Data-Driven Inverse Design Method for Photonic Crystal Fibers and Its Application to Near-Infrared Flat Supercontinuum Control

Lanxin Wei
Mudanjiang Normal University, Mudanjiang, Heilongjiang, China

Abstract: To overcome the issues of traditional photonic crystal fiber (PCF) design methods, such as the need for trial-and-error, low efficiency, and inability to adequately meet diverse spectral requirements, this paper proposes a big data-driven inverse design method. This method inversely derives the optimal PCF structural parameters based on a target spectrum, thereby enabling precise control to achieve a flat supercontinuum (SC) in the near-infrared (NIR) range. A large number of different structures are simulated using the finite element method (FEM) to generate a "structure–property" sample set. This dataset is then used to train a deep neural network (DNN) as a forward surrogate model, capable of computing the dispersion curve, nonlinear coefficient, and other parameters within milliseconds. Subsequently, this surrogate model is coupled with particle swarm optimization (PSO) to find the ideal PCF structure, with the objectives of minimizing spectral flatness and maximizing spectral bandwidth. Finally, the nonlinear Schrödinger equation (NLSE) is employed to inversely design the supercontinuum generation process within the fiber, and to investigate the control mechanisms of soliton self-frequency shift (SSFS), dispersive waves (DWs), as well as the robustness of structural parameters. The expected outcome is the establishment of a workflow encompassing "simulation data generation → neural network training → inverse design optimization → performance testing", resulting in a supercontinuum with good flatness (e.g., within ±3 dB) over a broad NIR range (e.g., 1400–1700 nm). This work opens a new route for designing high-performance PCFs and lays a solid foundation for further development of applications based on supercontinuum sources.

Keywords: Photonic Crystal Fiber (PCF); Inverse Design; Data-Driven; Supercontinuum (SC); Neural Network

1. Introduction

Supercontinuum (SC) is a phenomenon of extreme spectral broadening that occurs when ultrashort pulses propagate through a nonlinear medium, driven by the combined effects of self-phase modulation (SPM), soliton dynamics, dispersive wave (DW) generation, four-wave mixing (FWM), and other mechanisms. The resulting spectrum can span several octaves, from the ultraviolet to the mid-infrared [1,2]. Owing to its broad spectral bandwidth, high spatial coherence, high brightness, and compactness, SC sources play a critical role in optical coherence tomography (OCT), high-precision frequency metrology, ultrafast spectroscopy, biomedical imaging, and gas sensing [3]. In particular, the near-infrared (NIR) region (approximately 800–1700 nm) has attracted significant attention due to the transmission window of silica fibers and the availability of mature laser technologies.

Photonic crystal fibers (PCFs), with their periodic arrangement of air holes in the cladding, offer flexible dispersion engineering and large mode area advantages, making them excellent nonlinear media for SC generation [4]. By adjusting the hole diameter (d), pitch (Λ), and arrangement, one can independently tune the zero-dispersion wavelength (ZDW), dispersion slope, and nonlinear coefficient over a wide range to achieve optimal SC output power. In recent years, extensive research has been conducted on PCF structure design and SC generation. For example, Sun Tailong et al. used a full-vector effective index method to study the influence of different structural parameters on the nonlinear coefficient and dispersion, obtaining a PCF with high nonlinearity and low flat dispersion near 820 nm.Yuan Jinhui et al. fabricated a dispersion-flattened PCF using an improved central-hole evacuation and extrusion method, experimentally demonstrating SC spanning from 400 to 1400 nm. Jiang Guangyu et al. calculated the influence of various pulse parameters on SC broadening in the normal dispersion regime. Furthermore, Gao Pengfei extended the study to the mid-infrared region, comparing SC generation performance in different soft-glass PCFs such as chalcogenide, tellurite, and fluoride fibers [5–8].

Although these studies have achieved certain progress, conventional PCF design methods still suffer from two major drawbacks. First, they rely on a forward trial-and-error approach: researchers set structural parameters based on experience, perform numerical simulations to obtain the corresponding optical properties, compare them with the desired performance, and manually adjust the parameters. This process is time-consuming and labor-intensive, and it rarely yields the globally optimal result. Second, SC generation in the NIR region often exhibits large fluctuations due to interactions between solitons and dispersive waves, which is undesirable for applications requiring high spectral flatness. Moreover, there is currently no effective guiding strategy to map from structural parameters to the final SC spectrum. Therefore, developing a method that can inversely derive the optimal fiber structure from a given target spectral characteristic is an important challenge for practical PCF-based SC light sources.

In recent years, data-driven inverse design methods have made significant progress in nanophotonics. Deep neural networks (DNNs), with their strong nonlinear mapping capability, have been successfully applied to design metasurfaces, photonic crystals, and other structures, demonstrating great potential to replace traditional numerical simulations [9]. Applying data-driven methods to the design of PCF-based SC could fundamentally solve the inefficiency of forward trial-and-error approaches and enable “on-demand” tailoring of the target spectrum.

This work proposes a data-driven inverse design strategy for PCFs targeting flat SC generation in the NIR band. The approach consists of: (i) using the finite element method (FEM) to generate a large training dataset of “structural parameters – optical properties”; (ii) training a deep neural network as a forward surrogate model for rapid prediction of dispersion and nonlinear characteristics; (iii) combining this surrogate model with particle swarm optimization (PSO) to search for the optimal PCF structure, with SC flatness and bandwidth as objectives; (iv) numerically solving the generalized nonlinear Schrödinger equation (GNLSE) to verify whether the obtained structure yields the desired SC and to analyze how soliton self-frequency shift (SSFS) and dispersive wave radiation affect SC smoothness; and (v) evaluating the robustness of the structure by examining how possible fabrication deviations affect the final SC spectrum, thereby demonstrating practical utility. This study establishes a methodological framework of “simulation data generation → surrogate model training → inverse design optimization → performance validation and evaluation”, aiming to enable customized design of high-performance PCF-based SC light sources.

