Optics Letters

Infrared bound states in the continuum: random forest method

M. S. Molokeev,[1,2,3] A. S. Kostyukov,[1] A. E. Ershov,[1,4] D. N. Maksimov,[1,2] V. S. Gerasimov,[1,4] and S. P. Polyutov[1,*]

[1]IRC SQC, Siberian Federal University, Krasnoyarsk, 660041, Russia
[2]Kirensky Institute of Physics, Federal Research Center KSC SB RAS, Krasnoyarsk, 660036, Russia
[3]Laboratory of Theory and Optimization of Chemical and Technological Processes, University of Tyumen, Tyumen, 625003, Russia
[4]Institute of Computational Modelling SB RAS, Krasnoyarsk, 660036, Russia
*spolyutov@sfu-kras.ru

Received 5 May 2023; revised 26 June 2023; accepted 31 July 2023; posted 1 August 2023; published 17 August 2023

In this Letter, we consider optical bound states in the continuum (BICs) in the infrared range supported by an all-dielectric metasurface in the form of subwavelength dielectric grating. We apply the random forest machine learning method to predict the frequency of the BICs as dependent on the optical and geometric parameters of the metasurface. It is found that the machine learning approach outperforms the standard least square method at the size of the dataset of ≈4000 specimens. It is shown that the random forest approach can be applied for predicting the subband in the infrared spectrum into which the BIC falls. The important feature parameters that affect the BIC wavelength are identified.

https://doi.org/10.1364/OL.494629

Introduction. Optical bound states in the continuum (BICs) are localized eigenmodes of Maxwell's equations embedded in the continuum of scattering states [1]. In the last decade, optical BICs [2,3] have become an important instrument for designing nanophotonic devices with enhanced light–matter interaction. The optical BIC can be applied whenever one requires a pronounced resonant response [4–6] accompanied by critical field enhancement [7,8] in the near zone. In particular, the BICs have already found applications in such fields as second harmonic generation [9–12], light absorbers [13–17], sensors [18,19], and lasers [20–23].

The optical BICs are typically engineered by numerically solving Maxwell's equations with application of various numerical techniques such as the finite-element method (FEM) and finite-difference time-domain (FDTD) method. Once an optical BIC is predicted in a certain nanophotonic structure, its frequency can be tuned to any desirable wavelength using the scale invariance of electrodynamics. This approach, however, neglects the dispersion of material parameters and, most importantly, can be limited by nanofabrication capabilities which could dictate the geometry and sizes of building blocks of the nanodevice.

In this Letter, we examine the performance of the random forest (RF) machine learning (ML) method for predicting the frequencies of optical BICs supported by an all-dielectric metasurface. Recently, we have seen a surge of interest in the application of ML techniques to various problems of nanophotonics [24–26]. The ML algorithms have been proved useful in topological photonics [27] and for design of integrated photonic circuits [28]. So far, for analyzing optical BICs, researchers have mostly applied neural networks (NNs) [29–33]. Here we take a different route by using the RF supervised learning algorithm. The RF is a method that constructs an ensemble of decision trees which are used to return the prediction either via majority voting for classification problems or average values for regression problems. In comparison with NNs, the RF features a small number of hyperparameters and allows to estimate the importance of the parameters, which can pave a way to useful physical insights [34,35].

BIC in a dielectric grating and dataset description. We consider optical BICs in dielectric grating, as shown in Fig. 1. The grating consists of a lossless dielectric substrate with dielectric bars periodically placed on top. The refractive index of the substrate is denoted by n_s while the refractive index of the bars by n_b. The refractive index in the upper half-space is taken as $n_0 = 1$. All geometric parameters are defined in Fig. 1. We keep the structure period $p = 693.3$ nm. All other geometric parameters, namely bar width w and bar height h, are normalized to the period of the grating p. To generate the dataset, we solved the eigenvalue problem for Maxwell's equation with application of the FEM implemented in COMSOL. The optical and geometric parameters were randomly equidistributed as $w \in [0.2, 0.8], h \in [0.2, 0.8], n_b \in [1.5, 5], n_s \in [1, 4]$ with the goal to obtain an optical BIC in the IR spectral range (1.1–2.0 μm). In this work, we focus on symmetry protected dipole BICs [36], as shown in the inset of Fig. 1.

