Dispersion Metrology with Machine Learning

Ergun Simsek[1], Raonaqul Islam[1], Shao-Chien Ou[2,3], Gregory Moille[2,3], and Kartik Srinivasan[2,3]

[1]University of Maryland Baltimore County, 1000 Hilltop Circle, 21250 Baltimore, MD, USA
[2]National Institute of Standards and Technology, Gaithersburg, USA
[3]Joint Quantum Institute, NIST/University of Maryland, College Park, MD, USA

ABSTRACT

Precise knowledge of microring resonator dispersion is essential for high-performance nonlinear photonic devices, including optical parametric oscillators and integrated frequency combs. Yet current characterization methods such as destructive cross-section imaging and non-destructive ellipsometry are limited in scope or throughput. In this work, we develop machine learning (ML) frameworks addressing three complementary problems: (i) predicting resonator geometric dimensions from integrated dispersion data, (ii) classifying the Sellmeier model of the core material, and (iii) forward-predicting the integrated dispersion spectrum directly from ring dimensions. We validate our approach by comparing neural network-predicted dimensions against SEM-confirmed dimensions from a publicly available dataset.

Keywords: dispersion engineering, machine learning, neural networks, integrated photonics, process control.

1. INTRODUCTION

Silicon nitride (Si_3N_4) microring resonators have emerged as a leading platform for nonlinear integrated photonics, enabling octave-spanning dissipative Kerr soliton (DKS) microcombs with applications in precision metrology, optical atomic clocks, and telecommunications.[1–5] The performance of such devices depends critically on the resonator's integrated dispersion (D_{int}), which describes the deviation of resonance frequencies from a uniformly spaced grid and is governed by both geometric (ring radius, R, width RW, and height RH) and material contributions.[6,7]

Fabrication tolerances at the nanometer scale can shift dispersive-wave frequencies by more than 5 THz,[8] while material dispersion varies with the deposition gas ratio in Si_3N_4 growth.[9] Current in-line metrology is limited: ellipsometry addresses material dispersion only, while cross-sectional SEM imaging is destructive. Full optical characterization via D_{int} measurement is accurate but time-consuming and difficult to scale to wafer-level monitoring.

Machine learning (ML) offers a compelling path forward.[10–13] Prior work has demonstrated ML-based bus waveguide - ring resonator coupling efficiency prediction,[10] inverse design for waveguide dispersion engineering,[11] geometry inference from dispersion data,[12] and classification of DKS soliton states.[13] Building on these foundations, this work presents three ML frameworks that (i) inversely predicts ring dimensions from D_{int} data, (ii) classifies the Sellmeier material model, and (iii) forward-predicts D_{int} spectra from ring dimensions as illustrated in Fig. 1. We quantify the impact of measurement noise and spectral sampling density on prediction accuracy and validate the framework on an independent experimental dataset.

Further author information: (Send correspondence to E. S.)
E-mail: simsek@umbc.edu, Telephone: +1 (410) 455 3540