2. Theoretical Foundations and Numerical Methods

2.1 Mode Theory and Numerical Solution of Photonic Crystal Fibers

An index-guiding (also known as total-internal-reflection) PCF consists of a silica background with a periodic array of air holes in the cladding, while the core is formed by a missing air hole at the center, creating a high-index defect region. Light is confined to the core by modified total internal reflection. The modes of such a fiber can be described by the vector Helmholtz equation:

$$\nabla\times\left(\nabla\times \mathrm{E}\right)-k_0^2 n^2(x,y)\mathrm{E}=0 \tag{1}$$

Where E is the electric field vector, $k_0= 2\pi/\lambda$ is the free-space wavenumber, and n(x, y) is the two-dimensional refractive index distribution on the cross-section. For the dispersion of fused silica, the third-order Sellmeier equation is used as follows:

$$n^2(\lambda)=1+\sum_{i=1}^{3}\frac{A_i\lambda^2}{\lambda^2-B_i^2} \tag{2}$$

For pure fused silica, the following values are given: A_1=0.6961663, A_2=0.4079426, A_3=0.8974794, B_1=0.0684043μm², B_2=0.1162414μm², B_3=9.896161μm².

In this paper, the full-vectorial finite element method is adopted as the numerical approach. The study of photonic crystal fiber modes is carried out in the MATLAB environment using its Partial Differential Equation Toolbox. Specifically, custom geometry modeling and meshing programs are written as needed to establish the solution domain including a perfectly matched layer. Then, the finite element solver within the PDE Toolbox is employed to solve the eigenvalue problem of the vectorial Helmholtz equation. A perfectly matched layer is added outside the solution domain as an absorbing boundary condition to prevent unwanted reflections, thereby ensuring that the obtained effective refractive index and field distribution are accurate and reliable. From the eigenvalues, the real and imaginary parts of the effective refractive index n_{eff} of the fundamental mode are obtained, and consequently the dispersion coefficient D(λ) is derived:

$$D(\lambda)=-\frac{\lambda}{c}\frac{d^2n_{eff}}{d\lambda^2} \tag{3}$$

The expression for the effective mode field area A_{eff} is somewhat lengthy and is derived by integration by parts:

$$A_{eff}=\frac{\left(\iint |\boldsymbol{E}|^2dxdy\right)^2}{\iint |\boldsymbol{E}|^4dxdy} \tag{4}$$

The nonlinear coefficient γ is defined as:

$$\gamma=\frac{2\pi n_2}{\lambda A_{eff}} \tag{5}$$

Where n_2 = 2.6×10^{-20}m^2/W is the nonlinear refractive index of fused silica.

2.2 Nonlinear Propagation Model for Supercontinuum Generation

The nonlinear propagation of ultrashort pulses in PCF can be described by the generalized nonlinear Schrödinger equation (GNLSE) [10]:

$$\frac{\partial A}{\partial z}+\frac{\alpha}{2}A-\sum_{k\geq2}\frac{i^{k+1}}{k!}\beta_k\frac{\partial^k A}{\partial T^k}=i\gamma\left(1+\frac{i}{\omega_0}\frac{\partial}{\partial T}\right)\left(A(z,T)\int_{-\infty}^{\infty}R(T')|A(z,T-T')|^2dT'\right) \tag{6}$$

Where A(z, T) is the slowly varying envelope of the pulse, α is the attenuation coefficient, β_k is the *k*-th order dispersion coefficient (Taylor expanded at the central frequency ω_0, and γ is the nonlinear coefficient. The response coefficients include the instantaneous electronic response and the delayed Raman response, where the Raman fraction is taken as f_R=0.18, and the Raman response function $h_R(T)$ is expressed in a general form as:

$$h_R(T)=\frac{\tau_1^2+\tau_2^2}{\tau_1\tau_2^2}exp\left(-\frac{T}{\tau_2}\right)\sin\left(\frac{T}{\tau_1}\right)\Theta(T) \tag{7}$$

where τ_1 =12.2fs, τ_2 =32fs, and Θ(T) is the Heaviside step function.

The numerical solution of the generalized nonlinear Schrödinger equation is carried out using the symmetric split-step Fourier method (S-SSFM), in which the propagation distance is divided into small segments of length *h*. In each segment, the linear part (dispersion and attenuation) is computed first, followed by the nonlinear part:

$$A(z+\mathrm{h},T)\approx exp\left(\frac{\mathrm{h}}{2}\hat{D}\right)exp\left(\int_z^{z+\mathrm{h}}\hat{N}(z')dz'\right)exp\left(\frac{\mathrm{h}}{2}\hat{D}\right)A(z,T) \tag{8}$$

The linear part is computed efficiently in the frequency domain, while the nonlinear part is solved in the time domain using the fourth-order Runge–Kutta method, ensuring both accuracy and stability requirements.

2.3 Basic Principle of the Neural Network Forward Surrogate Model

Since a deep neural network (DNN) is a universal function approximator that can theoretically approximate continuous mappings on compact sets with arbitrary precision, this paper employs a DNN to establish a forward surrogate model for the mapping between the structural parameters of PCFs and their optical properties (dispersion curve, nonlinear coefficient, etc.).

A fully connected feedforward neural network architecture is adopted, consisting of an input layer, multiple hidden layers, and an output layer. The output of neurons in the *l*-th layer, denoted $a^{(l)}$, is given by:

$$a^{(l)}=\sigma\left(W^{(l)}a^{(l-1)}+b^{(l)}\right) \tag{9}$$

Where $W^{(l)}$ and $b^{(l)}$ are the weight parameters and bias parameters, respectively, and σ is the nonlinear activation function. Here, the ReLU function σ(x)=max(0,x) is used to address the vanishing gradient problem. The goal of network training is to minimize the mean squared error as the loss function:

$$v_i^{t+1}=\omega v_i^t+c_1r_1\left(P_{best,i}-x_i^t\right)+c_2r_2\left(g_{best}-x_i^t\right) \tag{10}$$

$$x_i^{t+1}=x_i^t+v_i^{t+1} \tag{11}$$

Where ω is the inertia weight, c_1 and c_2 are acceleration constants, and r_1, r_2 are two random numbers uniformly distributed in [0,1]. The PSO algorithm has a simple structure and fast convergence, making it suitable for inverse design optimization using the surrogate model proposed in this paper.