We obtained the dataset of 21,090 samples of the four feature parameters (w, h, n_b, n_s) resulting in the property parameter (BIC wavelength) sitting in the IR range. The descriptive statistics and the plot of the dataset are summarized in Table S1 and Fig. S1 in Supplement 1, respectively. Our aim is to build a model that can predict the BIC wavelength and find the most important feature parameters to influence it. Preliminary data analysis revealed that all feature parameters and wavelength values showed almost uniform distributions, see Fig. S1, which means that all representative cases were selected. The correlation matrix shown in Fig. 2 proves the absence of linear relationships between the

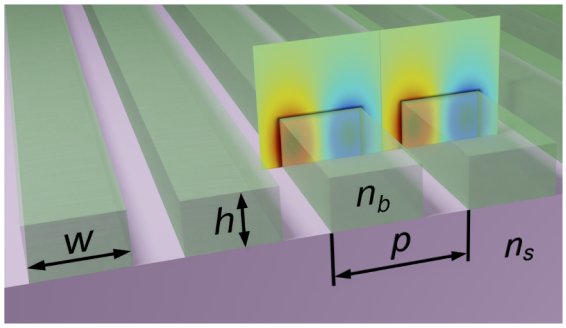

Fig. 1. Schematic of the metasurface with the electric field of the TE BIC mode shown in the inset.

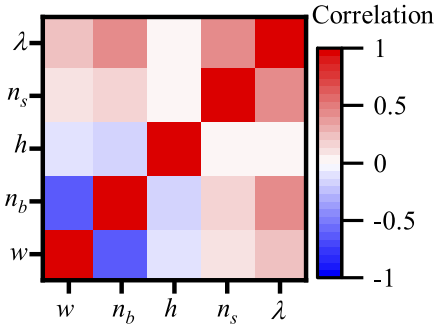

Fig. 2. Correlation matrix of the feature and property parameters.

features and the property, which justifies application of ML methods.

Nonlinear lest squares method. Before proceeding to the RF method, we introduce a benchmark approach based on the nonlinear least squares method (LSM). We have seen that there are no linear correlations among d, n_1, h, n_0 and the wavelength. Thus, we apply the following LSM formula for predicting the wavelength of BIC:

$$\lambda_{\mathrm{BIC}} = a + \sum_{j=1}^{4} a_j x_j + \sum_{j=1, j' \geq j}^{4} a_{jj'} x_j x_{j'}, \qquad (1)$$

where x_j is any of the four feature parameters w, h, n_b, n_s and $a, a_j, a_{jj'}$ are the coefficients fit by minimizing the sum $\sum(\lambda_{\mathrm{BIC}} - \lambda_n)^2$ with λ_n being the BIC wavelength obtained from solving the eigenvalue problem.

Regression random forest method. The regression trees are built by recursive binary partitioning of the multidimensional predictor space into domains by constructing a multitude of decision trees at a time and outputting mean/average prediction of the individual trees [35]. Predictions are done by passing new data parameters from the root through the internal nodes until a terminal node is reached. In accordance with the nonlinear model in Eq. (1), we extended the set of feature parameters by complementing the set of w, h, n_b, n_s with all possible products of its elements. Thus, the size of the extended feature parameter set is 14 (see Dataset 1, Ref. [37]). We used a self-written python script for Python 3.6 [38] which is available in Code 1, Ref. [39]. The libraries invoked in the script are *numpy, pandas, sklearn, matplotlib,* and *mpl_toolkits*. Since the RF algorithm is stochastic, we used it with averaging performance across ten repeats of cross-validation. Each time the data were split into the two random data sets: a set for training procedure (70% of total data) and another set for test (30% of total data). The mean absolute error (MAE) of the training set and the test set of wavelength values are 3.01 nm and 7.71 nm, respectively. We performed a comparison between the RF model and the benchmark LSM models for different sizes of the dataset. The results are shown in Fig. 3(a). The numerical values plotted in Fig. 3(a) are collected in Table S2 in Supplement 1. Figure 3(a) clearly shows the advantage of the RF over the LSM as the size of the training dataset is increased. We performed a 5-fold cross-validation test on the whole dataset, which showed MAE= 7.35±0.14 nm. The previously obtained MAE values are within three estimated standard deviations from the mean of this value. Thus we can conclude that the correlations between experimental features and the wavelength are captured by the RF method. The RF predicted wavelength values are plotted in Fig. 3(b) against the calculated ones. In total, the RF provides a good fit with 0.6% average relative error.