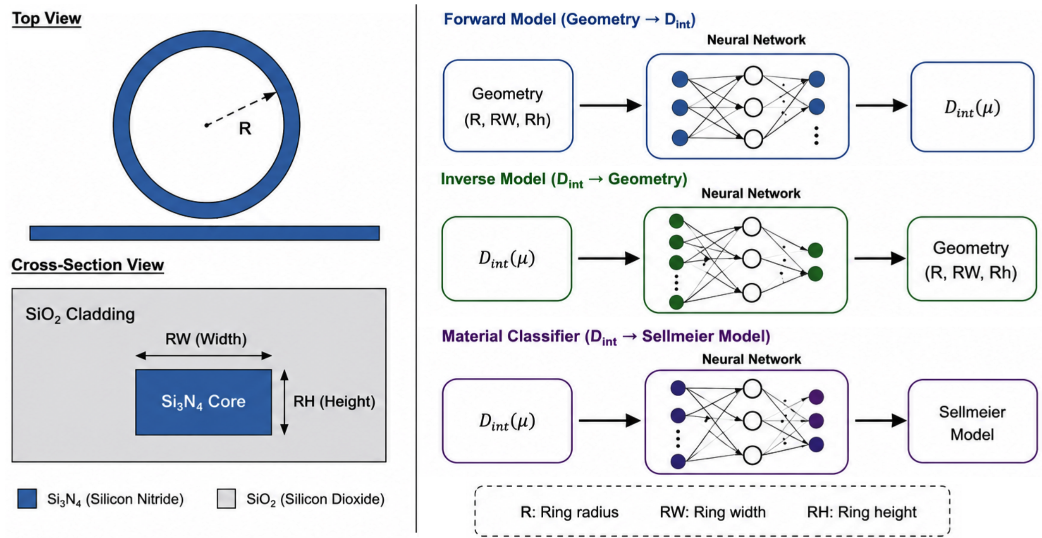

Figure 1. Schematic of the Si_3N4 microring resonator and the proposed machine learning framework. Left: Top view of the ring resonator showing the ring radius (R) and the coupling gap to the bus waveguide, together with the waveguide cross section illustrating the ring width (RW) and height (RH) embedded in a SiO2 cladding. Right: Bidirectional machine learning workflow consisting of (top) a forward neural network that predicts the integrated dispersion spectrum Dint from the ring geometry, (middle) an inverse neural network that infers the ring dimensions from measured Dint, and (bottom) a neural network classifier that identifies the Si_3N_4 Sellmeier material model from the measured integrated dispersion.

2. DATASET GENERATION

We consider Si_3N_4-core, SiO_2-clad microring resonators with radii selected in two different ranges. For the first group, we choose radii between 22 μm and 24 μm in 200 nm steps to obtain free spectral ranges (FSRs) near 100 GHz; the ring height is swept from 800 nm to 900 nm in 10 nm steps and the width from 1.8 μm to 2.5 μm in 100 nm steps. For the second group, the target FSR is 1 THz and we choose radii between 225 μm and 230 μm in 250 nm steps; the ring height is swept from 620 nm to 720 nm and width from 750 nm to 950 nm in 10 nm steps. In addition to the Sellmeier model developed by Luke et al.,[14] four additional Si_3N_4 material models (SM_1–SM_4), corresponding to NH_3:SiH_2Cl_2 gas ratios of 3:1, 5:1, 7:1, and 15:1, are implemented with a Sellmeier-type permittivity formula.[15] Effective indices are computed at 171 wavelengths (750 nm–1600 nm) with a frequency-domain finite-differences solver,[6] resonance frequencies are obtained via the azimuthal resonance condition $f_i = c_0\, m_i/(n_{\mathrm{eff},i}\, L)$, and D_{int} is evaluated over 201 mode indices relative to two different pump lasers near 1060 nm and 1550 nm using

$$D_{\mathrm{int}}(\mu) = f_\mu - (f_0 + D_1\mu), \tag{1}$$

where c_0 is speed of electromagnetic waves in vacuum, m_i is the azimuthal mode number, $n_{\mathrm{eff},i}$ is the effective index of the ring, $L = 2\pi R$, D_1 is the free spectral range at the pump, μ is the mode number relative to resonant mode at pump frequency f_0, and f_μ is the resonance frequency of the relative mode number μ.[7] Physically, $D_{\mathrm{int}}(\mu)$ quantifies how far the μ^{th} resonance deviates from a perfectly evenly spaced (linear) frequency grid. This deviation arises from the combined effects of geometric and material dispersions. A positive (anomalous) D_{int} profile is required for soliton mode-locking, since it allows the Kerr nonlinearity to balance dispersion and sustain a stable pulse circulating in the cavity. The detailed shape of D_{int}, including its curvature near dispersive-wave features, is therefore both a target for dispersion engineering and, as exploited in this work, a fingerprint that encodes the ring's geometry and material composition.

3. NEURAL NETWORK ARCHITECTURES AND RESULTS

3.1 Geometry Prediction

The first network takes 201 normalized D_{int} values as input and outputs ring width and height. It comprises 20 fully connected hidden layers of 64 neurons each with tanh activations, trained using the Levenberg–Marquardt algorithm with mean squared error (MSE) loss. In noiseless conditions, the network achieves mean absolute errors (MAE) below 0.3 nm for both dimensions across all Sellmeier models, with 49 of 50 test cases below 1 nm.

