



OPEN

Machine learning analysis of rogue solitons in supercontinuum generation

Lauri Salmela¹✉, Coraline Lapre², John M. Dudley² & Goëry Genty¹

Supercontinuum generation is a highly nonlinear process that exhibits unstable and chaotic characteristics when developing from long pump pulses injected into the anomalous dispersion regime of an optical fiber. A particular feature associated with this regime is the long-tailed “rogue wave”-like statistics of the spectral intensity on the long-wavelength edge of the supercontinuum, linked to the generation of a small number of “rogue solitons” with extreme red-shifts. Whilst the statistical properties of rogue solitons can be conveniently measured in the spectral domain using the real-time dispersive Fourier transform technique, we cannot use this technique to determine any corresponding temporal properties since it only records the spectral intensity and one loses information about the spectral phase. And direct temporal characterization using methods such as the time-lens has resolution of typically 100's of fs, precluding the measurement of solitons which possess typically much shorter durations. Here, we solve this problem by using machine learning. Specifically, we show how supervised learning can train a neural network to predict the peak power, duration, and temporal walk-off with respect to the pump pulse position of solitons at the edge of a supercontinuum spectrum from only the supercontinuum spectral intensity without phase information. Remarkably, the network accurately predicts soliton characteristics for a wide range of scenarios, from the onset of spectral broadening dominated by pure modulation instability to near octave-spanning supercontinuum with distinct rogue solitons.

The broadband fiber supercontinuum (SC) has developed into a widespread and versatile light source that has found many important applications in areas such as spectroscopy, imaging, and precision frequency metrology¹. The noise properties of broadband SC spectra have been a subject of much interest, not only from an applications perspective, but also as an example of the rich instability dynamics that can arise in a highly nonlinear system. Of particular focus has been the study of SC generation seeded by picosecond (or longer) pulses, where the spectral broadening is triggered by modulation instability (MI) which exponentially amplifies noise components outside the pump spectral bandwidth¹. In the time domain, this instability induces the break-up of the pump pulse envelope into a series of high-intensity breathers with random characteristics². With further propagation and the influence of higher-order dispersion and stimulated Raman scattering, these temporal breathers evolve into localized soliton structures whose subsequent dynamics seed the development of a broadband SC spectrum^{1,3}. The noise-seeded nature of the overall process leads to an incoherent spectrum with large shot-to-shot spectral fluctuations particularly pronounced on the long-wavelength edge^{1,4}.

There has been much interest in studying the spectral fluctuations on the long-wavelength edge of the broadband SC⁴, as these have been shown to exhibit highly skewed statistics associated with extreme events and the emergence of rogue solitons of tens of femtosecond duration^{4–8}. Originally characterized using long-pass filtering and the real-time dispersive Fourier transform⁴, the observation of these fluctuations stimulated a large number of subsequent theoretical, numerical and other experimental investigations, linking the emergence of rogue solitons with collision dynamics during the initial SC development phase^{5,7,9}. Related work has since extended the study of the properties of rogue solitons to a wider range of SC generation regimes and input conditions^{6,10–13}.

The dispersive Fourier transform technique has now become a standard tool to study ultrafast instabilities in the spectral domain^{14,15}, yielding significant insight into the dynamics of many complex nonlinear systems^{11,12,16–20}. However, a drawback of the dispersive Fourier transform is that it only measures the spectral intensity but not the spectral phase, which prevents extracting quantitative information about the associated

¹Photonics Laboratory, Tampere University, Tampere, FI-33014, Finland. ²Institut FEMTO-ST, Université Bourgogne Franche-Comté CNRS UMR 6174, Besançon, 25000, France. ✉e-mail: lauri.salmela@tuni.fi

single-shot temporal intensity profiles. In order to characterize real-time temporal fluctuations associated with incoherent dynamics, more complex techniques using time-lens or heterodyne approaches need to be used^{19,21–24}, yet measurements in this case are generally restricted to specific (narrow bandwidth) wavelength ranges with sub-ps timescale resolution. These limitations preclude the characterization of solitons or localized structures with 10's of femtosecond duration which is typically the duration of solitons emerging in supercontinuum generation.

Machine learning (ML) is an ensemble of numerical techniques specifically developed for classification, pattern recognition, prediction, and system optimization from large data sets²⁵. Recently, there has been a growing interest in applying the techniques of machine learning to optical systems, and in particular for the control of ultrafast dynamics with applications to pulse compression and shaping^{26,27} using neural networks, or supercontinuum spectrum customization using genetic algorithms^{28,29}. Machine learning is also a powerful tool to correlate quantitative characteristics in a complex system with multiple data features, a strength which has been successfully exploited to determine the maximum intensity of temporal peaks in modulation instability based only on spectral measurements³⁰.

