# A feed-forward neural network as a nonlinear dynamics integrator for supercontinuum generation

LAURI SALMELA[1,*], MATHILDE HARY[1,2], MEHDI MABED[2], ALESSANDRO FOI[3], JOHN M. DUDLEY[2], AND GOËRY GENTY[1]

[1] Photonics Laboratory, Physics Unit, Tampere University, 33014 Tampere, Finland
[2] Institut FEMTO-ST, Université Bourgogne Franche-Comté CNRS UMR 6174, 25000 Besançon, France
[3] Laboratory of Signal Processing, Tampere University, 33014 Tampere, Finland
[*] Corresponding author: lauri.salmela@tuni.fi

Compiled February 7, 2023

**The nonlinear propagation of ultrashort pulses in optical fibers depends sensitively on the input pulse and fiber parameters. As a result, optimizing propagation for specific applications generally requires time-consuming simulations based on the sequential integration of the generalized nonlinear Schrödinger equation (GNLSE). Here, we train a feed-forward neural network to learn the differential propagation dynamics of the GNLSE, allowing emulation of direct numerical integration of fiber propagation, and particularly the highly complex case of supercontinuum generation. Comparison with a recurrent neural network shows that the feed-forward approach yields faster training and computation, and reduced memory requirements. The approach is generic and can be extended to other physical systems.** 

http://dx.doi.org/10.1364/ao.XX.XXXXXX

Neural networks (NNs) are a subset of machine learning techniques widely used in data analysis, classification and prediction [1]. NNs possess the ability to link the input and output of a multidimensional system, of particular benefit for modeling complex relationships as is typically the case in the presence of nonlinearity [2–8]. NNs are being increasingly applied in optics [9], with recent results including mode-locked laser optimization [10–12] and the analysis of ultrafast instabilities [13–15].

A particular focus of NNs in optics has been the study of fiber nonlinear propagation and supercontinuum (SC) generation [13, 14, 16], complex processes involving multiple effects [17]. The dynamics depend sensitively on the injected pulse and fiber parameters, and matching an input to achieve a desired spectral or temporal output is a complex multivariate problem. The traditional approach for optimization is based on parameter scanning using step-by-step integration of the generalized nonlinear Schrödinger equation (GNLSE) [18]. Yet whilst the GNLSE can accurately model fiber nonlinear dynamics, direct simulations are time-consuming, especially with a large parameter space of potential boundary conditions (input pulse power, duration and chirp; fiber dispersion, nonlinearity, length).

To overcome this, machine learning techniques have been applied to optimize fiber dynamics, including the use of genetic algorithms to tailor broadband spectra [19, 20]. More recently, recurrent neural networks (RNNs) using only input temporal (or spectral) intensity profiles have successfully emulated fiber propagation [14], accurately predicting SC evolution maps in computation times as short as one second. A limitation, however, is the initial training phase of several hours due to the multiple loops associated with the RNN internal memory.

Here, we show how a faster and simpler feed-forward neural network (FNN) can model the full-field (intensity and phase) evolution of ultrashort pulses in optical fiber over a wide range of input pulse properties and fiber parameters. The key conceptual novelty is to train the network to learn GNLSE differential propagation dynamics i.e. to replicate the change in intensity and phase of the electric field between elementary longitudinal steps. Once trained, the network can model the long-term evolution from a given input. We also perform a detailed comparison with a RNN model, highlighting the benefits of the FNN approach in terms of speed and memory.

Figure 1 shows the principle. We first generate an ensemble of data for broadband coherent SC generation (using Matlab on a 3.4 GHz 4-core Intel Core i7). We include a one photon per spectral bin noise model [17], but its influence was found to be negligible. This data is generated by numerically integrating the GNLSE with the split-step method. A summary of input pulse and fiber parameter ranges used in all cases below is given in the Supplementary documentation.

