Research Article Vol. 33, No. 16 / 11 Aug 2025 / Optics Express 33421
Optics EXPRESS

Towards scalable and accurate property prediction for photonic crystal fibers with federated learning

GUANJUN WANG,[1] JIANXUN LIU,[2] SHENGCHAO CHEN,[3,4] AND SUFEN REN[2,5]

[1]School of Electronic Science and Technology, Hainan University, Haikou 570228, China
[2]School of Information and Communication Engineering, Hainan University, Haikou 570228, China
[3]Australian Artificial Intelligence Institute, University of Technology Sydney, NSW 2007, Australia
[4]shengchao.chen.uts@gmail.com
[5]Aug_0815@163.com

Abstract: The inverse design of photonic crystal fibers (PCFs) is essential for optimizing fiber optic devices. Although neural network-based methods have shown promise in predicting optical properties from structural parameters, most existing models are tailored to single PCF geometries and lack generalization across diverse structures. To overcome this limitation, we propose a federated learning (FL) framework for predicting optical properties across multiple PCF types. Our approach aggregates data from structurally diverse PCFs and explores two settings: (1) a heterogeneous FL scenario where each client holds data from a distinct PCF geometry, and (2) a non-IID scenario where data from a unified PCF structure is randomly distributed among clients. To handle variations in input/output dimensions across structures, we design a flexible architecture featuring a shared hidden layer with personalized input/output layers, coupled with a dynamic padding strategy to enhance model adaptability. Experiments demonstrate that our method achieves high prediction accuracy (Setting 1: MSE < 0.0085; Setting 2: MSE < 0.0057), with strong convergence and robustness under non-IID conditions. This framework effectively addresses structural diversity and data heterogeneity, mitigating distribution drift and breaking the "data island" limitations of centralized approaches, thereby enabling generalized optical property prediction for a wide range of PCF designs.

1. Introduction

The design and optimization of photonic crystal fiber (PCF) structures are vital for advancing optical sensing [1,2], telecommunication [3,4], and spectroscopy applications [5,6]. Accurate prediction of PCF optical properties—such as dispersion, birefringence, and confinement loss—is essential for enabling performance-driven fiber design. Traditionally, two main approaches have been used for this task: physics-based numerical modeling [7–9] and data-driven learning [10–12].

Physics-based methods, including the finite element method (FEM) [13,14], finite difference method (FDM) [15,16], effective refractive index method (ERIM) [17,18], and plane wave expansion method (PWEM) [19,20], simulate electromagnetic behaviors based on Maxwell's equations. While these methods offer strong interpretability and high accuracy, they are resource-intensive, require expert knowledge for parameter tuning, and scale poorly in high-dimensional design spaces, making them inefficient for iterative design in real-world scenarios.

Recently, data-driven methods based on machine learning and deep learning have emerged as efficient alternatives for PCF property prediction [10–12,21–25]. These models learn direct mappings from structural parameters to optical properties, enabling rapid inference and scalable optimization. For example, Chugh et al. [26] used feedforward neural networks for real-time

#567187 https://doi.org/10.1364/OE.567187
Journal © 2025 Received 6 May 2025; revised 4 Jun 2025; accepted 11 Jun 2025; published 30 Jul 2025

Optics EXPRESS

property prediction, while Jabin et al. [27] developed a deep learning model capable of predicting twelve optical characteristics of silica-based PCFs within milliseconds, achieving a significant speedup over conventional solvers. Despite their success, most of these models are trained on fixed structural templates and lack the ability to generalize across varying PCF geometries. The diversity in structural configurations and physical mechanisms among PCFs leads to distribution shifts that degrade model performance when transferred to unseen structures.

To tackle this generalization challenge, our recent work explores the use of federated learning (FL) [28–30] as a decentralized paradigm for collaborative PCF modeling. In prior studies [31,32], we introduced FL into the PCF design pipeline, demonstrating its ability to preserve data privacy, enable cross-institutional model training, and enhance robustness under data-scarce or resource-constrained settings. However, these works focused on single-structure scenarios, and the problem of multi-structural generalization remained unresolved.

In this paper, we propose a novel FL-based prediction framework specifically designed for multi-structured PCFs. By leveraging heterogeneous data from different PCF configurations, the proposed model addresses the structural generalization gap in existing methods. We formulate two realistic federated settings: (1) **Structurally Heterogeneous FL**, where each client holds data from a distinct PCF geometry, simulating collaboration among design institutions; and (2) **Non-IID data setting**, where samples from a unified PCF structure are randomly distributed across clients, reflecting common deployment challenges. To handle structural inconsistencies—including varying input/output dimensions—we develop a modular neural architecture with a shared hidden layer and personalized input/output layers, supported by a dynamic padding strategy for seamless integration of diverse structures. This architecture enhances both adaptability and scalability, allowing the model to generalize across PCFs with different topologies and optical behaviors. Experiments validate the effectiveness of our approach, showing high predictive accuracy (e.g., MSE < 0.0085 in Setting 1; MSE < 0.0057 in Setting 2) and strong robustness under non-IID conditions. Beyond performance, our method alleviates the "data island" problem of centralized modeling by enabling knowledge aggregation without direct data sharing, offering a practical path toward unified and extensible PCF property prediction across diverse platforms and applications.

The remainder of this paper is organized as follows: Section 2 introduces the dataset construction and preprocessing. Section 3 describes the proposed FL-based methodology. Section 4 presents experimental evaluations and discussions. Finally, Section 5 concludes the paper and outlines future directions.

2. Establish data sets

To support the prediction of optical properties from structural parameters, we construct a comprehensive dataset comprising four representative types of microstructured optical fibers: solid-core photonic crystal fibers (SC-PCF) [26], suspended-core fibers (SCF) [33], air-hole photonic crystal fibers (AH-PCF) [31], and porous-core photonic crystal fibers (P-PCF) [34]. These fiber types were selected for their distinct structural characteristics and widespread applications in sensing and communication. The cross-sectional schematics of each fiber type are illustrated in Fig. 1, highlighting the geometric diversity considered in our framework.