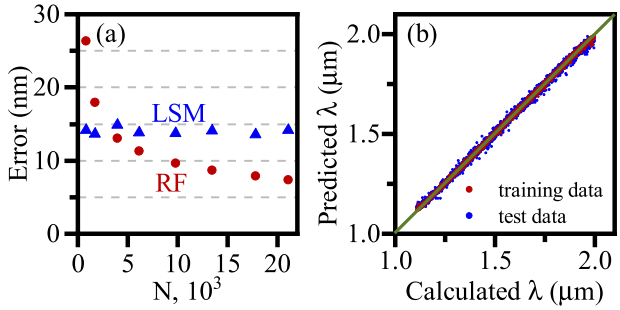

Fig. 3. (a) Prediction error calculated for 30% randomly selected test dataset for the LSM (blue triangles) and RF (red circles). Starting from N ~ 4000, the RF model has smaller prediction error than the LSM. (b) Comparative plot of the observed wavelength values against the calculated wavelengths obtained from RF model. For training dataset MAE= 3.01 nm, test dataset MAE= 7.71 nm, cross-validation MAE= 7.35± 3.14 nm.

The RF is notorious for allowing to rank the feature parameters according their importance after training. The selected value is permuted among the training data and the error is computed on this perturbed data set. The importance score for the selected feature is computed by averaging the difference in error before and after the permutation over all trees [40]. The score is normalized by the standard deviation of these differences. The features which produce large values for this score are ranked as more important than features which produce small values. In our case, the $w \cdot n_b$, $h \cdot n_s$, and n_b^2 are the three most important parameters, see Fig. 4(a). Now we can plot the BIC wavelengths in space of the three most important parameters which is presented in Fig. 4(b). One can see from Fig. 4(a) that the ranks of the two most important parameters, $w \cdot n_b$ and $h \cdot n_s$, add up to 90%. These parameters are nothing but the optical path lengths across the dielectric bars, see Fig. 1.

Classification random forest method. The BIC wavelengths calculated in our work are in the simulation range of 1100–2000 nm. The simulation range embraces the telecommunication 1260–1675-nm wavelength range which is conventionally subdivided into six subbands, see Table 1. Here we address the question whether it is possible to predict into which subband the BIC wavelength falls. To solve the classification problem, we additionally designated six subbands $X_1, \ldots, X_6$ at the edges of the simulation range. The auxiliary designations are also

