e., span number in Fig. 9. Therefore, the ML algorithm is expected to be capable of calibrating the deviation by removing the correlations between the deviation and input features. Indeed, with the ANN, the deviations are within ±1.1 dB. There are still some residual correlations after the ANN is applied, which means further improvement is possible with more training data and/or stronger ML algorithms.

D. Results of GN+Monitoring+ANN

Next, the results of the ML-aided combination of modeling and monitoring are shown in Fig. 10. For comparison, the

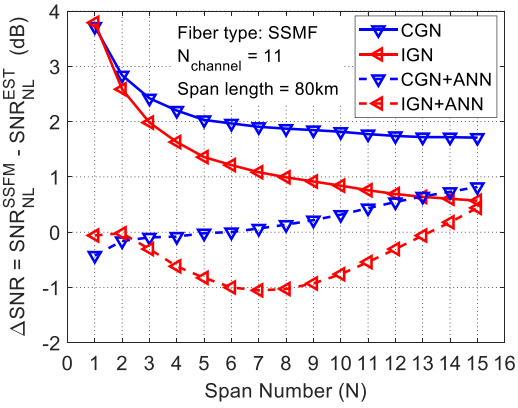

Fig. 9. SNR_{nl} deviation versus number of spans for various model schemes.

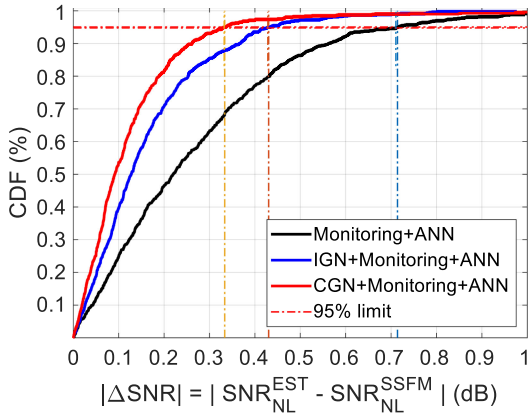

Fig. 11. CDF of absolute SNR_{nl} deviation for the Monitoring+ANN, IGN+Monitoring+ANN, and CGN+Monitoring+ANN schemes.

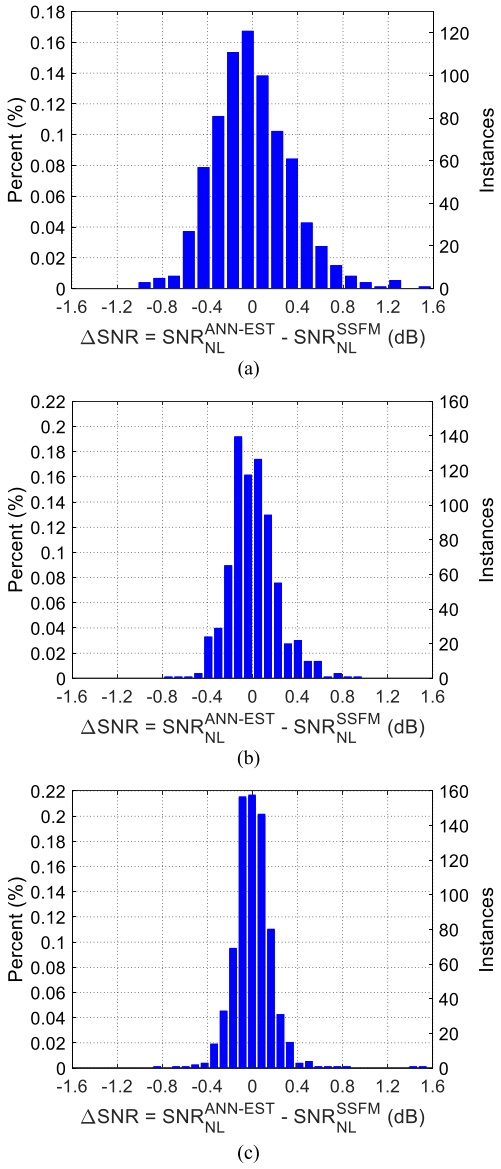

Fig. 10. Histogram of SNR_{nl} deviation for (a) Monitoring+ANN [14], (b) IGN+Monitoring+ANN and (c) CGN+Monitoring+ANN.