For each fiber type, we construct datasets by pairing structural design parameters with corresponding optical properties derived from numerical simulations. The input features consist of key geometric and material parameters, which vary depending on the specific PCF configuration. For instance:

- Solid-core PCFs (SC-PCF): ir-hole diameter (d), lattice pitch (Λ), core refractive index (n_c), operating wavelength (λ), and number of cladding rings (N_r).

Optics EXPRESS

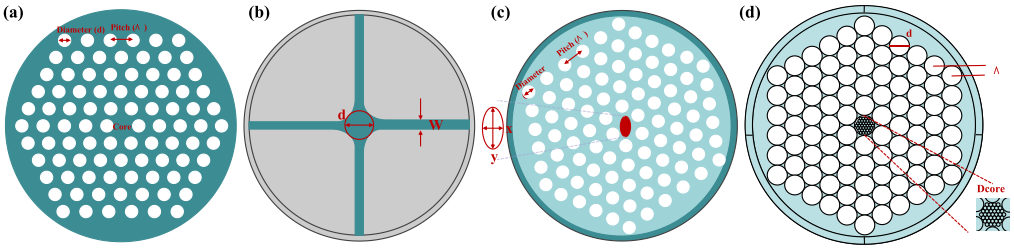

Fig. 1. Schematic cross-sections of four representative photonic crystal fiber (PCF) structures: (a) solid-core PCF (SC-PCF), (b) suspended-core fiber (SCF), (c) air-hole PCF (AH-PCF), and (d) porous-core PCF (P-PCF).

- Suspended-core fibers (SCF): number of supporting bridges (n), bridge width (W), taper radius (r), and wavelength (λ).
- Air-hole PCFs (AH-PCF): centrifugal ratio (e), elliptical core semi-axis (R_m), air-hole diameter (d), and pitch (Λ).
- Porous-core PCFs (P-PCF): core diameter (D_{core}), porosity, and operating frequency (f).

The target outputs include a set of optical properties computed via full-vector numerical solvers. Commonly predicted quantities across all fiber types including Effective refractive index (n_{eff}), Effective mode area (A_{eff}), Chromatic dispersion (D), and Confinement loss (α_c). These properties are defined as follows:

$$\begin{aligned} A_{\text{eff}} &= \frac{\left(\iint_{\Omega} |H_t|^2, dx, dy\right)^2}{\iint_{\Omega} |H_t|^4, dx, dy}, \\ D &= -\frac{\lambda}{c} \cdot \frac{d^2, \text{Re}(n_{\text{eff}})}{d\lambda^2}, \\ \alpha_c &= 8.686 \times 10^6 \cdot k_0 \cdot \text{Im}(n_{\text{eff}}) \quad [\text{dB/m}], \end{aligned} \tag{1}$$

where H_t is the transverse magnetic field, Ω denotes the computational domain, and $k_0 = 2\pi/\lambda$ is the free-space wavenumber. In addition to these fundamental metrics, certain structures require the prediction of nonlinear or polarization-sensitive characteristics. Two commonly used properties are:

$$\begin{aligned} \gamma &= \left(\frac{2\pi}{\lambda}\right) \cdot \left(\frac{n_2}{A_{\text{eff}}}\right), \\ B &= \text{Re}(n_{\text{eff}}^x) - \text{Re}(n_{\text{eff}}^y), \end{aligned} \tag{2}$$

where γ is the nonlinear coefficient and B represents birefringence, computed from the effective indices of orthogonal polarization modes.

Optics EXPRESS

For porous-core PCFs, additional material-related metrics are essential for modeling light–matter interaction. These include:

$$\begin{aligned} \mathrm{EMA} &= \frac{\left(\iint |E(x,y)|^2, dx, dy\right)^2}{\iint |E(x,y)|^4, dx, dy}, \\ \mathrm{EML} &= \sqrt{\frac{\varepsilon_0}{\mu_0}} \cdot \frac{\int_{\mathrm{mat}} n_{\mathrm{mat}} |E|^2 \alpha_{\mathrm{mat}}, dA}{2\left|\int_{\mathrm{all}} S_z, dA\right|}, \\ P_{\mathrm{core}} &= \frac{\int_X S_z, dA}{\int_{\mathrm{all}} S_z, dA}, \\ V &= \frac{2\pi r f}{c} \cdot \sqrt{n_{\mathrm{co}}^2 - n_{\mathrm{cl}}^2} \leq 2.405, \end{aligned} \tag{3}$$

where $E(x, y)$ is the electric field distribution, S_z denotes the longitudinal component of the Poynting vector, and α_{mat} is the absorption coefficient of the core material. The V-parameter characterizes the single-mode condition in waveguides. To improve numerical stability during training, certain properties with large dynamic ranges, such as α_c and EMA, are logarithmically scaled using $\log_{10}$.

Table 1 summarizes the key structural parameters, value ranges, associated optical properties, and dataset sizes for each fiber type. The resulting labeled datasets serve as the foundation for training and evaluating the proposed prediction model, enabling accurate mapping from geometric design to optical performance across structurally diverse PCFs.