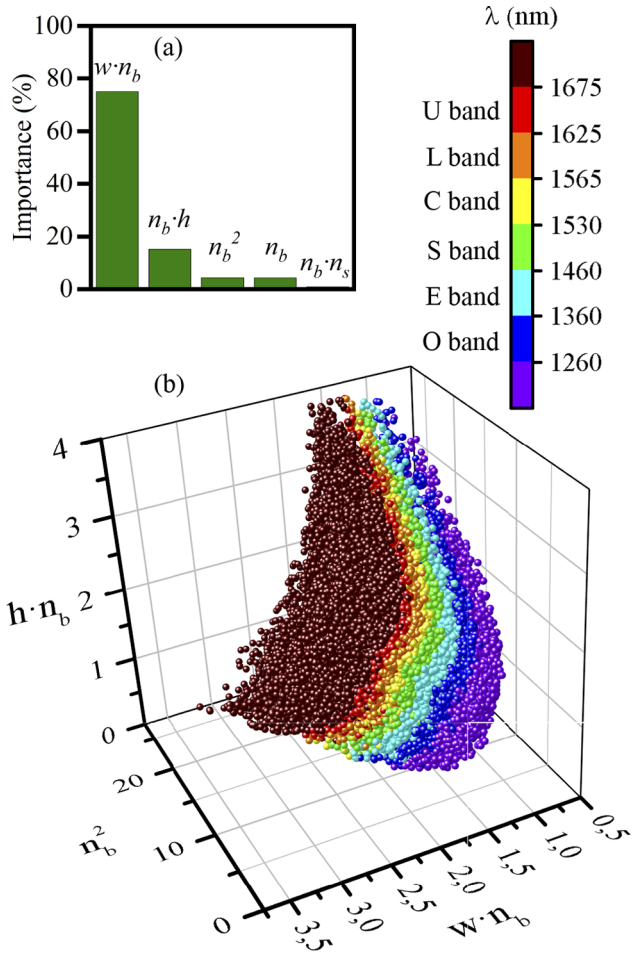

Fig. 4. (a) Importance of all feature parameters on BIC wavelength values in the RF model. The $w \cdot n_b$ parameter has the major influence. (b) Samples with low wavelength values (blue points) to the highest wavelengths (red points) are segregated into telecom bands (Table 1) in the 3D space spanned on the three most important parameters $w \cdot n_b$, n_b^2, and $h \cdot n_b$.

Table 1. Infrared Band Designations

Band	Description	Wavelength Range
X_1 band		1100–1160 nm
X_2 band		1160–1260 nm
O band	original	1260–1360 nm
E band	extended	1360–1460 nm
S band	short wavelengths	1460–1530 nm
C band	conventional ("erbium window")	1530–1565 nm
L band	long wavelengths	1565–1625 nm
U band	ultralong wavelengths	1625–1675 nm
X_3 band		1675–1750 nm
X_4 band		1750–1850 nm
X_5 band		1850–1950 nm
X_6 band		1950–2000 nm

explained in Table 1. The Python script for classification is available in Code 2, Ref. [41].

Thus, the simulation range is down to 12 classes (see Dataset 2, Ref. [42]). An RF containing ten decision trees was used to build the model for the classification problem. The five-fold

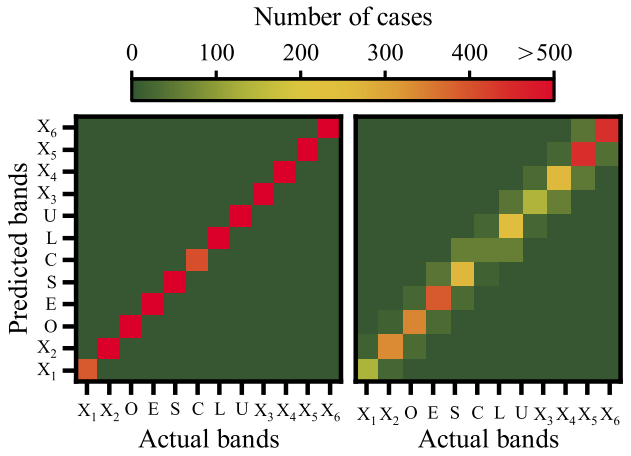

Fig. 5. Confusion matrix plotted for predicted and actual classes: (a) training dataset (70% randomly chosen); (b) test dataset (30% randomly chosen from all dataset).

cross-validation test on the whole dataset showed an accuracy of 83.1%. The confusion matrices for the training dataset (14,763 samples) and the test dataset (6327 cases) are presented in Figs. 5(a) and 5(b), respectively, which show good classification prediction of all classes.

