08

Machine learning-based predictive modeling for SiC/Si thin film growth

© A.V. Redkov[1], D.V. Rozhentsev[1], A.S. Grashchenko[1], A.V. Osipov[2], S.A. Kukushkin[1]

[1] Institute for Problems in Mechanical Engineering of the Russian Academy of Sciences, St. Petersburg, Russia
[2] St. Petersburg State University, St. Petersburg, Russia
E-mail: avredkov@gmail.com

Received July 14, 2025
Revised August 14, 2025
Accepted August 14, 2025

We demonstrate the application of machine learning methods for predicting the properties of epitaxial structures in multi-parameter technological processes characterized by complex nonlinear dependencies. The synthesis of silicon carbide thin films on silicon substrates via atomic substitution method was investigated as a model system. A neural network model capable of predicting key characteristics of the resulting SiC films based on synthesis process parameters, including pressure, temperature, substrate type, and other additional synthesis conditions, was developed. Comprehensive optimization of the model architecture was performed followed by validation of prediction accuracy. The high efficiency of machine learning algorithms for analyzing and controlling complex epitaxial processes was demonstrated.

Keywords: machine learning, neural network model, epitaxial growth, SiC, Si, atomic substitution method.

DOI: 10.61011/TPL.2025.11.62214.20441

Machine learning (ML) and artificial intelligence methods, which have undergone rapid development in recent years, are used more and more often to solve scientific and technological problems that are related to the growth of crystals and epitaxial thin films [1,2] and range from real-time control of growth parameters and optimization of synthesis processes [2,3] to analysis of defects in growing samples, RHEED images, and XRD spectra [4–6]. In addition, machine learning methods find wide application in optimization and prediction of the characteristics of growing thin films based on large experimental data sets [7,8]. This line of research seems to be especially relevant, since increasingly complex materials and processes, such as metalorganic or chloride-hydride epitaxy of gallium and aluminum nitrides [9], are utilized in the semiconductor industry. An analytical description of such growth processes in multicomponent systems with chemical reactions, which are affected by both thermodynamics and gas dynamics in the reactor volume and the kinetics of surface processes [10–13], is often impossible to compile or extremely inaccurate due to the lack of data on material constants. In turn, machine learning methods applied to experimental data directly from the growth setup provide an opportunity to overcome this obstacle, identify key dependences, and construct predictive ML models without any additional assumptions inherent in theoretical models and without the need for preliminary measurements of material constants. This makes machine learning methods a versatile tool for improving the accuracy of description of synthesis processes.

The present study is focused on further development of these methods and the construction of a predictive ML model for characterizing the growth of thin films of silicon carbide (SiC) on Si by the atomic substitution method, which was discussed in detail in [14]. A film synthesized in accordance with this procedure grows through the direct reaction of carbon monoxide (CO) gas with a silicon substrate. The key growth parameters are temperature, pressure in the reactor chamber, component (silane, SiH_4, and CO) fluxes, and synthesis time. The process is also influenced by the conductivity type of the silicon substrate, its specific conductivity and crystallographic orientation, and a number of other factors. It has been demonstrated experimentally that the discussed procedure results in the formation of a thin SiC film with a porous silicon sublayer on the substrate surface. The scanning electron microscopy (SEM) image of a typical cleaved sample face is shown in Fig. 1, *a*.

As it was demonstrated in [14], the thickness of the single-crystal SiC layer has an extremely non-monotonic dependence on conditions and normally does not exceed 200 nm, since the film forms inside the substrate as a result of penetration of CO gas and, if additional techniques are not used [15], its formation ceases at a certain point. The maximum thickness of SiC is achieved at a certain pressure; if the pressure increases further, the thickness decreases [14]. Under fixed conditions, the thickness of a porous layer has a root dependence on time, and this dependence differs in nature from the common diffusion one. A detailed description of the mechanisms and causes of the emergence of a porous layer and a model of its formation were presented in [14] and [16]. It is extremely hard (if not impossible) to devise an analytical formula that characterizes accurately the dependence of thickness, crystalline perfection, pore density, or other characteristics on synthesis conditions at any arbitrary combination of growth conditions. At the same time, as was noted above, the use of machine learning methods allows one to solve