Table 1. Summary of datasets for different types of microstructured optical fibers

Fiber Type	Structural Parameters	Optical Properties	Samples
SC-PCF	Hole diameter d: 0.8–2.0 μm (step: 0.2 μm) Pitch Λ: 0.6–0.9 μm (step: 0.1 μm) Wavelength λ: 0.5–1.8 μm (step: 0.0684 μm) Core refractive index $n_c = 1$, rings $N_r \in \{4, 5\}$	$n_{\mathrm{eff}}, A_{\mathrm{eff}}, D, \alpha_c$	1,160
SCF	Wavelength λ: 1–4 μm (step: 0.01 μm) Taper radius r: 0.4–1.1 μm (step: 0.05 μm) Bridge width W: 0.6–1.1 μm (step: 0.1 μm) Number of bridges $n \in \{3, 4, 6\}$	$A_{\mathrm{eff}}, \gamma, D$	71,760
AH-PCF	Wavelength λ: 1–2 μm (step: 0.1 μm) Centrifugal ratio $e \in \{0.33, 0.38, 0.43\}$ Elliptical core semi-axis R_m: 0.4–0.8 μm (step: 0.1 μm) Pitch $p = 2.8$ μm, Hole diameter $d = 1.35$ μm	$A_{\mathrm{eff}}, D, B, \gamma$	94,185
P-PCF	Core diameter D_{core}: 200–300 μm (step: 10 μm) Porosity: 30%–70% (step: 5%) Frequency f: 0.6–1.6 THz (step: 0.05 THz)	EML, P_{core}, $\log_{10}$ CL, $\log_{10}$ EMA, V_{param}	2,859

3. Methodology

3.1. Overall framework

Given the diversity in dimensionality and statistical distribution across different PCF structures, we propose a unified federated learning framework to predict optical properties based on structural parameters, without requiring centralized or mixed data. The method consists of four key stages: data preprocessing, local model construction, global aggregation, and iterative optimization. A

Optics EXPRESS

central feature of the approach is a federated network architecture comprising globally shared hidden layers and structure-specific input/output layers, enabling effective knowledge sharing across heterogeneous PCF datasets.

The overall pipeline is illustrated in Fig. 2. Initially, structural data for each PCF type (refer to Table 1) are assigned to separate clients, forming multiple local datasets. Since the input and output dimensions vary significantly across PCF types, a dynamic zero-padding strategy is applied to unify feature and label dimensions while preserving structural integrity. Each client then trains a local model to capture the nonlinear mapping from PCF geometry to optical characteristics. During training, only the parameters of the shared hidden layers are communicated. These shared layers capture transferable knowledge across clients, while the private input/output layers adapt to the unique structure-specific characteristics of each fiber type. The training procedure unfolds as follows:

- **Step 1: Client initialization.** The available data are split into training and test subsets. Each client is assigned a specific PCF dataset. Zero-padding is used to align the dimensionality of input features $\mathbf{x}$ and target outputs $\mathbf{y}$ across clients.
- **Step 2: Model setup.** Initialize the global model parameters W_0. Set training hyperparameters, including the learning rate η, batch size B, and the loss function $\mathcal{L}$ (e.g., mean squared error for regression).
- **Step 3: Local training.** Each client k receives the current shared model parameters W_t, then trains its local model on its dataset D_k using gradient descent. The model consists of a shared deep hidden backbone and private input/output projection layers. The local weights are updated via:
$$W_t^k := W_t - \eta \nabla \mathcal{L}(W_t, D_k). \tag{4}$$
- **Step 4: Upload.** After local training, each client sends the updated shared parameters W_t^k (excluding private layers) to the central server.
- **Step 5: Global aggregation.** The server performs weighted averaging of received parameters:
$$W_{t+1} = \sum_{k=1}^{K} \frac{n_k}{n} W_t^k, \tag{5}$$
where n_k is the number of samples at client k, and $n = \sum_k n_k$ is the total number of training samples.
- **Step 6: Iterative optimization.** The updated global parameters W_{t+1} are distributed back to all clients for the next round of local training. This loop continues until a convergence criterion is met (e.g., validation performance or maximum rounds).

This federated optimization process ensures structural privacy and scalability while enabling cross-structure generalization. The local and global models are implemented using deep neural networks (DNNs), which are capable of capturing the complex nonlinear dependencies between PCF geometries and their corresponding optical responses. Further architectural details are provided in the next section. The detailed implementation of the proposed federated learning framework for PCF optical property prediction is described in Algorithm 1.

3.2. Dense neural network as local PCF predictors

A key challenge in multi-structure PCF modeling lies in the heterogeneity of input and output dimensions across different fiber types. Directly applying a unified neural network model is infeasible due to dimension mismatch. To overcome this, we design a flexible neural network

Optics EXPRESS

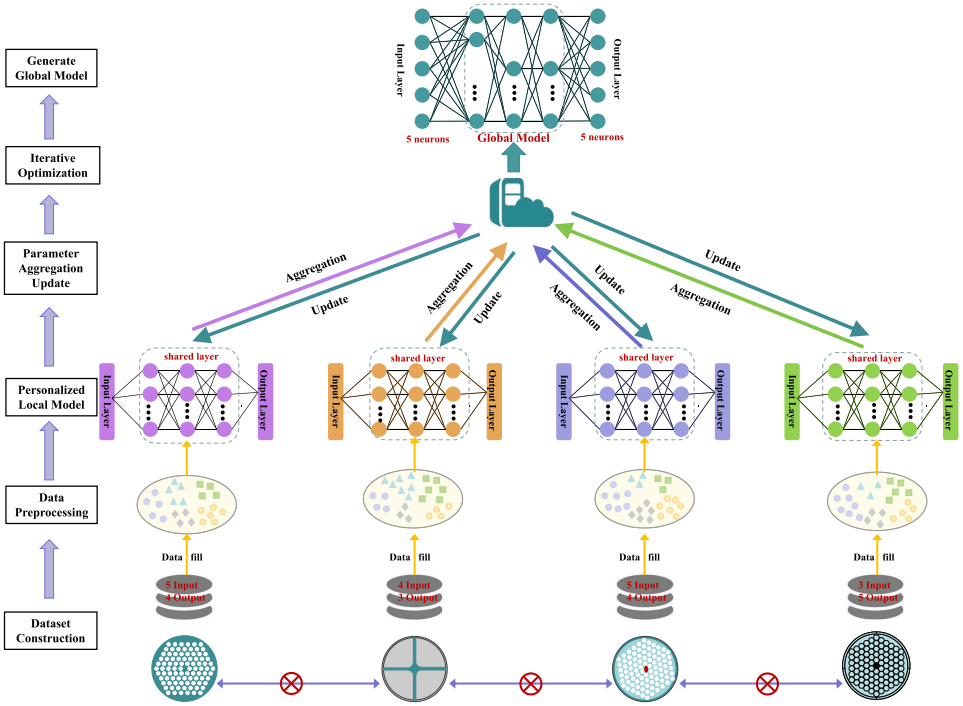

Fig. 2. Overall framework of the FL architecture for optical property prediction.