The dynamical maps are characterized by vector $[I_n(z_i, X), \Phi_n(z_i, X)]$, where $I_n$ and $\Phi_n$ represent intensity and phase at distance $z_i$, expressible either in temporal ($X = T$) or spectral domains ($X = \omega$), and from which the complex electric field can be reconstructed. Subscript $n = 1 \ldots N$ indicates a particular map for a given set of input pulse and fiber parameters. The key idea in Fig. 1(a) is to teach the network the differential change in intensity and phase associated with an elementary propagation distance $\Delta z$. To achieve a performance advantage relative to direct integration, the aim is to use a significantly larger step in the FNN compared to that used in GNLSE integration. To this end, the intensity and phase evolution are downsampled at distances $z_i = (i-1)\Delta z$

($i = 1..M$), where $\Delta z = L/M = 0.1$ cm is e.g. 50 times larger than in the GNLSE simulations used to generate the data. The downsampled vectors are then used as the FNN input. The network output vectors after an elementary step $\Delta z$ are $[I_n(z_{i+1}, X), \Phi_n(z_{i+1}, X)]$. The change in the intensity and phase modeled by the FNN is then compared to that from the GNLSE via an error function [13].

Once trained, the neural network acts as a very fast and memory-efficient GNLSE integrator. It predicts intensity and phase $[I(z + \Delta z, X), \Phi(z + \Delta z, X)]$ after an elementary propagation distance $\Delta z$ given the complex field $[I(z, X), \Phi(z, X)]$ at distance $z$, from which the dynamical evolution of the complex electric field can be reconstructed. The trained FNN can then be used to predict propagation dynamics over an extended distance using an iterative loop (see Fig. 1(b)) such that intensity and phase $[I(z_{i+1}, X), \Phi(z_{i+1}, X)]$ are fed back to the network as a new input to predict $[I(z_{i+2}, X), \Phi(z_{i+2}, X)]$ at distance $z + 2\Delta z$. This operation is performed over the full propagation distance.

The neural network itself consists of 3 hidden layers of 2000 nodes with ReLU activation ($f(x) = \max(0, x)$) and a sigmoid output layer with 2048 nodes. The codes were written in Python using Keras with Tensorflow backend [21]. The FNN is trained on a single GPU (NVIDIA Quadro K620) for 80 epochs with RMSprop optimizer and adaptive learning rate. The network can be trained in the temporal or spectral domain, and with data input on either linear or logarithmic (dB) scales. In the results below we used ensembles of spectral evolution maps in logarithmic scale. Examples of time domain evolution using linear input are in the Supplementary documentation (Fig. S1). The network accuracy is tested with a separate set of propagation maps not used in the training phase. We quantify performance using the average (normalized) root mean squared (RMS) error computed over the test evolution map at all propagation steps:

$$R = \sqrt{\frac{\sum_{d,i}(x_{n,d,i} - \hat{x}_{n,d,i})^2}{\sum_{d,i}(x_{n,d,i})^2}}, \qquad \textbf{(1)}$$

where $\mathbf{x}_n$ and $\hat{\mathbf{x}}_n$ denote GNLSE simulation and FNN prediction for a particular realization $n$. Variables $d$ and $i$ indicate summation over intensity (spectral or temporal) and propagation steps, respectively. When evaluating performance over an ensemble, the error is calculated over $N$ evolution maps.