To assess robustness under realistic measurement uncertainty, random noise is added in $[-\Delta f,\ \Delta f]$ to D_{int} values. At $\Delta f = 50$ MHz (typical laser wavemeter precision), MAEs remain below 2.2 nm for width and 1.4 nm for height ($R^2 > 0.99$). At $\Delta f = 200$ MHz, errors increase to $\approx$8.5 nm (width) and $\approx$3.4 nm (height), with $R^2 > 0.95$. Height predictions are consistently more accurate than width, reflecting the higher sensitivity of D_{int} to ring height variations for the first quasi transverse electric (TE) mode, where the ρ component of the electric field is the strongest component with respect to the ϕ and z components.

Table 1 benchmarks the neural network against standard ML models on the noiseless SM_1 dataset. Neural networks deliver the best accuracy–efficiency trade-off, while Gaussian process regression achieves slightly lower errors at roughly 3× the training cost.

Table 1. Comparison of ML models' accuracies and training times (SM_1, noiseless).

ML Model	Width MAE (nm)	Height MAE (nm)	Train time (s)
Linear Regression	1.25	1.73	0.01
Decision Tree	8.76	2.37	0.05
Support Vector Machine	3.09	1.56	18.7
Random Forest	5.20	1.30	19.3
Neural Network	0.278	0.232	84.1
Gaussian Process	0.066	0.053	277

We further investigated how many D_{int} samples are needed for reliable predictions by randomly selecting M samples from each of three spectral regions (far-left, center, far-right), for a total of $N = 3M$ samples, repeated 40 times. With $N = 45$ samples and $\pm$50 MHz noise, width errors range from 2 nm to 8 nm and height errors from 1 nm to 3 nm across all gas ratios. Critically, samples taken far from the pump resonance, where D_{int} exhibits stronger curvature, are most informative. When a 25-sample window is centered near the pump, errors spike sharply; shifting away from the pump reduces errors to the 8–15 nm range. Covering $\approx$20% of the full octave-spanning spectrum (45 samples) ensures that at least some samples always fall in informative spectral regions.

3.2 Sellmeier Model Classification

The second network identifies which of the four Sellmeier material models underlies the D_{int} data. It uses the same 201-element input and employs four hidden layers (128→64→32→16 neurons) with tanh/ReLU activations and a 4-class softmax output, trained with the Adam optimizer and categorical cross-entropy loss.

The classifier achieves >99 % accuracy across all noise conditions. In the full 201-sample setting, only a single misclassification occurs across 200 test cases, regardless of the noise level. With as few as $N = 21$ randomly sampled points, classification accuracy already exceeds 90% even at $\pm$200 MHz noise, and surpasses 97 % for $N \gtrsim 45$ samples. This robustness reflects the large spectral contrast between Sellmeier models: different gas ratios produce qualitatively distinct D_{int} profiles.

This classifier is suited to a practical foundry scenario: when a small, pre-characterized set of deposition recipes exists, it identifies which recipe was realized for a given device using dispersion data already collected for geometric metrology, without requiring a separate ellipsometry step.

3.3 Forward D_{int} Spectrum Prediction

The third network performs the forward mapping: given ring width and height as inputs, it predicts the seven coefficients of a 6^{th}-order polynomial fit to the D_{int} spectrum. The architecture uses three fully connected layers (64–128–128 nodes) with leaky ReLU activations, batch normalization, and dropout, trained with the Adam optimizer, early stopping, and L_2 regularization. Training covers $RH \in [620, 720]$ nm and $RW \in [800, 900]$ nm in 10 nm steps (121 designs).

For a representative test case ($RW = 867.9$ nm, $RH = 664.2$ nm), the predicted spectrum closely matches the numerically computed reference as shown in Fig. 2. Minor discrepancies at low relative mode indices are attributable to the large dynamic range of higher-order polynomial coefficients. This forward model enables rapid assessment of the dispersion regime without repeated electromagnetic simulations, complementing the inverse networks in a bidirectional design workflow.