performance of the "Monitoring+ANN" case is also shown in Fig. 10(a) with the same ANN inputs as [14], i.e., noise covariances plus net CD and channel number. We can see that the accuracy is already quite good since most of the cases are within ± 1 dB. After we combine the model and monitoring through an ANN, the accuracy is further improved. To quantify the improvement, the CDF of the absolute SNR_{nl} deviation is plotted in Fig. 11. At 95% cumulative probability, the absolute SNR_{nl} deviation is 0.71 dB, 0.43 dB and 0.33 dB for the Monitoring+ANN, IGN+Monitoring+ANN and CGN+Monitoring+ANN models, respectively. This result indicates that using modeling and monitoring together through ML is a promising framework for the development of future optical network planning tools.

E. Results With Optical Power Uncertainty

In practical systems, link parameters normally contain uncertainties due to inaccurate measurement and other reasons. In the last part, we evaluate the performance of the proposed ML-aided schemes in the presence of random offsets in optical powers. The power offsets were generated independently for each channel, following a Gaussian distribution with a standard deviation of 0.7 dB. Then they were added to the true launch powers (in dB) before being used as the input of the SNR_{nl} estimation schemes. We only evaluated the IGN based schemes because the IGN model is more suitable for practical deployment [27]. Note that in the training of the IGN+Monitoring+ANN scheme, the launch power uncertainty was also included. This delivered an improved performance since the ANN learned to rely more on the monitored ANC and PNC, which were not affected by the power uncertainty. As expected, such improvement did not exist in the IGN+ANN scheme, so the ANN trained without power uncertainty was used.

First, the IGN+ANN scheme is expected to be less tolerant to the launch power uncertainty because it purely relies on the link parameters to estimate the nonlinear noise variance. As shown in Fig. 12(a), the absolute deviation at 95% is increased by 0.67 dB with power uncertainty. On the contrary, the link

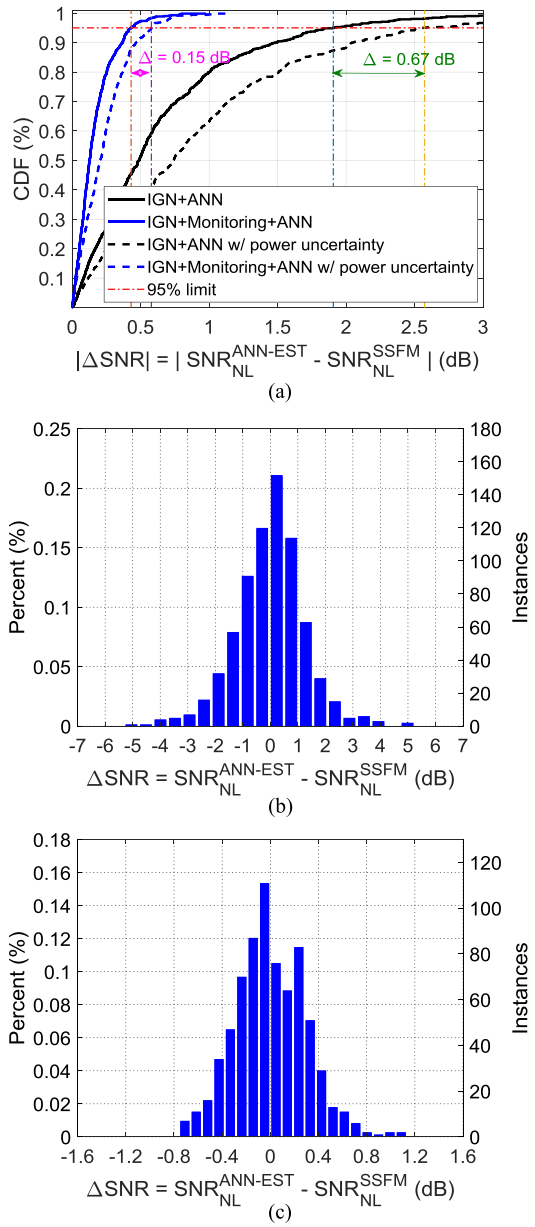

Fig. 12. (a) CDF of absolute SNR_{nl} deviation for the IGN+ANN and IGN+Monitoring+ANN schemes with launch power uncertainty. Histogram of SNR_{nl} deviation for (b) IGN+ANN, (c) IGN+Monitoring+ANN with launch power uncertainty.

monitored covariances are not affected by the launch power uncertainties. Therefore, we expect a higher tolerance for the combined scheme. Fig. 12(a) shows that the absolute deviation of the IGN+Monitoring+ANN scheme at 95% is only increased by 0.15 dB. The error histograms of the two schemes are plotted in Fig. 12(b) and (c) for reference.