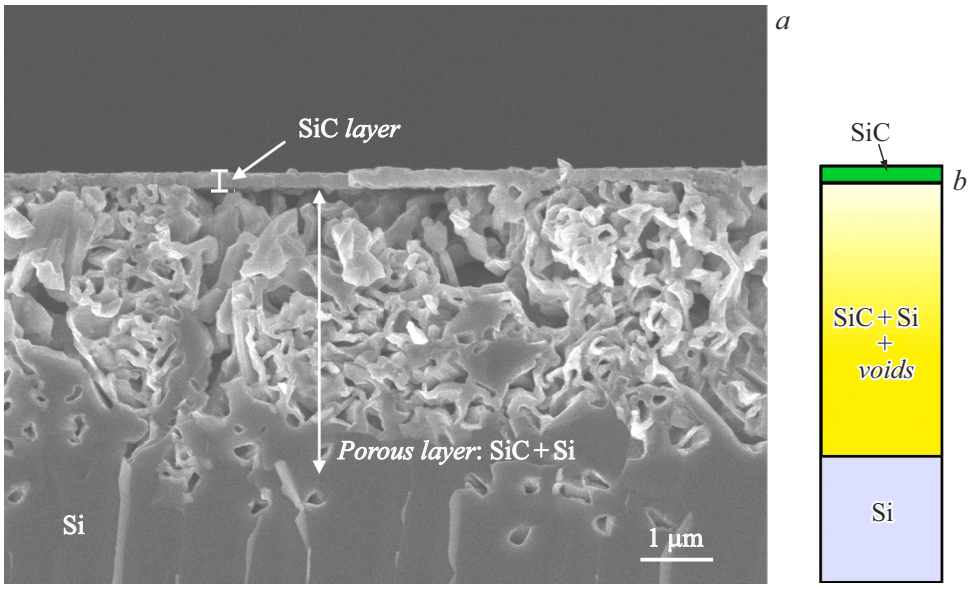


Figure 1. *a* — SEM image of a typical cleaved face of a SiC/Si sample grown by atomic substitution; *b* — diagram of the ellipsometric model [14] used to measure the thickness of the single-crystal (top) SiC layer.

such multidimensional regression problems efficiently and reveal hidden $R(\mathbf{X})$ dependences in sets of experimental data on the epitaxial process (R is the parameter of interest (e. g., the thickness of a film, its defect density, porosity, or any other characteristic that may be measured by examining the sample after growth) and $\mathbf{X}$ is a vector characterizing the synthesis conditions, substrate properties, or other important factors that may affect the studied parameter).

The availability of a sufficient amount of data is a necessary condition for constructing an adequate predictive model. An extensive set of experimental data on growth, which includes more than 1000 experiments with various combinations of growth parameters, was compiled in the course of development of the atomic substitution method [14]. In each experiment, the properties of the sample (including the thickness of the single-crystal SiC layer) were determined in accordance with the model shown in Fig. 1, *b* using a J.A. Woollam M-2000 ellipsometer. A total of 451 experiments, where, according to the ellipsometric model, the thickness of SiC was determined as accurately as possible (with a mean square error (MSE) less than 0.5), were selected from this set. Thus, the final data set used to train the ML model consisted of 451 rows, which included 11 numerical and categorical parameters of the synthesis process: type of a silicon substrate (p or n); its orientation ((100), (110), (111)); specific conductivity ($3.5-10\,000\,\Omega^{-1}\cdot\mathrm{cm}^{-1}$); pre-annealing time without gas (0–30 min); growth time (0–40 min); temperature (900–1330 °C); pressure (0.05–8 Torr); gas flow rates: total (5–20 sccm), CO flow rate (5–20 sccm), and SiH_4 flow rate (0–1 sccm); CO gas temperature (25–1200 °C); and ellipsometric thickness of a SiC film on the sample grown under the given conditions. The data were normalized to ensure stability of the subsequent training process of the neural network ML model. Prior to training, the data were divided into training and test samples in a 90/10 % ratio.

Python and the PyTorch library [17] were used to construct the model. The diagram of the model is shown in Fig. 2. The optimal hyperparameters were determined using the TPE algorithm [18]. Both the neural network architecture and the training parameters, such as number N of hidden layers of the neural network (2–5), the number of neurons in hidden layers (32–512), dropout rates (0.1–0.5), activation functions (ReLU, Tanh), learning rates (from 10^{-5} to 10^{-2}), and weight decay factors for neurons (from 10^{-6} to 10^{-2}) [19], were optimized. When the optimal combination of hyperparameters was found, the model was trained with them until the minimum of the error function in the test sample was reached.

The trained neural network model had fairly good predictive power with coefficient $R^2 = 0.57$ in an independent test set; i. e., it accounts for 57 % of variability of the quantity of interest. The remaining variability may be attributed both to noise in the data set (fluctuations and jumps in pressure, temperature during synthesis, etc.) and to unaccounted factors, such as different regimes of wafer preparation before the synthesis process. The mean absolute error (MAE) of the model prediction was 22 nm, which, with an overall range of SiC film thickness variation extending from 0 to $\sim$ 200 nm, is close to 10 % of the upper limit. Figure 3, *a* illustrates the predictive power of the model with scatterplots of predicted values against actual values. The figure reveals a strong linear correlation between them. If one takes into account the complexity, multifactorial nature, and nonlinearity of the SiC/Si growth process by atomic substitution [14] and the fact that no constants or a priori knowledge about the process or the reactor were used to