We first show how the FNN can predict SC evolution from transform-limited (TL) input pulses. We used an ensemble of 1400 simulations for training, and 100 for testing. The ensemble of SC maps correspond to hyperbolic-secant input pulses at $\lambda_0 = 830$ nm, with peak power and duration (FWHM) in the range $P_0 = 0.77$–$1.43$ kW and $T_{\mathrm{FMHM}}$ =70–130 fs (±30% variation). The dispersion parameters are $\beta_2 = -5.90 \times 10^{-27}$ s$^2$m$^{-1}$, $\beta_3 = 4.21 \times 10^{-41}$ s$^3$m$^{-1}$, $\beta_4 = -1.25 \times 10^{-55}$ s$^4$m$^{-1}$, and $\beta_5 = -2.45 \times 10^{-70}$ s$^5$m$^{-1}$ (zero-dispersion wavelength (ZDW) at 767 nm), and the nonlinear coefficient is: $\gamma = 0.1\,\mathrm{W}^{-1}\mathrm{m}^{-1}$. The fiber length is $L$ = 20 cm. FNN predicted spectral evolution for input peak power and pulse duration of 1.32 kW and 120 fs is shown in Fig. 2a. For comparison, we also plot the map from direct GNLSE integration. The RMS error is R = 0.098, while the average error computed over the 100 test maps is R = 0.094. The FNN accurately predicts the SC development, with dispersive wave and soliton dynamics reproduced over ∼ 40 dB dynamic range.

We next tested modeling of SC development from chirped pulses. We performed 3000 simulations with parameters as above, except with peak power variation of ±20% and input

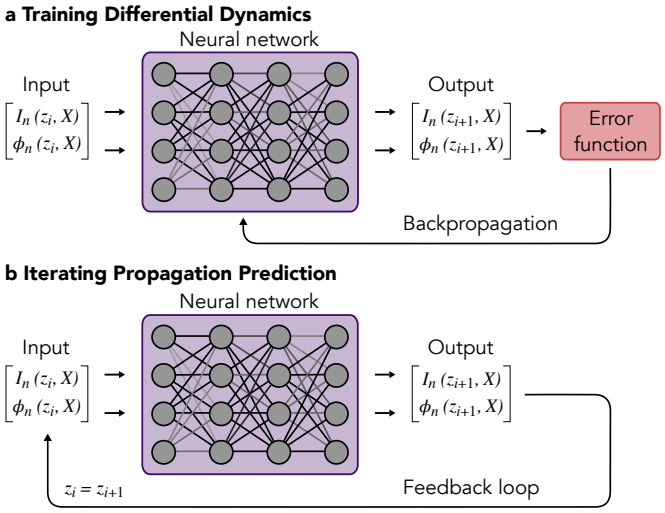


**Fig. 1.** Principle of the FNN integrator. **a** Training Differential Dynamics. Training uses multiple input/output pairs generated by direct integration of the GNLSE and corresponding to the temporal ($X = T$) or spectral ($X = \omega$) intensity $I$ and phase $\Phi$ of the propagating field at distances separated by $\Delta z$ (see text). The network variables are adjusted via gradient descent backpropagation. **b** Iterating Propagation Prediction. Once trained, the network predicts iteratively the intensity and phase evolution via a feedback loop. The prediction is initialized from the intensity and phase at the fiber input.

spectral bandwidth varying from TL to twice the TL with random sign of chirp. The predicted spectral evolution for pulses with 942 W peak power, 84 fs duration, and positive chirp of 1.53 times the TL bandwidth are shown in Fig. 2b. Again we see how the main features (including spectral interference) are well-reproduced by the FNN (R = 0.190) although we note a small discrepancy in the distance of maximum compression at -20 dB bandwidth. The RMS error R = 0.383 (0.242 median) computed over the 100 test ensemble shows that the network accurately models chirped pulse dynamics.

The results above correspond to typical anomalous dispersion dynamics, but the network can be trained over a much wider range of parameters using the normalized form of the GNLSE to generate the training ensemble (see Supplementary documentation). Here we map a specific set of dimensional parameters to normalized values to predict the corresponding evolution. Figure 3 plots predicted SC evolution (over 100 longitudinal steps) for a pump wavelength in the normal dispersion region (see caption for parameters). Specifically, Fig. 3a corresponds to a TL limited pulse injected near the ZDW while Fig. 3b is for a pump wavelength further detuned into the normal dispersion regime. We see very good accuracy with R = 0.141 for Fig. 3a and R = 0.043 for Fig. 3b. The RMS error over an ensemble of 200 realizations is R = 0.060. Time domain predictions are given in the Supplementary documentation (Fig. S2).