Algorithm 1. Federated training for PCF optical property prediction

Input: $\{D_k\}_{k=1}^{K}$: local datasets, T: number of global iterations, η: learning rate
Output: W_T: final global model parameters

Initialize global model parameters W_0;
for $t = 0$ **to** $T - 1$ **do**
  foreach *client* $k \in \{1, \ldots, K\}$ ***in parallel*** **do**
    Receive W_t from server;
    $W_t^k \leftarrow W_t$;
    for *each local epoch* **do**
      $W_t^k \leftarrow W_t^k - \eta \nabla \mathcal{L}(W_t^k, D_k)$;
    end
    Send W_t^k to server;
  end
  Server aggregates updates:;
  $W_{t+1} \leftarrow \sum_{k=1}^{K} \frac{n_k}{n} W_t^k$;
end
return W_T

Optics EXPRESS

architecture featuring a shared hidden backbone and personalized input/output layers. This configuration enables knowledge transfer across structures while preserving compatibility with their respective feature spaces.

As depicted in Fig. 3, each local model is composed of three components: an input projection layer, a shared hidden feature extractor, and an output projection layer. To standardize data representations, we introduce a dynamic zero-padding mechanism that expands both input vectors $\mathbf{x}$ and output labels $\mathbf{y}$ to a common dimensionality:

$$D_{\max} = \max(d_1, d_2, \ldots, d_K), \tag{6}$$

where d_k denotes the original dimension of structure k, and K is the total number of distinct PCF types. Shorter feature vectors are padded with zeros to align with $D_{\max}$, ensuring architectural uniformity across clients.

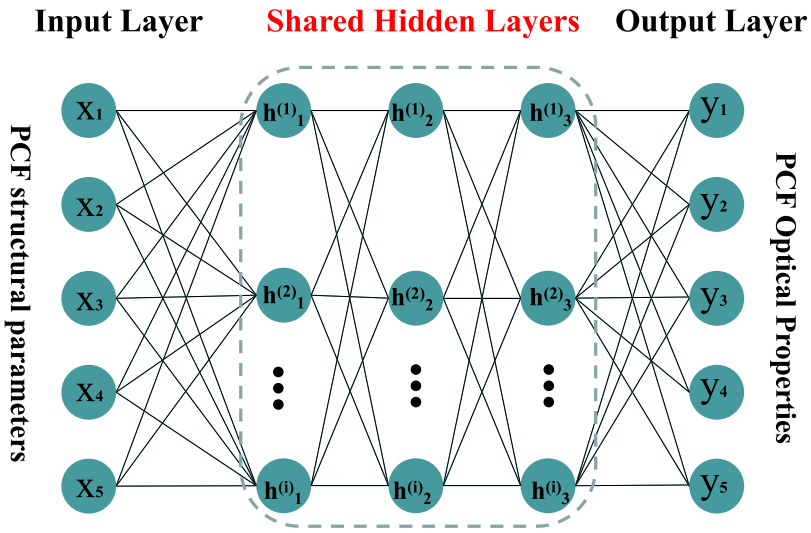

Fig. 3. Architecture of local model on each client.

The shared hidden layers consist of a deep feedforward network with L fully connected layers. Each layer learns a hierarchical representation of structural-to-optical mappings that is generalizable across different PCF designs. The forward propagation through the shared backbone is given by:

$$\begin{aligned} \mathbf{h}^{(1)} &= \sigma\left(W^{(1)}\mathbf{x}' + \mathbf{b}^{(1)}\right), \\ \mathbf{h}^{(l)} &= \sigma\left(W^{(l)}\mathbf{h}^{(l-1)} + \mathbf{b}^{(l)}\right), \quad l = 2, \ldots, L, \end{aligned} \tag{7}$$

where $\mathbf{x}' \in \mathbb{R}^{D_{\max}}$ is the padded input vector, $\sigma(\cdot)$ is a nonlinear activation function (ReLU is used in our implementation), and $W^{(l)}, \mathbf{b}^{(l)}$ are the shared weights and biases at layer l. The output of the final hidden layer $\mathbf{h}^{(L)}$ is fed into a client-specific output projection layer to map the shared representation back to the corresponding output space, accommodating variations in the number and type of optical properties for each PCF structure.

This architecture design offers the following advantages: (1) The shared hidden layers extract a latent representation that captures universal physical relationships between geometry and optics; (2) The personalized input/output layers, combined with zero-padding, allow seamless handling of arbitrary dimensional configurations; and (3) The model naturally integrates into the federated

Optics EXPRESS

learning pipeline, where only shared parameters are synchronized globally, ensuring both privacy and structural decoupling.

4. Experiment and results

4.1. Experiment design

The parameter and hyperparameter values used in this study are determined through extensive empirical tuning and reference to prior work [31,32]. To systematically investigate the influence of network architecture and client participation strategies on the performance of the global model, controlled experiments are conducted under a unified set of training configurations. Specifically, the network architecture and optimizer settings are kept consistent across all tests to isolate the effect of the studied variables. The detailed hyperparameter configurations are summarized in Table 2.