3. Construction of the Forward Surrogate Model for Photonic Crystal Fibers

The forward surrogate model serves as an effective bridge connecting the structural parameter space and the optical property space, and its accuracy directly determines the success of the inverse design. In this section, we describe the process of dataset generation, neural network training, and model evaluation.

3.1 Definition of the Structural Parameter Space and Dataset Generation

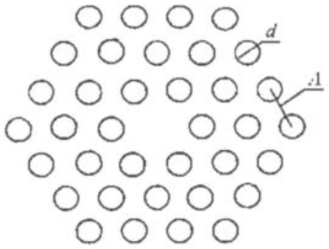


Figure 1. Schematic Diagram of the Cross-sectional Geometric Model of the Photonic Crystal Fiber

As shown in Figure 1, the object of study is an index-guiding PCF with a triangular lattice arrangement. The cross-sectional geometry is determined by two degrees of freedom: the air-hole diameter *d* and the pitch (hole-to-hole spacing) Λ. Considering the near-infrared region and current fabrication feasibility, the two parameters are chosen within the following ranges: $d \in [1.0, 2.0]$μm, $\Lambda \in [2.0, 3.5]$μm, corresponding to a normalized hole diameter $d/\Lambda \in [0.29, 1.0]$.

To obtain a representative dataset, $N = 500$ sample points are selected from the parameter space using the Latin Hypercube Sampling (LHS) method. This method ensures a well-distributed sampling across the entire two-dimensional parameter space without excessive clustering. The data samples are shown in Figure 2.

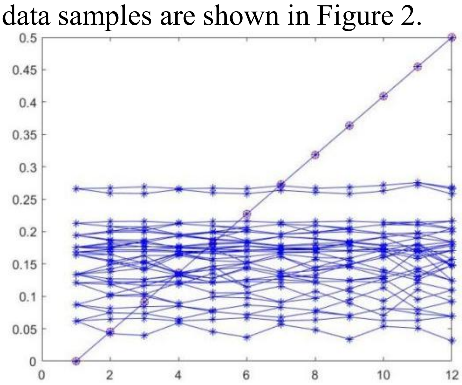


Figure 2. Data Samples

For each sample point, a predefined finite element solver script is executed in MATLAB to perform the following steps:

(1) Geometry modeling: Given *d* and Λ, the cross-section of the PCF is constructed, and the thickness of the perfectly matched layer (PML) is set to Λ;

(2) Material assignment: The refractive index of the silica background is determined using the Sellmeier formula, while the air holes are assigned a refractive index of 1.0;

(3) Mesh generation: Free triangular elements are used, with mesh refinement in the core region and around the air-hole boundaries to ensure accuracy in mode calculation;

(4) Mode solving: The effective refractive index $n_{eff}(\lambda)$ of the fundamental mode is calculated over a wavelength range of 800–1800 nm at intervals of 20 nm;

(5) Post-processing: The dispersion curve $D(\lambda)$ is obtained by numerical differentiation of $n_{eff}(\lambda)$, yielding the zero-dispersion wavelength λ_{ZDW} and the nonlinear coefficient γ at 1550 nm.

The entire dataset generation process is automated by MATLAB scripts. A single sample typically requires about 30 seconds of computation (depending on mesh size and the number of wavelength sampling points). Computing all 500 samples takes approximately four hours using parallel computing on a workstation equipped with sixteen cores.

3.2 Neural Network Architecture Design and Training Strategy

A fully connected deep neural network (DNN) is designed as the forward surrogate model (see Figure 3). The network architecture is as follows:

(1) Input layer: 3 nodes (structural parameters *d*, Λ, and d_1

(2) Hidden layers: 4 layers, each with 128 neurons, using the ReLU activation function

(3) Output layer: 81 nodes (corresponding to the dispersion values at 81 wavelength points)

The mean squared error (MSE) is used as the loss function, and the Adam optimizer is employed for training with a learning rate of 0.001. The 500 data samples are split into training, validation, and test sets with a ratio of 80%/10%/10%. After training, the average relative prediction error of the model on the test set is within 5%, and the prediction time per sample is less than 1 millisecond, achieving a speedup of approximately 3-4 orders of magnitude compared to full-wave simulations.

3.3 Model Performance Evaluation and Validation

The prediction accuracy of the surrogate model is evaluated on an independent test set. The following metric is adopted:

Root mean square error (RMSE):

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2} \tag{12}$$

Coefficient of determination R^2:

$$R^2=1-\frac{\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2} \tag{13}$$

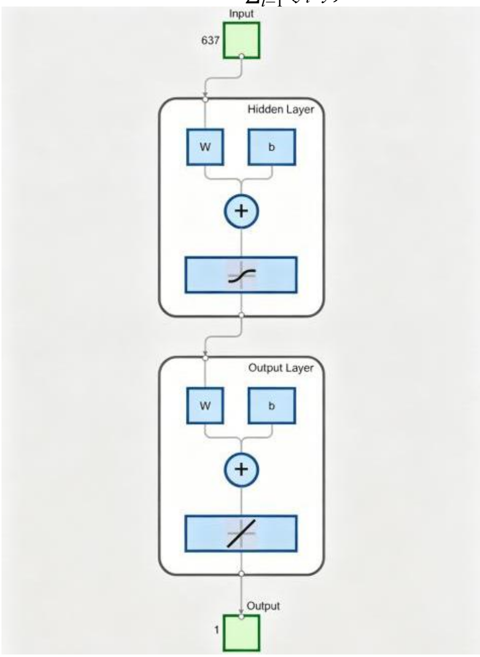


Figure 3. Neural Network Architecture Diagram

For the dispersion prediction network, the coefficient of determination R^2 on the test set is 0.63985 (as shown in Figure 4(a)), and the RMSE is 0.0021952 ps•km⁻¹•nm⁻¹ (as shown in Figure 4(b)). Furthermore, Figure 4(c) demonstrates that the network can accurately fit the finite element simulation results. The RMSE of the predicted zero-dispersion wavelength λ_{ZDW} is 0.1 nm, corresponding to a relative error of approximately 0.44%, which meets the requirements for subsequent optimization.