To reduce computational memory and increase the speed in the training phase, one can train the network from convolved spectral intensity and phase evolution maps. At first sight, a disadvantage of convolved data is that the resulting wavelength/frequency grid is no longer on a Fourier grid, requiring separate training to predict spectral and temporal evolution. However, this is in fact a major benefit, because it allows the appropriate selection of resolution in spectral or temporal domains

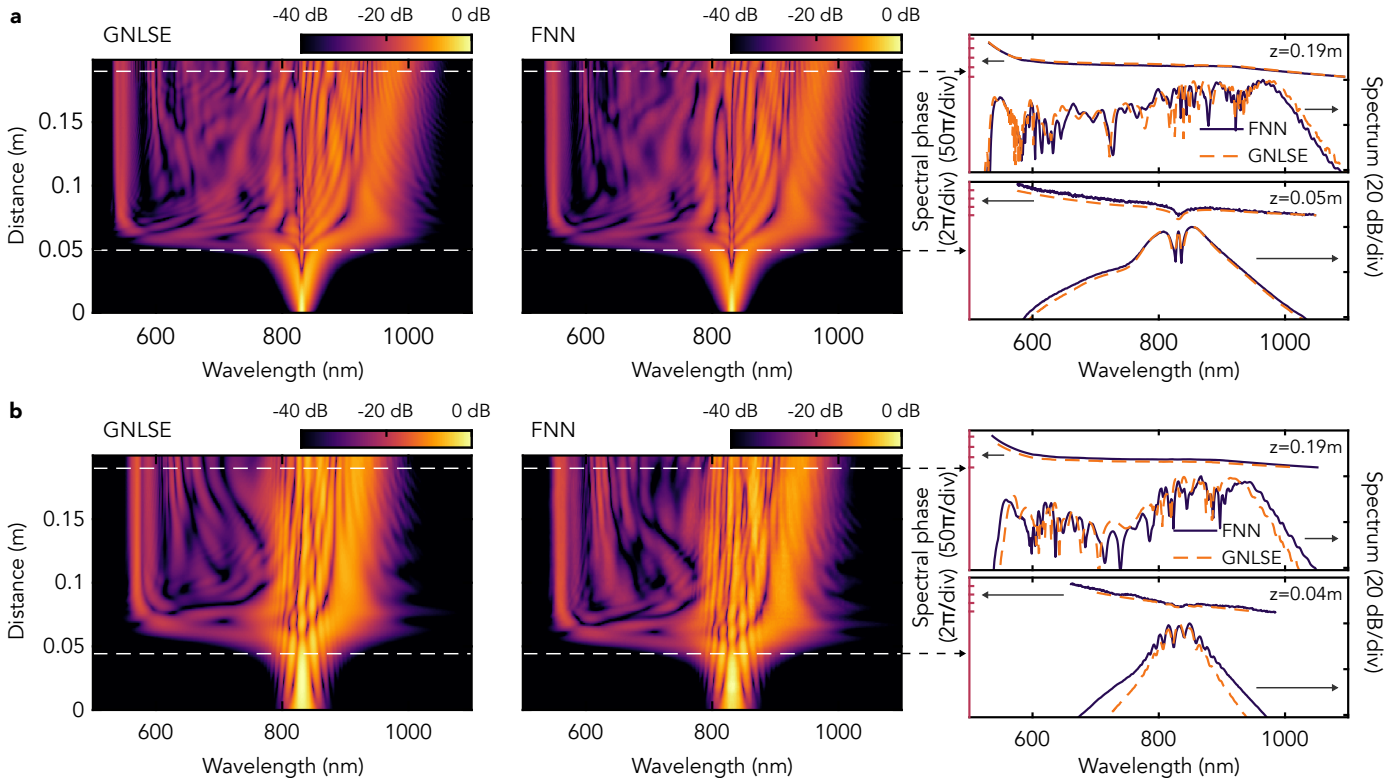


**Fig. 2.** Comparing supercontinuum spectral evolution from GNLSE integration (left panel) and FNN prediction (middle panel). The right panel shows the spectral intensity and phase at distances indicated by the arrows. **a** shows results for transform limited (TL) input pulses of peak power and pulse duration $P_0 = 1.32\,\mathrm{kW}$ and $T_{FWHM} = 120\,\mathrm{fs}$. **b** shows results for a chirped input pulse of peak power and pulse duration $P_0 = 942\,\mathrm{W}$ and $T_{FWHM} = 84\,\mathrm{fs}$, and positive chirp of 1.53 times the TL bandwidth. The GNLSE simulations and FNN predictions were performed using 10000 and 200 longitudinal steps, respectively.

to optimally capture the relevant physical structure.

Results of predicted spectral evolution using convolved training data (using an 8 nm FWHM super-Gaussian spectral filter) are shown in Fig. 4a,b. These results use the same input pulse and fiber parameters as in Fig. 2a,b. We see how the FNN predictions remain accurate with a mean convolved (logarithmic) spectral intensity RMS errors of 0.06 and 0.16 for TL and chirped cases respectively (calculated over 100 distinct test evolution maps). Predictions using other spectral resolutions are in the Supplementary documentation Fig. S3.

We then compared the computation resources and performance of the FNN model and a RNN similar to that used in Ref. [14]. The comparison was performed over an ensemble of 12000 (11800 for training and 200 for testing) convolved SC evolution maps for anomalous dispersion dynamics with variations in peak power, pulse duration and dispersion (see Supplementary documentation) and using 50 longitudinal prediction steps. Table 1 summarizes the results, with examples