In this paper, we extend the scope of machine learning applications to the analysis of SC instabilities and show how machine learning can overcome the current limitations of real-time experimental techniques by analyzing real-time spectral intensity measurements in a way that allows key temporal characteristics of SC rogue solitons to be determined. More specifically, we train a supervised neural network (NN) using numerical simulations of the generalized nonlinear Schrödinger equation (GNLSE) to correlate key temporal characteristics (peak power, duration, temporal walk-off with respect to the pump pulse position) of the most red-shifted rogue soliton in a SC with the corresponding complex supercontinuum spectral intensity profile. Despite the absence of any phase information at the network input, the trained network is able to infer the red-shifted soliton characteristics with excellent accuracy, exceeding that obtained when using a single-parameter metric such as e.g. the spectral bandwidth or energy in the long-wavelength edge of the SC spectrum. Significantly, the NN analysis remains accurate over a wide variety of dynamical regimes, from the onset of spectral broadening dominated by modulation instability, to near octave-spanning supercontinuum with distinct rogue solitons.

Results

Rogue soliton generation and machine learning. We begin by illustrating in Fig. 1(a) a schematic of our numerical experiment. Ultrashort pulses from a pulsed laser are injected into a highly nonlinear fiber triggering the development of a broadband supercontinuum. The shot-to-shot spectral intensity fluctuations are captured using the dispersive Fourier transform technique, stretching consecutive supercontinuum pulses in a strongly dispersive fiber and measuring the dispersed spectra with a fast photodetector and oscilloscope^{14,15}. A long-pass filter can also be used to isolate the long-wavelength components⁴. The noise sensitivity of the SC generation process and rogue soliton emergence when spectral broadening is seeded from a long pump pulse injected into the anomalous dispersion regime is shown in Fig. 1(b,c). The input conditions here correspond to hyperbolic-secant pulses with 2 ps duration (full-width at half-maximum FWHM), 400 W peak power and 810 nm central wavelength injected into an 85 cm long photonic crystal fiber (PCF) with zero-dispersion at 750 nm. Input noise is included in the spectral domain using a standard one-photon-per-mode background with random phase¹. See Methods for additional simulation parameters.

We performed simulations to generate an ensemble of 30,000 pairs of spectra and temporal intensity profiles using different random noise seeds. The middle panels in Fig. 1 plot results from one realization in the ensemble to show the spectrum (b) and the corresponding temporal intensity profile (c) associated with rogue soliton characteristics. In particular, we clearly identify the initial stage of noise-seeded modulation instability after a propagation distance of about 30 cm, with the generation of distinct spectral sidebands and the break-up of the pump pulse envelope into multiple breather structures². With further propagation, the breathers evolve into fundamental solitons experiencing Raman self-frequency shift and separating in time from the residual background envelope³¹. After a distance of ~50 cm, collisions between multiple solitons increase the frequency-shift rate of one of the solitons emerging from the envelope center, leading to the generation of a distinct rogue soliton with extreme red-shift^{5,7,9}.

The results in Fig. 1(d) show a selection of 5 supercontinuum spectra from the ensemble, and the black lines in (e) show the corresponding temporal intensity profiles. Note that we have sorted the results in (d) and (e) to show increasing spectral width from top to bottom but of course these results occur randomly in the ensemble. From these results, we clearly see how the noise-sensitive dynamics lead to large shot-to-shot spectral variations between different results in the ensemble, and we see particularly how the shot-to-shot spectral variations in the long-wavelength edge correspond to isolated solitons with different peak power, duration and temporal walk-off.

The numerical results in Fig. 1(d,e) shows how the dispersive Fourier transform approach indeed isolates the temporal solitons⁴, where the blue region in (d) shows the filter cutoff region, and the blue curves in (e) shows the corresponding temporal intensity profiles after numerical Fourier transform. However, the temporal duration of these rogue solitons (10's of femtoseconds) is too short for direct real-time measurements.

But it is here that we can apply machine learning to extract the temporal characteristics of the rogue solitons only from simple unfiltered spectral measurements [see Fig. 1(a)]. Specifically, we apply a supervised feed-forward neural network to correlate the full spectrum and the temporal characteristics of the most red-shifted soliton of a given (noisy) SC spectrum. A schematic of the neural network is shown in Fig. 2. The NN input is a vector ($\mathbf{X}_n = [x_1, x_2, \dots, x_N]$) corresponding to the SC spectrum, whilst the NN output is a scalar value equal to the maximum temporal intensity, duration or temporal walk-off or delay (defined with respect to the pump pulse center) of the most red-shifted soliton. The NN is trained from an ensemble of 20,000 simulated single-shot SC using a conjugate gradient back-propagation method³². See Methods for further details.