Table 2. Hyperparameter settings for the proposed federated learning framework

Hyperparameter	Value	Description
Hidden layer size	64 / 128 / 256	Number of neurons in shared hidden layers
Epochs	1000	Number of local training epochs per client
Batch size	64	Mini-batch size used during local training
Learning rate (η)	0.0005	Step size for gradient descent optimization
Adam betas	(0.9, 0.99)	Momentum coefficients for Adam optimizer
Communication rounds	10	Total number of global aggregation rounds
Test split ratio	20%	Proportion of each client's data for testing
Random seed	3	Seed for reproducibility in data partitioning
Activation function	ReLU	Nonlinear activation used in all layers
Loss function	MSE	Mean squared error for regression training

To evaluate model performance, three widely adopted regression metrics are employed: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE). These metrics are defined as follows:

$$\begin{aligned} MSE &= \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2, \\ RMSE &= \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}, \\ MAE &= \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right|, \end{aligned} \tag{8}$$

where y_i and $\hat{y}_i$ denote the true and predicted values of the i-th test sample, respectively, and n represents the total number of samples.

4.2. Experiment results and analysis

This section presents a comparative evaluation of federated learning (FL) performance under two distinct experimental settings. **Setting 1 – Structural Heterogeneity:** each client possesses data from a different photonic crystal fiber (PCF) geometry, simulating a scenario where clients operate on structurally diverse physical systems. **Setting 2 – Non-IID Distribution:** all clients use the same PCF geometry, but their local datasets are randomly shuffled to introduce statistical heterogeneity across clients.

Optics EXPRESS

In both settings, parameter aggregation is restricted to shared hidden layers to preserve model flexibility and accommodate structural or distributional differences. Model performance is assessed using MSE, RMSE, and MAE. We also investigate how hidden layer configurations and client participation ratios affect convergence behavior and prediction accuracy.

4.2.1. Results for Setting 1: structural heterogeneity

In this setting, four clients (Client 1–Client 4) are associated with four distinct PCF structures (PCF(1)–PCF(4)). Each local model shares a unified three-layer feedforward neural network architecture, with each hidden layer containing an identical number of neurons. Local datasets are partitioned into training and testing subsets using an 80/20 split.

To investigate the influence of model complexity, we evaluate three hidden layer sizes: 64, 128, and 256 neurons. As shown in Table 3, the best performance is achieved with 64 neurons per layer, offering a strong trade-off between expressiveness and generalization. For example, Client 4 attains an RMSE of 0.0546 and an MAE of 0.0296 under this configuration. Conversely, a smaller model with only 16 neurons (not shown in table) suffers from underfitting—evidenced by the poor performance on Client 3 (RMSE = 0.2426)—highlighting its limited capacity to capture structural complexities. Interestingly, increasing the neuron count to 128 or 256 does not improve performance; instead, it leads to performance degradation, likely due to overfitting in federated settings where each client has limited data. Thus, a configuration of 64 neurons per layer is adopted in all subsequent experiments.

Table 3. Performance of FL with varying hidden layer sizes across clients

Hidden Units	Client	MSE ($\times 10^{-3}$)	RMSE ($\times 10^{-2}$)	MAE ($\times 10^{-2}$)	Time/Iter. (s)
16	Client 1	12.13	10.95	6.22	19.63
	Client 2	9.82	9.89	6.16	
	Client 3	58.94	24.26	17.41	
	Client 4	24.91	15.78	11.15	
32	Client 1	9.33	9.64	7.21	35.14
	Client 2	7.75	8.77	6.28	
	Client 3	42.41	20.61	11.95	
	Client 4	8.83	9.41	6.74	
64	Client 1	5.72	7.58	5.28	93.16
	Client 2	4.14	6.35	3.59	
	Client 3	8.43	9.13	6.91	
	Client 4	2.95	5.46	2.96	
128	Client 1	15.26	12.36	8.53	226.37
	Client 2	12.11	11.01	6.93	
	Client 3	64.15	25.31	15.17	
	Client 4	10.37	10.17	6.43	

To further analyze training dynamics, we plot the loss curves for each client over 10 rounds of federated learning in Fig. 4. Each local model is trained for 1000 steps per round, with loss values recorded every 100 steps, yielding 100 data points per client.

The training curves indicate a rapid decline in loss during the initial rounds, followed by a gradual flattening after round 5. This behavior suggests the model is approaching convergence. Figure 5 presents the error distribution between predicted and true values across the test sets of all four PCF structures. Despite variations among clients, prediction errors are generally confined