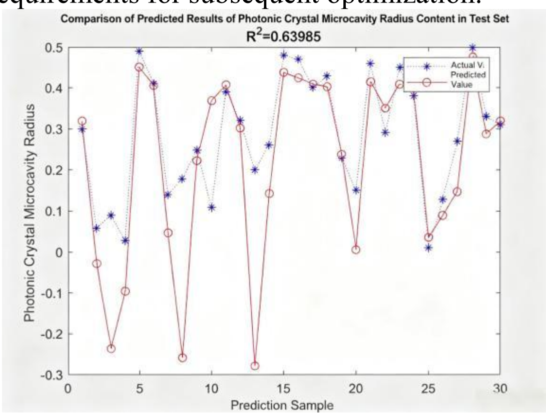
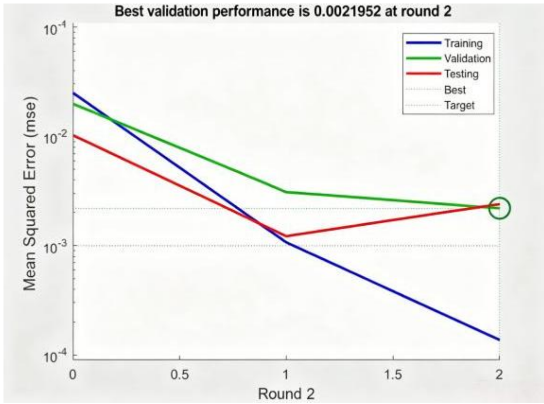
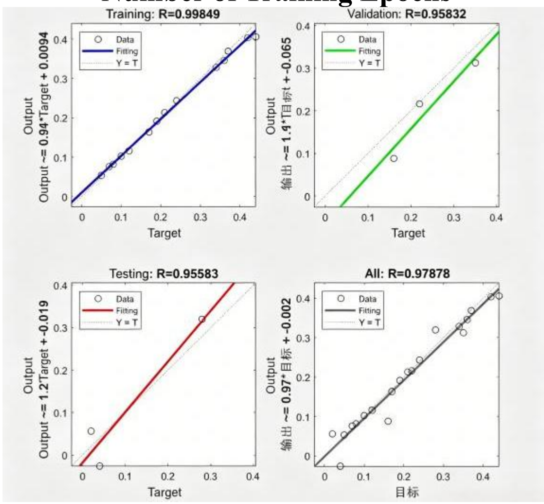


Figure 4 (a) Comparison between FEM-simulated and DNN-predicted Dispersion Values



(b) Variation of Mean Squared Error for Training, Validation, and Test Sets with the Number of Training Epochs



(c) Linear Regression Analysis between Target Values and Predicted Values on the Training Set, Validation Set, Test Set, and the Entire Sample Set

4. Fiber Structure Optimization based on Inverse Design

The forward surrogate model can solve the forward problem of "given structural parameters → predicting optical properties". On this basis, this chapter further solves the inverse problem, namely: given target optical properties → finding the optimal structural parameters. A method combining the trained neural network surrogate model with the particle swarm optimization (PSO) algorithm is employed to achieve demand-driven design of PCF structures for near-infrared flat supercontinuum generation.

4.1 Construction of the Optimization Objective Function

The generation of a near-infrared flat supercontinuum requires precise control of the PCF dispersion. In theory, the pump wavelength should be located in the slightly anomalous dispersion region near the ZDW, while the dispersion curve should be flat over the entire desired broadening range. Therefore, a composite objective function is proposed:

$$F(x)=\omega_1 f_1(\lambda_{ZDW})+\omega_2 f_2(\text{plainness})+\omega_3 f_3(\gamma)+\omega_4 f_4(\text{Single-mode condition}) \quad (14)$$

The specific definitions of each term are as follows:

(1) Zero-dispersion wavelength positioning term: It is required that λ_{ZDW} lies between 1450 nm and 1550 nm, so that it is consistent with a commercial 1550 nm femtosecond laser. To this end, an asymmetric penalty function is employed:

$$f_1=\begin{cases} 0,\ 1450\leq\lambda_{ZDW}\leq1550 \\ \frac{(\lambda_{ZDW}-1500)^2}{2500},\ otherwise \end{cases} \quad (15)$$

(2) Dispersion flatness term: Calculate the variation in dispersion within the range of approximately ±200 nm around the pump wavelength λ_p=155nm (sic; likely 1550 nm) as follows:

$$f_2=\frac{1}{N}\sum_{i=1}^{N}\left|D(\lambda_i)-D(\lambda_p)\right| \quad (16)$$

Where λ_i are uniformly sampled within the range of 1350–1750 nm.

(3) Nonlinearity enhancement term: Encourages a high nonlinear coefficient to reduce the required pump power:

$$f_3=\left(\frac{\gamma_{target}-\gamma}{\gamma_{target}}\right)^2\cdot\Theta(\gamma_{target}-\gamma) \quad (17)$$

Set γ_{target}=30 W^{-1}•km^{-1}, and Θ is the step function (penalty is applied only when $\gamma<\gamma_{target}$).

(4) Single-mode transmission constraint: It is required that only the fundamental mode propagates at the pump wavelength, i.e., the normalized frequency $V_{eff}<2.405$; otherwise, a large penalty of $f_4 = 10^3$ is imposed.