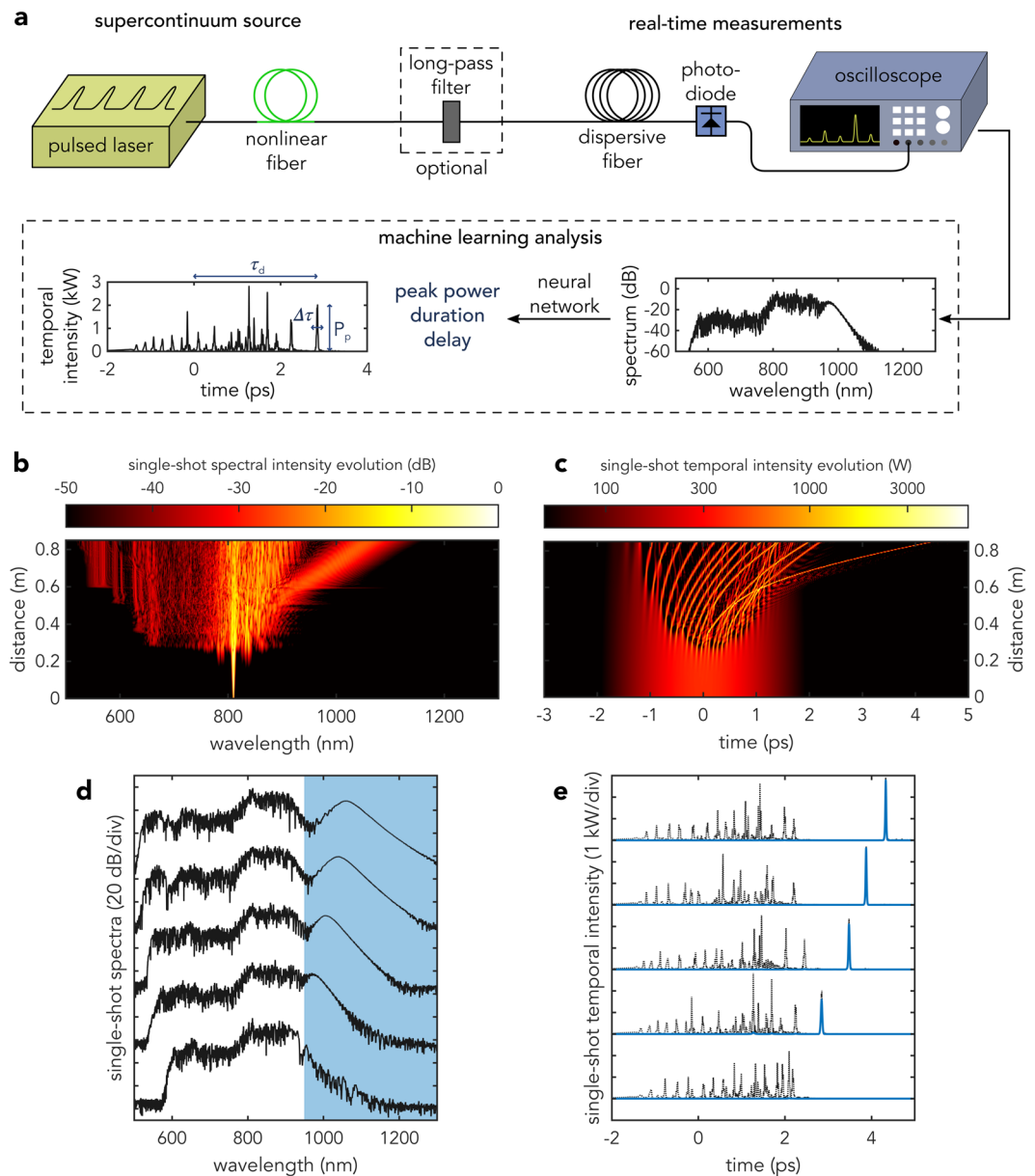


Figure 1. (a) Schematic of machine learning application to the analysis and prediction of rogue soliton temporal characteristics in supercontinuum generation based only on single-shot spectral measurements. (b–e) Dynamics of noisy supercontinuum generation. Single-shot spectral (b) and temporal (c) evolution where we see enhanced spectral broadening associated with the emergence of a rogue soliton. (d,e) show a selection of 5 random supercontinuum spectra and corresponding temporal intensity profiles in black. The blue curve in (e) shows the rogue soliton temporal intensity associated with the long-pass filtered spectral components (blue area) in (d) while the black dotted line shows the full time-domain intensity.

After the training, the NN is tested using an independent ensemble of 10,000 simulated single-shot SC spectra that were not used in the training phase. The results are shown in Fig. 3, where we compare the peak power [Fig. 3(a)], duration [Fig. 3(b)] and temporal delay [Fig. 3(c)] of the most red-shifted soliton predicted by the NN from the single-shot SC spectra and the expected (“ground truth”) value extracted directly from the simulated time-domain profiles. For convenient visualization, the comparison is plotted as a false colour representation of a histogram (using a logarithmic scale) where the predictions are grouped into bins of constant area, and the histogram shows the normalized density of data points grouped into each bin. For all different characteristics, we can see near-perfect clustering around the ideal “ $x=y$ ” relation (indicated by the white dashed line) with a Pearson correlation coefficient of $\rho = 0.91$, $\rho = 0.84$ and $\rho = 0.91$ for the peak power, duration and temporal delay, respectively (yellow dashed line).