Optics EXPRESS

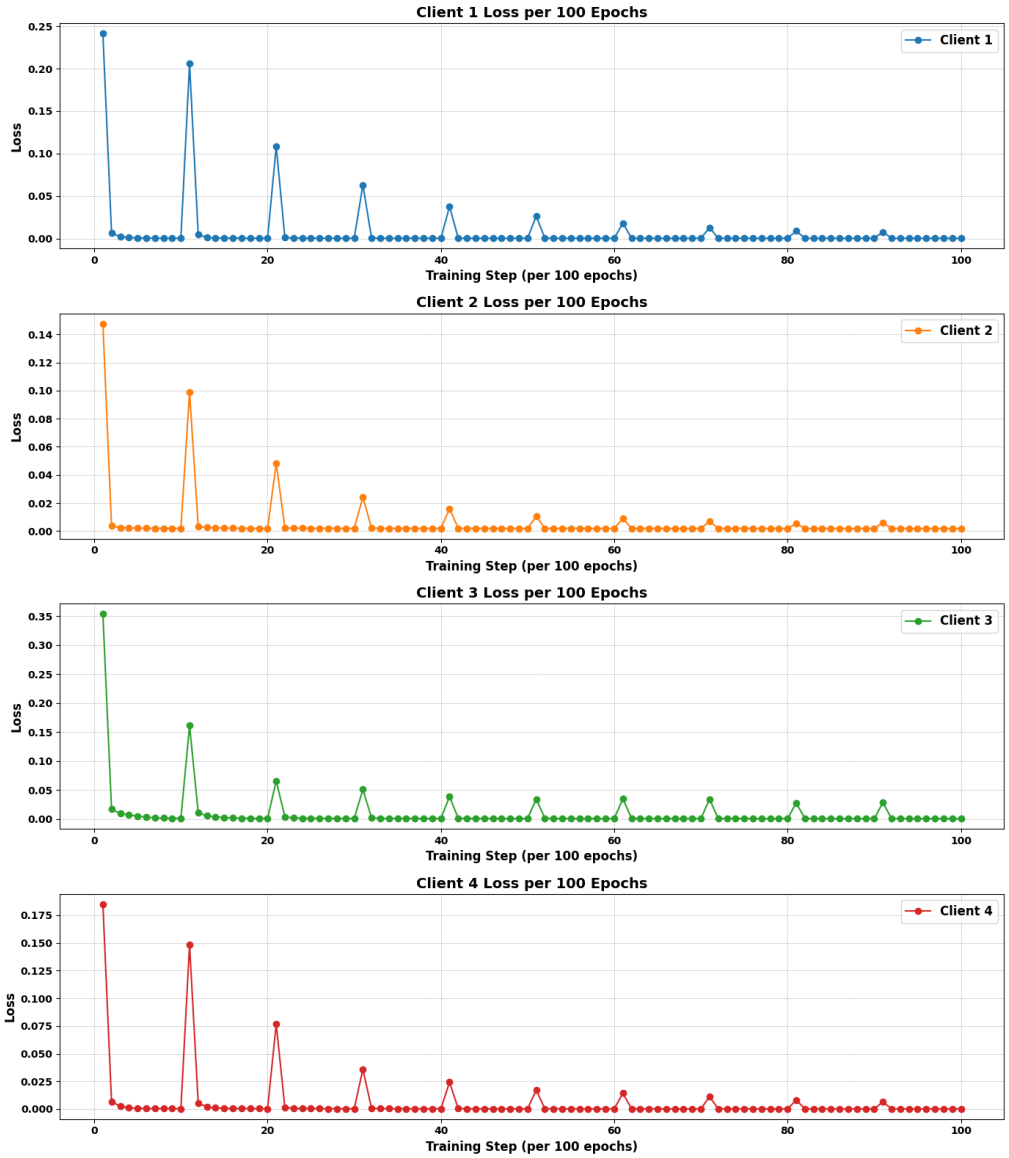

Fig. 4. The local training loss variation curve of each client during federated training process.

Optics EXPRESS

within a small range (±0.06), with the majority falling within ±0.03, demonstrating the global model's robustness and strong generalization across heterogeneous structural data.

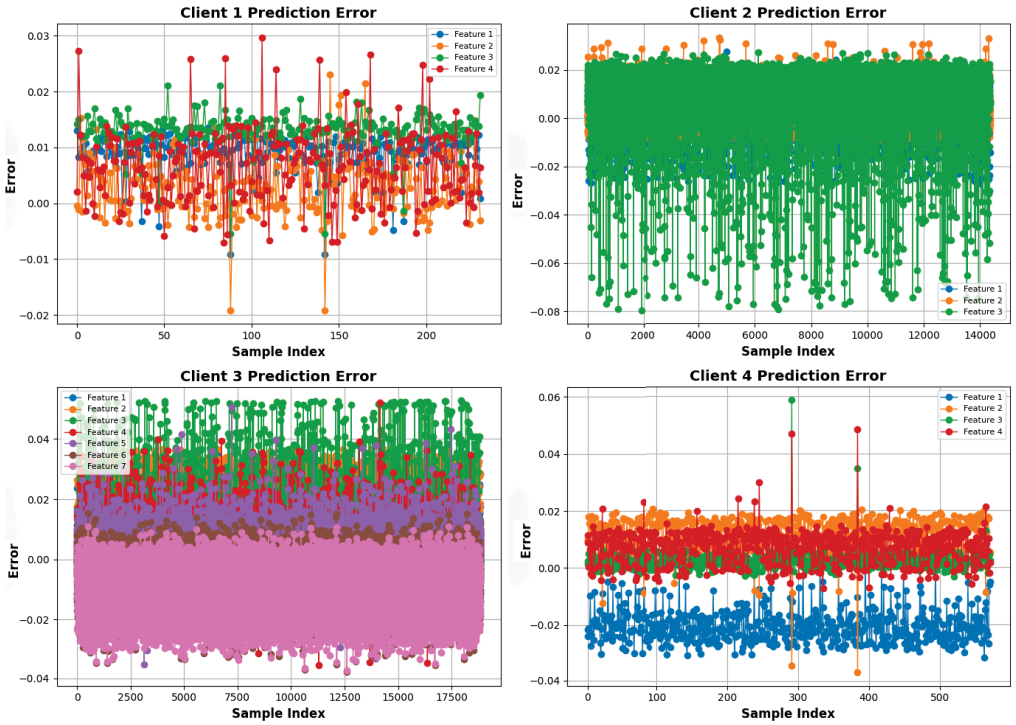

Fig. 5. The prediction error of PCF structure for different customers.

4.2.2. Results for Setting 2: non-IID distribution

Table 4. Performance of the global model under different client participation ratios in Setting 2

Participation Ratio	MSE ($\times 10^{-3}$)	RMSE ($\times 10^{-2}$)	MAE ($\times 10^{-2}$)	Time per Iteration (s)
2	5.64	7.51	3.39	15.40
5	5.03	7.09	3.11	39.24
8	4.64	6.81	3.09	62.48
10	**4.04**	**6.36**	**3.03**	**81.23**

In Setting 2, the dataset is preprocessed and dimensionally aligned before being randomly distributed among 10 clients. Each client thus receives a mixture of samples from all four PCF structures, introducing statistical heterogeneity while maintaining structural consistency. The optimal architecture (64 neurons per layer) from Setting 1 is reused here.

To examine the effect of client participation ratio, we vary the number of clients selected per communication round and track the average training loss across iterations (Fig. 6). Regardless of participation level, the global model shows consistent convergence. However, higher participation (e.g., 8 or 10 clients per round) results in faster convergence and more stable loss reduction. In contrast, lower participation (e.g., 2 clients) leads to slower convergence and higher variance, likely due to limited representativeness of the global data distribution.