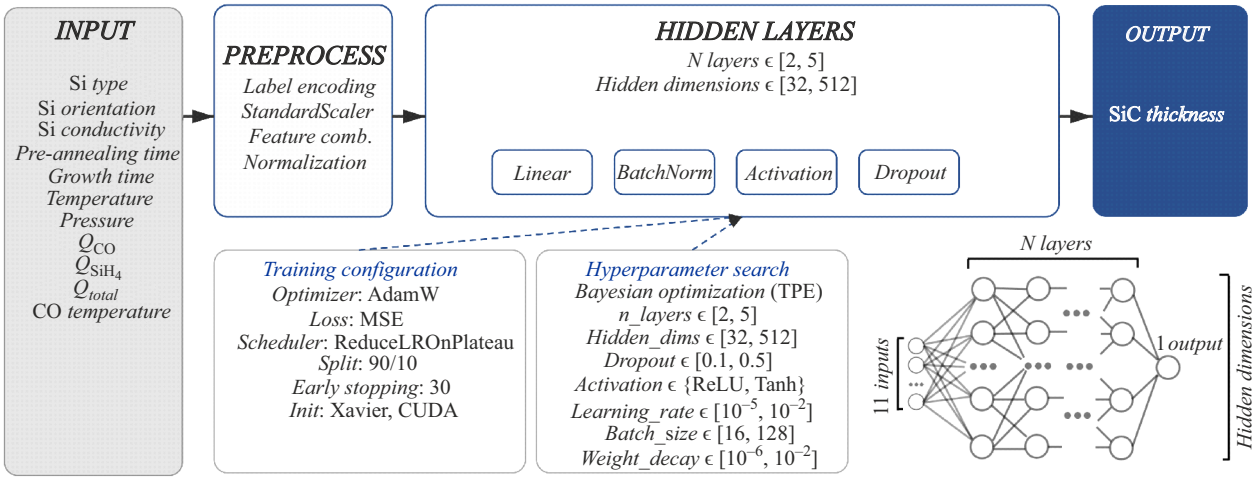


Figure 2. Schematic diagram of the neural network model, which is used to describe the SiC/Si epitaxial process, and its training: data reduction and scaling, selection of optimal hyperparameters providing the most accurate description, final model training, and accuracy assessment.

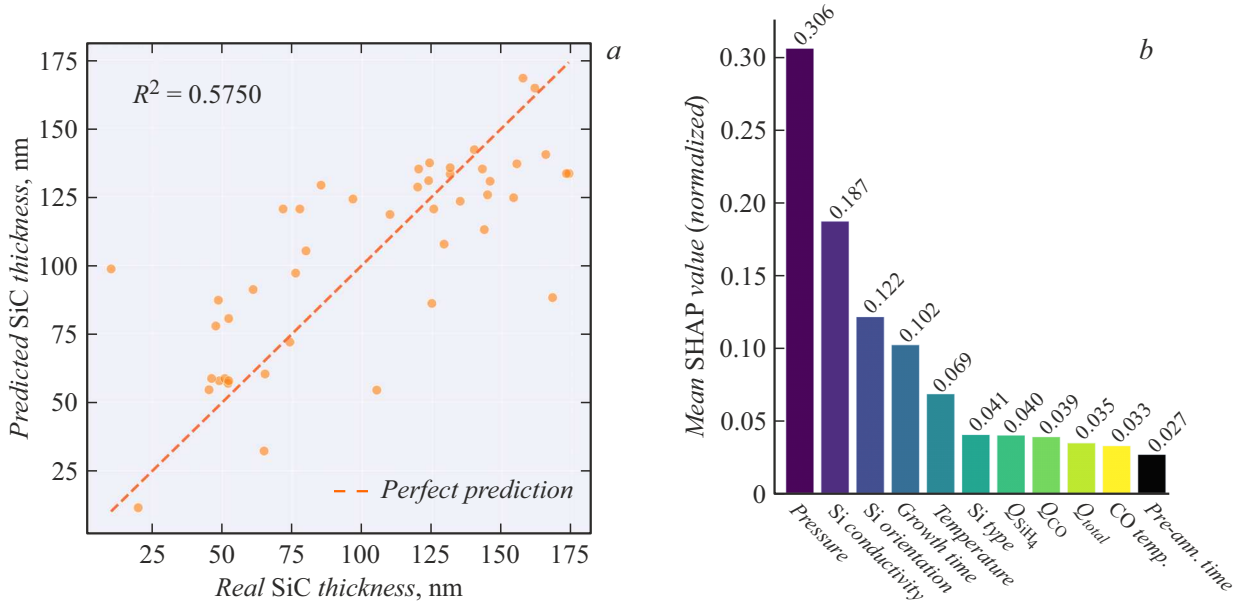


Figure 3. *a* — Comparison of true and predicted (by the trained model) values of the SiC layer thickness from the test part of the data set (not used for training); *b* — assessment of the influence of synthesis process parameters on SiC thickness performed using the SHAP method [20].