We also plot the peak power of the most red-shifted solitons against simpler possible “predictive” metrics such as the wavelength corresponding to the -30 dB SC bandwidth on the long-wavelength edge [Fig. 3(d)], and the integrated energy in the long-wavelength edge beyond 950 nm [Fig. 3(e)]. In contrast to the strong correlations

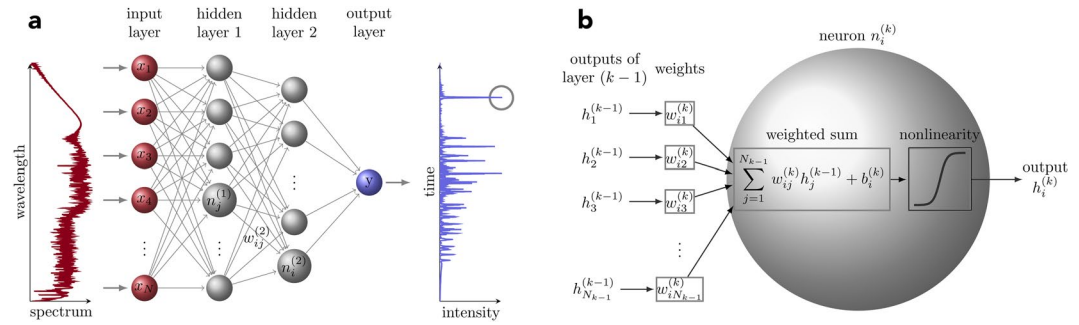


Figure 2. Schematic of the feed-forward neural network used in this work. **(a)** The input of the network is a single-shot supercontinuum spectral intensity vector $\mathbf{X}_n = [x_1, x_2, \dots, x_N]$ yielding the output of the network y that corresponds to the peak power, duration or temporal delay of the rogue solitons. The network consists of two fully connected hidden layers and a single output neuron. **(b)** shows the operation of a single neuron. The output of a generic neuron $n_i^{(k)}$ (i^{th} neuron in layer k) is calculated as a weighted sum between the outputs from the previous layer $k - 1$ and the weights of each connection $w_{ij}^{(k)}$ which is followed by adding a bias term $b_i^{(k)}$ and nonlinearity. See Methods for more details.

obtained using the NN, we see that the correlation between these simple metrics and the most red-shifted soliton peak power is extremely poor with correlation coefficients of $\rho = 0.39$ and $\rho = 0.21$, respectively. These results clearly illustrate the trained NN's capability to accurately predict the characteristics of the most red-shifted soliton from single shot SC spectra without any spectral phase information, and its superiority compared to simpler metrics such as the spectral bandwidth or filtered spectral energy.

To further evaluate the performance of the NN, we ran a series of additional tests where the NN was trained from an ensemble of SC spectra with large variations in the input pulse parameters. More specifically, we added a uniform $\pm 50\%$ variation in the input pulse peak power and duration from 200 W to 600 W and from 1 ps to 3 ps such that the training ensemble now contained a large variety of SC development scenarios from essentially pure modulation instability to broad octave-spanning spectra. Using an ensemble of 20,000 simulations, the neural network was again trained to relate the temporal properties of the most red-shifted soliton with the full SC spectral intensity profile, and the network was tested on a separate set of 10,000 simulations not used in the training phase. The comparison between the peak power, duration and delay of the most red-shifted soliton predicted by the NN from the single-shot SC spectra and the “ground truth” value extracted directly from the time-domain profiles are shown in Fig. 4. Despite the large variation in the propagation dynamics, the peak power and delay are again predicted with a very high degree of accuracy with a correlation coefficient of $\rho = 0.90$, and $\rho = 0.99$, respectively, and the soliton duration with a lower value of $\rho = 0.75$ nonetheless still indicates strong correlation. For completeness, we also plot the most red-shifted soliton peak power vs. the SC spectral bandwidth and energy beyond a wavelength of 950 nm. Again, one can see how the NN performs extremely well, and again with a prediction capacity far superior to the use of a simpler metrics ($\rho = 0.60$ and $\rho = 0.39$, respectively).

Conclusion

We have applied machine learning to the analysis of supercontinuum instabilities in the long pulse regime and our results expand previous studies of nonlinear dynamics using machine learning³⁰. Using a feed-forward neural network trained on numerical simulations of the GNLS, we have shown how the temporal characteristics of solitons with extreme red-shift (peak power, duration and temporal delay) can be predicted with high accuracy based only on single-shot SC spectral intensity profiles without any spectral phase information. The network is able to accommodate and maintain high accuracy for a wide range of regimes, with far superior performance when compared to using spectral metrics such as bandwidth or energy over a specific wavelength range. This shows the potential for machine learning to overcome the limitations of current experimental techniques by allowing the temporal characteristics of ultrashort localized structures to be determined from only spectral intensity measurements.

Machine learning is particularly suited to the study of complex nonlinear systems, not only because of its ability to extract specific properties from hidden features in a data ensemble, but also because of the possibility to link the system parameters with particular dynamical behavior. This is of fundamental interest for nonlinear optical systems governed by NLSE-type equations where localized structures and short pulses are challenging to capture in real-time, but also for studies of rogue waves and extreme events in other physical systems where important quantitative information may be masked in noisy or partial measurements. Moreover, machine learning provides new opportunities to control and optimize complex nonlinear systems and we anticipate it will become a key approach for the design and optimization of supercontinuum spectra and frequency combs tailored to specific applications.