Optics EXPRESS

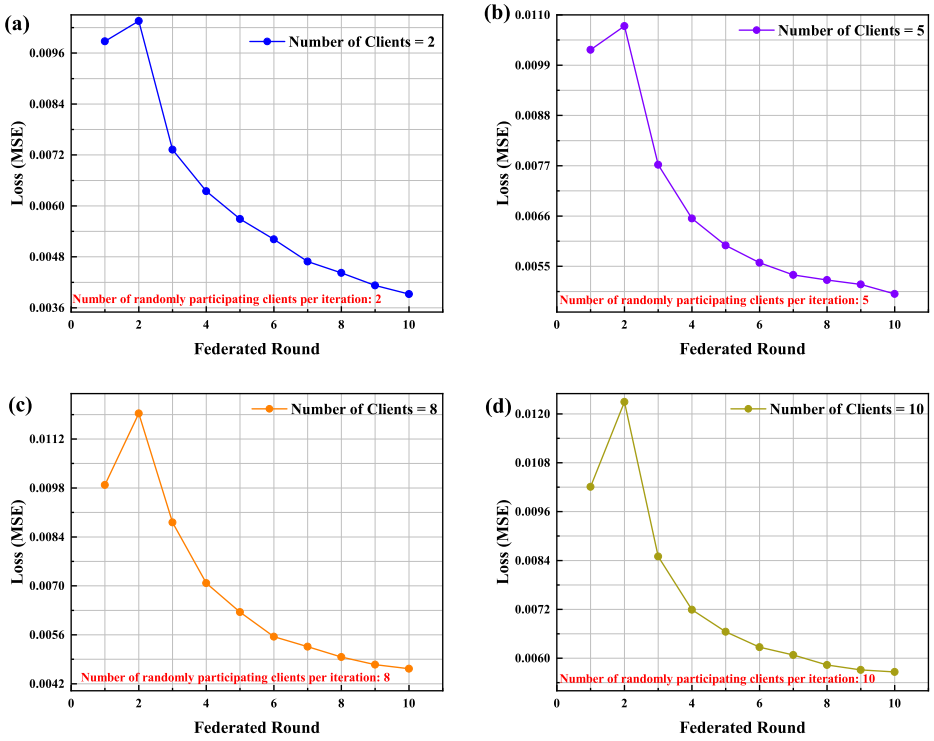

Fig. 6. Training loss curves under varying client participation ratios. Specifically, for each round, the average of the final loss values (after 1000 local training steps) across the participating clients was computed. Subfigures (a)–(d) correspond to scenarios with 2, 5, 8, and 10 participating clients, respectively.

Table 4 summarizes the final prediction performance and average training time for each participation ratio. Full participation (10 clients) yields the lowest MSE, representing a 28.4% improvement compared to the 2-client setting. However, this comes at the cost of increased computation. Notably, using 8 clients achieves 95% of the accuracy of full participation while reducing training time by 23%, suggesting that moderate participation strikes a favorable balance between accuracy and efficiency.

Overall, the proposed framework exhibits stable performance across varying degrees of client involvement. It maintains low prediction errors even under limited participation, and benefits from the increased data diversity introduced by randomized sample distribution. Compared to Setting 1, this configuration enables better generalization due to mixed data exposure. However, one trade-off is that data privacy at the structural level (i.e., per PCF type) is not strictly preserved in this mixed setup.

5. Conclusion

In conclusion, the proposed federated learning framework effectively tackles the challenges of dimensional inconsistency and distributional drift arising from structural heterogeneity in photonic crystal fiber (PCF) datasets. By integrating a shared hidden-layer architecture with personalized input/output layers and a dynamic feature alignment strategy, the framework ensures both model adaptability and representation consistency across clients. Experimental results under structurally heterogeneous and non-IID conditions demonstrate that the approach delivers

Optics EXPRESS

high prediction accuracy with strong generalization and computational efficiency. Beyond performance, this method marks a significant step forward for inverse PCF design by alleviating the reliance on centralized datasets and enabling privacy-preserving, collaborative learning across institutions. Overall, it contributes to bridging the gap between data-driven photonic design methodologies and scalable, real-world engineering deployment.

Funding. Hainan Province Science and Technology Special Foundation (ZDYF2023GXJS013, ZDYF2022-SHFZ304); National Natural Science Foundation of China (62175054); Collaborative Innovation Center Foundation of Hainan University (XTCX2022XXC06).