Conclusion. The data analysis revealed the absence of linear relationships between the features (w, h, n_b, n_s) and the BIC wavelength, and the error of the RF model prediction is twice as small as the error of the nonlinear LSM model, which justifies the use of supervised ML methods for the problem under scrutiny. The regression RF model was able to predict the wavelength with a small error of 7.35 nm which is enough to classify such narrow bands as C-band (1530–1565 nm). The cross-validation accuracy of class prediction is 83.1% which means that only 16.9% of cases were wrongly classified. It should be noted that the confusion matrix of the test dataset showed that only the closest adjacent classes were chosen in the wrong classification cases. Therefore, the error in classification is minor. Meanwhile, the speed of predicting calculation is much higher than direct calculation using COMSOL, which opens an opportunity to quickly screen the geometry of the bars and the indices of dielectric materials to obtain a BIC in the desired telecom window. The importance ranks of the feature parameters indicated that the optical path lengths across the dielectric bars are the most important parameters affecting the BIC wavelengths. This allows for a certain freedom in the design of dielectric metasurfaces as, for instance, a variation of parameters under the constraint that the two features are constant does not significantly affect the BIC wavelength. In summary, we believe that the RF method may prove instrumental for engineering optical BICs at a given wavelength in the telecom band. The RF can be potentially applied for the reverse design by using the BIC frequency as one of the feature parameters whereas one of the geometric parameters, say the thickness of the bars, is used as one of the property parameters. It could also be interesting to test the RF method for predicting the wavelengths and quality factors of the high-quality resonances which occur in the BIC metasurfaces subject to the breaking of symmetry as suggested in [33]. The above problems are to be the subject of future studies.

Funding. Ministry of Science and Higher Education of the Russian Federation (FSRZ-2023-0006).

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are available in Dataset 1, Ref. [37] and Dataset 2, Ref. [42].

Supplemental document. See Supplement 1 for supporting content.