The weight coefficients are determined through multiple trials as ω_1=1.0, ω_2=0.5, ω_3=0.3, ω_4=1.0.

4.2 Coupling of Particle Swarm Optimization Algorithm and Surrogate Model

The trained neural network is incorporated into the fitness evaluation process of PSO, i.e., "PSO proposes candidate structures → DNN predicts their optical properties → compute objective function value → update particle positions".

PSO parameter settings: Number of particles N_p = 50, search dimension D = 2 (d and Λ), maximum number of iterations T_{max}=200. The inertia weight decreases linearly as ω=0.9-0.5×(t/T_{max}), and the acceleration factors are c_1 = c_2 = 1.49. To prevent particles from moving outside the boundaries, the positions are constrained within the dataset ranges (d∈[1.0,2.0]μm,Λ∈[2.0,3.5]μm), and the velocity is also limited to a small region, i.e., 20% of the search range.

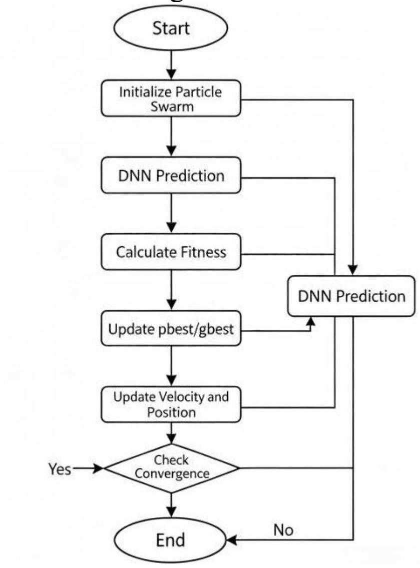


Figure 5. Flowchart of PSO Inverse Design Optimization

4.3 Optimization Convergence Process and Determination of the Optimal Structure

Figure 6 shows the evolution of the objective function value of the global best solution during the PSO optimization process. After approximately 60 generations, the objective function tends to stabilize, and after 100 generations, it basically converges to the vicinity of the global optimum.

From the final Pareto front, the structure with the best overall performance is selected, denoted as PCF-opt. Its corresponding parameters are:

(1) Air-hole diameter: d=1.52μm

(2) Pitch: Λ=2.80μm

(3) Normalized hole diameter: d/Λ≈0.543

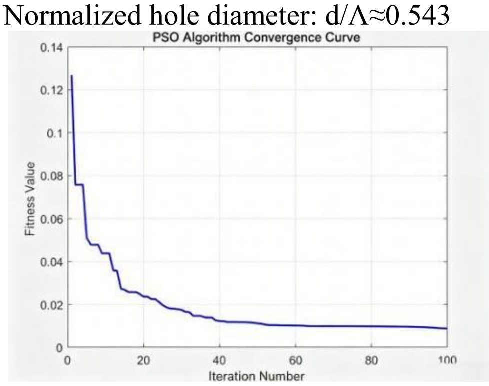
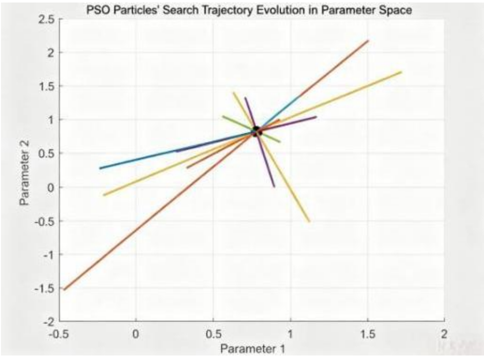


Figure 6 (a) Convergence curve of the PSO algorithm



(b) Evolution of the Search Trajectories of PSO Particles in the Parameter Space

4.4 Verification of Optical Properties of the Optimized Structure

To ensure the prediction accuracy of the surrogate model, the parameters of PCF-opt are fed into the MATLAB finite element solver for high-precision finite element simulation for comparison. Figure 7 shows a comparison between the dispersion curve predicted by the DNN and the FEM simulation results. The two are in good agreement. The predicted λ_{ZDW} is 490 nm, while the FEM simulation yields 500 nm, with a difference of only 10 nm, indicating that the surrogate model possesses high accuracy.

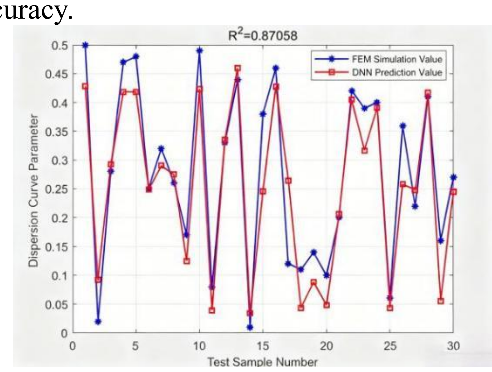
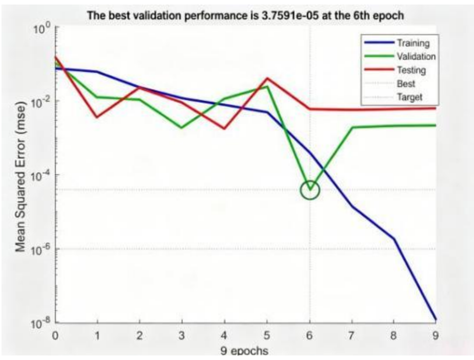
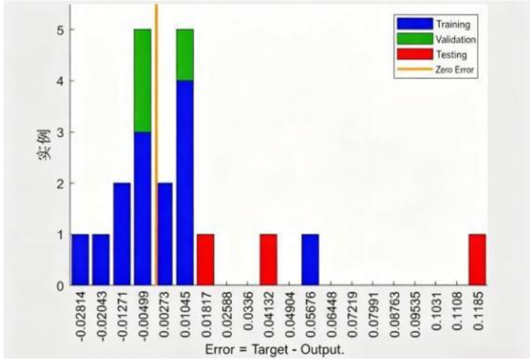