Finally, the absolute $\mathrm{SNR_{nl}}$ deviations of the evaluated schemes at 95% cumulative probability are summarized in Table IV. The deviation of the IGN model is only increased by 0.22 dB, which implies that the ANN might be more sensitive to the power uncertainty. The Monitoring +ANN scheme is insensitive to the power uncertainty, but its deviation is still 0.13 dB higher than the IGN+Monitoring+ANN scheme. The results show that the proposed ML-aided combination of the link modeling and monitoring delivers a superior performance.

TABLE IV
SUMMARY OF ABSOLUTE DEVIATIONS AT 95% CUMULATIVE PROBABILITY

Scheme	Without power uncertainty	With power uncertainty
IGN/CGN	4.34 / 5.41	4.56/ n/a
IGN/CGN+ANN	1.91 / 1.76	2.58 / n/a
Monitoring+ANN [14]	0.71	0.71
IGN/CGN+ Monitoring+ANN	0.43 / 0.33	0.58 / n/a

V. CONCLUSION

In this paper, we report the application of machine learning (ML) techniques in fiber nonlinearity modeling and monitoring. In particular, we propose to use ML to calibrate the deviations that exist in nonlinear models, and then propose to use ML to combine link modeling and monitoring for a better estimation of fiber nonlinear noise variance. Gaussian-noise (GN) model, nonlinear noise covariance, and artificial neural network (ANN) are used as illustrative examples to investigate and compare the performance of different approaches with extensive simulations.

REFERENCES

[1] D. Qian *et al.*, "101.7-Tb/s (370 × 294-Gb/s) PDM-128QAM-OFDM transmission over 3 × 55-km SSMF using pilot-based phase noise mitigation," in *Proc. Int. Conf. Opt. Fiber Commun. Nat. Fiber Optic Eng.*, Los Angeles, CA, USA, 2011, Paper PDPB5.

[2] K. Roberts, Q. Zhuge, I. Monga, S. Gareau, and C. Laperle, "Beyond 100 Gb/s: Capacity, flexibility, and network optimization," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 9, no. 4, pp. C12–C24, Apr. 2017.

[3] P. J. Winzer and D. T. Neilson, "From scaling disparities to integrated parallelism: A decathlon for a decade," *IEEE J. Lightw. Technol.*, vol. 35, no. 5, pp. 1099–1115, Mar. 2017.

[4] O. Gerstel, M. Jinno, A. Lord, and S. Yoo, "Elastic optical networking: A new dawn for the optical layer?," *IEEE Commun. Mag.*, vol. 50, no. 2, pp. s12–s20, Feb. 2012.

[5] L. Yan, E. Agrell, H. Wymeersch, and M. Brandt-Pearce, "Resource allocation for flexible-grid optical networks with nonlinear channel model [invited]," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 7, no. 11, pp. B101–B108, Nov. 2015.

[6] I. Roberts and J. M. Kahn, "Efficient discrete rate assignment and power optimization in optical communication systems following the Gaussian noise model," *IEEE J. Lightw. Technol.*, vol. 35, no. 20, pp. 4425–4437, Oct. 2017.

[7] G. P. Agrawal, "Nonlinear fiber optics," in *Nonlinear Science at the Dawn of the 21st Century*. Berlin, Germany: Springer, 2000, pp. 195–211.

[8] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "The GN-model of fiber non-linear propagation and its applications," *IEEE J. Lightw. Technol.*, vol. 32, no. 4, pp. 694–721, Feb. 2014.

[9] F. Zhang *et al.*, "Blind adaptive digital backpropagation for fiber nonlinearity compensation," *IEEE J. Lightw. Technol.*, vol. 36, no. 9, pp. 1746–1756, May 2018.