of predicted maps shown in the Supplementary documentation (Fig. S4). We also list the resources used by the GNLSE simulations. Both FNN and RNN used the same number of free parameters/network variables, but the RNN is trained from spectral intensity maps which reduces by half the number of grid points compared to the FNN that includes both intensity and phase. The computational advantage of the FNN is clear with training and prediction times reduced by a factor of four and five respectively, while memory usage during training is decreased by a factor of two. As might be expected, the FNN does show increased error compared to the RNN, but this does not lead to significant visual differences in the evolution maps obtained.

In general, comparing direct GNLSE integration and the FNN (RNN) approach is a complex problem involving multiple variables such as the number of grid points and integration steps (GNLSE), and the propagation sampling interval and training ensemble (FNN/RNN). The comparison also depends on the particular simulation regime being considered and it is not possible to provide a "one size fits all" estimation, although a comparison over a range of parameters is given in the Supplementary documentation. Key advantages of the FNN are that training data can be discarded and only hyper-parameters need to be stored, and that a trained FNN can perform predictions with extremely short run-times. However, attempting FNN prediction too far outside the parameter range used in training can lead to errors, and re-training may be necessary for new simulation regimes.

These results have shown model-free prediction of the full-field dynamics of ultrashort pulse propagation in optical fiber based on a feed-forward neural network trained to recognize differential propagation dynamics within a GNLSE model. As compared to the recently introduced RNN approach, this FNN method is simpler and possesses significant advantages in speed and memory. We expect our results to be of significance for real-time optimization and control of nonlinear dynamics and we anticipate this approach could become a standard tool in nonlinear physics. As a field of further study, it could be interesting to study transfer learning in other NLSE -like systems.