Methods

Numerical modelling. Numerical modelling is based on the generalized NLSE that describes the propagation of the envelope of an optical field¹. We model the propagation of 2 ps duration (FWHM) and 400 W peak power hyperbolic-secant optical pulse centered at 810 nm. The pulse is injected into the anomalous dispersion regime of a 85 cm long photonic crystal fiber (similar to NKT Photonics NL-PM-750) with zero-dispersion

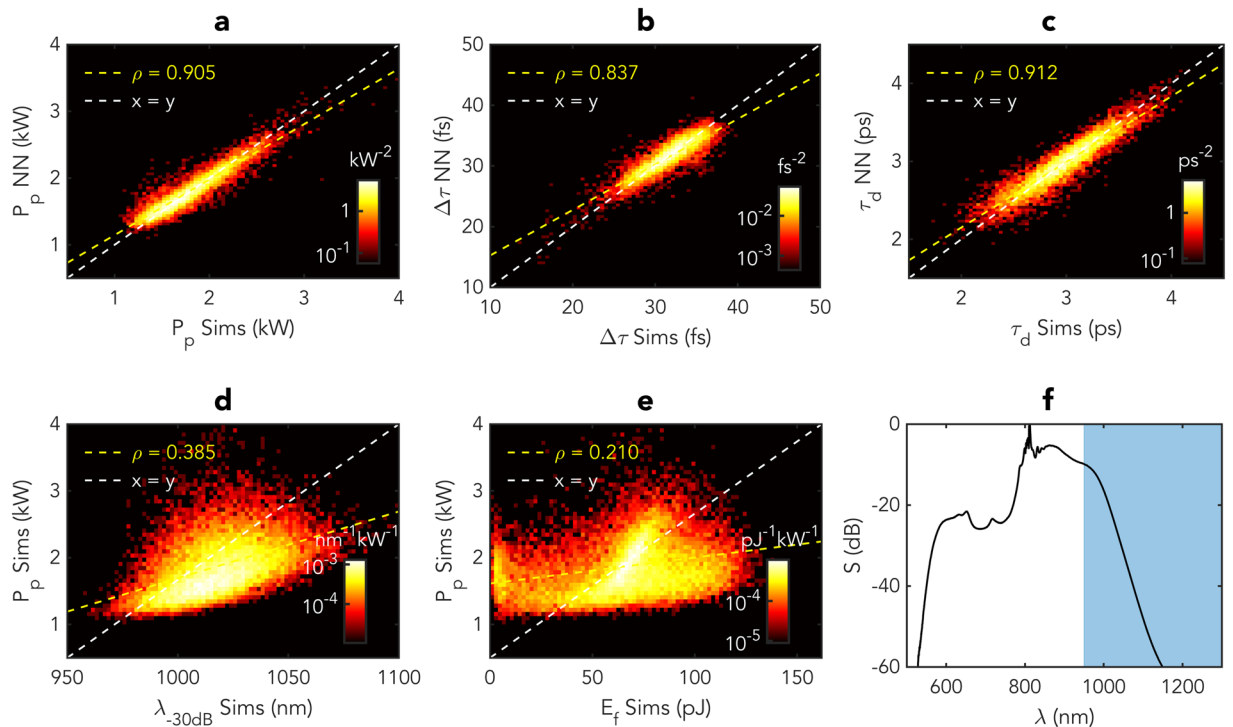


Figure 3. Results showing prediction of the rogue soliton characteristics by the neural network for an ensemble of 10,000 SC ensemble generated with identical input pulses (except for the noise seed). **(a)** compares the predicted maximum intensity of the most red-shifted soliton with the exact value simulated time-domain profiles. **(b)** compares the predicted duration $\Delta\tau$ of the solitons with the exact value from the simulations. **(c)** compares the predicted delay τ_d of the solitons with the exact value from the simulations. **(d)** plots the relation between the SC spectral bandwidth and peak power of the most red-shifted soliton. **(e)** plots the relation between the SC energy beyond 950 nm corresponding to the shaded area in **(f)** and peak power of the most red-shifted soliton. **(f)** shows the average spectrum of the data ensemble. For all sub-figures, the logarithmic histograms show the normalized density of points grouped into bins of constant area. The dashed white line in each case marks the 1-to-1 correspondence and the yellow dashed line is a linear fit to the predictions along with the Pearson correlation coefficient ρ .