[10] Z. Dong, F. N. Khan, Q. Sui, K. Zhong, C. Lu, and A. P. T. Lau, "Optical performance monitoring: A preview of current and future technologies," *IEEE J. Lightw. Technol.*, vol. 34, no. 2, pp. 525–543, Jan. 2016.

[11] E. Seve, J. Pesic, C. Delezoide, S. Bigo, and Y. Pointurier, "Learning process for reducing uncertainties on network parameters and design margins," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 10, no. 2, pp. A298–A306, Feb. 2018.

[12] M. Bouda *et al.*, "Accurate prediction of quality of transmission based on a dynamically configurable optical impairment model," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 10, no. 1, pp. A102–A109, Jan. 2018.

[13] I. Sartzetakis, K. Christodoulopoulos, and E. Varvarigos, "Accurate quality of transmission estimation with machine learning," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 11, no. 3, pp. 140–150, Mar. 2019.

[14] A. S. Kashi *et al.*, "Fiber nonlinear noise-to-signal ratio monitoring using artificial neural networks," in *Proc. Eur. Conf. Opt. Commun.*, Gothenburg, Sweden, 2017, Paper M.2.F.2.

[15] F. J. V. Caballero *et al.*, "Machine learning based linear and nonlinear noise estimation," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 10, no. 10, pp. D42–D51, Oct. 2018.

[16] R. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *IEEE J. Lightw. Technol.*, vol. 28, no. 4, pp. 662–701, Feb. 2010.

[17] D. Marcuse, C. R. Manyuk, and P. K. A. Wai, "Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *IEEE J. Lightw. Technol.*, vol. 15, no. 9, pp. 1735–1746, Sep. 1997.

[18] 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. 16335–16362, 2014.

[19] F. N. Hauske, M. Kuschnerov, B. Spinnler, and B. Lankl, "Optical performance monitoring in digital coherent receivers," *IEEE J. Lightw. Technol.*, vol. 27, no. 16, pp. 3623–3631, Aug. 2009.

[20] J. A. Jargon, X. Wu, and A. E. Willner, "Optical performance monitoring by use of artificial neural networks trained with parameters derived from delay-tap asynchronous sampling," in *Proc Int. Conf. Opt. Fiber Commun.*, 2009, pp. 1–3.

[21] F. N. Khan, Y. Yu, M. C. Tan, C. Yu, A. P. T. Lau, and C. Lu, "Simultaneous OSNR monitoring and modulation format identification using asynchronous single channel sampling," in *Proc. Asia Commun. Photon. Conf.*, Hong Kong, Nov. 2015, Paper AS4F.6.

[22] F. N. Khan *et al.*, "Joint OSNR monitoring and modulation format identification in digital coherent receivers using deep neural networks," *Opt. Express*, vol. 25, no. 15, pp. 17767–17776, 2017.

[23] Z. Dong, A. P. T. Lau, and C. Lu, "OSNR monitoring for QPSK and 16-QAM systems in presence of fiber nonlinearities for digital coherent receivers," *Opt. Express*, vol. 20, no. 17, pp. 19520–19534, 2012.

[24] T. Fehenberger, M. Mazur, T. A. Eriksson, M. Karlsson, and N. Hanik, "Experimental analysis of correlations in the nonlinear phase noise in optical fiber systems," in *Proc. 42nd Eur. Conf. Opt. Commun.*, Düsseldorf, Germany, 2016, Paper W.1.D.4.

[25] J. Mata *et al.*, "Artificial intelligence (AI) methods in optical networks: A comprehensive survey," *Opt. Switching Netw.*, vol. 28, pp. 43–57, 2018.

[26] R. M. Morais and J. Pedro, "Machine learning models for estimating quality of transmission in DWDM networks," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 10, no. 10, pp. D84–D99, Feb. 2018.

[27] M. Filer, M. Cantono, A. Ferrari, G. Grammel, G. Galimberti, and V. Curri, "Multi-vendor experimental validation of an open source QoT estimator for optical networks," *IEEE J. Lightw. Technol.*, vol. 36, no. 15, pp. 3073–3082, Aug. 2018.