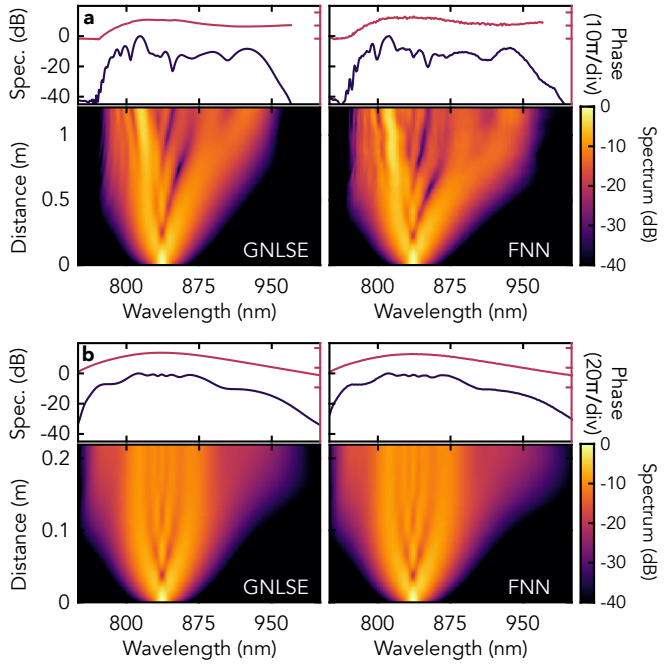


**Fig. 3.** Spectral evolution from GNLSE simulations (left panels) and FNN predictions (right panels) for: **a** normal near-ZDW pumping ($\gamma = 0.01$ W$^{-1}$m$^{-1}$, $\beta_2 = 1.3 \times 10^{-27}$ s$^2$m$^{-1}$, $\beta_3 = 2 \times 10^{-41}$s$^3$m$^{-1}$, $P_0 = 2.0$ kW, $\lambda_0 = 835$ nm, $T_{\mathrm{FWHM}} = 100$ fs); and **b** far-normal pumping ($\gamma = 0.01$ W$^{-1}$m$^{-1}$, $\beta_2 = 7.2 \times 10^{-27}$ s$^2$m$^{-1}$, $\beta_3 = 2 \times 10^{-41}$ s$^3$m$^{-1}$, $P_0 = 13.6$ kW, $\lambda_0 = 835$ nm, $T_{\mathrm{FWHM}} = 100$ fs). The top panels show spectral intensity and phase at the fiber output.

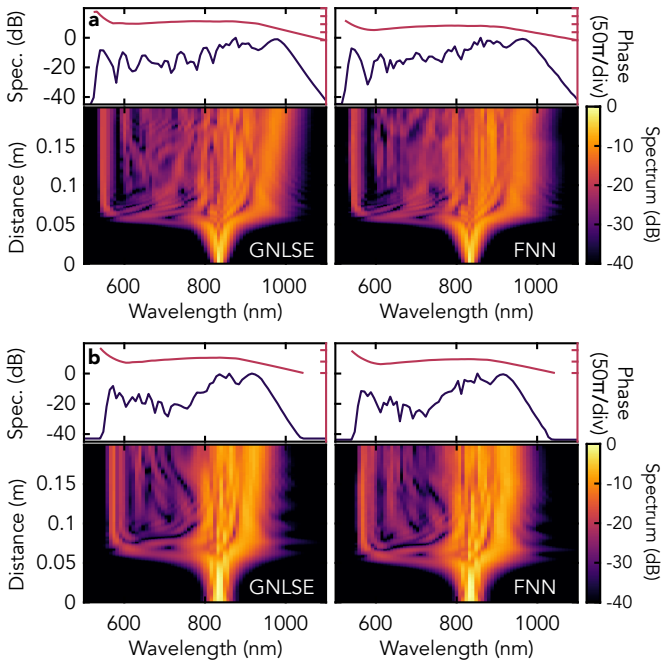


**Fig. 4.** Comparison between the spectral evolution from simulations (GNLSE, left panel) and predicted by the neural network (FNN, right panel) when using convolved evolution maps for training with parameters identical to those in Fig. 2**a** and **b**. The top panels show the spectral intensity and phase at the fiber output.

| | GNLSE | RNN | FNN |
|---|---|---|---|
| RMS error | N/A | R = 0.09 | R = 0.19 |
| Training time* | N/A | 7.7 h | 1.9 h |
| Simulation time** | 38 min | 1.6 s | 0.35 s |
| Memory* | 79 GB | 7.7 GB | 3.2 GB |
| Network var. | N/A | 600k | 600k |
| Number of points | 8,192 | 132 | 264 |

**Table 1.** Comparison between normalized GNLSE numerical simulations, RNN model [14], and FNN model for convolved spectral data. *11800 simulations **200 simulations.