wavelength (ZDW) at 750 nm. The fiber has a second ZDW beyond 1200 nm. The nonlinear coefficient of the fiber is $\gamma = 0.1 \text{ W}^{-1} \text{ m}^{-1}$ and the Taylor-series expansion coefficients of the dispersion at 810 nm are $\beta_2 = -1.24 \times 10^{-26} \text{ s}^2 \text{ m}^{-1}$, $\beta_3 = 8.94 \times 10^{-41} \text{ s}^3 \text{ m}^{-1}$, $\beta_4 = -2.54 \times 10^{-56} \text{ s}^4 \text{ m}^{-1}$, $\beta_5 = -7.01 \times 10^{-70} \text{ s}^5 \text{ m}^{-1}$, $\beta_6 = 2.28 \times 10^{-84} \text{ s}^6 \text{ m}^{-1}$ and $\beta_7 = -2.21 \times 10^{-99} \text{ s}^7 \text{ m}^{-1}$. The simulations use 16384 spectral/temporal grid points with a temporal window of 20 ps. Noise is added in the frequency domain in the form of a one-photon-per-mode with random phase. An ensemble of 30,000 simulations corresponding to different input noise seeds was generated. The ensemble was split between two sub-ensembles of 20,000 and 10,000 realizations used for the training and testing, respectively. For the generalization of the training, both the pulse duration and peak power were randomly and uniformly distributed with $\pm 50\%$ variations from the nominal values above, resulting in dynamics from essentially pure MI to octave-spanning SC.

Deep learning. The neural networks relies on supervised learning where one has the knowledge of the relation between the input X_n and output Y_n of a specific system³³. The training is then based on feeding a large number of distinct example input and output pairs to create a predictive model that minimizes the prediction error ε between the desired Y and predicted Y^* values. Here we have a training set of 20,000 simulations, where the input is a spectral intensity profile X_n ($n = 1 \dots 20,000$) and the output Y_n is a temporal characteristics associated with this spectral intensity profile (most red-shifted soliton peak power, duration or temporal delay). The input spectra are pre-processed to a resolution of 1 nm to reduce the computational load in the training process and allow the possibility for reasonable future experimental applications such as in ref.³⁰.

Thus, the input consists of 801 uniformly distributed wavelength bins from 500 nm to 1300 nm. The input is sequentially fed through the NN with each of the layers operating on the data to yield the desired output at output layer. The connections between the nodes on following layers are weighted and the output of each node is computed as a weighted sum from the output of the previous layer. Additionally, an adjustable bias term is included for the sum and followed by a nonlinear activation function to yield the output of the node.

The NN consists of the input layer, two fully-connected hidden layers with 80 and 20 nodes, respectively, and a single output node. The output of a generic k th layer $\mathbf{h}^{(k)} \in \mathbb{R}^M$ can be calculated by