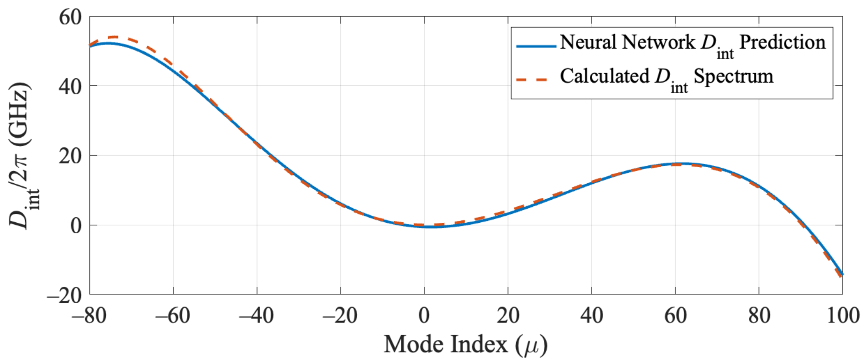

Figure 2. Comparison of dispersion spectra obtained from neural network prediction (blue) and numerical simulation (red dashed curve) for randomly chosen ring width and height values of 867.9 nm and 664.2 nm, respectively.

4. EXPERIMENTAL VALIDATION

As an independent validation, we applied the framework to the publicly available dataset of Ye et al.,[5] who fabricated Si_3N_4 ring resonators with SEM-measured dimensions of width = 2 μm and height = 810 nm. The ring radius was not reported, but since we know the FSR is 100 GHz, it has to be between 225 μm and 230 μm. Two independent ML models, a multilayer perceptron (MLP) and a deep kernel learning (DKL) model, were applied to infer geometry from the measured D_{int}. As shown in Table 2, both models predict dimensions very close to those measured by SEM values. Figure 3 shows the good agreement between the measured D_{int} and recalculated spectra from predicted dimensions closely matches the measured spectrum.

This cross-foundry validation, conducted on an independently fabricated device with an unconstrained ring radius, provides strong support for the accuracy and generalizability of the proposed framework. A more detailed version of this work has been recently submitted to the journal Nanophotonics.[16]

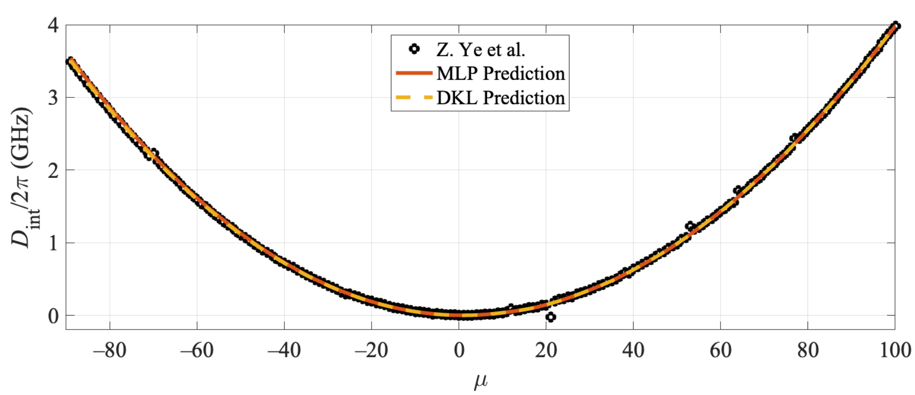

Figure 3. Integrated dispersions measured by Ye et al.[5] and calculated by us using the parameters determined by neural networks.

Table 2. ML-predicted vs. SEM-measured dimensions (Ye et al. dataset).

Source	Radius (μm)	Width (nm)	Height (nm)
SEM (reported)	—	2000	810
MLP (predicted)	226.9	2001.4	802.8
DKL (predicted)	225.3	2018.6	807.7

5. CONCLUSION

We have presented a bidirectional machine learning framework for non-destructive dispersion metrology of Si_3N_4 microring resonators. Three complementary neural networks (a regression network for geometry inference, a classifier for material model identification, and a forward-prediction network for D_{int} reconstruction) enable rapid, scalable characterization without cross-sectional imaging or dense spectral scans. Key findings are: (i) sub-nanometer geometry inference in noiseless conditions; (ii) >99% Sellmeier model classification accuracy under all noise conditions tested; (iii) fewer than 45 carefully chosen D_{int} samples taken away from the pump resonance suffice for <8 nm dimensional accuracy at ±50 MHz noise; and (iv) experimental validation on an independent dataset confirms practical applicability. The ML frameworks offer a generalizable, non-destructive pathway to wafer-scale quality control in photonic foundries, reducing reliance on multiple wideband tunable lasers and time-consuming full-spectrum characterization.