**Funding.** Tampere University Faculty of Engineering and Natural Sciences graduate school. Academy of Finland (298463, 318082, 320165). French Agence Nationale de la Recherche (ANR-15-IDEX- 0003, ANR-17-EURE-0002, ANR-20-CE30-0004).
**Disclosures.** The authors declare no conflicts of interest.
**Data availability.** The data in this paper may be obtained from the authors upon request.
See Supplementary documentation for supporting content.

## REFERENCES

1. I. Goodfellow, Y. Bengio, and A. Courville, Deep learning (MIT Press, 2016).
2. S. L. Brunton, J. L. Proctor, and J. N. Kutz, Proc. Natl. Acad. Sci. **113**, 3932 (2016).
3. M. Raissi, J. Mach. Learn. Res. **19**, 932 (2018).
4. M. Raissi, P. Perdikaris, and G. E. Karniadakis, J. Comput. Phys. **378**, 686 (2019).
5. P. R. Vlachas, J. Pathak, B. R. Hunt, T. P. Sapsis, M. Girvan, E. Ott, and P. Koumoutsakos, Neural Networks (2020).
6. J. Jiang and Y.-C. Lai, Phys. Rev. Res. **1**, 033056 (2019).
7. M. Sangiorgio and F. Dercole, Chaos, Soliton. & Fract. **139**, 110045 (2020).
8. Y. Lei, J. Hu, and J. Ding, arXiv preprint arXiv:2002.00799 (2020).
9. G. Genty, L. Salmela, J. M. Dudley, D. Brunner, A. Kokhanovskiy, S. Kobtsev, and S. K. Turitsyn, Nat. Photonics **15**, 91 (2020).
10. G. Pu, L. Yi, L. Zhang, and W. Hu, Optica **6**, 362 (2019).
11. U. Andral, R. S. Fodil, F. Amrani, F. Billard, E. Hertz, and P. Grelu, Optica **2**, 275 (2015).
12. A. Kokhanovskiy, A. Ivanenko, S. Kobtsev, S. Smirnov, and S. Turitsyn, Sci. Reports **9**, 2916 (2019).
13. M. Närhi, L. Salmela, J. Toivonen, C. Billet, J. M. Dudley, and G. Genty, Nat. Commun. **9**, 1 (2018).
14. L. Salmela, N. Tsipinakis, A. Foi, C. Billet, J. M. Dudley, and G. Genty, Nat. Mach. Intell. **3**, 344 (2021).
15. P. Amil, M. C. Soriano, and C. Masoller, Chaos **29**, 113111 (2019).
16. L. Salmela, C. Lapre, J. M. Dudley, and G. Genty, Sci. Reports **10**, 1 (2020).
17. J. M. Dudley, G. Genty, and S. Coen, Rev. Mod. Phys. **78**, 1135 (2006).
18. G. Agrawal, Nonlinear fiber optics (Academic Press, 2019), 6th ed.
19. B. Wetzel, M. Kues, P. Roztocki, C. Reimer, P.-L. Godin, M. Rowley, B. E. Little, S. T. Chu, E. A. Viktorov, D. J. Moss, A. Pasquazi, M. Peccianti, and R. Morandotti, Nat. Commun. **9**, 1 (2018).
20. L. Michaeli and A. Bahabad, J. Opt. **20**, 055501 (2018).
21. M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard et al., 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16), (2016), pp. 265–283.