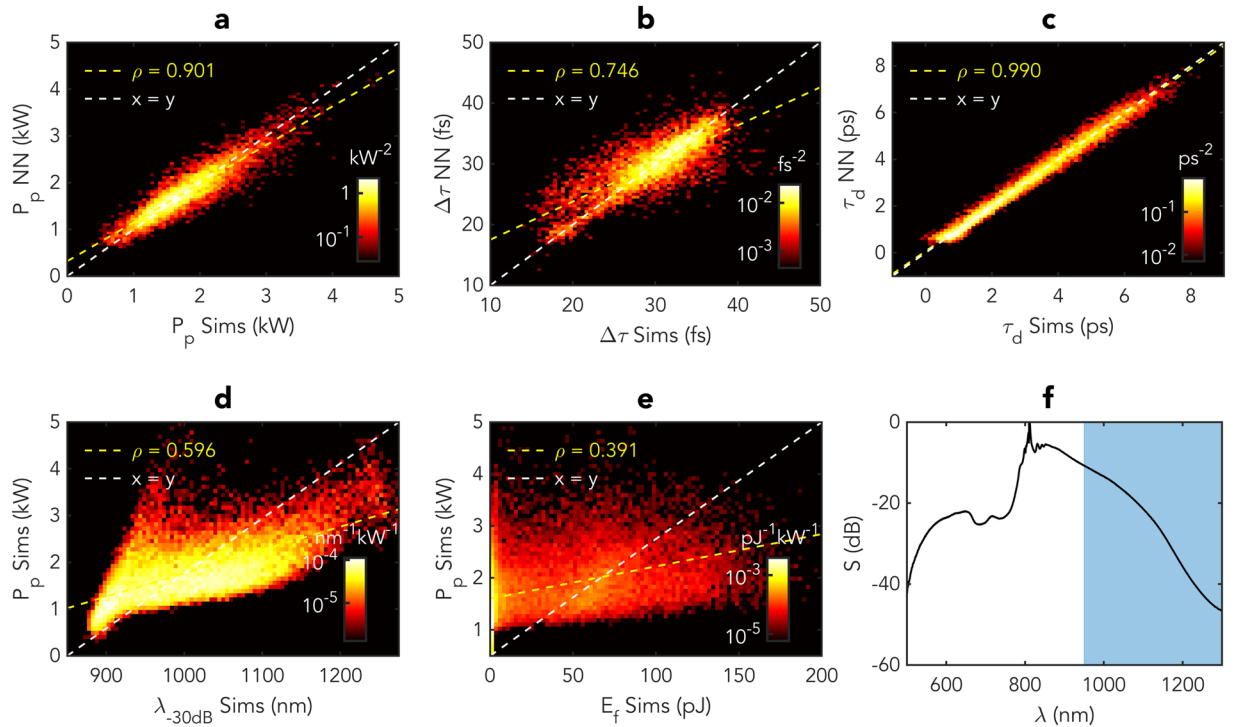


Figure 4. Results showing prediction of the rogue soliton characteristics by the neural network for an ensemble of 10,000 SC ensemble generated with large ($\pm 50\%$) variations in the input pulse peak power and pulse duration. **(a)** compares the predicted maximum intensity of the most red-shifted soliton with the exact value simulated time-domain profiles. **(b)** compares the predicted duration $\Delta\tau$ of the solitons with the exact value from the simulations. **(c)** compares the predicted delay τ_d of the solitons with the exact value from the simulations. **(d)** plots the relation between the SC spectral bandwidth and peak power of the most red-shifted soliton. **(e)** plots the relation between the SC energy beyond 950 nm corresponding to the shaded area in **(f)** and peak power of the most red-shifted soliton. **(f)** shows the average spectrum of the data ensemble. For all sub-figures, the logarithmic histograms show the normalized density of points grouped into bins of constant area. The dashed white line in each case marks the 1-to-1 correspondence and the yellow dashed line is a linear fit to the predictions along with the Pearson correlation coefficient ρ .

$$\mathbf{h}^{(k)} = f(\mathbf{g}^{(k)}) = f(\mathbf{W}^{(k)}\mathbf{h}^{(k-1)} + \mathbf{b}^{(k)}), \tag{1}$$

where $\mathbf{W}^{(k)} \in \mathbb{R}^{M \times D}$ is a matrix of weights between the layers $k - 1$ (D nodes) and k (M nodes). The vector $\mathbf{b}^{(k)} \in \mathbb{R}^M$ contains the bias terms for each node in layer k and $f()$ is the activation function. In this work, hidden layers were connected by hyperbolic tangent sigmoid activation function $f(x) = 2/[1 + \exp(-2x)] - 1$, and a single node with linear activation function was used for the output layer. The output for a single generic node in layer k can be written as

$$h_i^{(k)} = f\left(\sum_{j=1}^{N_{k-1}} w_{ij}^{(k)} h_j^{(k-1)} + b_i^{(k)}\right). \tag{2}$$

Here $w_{ij}^{(k)}$ are the weights between nodes i and j in layers k and $k - 1$, respectively. Variable $b_i^{(k)}$ is the bias term associates with node $n_i^{(k)}$. The summation includes all the N_{k-1} nodes in layer $k - 1$. We use mean squared error function

$$\varepsilon = \frac{1}{N} \sum_{n=1}^N (Y_n - Y_n^*)^2, \tag{3}$$

where N is the number of samples, and Y_n and Y_n^* are the target and predicted outputs, respectively. For the training, a conjugate gradient back-propagation³² is used. The weights and biases are adjusted relative to their current state according to their partial derivatives respect to the error function. The “speed” of the adjustment determined by the learning rate or step size η . The adjustment to weight $w_{ij}^{(k)}$ is taken towards the negative gradient of the error function, given by $\Delta w_{ij}^{(k)} = -\eta \partial \varepsilon / \partial w_{ij}^{(k)}$. Once all of the input and output pairs (\mathbf{X}_n, Y_n) in the training set have been passed through the network one *epoch* has passed. The networks were trained for 500 epochs until convergence. After training, the NN is tested with a separate set not used in the training phase to evaluate the performance of the network.

Received: 12 March 2020; Accepted: 30 April 2020;