Acknowledgment

E.S. and R.I.'s contributions were supported in part by the National Center for Manufacturing Sciences (cooperative agreements 2022138-142232 and 2023200-142386). S.-C.O., G.M., and K.S. acknowledge support from the NIST-on-a-chip program and the Air Force Research Laboratory Space Vehicles Directorate. K.S. further acknowledges support from the NIST CHIPS Metrology Program and the Microelectronics Commons QUPICS program.

REFERENCES

[1] T. J. Kippenberg, A. L. Gaeta, M. Lipson, and S. B. Spillane, "Dissipative Kerr solitons in optical microresonators," *Science*, vol. 361, p. eaan8083, 2018.

[2] A. L. Gaeta, M. Lipson, and T. J. Kippenberg, "Photonic-chip-based frequency combs," *Nature Photon.*, vol. 13, pp. 158–169, 2019.

[3] D. T. Spencer *et al.*, "An optical-frequency synthesizer using integrated photonics," *Nature*, vol. 557, pp. 81–85, 2018.

[4] G. Moille *et al.*, "All-optical stabilization of a soliton microcomb," *Optica*, vol. 12, pp. 391–398, 2025.

[5] Z. Ye *et al.*, "Foundry manufacturing of tight-confinement, dispersion-engineered, ultralow-loss silicon nitride photonic integrated circuits," *Photon. Res.*, vol. 11, no. 4, pp. 558-568, 2023.

[6] E. Simsek, A. Niang, R. Islam, L. Courtright, P. Shandilya, G. M. Carter, and C. R. Menyuk , "A mixed-field formulation for modeling dielectric ring resonators and its application in optical frequency comb generation," Scientific Reports, vol. 15, 35098, 2025.

[7] E. Simsek, A. Niang, R. Islam, P. Shandilya, L. Courtright, G. M. Carter, and C. R. Menyuk, "One-Stop-Shop for Modeling Optical Frequency Comb Generation," 26th International Conference on Electromagnetics in Advanced Applications (ICEAA 2025), Palermo, Italy, September 8-12, 2025.

[8] T. C. Briles *et al.*, "Interlocking Kerr-microresonator frequency combs for microwave to optical synthesis," *Optica*, vol. 5, p. 1, 2018.

[9] G. Moille *et al.*, "Broadband resonator-waveguide coupling for efficient extraction of octave-spanning microcombs," *Opt. Lett.*, vol. 46, pp. 5939–5942, 2021.

[10] M. Soroush, E. Simsek, G. Moille, K. Srinivasan, and C. R. Menyuk, "Predicting broadband resonator-waveguide coupling for microresonator frequency combs through fully connected and recurrent neural networks and attention mechanism," *ACS Photonics*, vol. 6, pp. 1795–1805, 2023.

[11] X. Wang *et al.*, “Agile dispersion-profile design for integrated photonics using machine learning,” *Laser Photon. Rev.*, vol. 16, p. 2100519, 2022.

[12] A. Pal *et al.*, “Machine learning assisted inverse design of microresonators,” *Opt. Lett.*, vol. 48, 2023.

[13] C. Li *et al.*, “High-accuracy classification of dissipative Kerr soliton states using hybrid deep learning,” *Laser Photon. Rev.*, vol. 20, p. 2400935, 2026.

[14] K. Luke, Y. Okawachi, M. R. E. Lamont, A. L. Gaeta, and M. Lipson, “Broadband mid-infrared frequency comb generation in a Si_3N_4 microresonator," *Opt. Lett.*, vol. 40, pp. 4823-4826, 2015.

[15] G. Moille, D. Westly, G. Simelgor, and K. Srinivasan, “Impact of the precursor gas ratio on dispersion engineering of broadband silicon nitride microresonator frequency combs," *Opt. Lett.*, vol. 46, no. 23, pp. 5970–5973, 2021.

[16] E. Simsek, S-C. Ou, G. Moille, and K. Srinivasan, “Microring Resonator Dispersion Metrology with Neural Networks," submitted to *Nanophotonics*, June 2026.