## FULL REFERENCES


1. I. Goodfellow, Y. Bengio, and A. Courville, Deep learning (MIT Press, 2016).
2. S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems," Proc. Natl. Acad. Sci. **113**, 3932–3937 (2016).
3. M. Raissi, "Deep hidden physics models: Deep learning of nonlinear partial differential equations," J. Mach. Learn. Res. **19**, 932–955 (2018).
4. M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," J. Comput. Phys. **378**, 686–707 (2019).
5. P. R. Vlachas, J. Pathak, B. R. Hunt, T. P. Sapsis, M. Girvan, E. Ott, and P. Koumoutsakos, "Backpropagation algorithms and reservoir computing in recurrent neural networks for the forecasting of complex spatiotemporal dynamics," Neural Networks (2020).
6. J. Jiang and Y.-C. Lai, "Model-free prediction of spatiotemporal dynamical systems with recurrent neural networks: Role of network spectral radius," Phys. Rev. Res. **1**, 033056 (2019).
7. M. Sangiorgio and F. Dercole, "Robustness of LSTM neural networks for multi-step forecasting of chaotic time series," Chaos, Soliton. & Fract. **139**, 110045 (2020).
8. Y. Lei, J. Hu, and J. Ding, "A hybrid model based on deep LSTM for predicting high-dimensional chaotic systems," arXiv preprint arXiv:2002.00799 (2020).
9. G. Genty, L. Salmela, J. M. Dudley, D. Brunner, A. Kokhanovskiy, S. Kobtsev, and S. K. Turitsyn, "Machine learning and applications in ultrafast photonics," Nat. Photonics **15**, 91–101 (2020).
10. G. Pu, L. Yi, L. Zhang, and W. Hu, "Intelligent programmable mode-locked fiber laser with a human-like algorithm," Optica **6**, 362–369 (2019).
11. U. Andral, R. S. Fodil, F. Amrani, F. Billard, E. Hertz, and P. Grelu, "Fiber laser mode locked through an evolutionary algorithm," Optica **2**, 275–278 (2015).
12. A. Kokhanovskiy, A. Ivanenko, S. Kobtsev, S. Smirnov, and S. Turitsyn, "Machine learning methods for control of fibre lasers with double gain nonlinear loop mirror," Sci. Reports **9**, 2916 (2019).
13. M. Närhi, L. Salmela, J. Toivonen, C. Billet, J. M. Dudley, and G. Genty, "Machine learning analysis of extreme events in optical fibre modulation instability," Nat. Commun. **9**, 1–11 (2018).
14. L. Salmela, N. Tsipinakis, A. Foi, C. Billet, J. M. Dudley, and G. Genty, "Predicting ultrafast nonlinear dynamics in fibre optics with a recurrent neural network," Nat. Mach. Intell. **3**, 344–354 (2021).
15. P. Amil, M. C. Soriano, and C. Masoller, "Machine learning algorithms for predicting the amplitude of chaotic laser pulses," Chaos **29**, 113111 (2019).
16. L. Salmela, C. Lapre, J. M. Dudley, and G. Genty, "Machine learning analysis of rogue solitons in supercontinuum generation," Sci. Reports **10**, 1–8 (2020).
17. J. M. Dudley, G. Genty, and S. Coen, "Supercontinuum generation in photonic crystal fiber," Rev. Mod. Phys. **78**, 1135 (2006).
18. G. Agrawal, Nonlinear fiber optics (Academic Press, 2019), 6th ed.
19. B. Wetzel, M. Kues, P. Roztocki, C. Reimer, P.-L. Godin, M. Rowley, B. E. Little, S. T. Chu, E. A. Viktorov, D. J. Moss, A. Pasquazi, M. Peccianti, and R. Morandotti, "Customizing supercontinuum generation via on-chip adaptive temporal pulse-splitting," Nat. Commun. **9**, 1–10 (2018).
20. L. Michaeli and A. Bahabad, "Genetic algorithm driven spectral shaping of supercontinuum radiation in a photonic crystal fiber," J. Opt. **20**, 055501 (2018).
21. M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard et al., "Tensorflow: A system for large-scale machine learning," in 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16), (2016), pp. 265–283.