Published online: 12 June 2020

References

- Dudley, J. M., Genty, G. & Coen, S. Supercontinuum generation in photonic crystal fiber. *Rev. Mod. Phys.* **78**, 1135 (2006).
- Dudley, J. M., Genty, G., Dias, F., Kibler, B. & Akhmediev, N. Modulation instability, akhmediev breathers and continuous wave supercontinuum generation. *Opt. Express* **17**, 21497–21508 (2009).
- Genty, G., Coen, S. & Dudley, J. M. Fiber supercontinuum sources. *J. Opt. Soc. Am. B* **24**, 1771–1785 (2007).
- Solli, D., Ropers, C., Koonath, P. & Jalali, B. Optical rogue waves. *Nature* **450**, 1054 (2007).
- Mussot, A. *et al.* Observation of extreme temporal events in cw-pumped supercontinuum. *Opt. Express* **17**, 17010–17015 (2009).
- Erkintalo, M., Genty, G. & Dudley, J. M. Rogue-wave-like characteristics in femtosecond supercontinuum generation. *Opt. Lett.* **34**, 2468–2470 (2009).
- Erkintalo, M., Genty, G. & Dudley, J. M. On the statistical interpretation of optical rogue waves. *The Eur. Phys. J. Special Top.* **185**, 135–144 (2010).
- Dudley, J. M., Dias, F., Erkintalo, M. & Genty, G. Instabilities, breathers and rogue waves in optics. *Nat. Photonics* **8**, 755 (2014).
- Antikainen, A., Erkintalo, M., Dudley, J. M. & Genty, G. On the phase-dependent manifestation of optical rogue waves. *Nonlinearity* **25**, R73–R83 (2012).
- Dudley, J. M., Genty, G. & Eggleton, B. J. Harnessing and control of optical rogue waves in supercontinuum generation. *Opt. Express* **16**, 3644–3651 (2008).
- Wetzel, B. *et al.* Real-time full bandwidth measurement of spectral noise in supercontinuum generation. *Sci. Reports* **2**, 882 (2012).
- Godin, T. *et al.* Real time noise and wavelength correlations in octave-spanning supercontinuum generation. *Opt. Express* **21**, 18452–18460 (2013).
- Akhmediev, N. *et al.* Roadmap on optical rogue waves and extreme events. *J. Opt.* **18**, 063001 (2016).
- Goda, K. & Jalali, B. Dispersive fourier transformation for fast continuous single-shot measurements. *Nat. Photonics* **7**, 102–112 (2013).
- Mahjoubfar, A. *et al.* Time stretch and its applications. *Nat. Photonics* **11**, 341 (2017).
- Solli, D., Herink, G., Jalali, B. & Ropers, C. Fluctuations and correlations in modulation instability. *Nat. Photonics* **6**, 463 (2012).
- Herink, G., Kurtz, F., Jalali, B., Solli, D. & Ropers, C. Real-time spectral interferometry probes the internal dynamics of femtosecond soliton molecules. *Science* **356**, 50–54 (2017).
- Krupa, K., Nithyanandan, K., Andral, U., Tchofo-Dinda, P. & Grell, P. Real-time observation of internal motion within ultrafast dissipative optical soliton molecules. *Phys. Rev. Lett.* **118**, 243901 (2017).
- Ryczkowski, P. *et al.* Real-time full-field characterization of transient dissipative soliton dynamics in a mode-locked laser. *Nat. Photonics* **12**, 221 (2018).
- Lapre, C. *et al.* Real-time characterization of spectral instabilities in a mode-locked fibre laser exhibiting soliton-similariton dynamics. *Sci. Reports* **9**, 13950 (2019).
- Närhi, M. *et al.* Real-time measurements of spontaneous breathers and rogue wave events in optical fibre modulation instability. *Nat. Commun.* **7**, 13675 (2016).
- Suret, P. *et al.* Single-shot observation of optical rogue waves in integrable turbulence using time microscopy. *Nat. Commun.* **7**, 13136 (2016).
- Tikan, A., Bielawski, S., Szwaj, C., Randoux, S. & Suret, P. Single-shot measurement of phase and amplitude by using a heterodyne time-lens system and ultrafast digital time-holography. *Nat. Photonics* **12**, 228–234 (2018).
- Liu, X., Yao, X. & Cui, Y. Real-time observation of the buildup of soliton molecules. *Phys. Rev. Lett.* **121**, 023905 (2018).
- Jordan, M. I. & Mitchell, T. M. Machine learning: Trends, perspectives, and prospects. *Science* **349**, 255–260 (2015).
- Farfan, C. A., Epstein, J. & Turner, D. B. Femtosecond pulse compression using a neural-network algorithm. *Opt. Lett.* **43**, 5166–5169 (2018).
- Finot, C., Gukov, I., Hammani, K. & Boscolo, S. Nonlinear sculpturing of optical pulses with normally dispersive fiber-based devices. *Opt. Fiber Technol.* **45**, 306–312 (2018).
- Wetzel, B. *et al.* Customizing supercontinuum generation via on-chip adaptive temporal pulse-splitting. *Nat. Commun.* **9**, 4884 (2018).
- Michaeli, L. & Bahabad, A. Genetic algorithm driven spectral shaping of supercontinuum radiation in a photonic crystal fiber. *J. Opt.* **20**, 055501 (2018).
- Närhi, M. *et al.* Machine learning analysis of extreme events in optical fibre modulation instability. *Nat. Commun.* **9**, 4923 (2018).
- Mahnke, C. & Mitschke, F. Possibility of an akhmediev breather decaying into solitons. *Phys. Rev. A* **85**, 033808 (2012).
- Fletcher, R. & Reeves, C. M. Function minimization by conjugate gradients. *The Comput. J.* **7**, 149–154 (1964).
- Schmidhuber, J. Deep learning in neural networks: an overview. *Neural Networks* **61**, 85–117 (2015).

Acknowledgements

G.G. acknowledges the Academy of Finland (298463, 318082, Flagship PREIN 320165). L.S. acknowledges the Faculty of Engineering and Natural Sciences graduate school of Tampere University. J.D. acknowledges the French Agence Nationale de la Recherche (ANR-15-IDEX-0003, ANR-17-EURE-0002).

Author contributions

L.S. performed the numerical simulations. G.G. and J.D. supervised the project. All authors contributed to the data analysis and writing of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to L.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020