REFERENCES

1. C. W. Hsu, B. Zhen, A. D. Stone, J. D. Joannopoulos, and M. Soljačić, Nat. Rev. Mater. **1**, 16048 (2016).
2. K. Koshelev, G. Favraud, A. Bogdanov, Y. Kivshar, and A. Fratalocchi, Nanophotonics **8**, 725 (2019).
3. A. F. Sadreev, Rep. Prog. Phys. **84**, 055901 (2021).
4. S. P. Shipman and S. Venakides, Phys. Rev. E **71**, 026611 (2005).
5. A. F. Sadreev, E. N. Bulgakov, and I. Rotter, Phys. Rev. B **73**, 235342 (2006).
6. C. Blanchard, J.-P. Hugonin, and C. Sauvan, Phys. Rev. B **94**, 155303 (2016).
7. J. W. Yoon, S. H. Song, and R. Magnusson, Sci. Rep. **5**, 18301 (2015).
8. V. Mocella and S. Romano, Phys. Rev. B **92**, 155117 (2015).
9. F. R. Ndangali and S. V. Shabanov, *Active Photonic Materials V*, Vol. 8808 (International Society for Optics and Photonics, 2013), p. 88081F.
10. T. Wang and S. Zhang, Opt. Express **26**, 322 (2018).
11. L. Carletti, K. Koshelev, C. De Angelis, and Y. Kivshar, Phys. Rev. Lett. **121**, 033903 (2018).
12. K. Koshelev, S. Kruk, E. Melik-Gaykazyan, J.-H. Choi, A. Bogdanov, H.-G. Park, and Y. Kivshar, Science **367**, 288 (2020).
13. M. Zhang and X. Zhang, Sci. Rep. **5**, 8266 (2015).
14. X. Wang, J. Duan, W. Chen, C. Zhou, T. Liu, and S. Xiao, Phys. Rev. B **102**, 155432 (2020).
15. T. Sang, S. A. Dereshgi, W. Hadibrata, I. Tanriover, and K. Aydin, Nanomaterials **11**, 484 (2021).
16. S. Xiao, X. Wang, J. Duan, T. Liu, and T. Yu, J. Opt. Soc. Am. B **38**, 1325 (2021).
17. Y. Cai, X. Liu, K. Zhu, H. Wu, and Y. Huang, J. Quant. Spectrosc. Radiat. Transfer **283**, 108150 (2022).
18. Y. Liu, W. Zhou, and Y. Sun, Sensors **17**, 1861 (2017).
19. S. Romano, G. Zito, S. Torino, G. Calafiore, E. Penzo, G. Coppola, S. Cabrini, I. Rendina, and V. Mocella, Photonics Res. **6**, 726 (2018).
20. A. Kodigala, T. Lepetit, Q. Gu, B. Bahari, Y. Fainman, and B. Kanté, Nature **541**, 196 (2017).
21. M.-S. Hwang, H.-C. Lee, K.-H. Kim, K.-Y. Jeong, S.-H. Kwon, K. Koshelev, Y. Kivshar, and H.-G. Park, Nat. Commun. **12**, 4135 (2021).
22. Y. Yu, A. Sakanas, A. R. Zali, E. Semenova, K. Yvind, and J. Mørk, Nat. Photonics **15**, 758 (2021).
23. J.-H. Yang, Z.-T. Huang, D. N. Maksimov, P. S. Pankin, I. V. Timofeev, K.-B. Hong, H. Li, J.-W. Chen, C.-Y. Hsu, Y.-Y. Liu, T.-C. Lu, T.-R. Lin, C.-S. Yang, and K.-P. Chen, Laser Photonics Rev. **15**, 2100118 (2021).
24. W. Ma, Z. Liu, Z. A. Kudyshev, A. Boltasseva, W. Cai, and Y. Liu, Nat. Photonics **15**, 77 (2021).
25. J. Jiang, M. Chen, and J. A. Fan, Nat. Rev. Mater. **6**, 679 (2020).
26. S. So, T. Badloe, J. Noh, J. Bravo-Abad, and J. Rho, Nanophotonics **9**, 1041 (2020).
27. L. Pilozzi, F. A. Farrelly, G. Marcucci, and C. Conti, Commun. Phys. **1**, 57 (2018).
28. Z. A. Kudyshev, V. M. Shalaev, and A. Boltasseva, ACS Photonics **8**, 34 (2021).
29. R. Lin, Z. Alnakhli, and X. Li, Photonics Res. **9**, B96 (2021).
30. X. Ma, Y. Ma, P. Cunha, Q. Liu, K. Kudtarkar, D. Xu, J. Wang, Y. Chen, Z. J. Wong, M. Liu, M. Cynthia Hipwell, and S. Lan, Laser Photonics Rev. **16**, 2100658 (2022).
31. F. Wang, Y. Chen, Z. Zhang, X. Zhang, X. Zhou, Y. Zuo, Z. Chen, and C. Peng, Opt. Express **31**, 12384 (2023).
32. Z. Wang, J. Sun, J. Li, L. Wang, Z. Li, X. Zheng, and L. Wen, Adv. Sci. **10**, 2206236 (2023).
33. W. Wang, Y. K. Srivastava, T. C. Tan, Z. Wang, and R. Singh, Nat. Commun. **14**, 2811 (2023).
34. L. Breiman, Machine Learning **45**, 5 (2001).
35. T. K. Ho, in Proceedings of 3rd International Conference on Document Analysis and Recognition, Vol. 1 (1995), p. 278.
36. D. N. Maksimov, V. S. Gerasimov, S. Romano, and S. P. Polyutov, Opt. Express **28**, 38907 (2020).
37. D. Maksimov, A. Kostyukov, A. Ershov, M. Molokeev, V. Gerasimov, and S. Polyutov, "Dataset: Regression," figshare (2023), https://doi.org/10.6084/m9.figshare.22736858.
38. L. P. Coelho, W. Richert, and M. Brucher, *Building Machine Learning Systems with Python : Explore Machine Learning and Deep Learning Techniques for Building Intelligent Systems Using Scikit-learn and TensorFlow*, 3 ed. (Packt Publishing, 2018).
39. D. Maksimov, A. Kostyukov, A. Ershov, M. Molokeev, V. Gerasimov, and S. Polyutov, "Python Code: Regression," figshare, 2023https://doi.org/10.6084/m9.figshare.22736909.
40. R. Zhu, D. Zeng, and M. R. Kosorok, J. Am. Stat. Assoc. **110**, 1770 (2015).
41. D. Maksimov, A. Kostyukov, A. Ershov, M. Molokeev, V. Gerasimov, and S. Polyutov, "Python Code: Classification," figshare, 2023https://doi.org/10.6084/m9.figshare.22736912.
42. D. Maksimov, A. Kostyukov, A. Ershov, M. Molokeev, V. Gerasimov, and S. Polyutov, "Dataset: Classification," figshare, 2023https://doi.org/10.6084/m9.figshare.22736855.