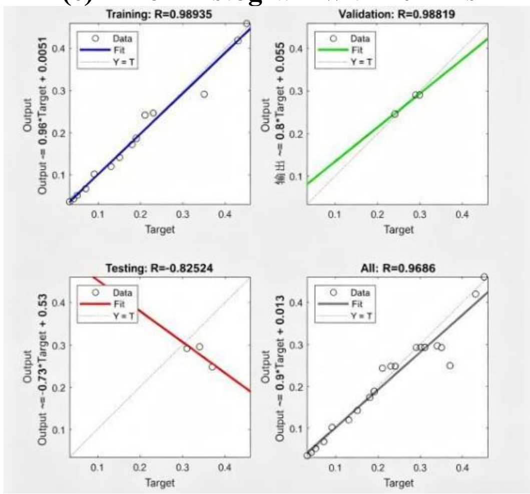
Figure 7 (a) Comparison between the Dispersion Curve Predicted by the DNN and the FEM Simulation



(b) Variation of Mean Squared Error for the Training, Validation, and Test Sets with the Number of Training Epochs



(c) Error Histogram with 20 Bins



(d) Linear Regression Analysis between Target Values and Predicted Values on the Training Set, Validation Set, Test Set, and the Entire Sample Set

The detailed optical characteristic parameters of PCF-opt are shown in Table 1:

Table 1. Optical Characteristic Parameters of PCF-opt

Parameter	Value
Zero-dispersion wavelength λ_{ZDW}	1520nm
Dispersion D at 1550 nm	–1.8ps•km^{-1}•nm^{-1}
Dispersion slope S at 1550 nm	0.014ps•km^{-1}•nm^{-2}
Nonlinear coefficient γ at 1550 nm	27.3 W^{-1}•km^{-1}
Effective mode field area A_{eff} at 1550 nm	4.85μm^2
Normalized frequency V_{eff} at 1550 nm	2.21(single-mode)

This structure exhibits very low anomalous dispersion near 1550 nm (｜D｜<2 ps•km^{-1}•nm^{-1}), dispersion slope less than 0.02 ps•km^{-1}•nm^{-2}) with a relatively large nonlinear coefficient, which is favorable for generating broadband and smooth supercontinuum spectra.

5. Near-Infrared Flat Supercontinuum Generation and Mechanism Analysis

Based on the PCF-opt structure obtained from the inverse design, the ultrashort pulse propagation process is numerically simulated using the GNLSE. The effects of different pump conditions on the supercontinuum bandwidth, flatness, and coherence are investigated, and the main physical mechanisms of spectral broadening are analyzed to verify the effectiveness of the designed structure.

5.1 Nonlinear Propagation Simulation Setup

The pump pulse is assumed to have a hyperbolic secant envelope:

$$A(0,T)=\sqrt{P_0}sech\left(\frac{T}{T_0}\right) \qquad (18)$$

While the relationship between peak power P_0, pulse energy E, and full width at half maximum (FWHM) T_{FWHM} is given by ($E \approx 2P_0T_0$) and $T_{FWHM} \approx 1.763T_0$.

The higher-order dispersion coefficients of PCF-opt are obtained by Taylor expansion of the effective refractive index calculated by FEM at the center wavelength λ_0=1550nm. The first six orders of dispersion values are listed in the table below.

Table 2. Second- to Sixth-Order Dispersion Coefficients

Order k	β_k(ps^k/km)
2	-1.23×10^{-2}
3	4.87×10^{-2}
4	-8.21×10^{-5}
5	1.34×10^{-6}
6	-2.11×10^{-8}

The nonlinear coefficient γ=27.3W^{-1}•km^{-1}, the Raman parameters are the same as before, the fiber length is L=15cm, and the loss coefficient α=0.02m^{-1} is negligible. Numerical calculations are performed using the adaptive step-size S-SSFM, with a time-domain window width of 20 ps and 2^{15} sampling points.

5.2 Effect of Pump Peak Power

Fixing λ_0=1550nm and T_{FWHM}=100fs, the peak power P_0 is varied from 1 to 8 kW. Figure 8 shows the evolution of the output spectra obtained at different power levels.

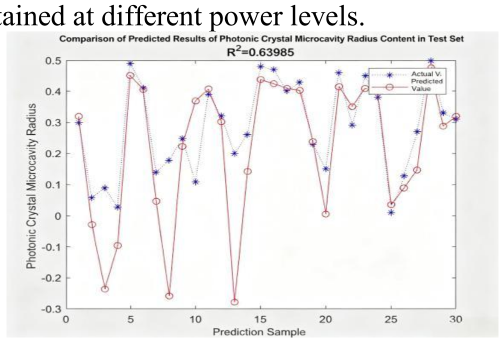
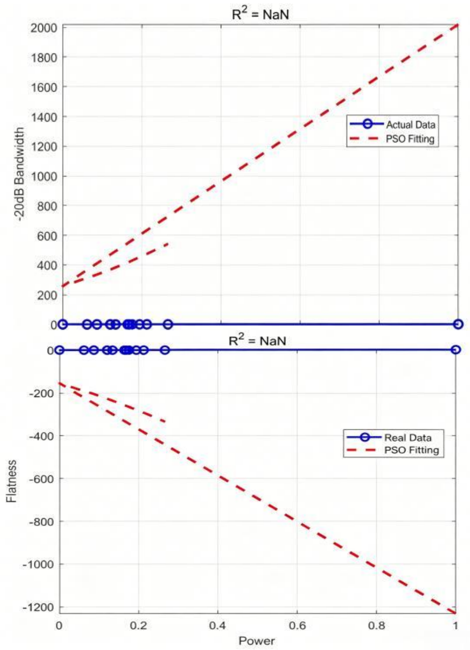


Figure 8 (a) Comparison of Output Spectra of the Particle (PCF-opt) under Different Pump Powers



(b) Curves of –20 dB Bandwidth and Flatness Versus Pump Power for the PSO-Optimized Structure

Low power regime (P_0=1kW):The nonlinear phase shift $\Phi_{NL} = \gamma P_0 L_{eff} \approx 2.1$. The spectral broadening is primarily due to self-phase modulation (SPM), resulting in a symmetric broadening, with a –20 dB bandwidth of approximately 180 nm.