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. S. O. Konorov, A. M. Zheltikov, and M. Scalora, "Photonic-crystal fiber as a multifunctional optical sensor and sample collector," Opt. Express **13**(9), 3454–3459 (2005).
2. B. K. Paul, K. Ahmed, S. Asaduzzaman, *et al.*, "Folded cladding porous shaped photonic crystal fiber with high sensitivity in optical sensing applications: design and analysis," Sens. Bio-Sensing Res. **12**, 36–42 (2017).
3. H. Pakarzadeh, M. Taghizadeh, and M. Hatami, "Designing a photonic crystal fiber for an ultra-broadband parametric amplification in telecommunication region," J. Nonlinear Optic. Phys. Mat. **25**(02), 1650023 (2016).
4. P. Goyal and N. Kathpal, "Potential of photonic crystal fiber for designing optical devices for telecommunication networks," Opt. Quantum Electron. **56**(2), 147 (2024).
5. M. S. Islam, J. Sultana, K. Ahmed, *et al.*, "A novel approach for spectroscopic chemical identification using photonic crystal fiber in the terahertz regime," IEEE Sens. J. **18**(2), 575–582 (2018).
6. A. Rahman, A. Khaleque, M. Y. Ali, *et al.*, "Thz spectroscopic sensing of liquid chemicals using a photonic crystal fiber," OSA Continuum **3**(11), 2982–2996 (2020).
7. F. Bréchet, J. Marcou, D. Pagnoux, *et al.*, "Complete analysis of the characteristics of propagation into photonic crystal fibers, by the finite element method," Opt. Fiber Technol. **6**(2), 181–191 (2000).
8. K. Saitoh and M. Koshiba, "Numerical modeling of photonic crystal fibers," J. Lightwave Technol. **23**(11), 3580–3590 (2005).
9. R. Mamtaz, K. Ahmed, B. K. Paul, *et al.*, "Design and fem analysis of pentagonal photonic crystal fiber for highly non-linear applications," Opt. Quantum Electron. **52**(10), 455 (2020).
10. A. da Silva Ferreira, G. N. Malheiros-Silveira, and H. E. Hernández-Figueroa, "Computing optical properties of photonic crystals by using multilayer perceptron and extreme learning machine," J. Lightwave Technol. **36**(18), 4066–4073 (2018).
11. S. M. Abdul Kaium and M. A. Mollah, "Optical properties estimation of photonic crystal fiber using gaussian process regression," Opt. Continuum **3**(8), 1369–1388 (2024).
12. M. Khatun and M. M. Hossain, "Efficient prediction of optical properties in hexagonal pcf using machine learning models," Optik **312**, 171929 (2024).
13. H. Mozafari, S. Khatami, H. Molatefi, *et al.*, "Finite element analysis of foam-filled honeycomb structures under impact loading and crashworthiness design," Int. J. Crashworthiness **21**(2), 148–160 (2016).
14. K. Anwar, N. Nowsher, and B. N. Islam, "Analysis of various pcf structures using finite element method," Ph.D. thesis, Dept. of Electrical, Electronic and Communication Engineering (2017).
15. C.-P. Yu and H.-C. Chang, "Applications of the finite difference mode solution method to photonic crystal structures," Opt. Quantum Electron. **36**(1-3), 145–163 (2004).
16. M. Karimi, "Analysis of photonic crystal fibers using finite difference frequency domain method," Applied Electromagnetics **6**, 33–42 (2018).
17. E. Haque, S. Mahmuda, M. A. Hossain, *et al.*, "Highly sensitive dual-core pcf based plasmonic refractive index sensor for low refractive index detection," IEEE Photonics J. **11**(5), 1–9 (2019).
18. G. Wang, Y. Lu, L. Duan, *et al.*, "A refractive index sensor based on pcf with ultra-wide detection range," IEEE J. Sel. Top. Quantum Electron. **27**(4), 1–8 (2021).
19. R. A. Norton and R. Scheichl, "Planewave expansion methods for photonic crystal fibres," Appl. Numer. Math. **63**, 88–104 (2013).
20. S. Omi, M. Hirose, M. Ameya, *et al.*, "Plane-wave synthesis employing propagating plane-wave expansion for 3-d and 2-d rcs prediction including the multiple scattering effects," IEEE Trans. Antennas Propag. **69**(8), 4827–4835 (2021).
21. J. Soto-Perdomo, E. Reyes-Vera, J. Montoya-Cardona, *et al.*, "Design of porous-core photonic crystal fiber based on machine learning approach," Opt. Eng. **63**(01), 015102 (2024).
22. M. Vijayan, S. Sridhar, and D. Vijayalakshmi, "A deep learning regression model for photonic crystal fiber sensor with xai feature selection and analysis," IEEE Trans. NanoBiosci. **22**(3), 590–596 (2023).
23. C. Kalyoncu, A. Yasli, and H. Ademgil, "Machine learning methods for estimating bent photonic crystal fiber based spr sensor properties," Heliyon **8**(11), e11582 (2022).

Optics EXPRESS

24. Z. Liu, R. Chen, J. Wen, *et al.*, "Optimization of low-loss, high birefringence parameters of a hollow-core anti-resonant fiber with back-propagation neural network assisted hyperplane segmentation algorithm," Opt. Express **32**(17), 29638–29655 (2024).
25. Z. Gu, T. Ning, H. Hou, *et al.*, "Tapered photonic crystal fiber based on artificial intelligence-design for pulse compression," Opt. Laser Technol. **181**, 111650 (2025).
26. S. Chugh, A. Gulistan, S. Ghosh, *et al.*, "Machine learning approach for computing optical properties of a photonic crystal fiber," Opt. Express **27**(25), 36414–36425 (2019).
27. M. A. Jabin and M. P. Fok, "Prediction of 12 photonic crystal fiber optical properties using mlp in deep learning," IEEE Photonics Technol. Lett. **34**(7), 391–394 (2022).
28. K. Bonawitz, H. Eichner, W. Grieskamp, *et al.*, "Towards federated learning at scale: System design," *Proceedings of Machine Learning and Systems* 1, 374–388 (2019).
29. A. Z. Tan, H. Yu, L. Cui, *et al.*, "Towards personalized federated learning," IEEE Trans. Neural Netw. Learning Syst. **34**(12), 9587–9603 (2023).
30. Y. Liu, Y. Kang, T. Zou, *et al.*, "Vertical federated learning: Concepts, advances, and challenges," IEEE Trans. Knowl. Data Eng. **36**(7), 3615–3634 (2024).
31. S. Chen, X. Wang, S. Ren, *et al.*, "Collaborative photonic crystal fiber property optimization: A new paradigm for reverse design," IEEE Photonics Technol. Lett. **35**(19), 1035–1038 (2023).
32. S. Ren, S. Chen, J. Wang, *et al.*, "A distributed photonic crystal fiber reverse design framework based on multi-source knowledge fusion," Opt. Fiber Technol. **84**, 103718 (2024).
33. G. Wang, S. Ren, S. Li, *et al.*, "Semi-supervised deep learning model for efficient computation of optical properties of suspended-core fibers," Sensors **22**(18), 6751 (2022).
34. A. Rahman and M. A. Mollah, "Data augmentation and simultaneous prediction of optical properties of a porous core single mode fiber: Ctgan in the domain of optics," J. Opt. Communications (2025).