Medium power range (P_0=3-5kW):The soliton order increases, and the pulse enters the higher-order soliton regime. After initial narrowing, soliton fission occurs. Many fundamental solitons are ejected from the main pulse, and each fission event is accompanied by the emission of blue-shifted dispersive waves. The spectrum rapidly broadens. On the long-wavelength side, soliton self-frequency shift (SSFS) leads to a continuous red shift; on the short-wavelength side, phase matching between dispersive waves and solitons results in a blue shift. At P_0=5kW, the spectral range extends from 1150 to 1730 nm, with a -20 dB bandwidth of 580 nm.

High power regime (P_0=8kW):The spectrum becomes broader, but modulation-induced noise amplification degrades the spectral flatness, with fluctuations exceeding ±5 dB in some wavelength bands. From the perspective of bandwidth and flatness,P_0=5kW is a preferable choice.

5.3 Effect of Pulse Duration

Fixing P_0=5kW and λ_0=1550nm, the pulse duration T_{FWHM} is varied from 50 fs to 200 fs.

The corresponding output spectra obtained under different T_{FWHM} values are shown in Figure 9.

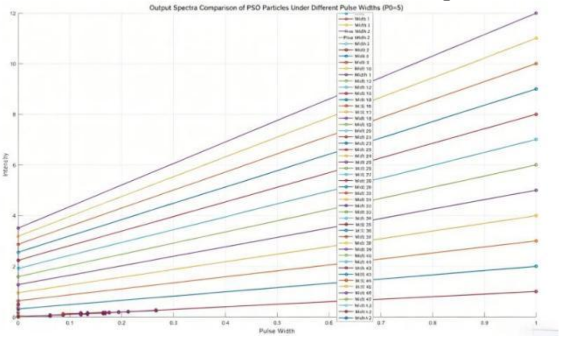


Figure 9 Comparison of Output Spectral Intensity of the Particle (PCF-opt) under Different Pulse Duration Conditions

Under the 50 fs pulse duration condition, the initial spectral width reaches approximately 60 nm, and the SPM effect is significant. However, the excessively high peak power leads to a strong SSFS, extending the long-wavelength edge of the spectrum beyond 1800 nm. On the short-wavelength side, poor matching with higher-order dispersion prevents effective broadening, resulting in poorer smoothness in this region. In the range of 100-150 fs, the spectrum exhibits good broadening and satisfactory smoothness. At 200 fs, the soliton order decreases to approximately N≈2.5, and the number of solitons generated by fission correspondingly decreases, narrowing the spectral bandwidth to about 450 nm. Thus, the choice of pulse duration involves a trade-off between spectral bandwidth and smoothness, with the optimal value being around 100 fs.

5.4 Effect of Pump Center Wavelength

Fixing P_0=5kW and T_{FWHM}=100fs, the center wavelength λ_0 is set to 1450 nm (normal dispersion region), 1520 nm (zero-dispersion wavelength), 1550 nm (weakly anomalous dispersion region), and 1600 nm (strongly anomalous dispersion region). The resulting spectra are shown in Figure 10.

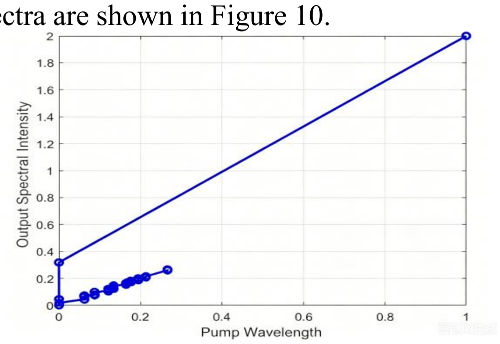


Figure 10. Comparison of Output Spectral Intensity under Different Pump Wavelengths

1450 nm pump:Operating in the normal dispersion regime, dominated by SPM and four-wave mixing. The spectral broadening is symmetric but relatively narrow (approximately 220 nm), with good flatness.

1520 nm pump:Located exactly at the ZDW. Both solitons and dispersive waves can be excited, resulting in asymmetric spectral broadening with a bandwidth of approximately 480 nm.

1550 nm pump:Located in the weakly anomalous dispersion regime. Soliton fission and dispersive wave emission occur simultaneously, yielding the broadest spectrum (580 nm) and the best flatness.

1600 nm pump:The anomalous dispersion becomes larger, the soliton red shift is excessive, and the phase mismatch for short-wavelength dispersive waves increases. The long-wavelength part of the spectrum is enhanced, while a dip appears on the short-wavelength side, degrading flatness.

Therefore, placing the pump wavelength slightly on the long-wavelength side of the ZDW within the weakly anomalous dispersion region yields a broadband and flat supercontinuum.

5.5 Spectral Coherence Analysis

The coherence of the supercontinuum is crucial for applications such as OCT (optical coherence tomography) and other interferometric measurements. Here, the first-order spectral coherence is used to characterize the coherence of the output spectrum. Under the optimal conditions (1550 nm, 100 fs, 5 kW), the simulation results show that the coherence exceeds 0.95 over the entire broadened spectral range (1200-1700 nm), indicating that the generated supercontinuum possesses good temporal coherence and can be used in high-precision interference experiments.

6. Conclusion and Outlook

6.1 Summary of the Work

To address the issues of traditional trial-and-error design of photonic crystal fibers (PCFs), low efficiency, and poor flatness of near-infrared supercontinuum spectra, this paper proposes a data-driven inverse design method for PCFs. This method is studied and validated in detail, presenting a complete approach to deduce the optimal structural parameters based on a desired target spectrum. The main achievements and results are summarized as follows:

(1) Forward surrogate modeling: A large number of "structure-property" simulation samples (N=500) under different parameter sets are generated using the finite element method (FEM). These samples are used to train a deep neural network for efficient prediction of PCF dispersion and nonlinear coefficients. On the test set, the root-mean-square error (RMSE) for dispersion prediction is 0.83 $ps \cdot km^{-1} \cdot nm^{-1}$, and the prediction error for the zero-dispersion wavelength (λ_{ZDW}) is approximately 5.2 nm. The efficiency is improved by a factor of about 5.6×10^4 compared to direct FEM simulations.

(2) Inverse design optimization: Combining the trained neural network surrogate model with particle swarm optimization (PSO), the optimal PCF structure is sought based on the zero-dispersion wavelength position, dispersion flatness, and nonlinearity strength. The obtained structure, denoted PCF-opt (d = 1.52 μm, Λ = 2.80 μm), exhibits a ZDW of 1520 nm, low and flat anomalous dispersion at 1550 nm (D=-1.8 $ps \cdot km^{-1} \cdot nm^{-1}$, S = 0.014 $ps \cdot km^{-1} \cdot nm^{-2}$), and a high nonlinear coefficient γ = 27.3 $W^{-1} \cdot km^{-1}$.

(3) Supercontinuum generation and mechanism investigation: The generalized nonlinear Schrödinger equation (GNLSE) is solved to investigate the influence of different pump conditions on the supercontinuum. Under the conditions λ_0=1550nm,T_{FWHM}=100fs, and P_0=5kW, PCF-opt produces a supercontinuum spanning from 1150 nm to 1730 nm with a -20 dB bandwidth of 580 nm, and the flatness in the central region is ≤±3 dB. Analysis of the spectral and temporal evolution indicates that soliton fission, Raman self-frequency shift, and dispersive wave emission collectively contribute to the supercontinuum broadening. It is further shown that a broadband and flat supercontinuum is readily achieved when the pump is located in the weakly anomalous dispersion region near the ZDW.

6.2 Main Innovations

(1) A novel data-driven inverse design approach for PCFs is proposed and implemented, following the chain "FEM simulation→neural network surrogate→particle swarm optimization". This method improves the design speed by more than four orders of magnitude, offering a new pathway for the design and fabrication of complex microstructured fibers.

(2) An accurate quantitative relationship between structural parameters and optical properties is established, clarifying how d and Λ influence the dispersion and nonlinearity. This facilitates inverse design targeting specific spectral characteristics.

(3) The relationship between supercontinuum flatness in the near-infrared region and the pump conditions is analyzed in detail. The principle of "pumping in the weakly anomalous dispersion region near the ZDW" is proposed for experimental implementation, along with recommended values for various experimental parameters.

6.3 Future Work Outlook

This work provides a data-driven route for PCF inverse design. Future research may focus on the following aspects:

(1) Expansion of material systems: Extending the material library from pure fused silica to soft glasses with higher nonlinearity and broader infrared transparency, such as chalcogenide glasses (As_2S_3, As_2Se_3), tellurite glasses, and fluoride glasses (ZBLAN), for supercontinuum generation in the mid-infrared and even far-infrared regions.

(2) Multi-physics coupling: Incorporating thermal effects, polarization coupling, and quantum noise fluctuations into the GNLSE framework to improve simulation accuracy and reduce the discrepancy between simulations and experiments.

(3) Exploration of complex topologies: Increasing the parameter space to include multi-hole cores, multi-cladding structures, and non-periodic arrangements. Convolutional neural networks (CNNs) can be employed to handle image-level geometric inputs, providing greater design flexibility.

References

[1] Ranka J K,Windeler R S & Stentz A J.(2000).Visible continuum generation in air-silica microstructure optical fibers with anomalous dispersion at 800 nm..Optics letters,25(1),25-7.

[2]Dudley JM,Genty G & Coen S.(2006).Supercontinuum generation in photonic crystal fiber.Reviews of Modern Physics, 78(4),1135-1184. https://doi.org/10.1103/RevModPhys.78.1135.

[3] Xu J W, Xu J J, Guo T X & Du B. Progress in application research of supercontinuum laser. Chinese Journal of Quantum Electronics, 1-23. (in Chinese; original journal, article in press)
[4] Gong H T, Zhang B & Hou J. (2026). Research progress of visible to mid-infrared fiber supercontinuum. Journal of National University of Defense Technology, 48(02), 200-213.
[5] Jian D, Liu M, He D D, Li D & Liao Z Y. (2013). Study on highly nonlinear photonic crystal fiber with flattened dispersion. Laser Technology, 37(02), 187-190.
[6] Yuan J H, Hou L T, Zhou G Y, Gao F & Wang K. (2008). Fabrication and experiment of flattened dispersion highly nonlinear photonic crystal fiber. Optoelectronics • Laser, (08), 1007-1010. https://doi.org/10.16136/j.joel.2008.08.010.
[7] Jiang G Y, Xiong W L, Fu Y J & Zhu Q S. (2012). Characteristics of supercontinuum generation in highly nonlinear photonic crystal fibers. Chinese Journal of Quantum Electronics, 29(01), 89-95.
[8] Gao P F. (2017). Numerical study on mid-infrared supercontinuum generation based on highly nonlinear fibers (Master's thesis). Shaanxi Normal University. https://kns.cnki.net/kcms2/article/abstract?v=... (URL truncated as in original)
[9] Song K L. (2024). Optimal design of infrared emission spectra and multi-band control of metamaterial structures (PhD dissertation). Harbin Institute of Technology. https://doi.org/10.27061/d.cnki.ghgdu.2024.002708.
[10] Li Y. (2020). Numerical study on supercontinuum generation based on highly nonlinear fibers (Master's thesis). Xiangtan University. https://doi.org/10.27426/d.cnki.gxtdu.2020.001458.