train the model (the model operates solely on the basis of available experimental data), this accuracy appears to be fairly high. However, a significant room for accuracy improvement, which may be achieved by increasing the size of the training data set and taking additional factors affecting growth into account, still remains. One of the advantages of ML models is the potential to assess the influence of a particular input parameter on the resulting prediction. Estimates of the influence of input parameters on the SiC layer thickness obtained using the SHAP method [20] are presented in Fig. 3, *b*. High mean SHAP values for such input parameters as pressure, substrate type, temperature, and synthesis time indicate that these factors are the ones governing the decisions of the model and account for the major part of explained variability. Other parameters, such as gas flow rates and pre-annealing time, exert a weaker influence (at least within the ranges of their variation examined in the present study). Note that such metrics are particularly valuable in the optimization of epitaxial processes and are the ones that need to be investigated further, since they allow one to identify key parameters in a limited set of expensive experiments.

Thus, a predictive neural network model, which allows for high-accuracy evaluation of the thickness of a grown single-crystal SiC layer based solely on growth conditions, was proposed and implemented. It was demonstrated that

artificial neural networks have a number of advantages over traditional analytical approaches when applied to epitaxial processes. Neural network models provide an opportunity to reveal complex nonlinear relations between process parameters and the properties of the resulting film while requiring no knowledge of mathematical expressions that characterize the entire set of phenomena in the reactor. This is especially valuable in epitaxial growth processes, where kinetic processes on the surface of a growing crystal, chemical reactions, and gas dynamics are intertwined. However, a number of limitations must be taken into account. The model is a „black box," which is a detriment to its interpretability, and requires a significant amount of experimental data, which may be hard to obtain for new materials. Its predictive power depends on data representativeness and is sensitive to measurement errors, which may lead to „retention" of spurious correlations instead of fundamental physical relations. However, despite these limitations, the approach is a valuable addition to traditional methods for characterizing epitaxial phenomena.

Acknowledgments

This study was carried out using the equipment of the unique scientific unit „Physics, Chemistry, and Mechanics of Crystals and Thin Films" (Institute for Problems in Mechanical Engineering of the Russian Academy of Sciences, St. Petersburg).

Funding

A.S. Grashchenko, S.A. Kukushkin, and A.V. Redkov performed their part of the work (sample synthesis, construction of a neural network model) as part of the state assignment of the Institute for Problems in Mechanical Engineering of the Russian Academy of Sciences (No. FFNF-2021-0001). A.V. Osipov performed his part of the work (data processing, measurement of the properties of samples) with support of the Ministry of Education and Science of the Russian Federation as part of the state assignment under contract No. 129360164 with the St. Petersburg State University.

Conflict of interest

The authors declare that they have no conflict of interest.

References

[1] N.G. Orji, M. Badaroglu, B.M. Barnes, C. Beitia, B.D. Bunday, U. Celano, R.J. Kline, M. Neisser, Y. Obeng, A.E. Vladar, Nat. Electron., **1** (10), 532 (2018). DOI: 10.1038/s41928-018-0150-9

[2] Y.K. Wakabayashi, T. Otsuka, Y. Krockenberger, H. Sawada, Y. Taniyasu, H. Yamamoto, APL Mater., **7** (10), 101114 (2019). DOI: 10.1063/1.5123019

[3] G. Wu, Y. Wang, Q. Gong, L. Li, X. Wu, IEEE Access, **10**, 9848 (2022). DOI: 10.1109/ACCESS.2022.3143811

[4] H. Shi, Z. Jin, W. Tang, J. Wang, K. Jiang, M. Xu, W. Xia, X. Xu, Knowledge-Based Syst., **280**, 110994 (2023). DOI: 10.1016/j.knosys.2023.110994

[5] T.C. Kaspar, S. Akers, H.W. Sprueill, A.H. Ter-Petrosyan, J.A. Bilbrey, D. Hopkins, A. Harilal, J. Christudasjustus, P. Gemperline, R.B. Comes, J. Vac. Sci. Technol. A, **43** (3), 032702 (2025). DOI: 10.1116/6.0004493

[6] R.K. Vasudevan, A. Tselev, A.P. Baddorf, S.V. Kalinin, ACS Nano, **8** (10), 10899 (2014). DOI: 10.1021/nn504730n

[7] I. Ohkubo, Z. Hou, J.N. Lee, T. Aizawa, M. Lippmaa, T. Chikyow, K. Tsuda, T. Mori, Mater. Today Phys., **16**, 100296 (2021). DOI: 10.1016/j.mtphys.2020.100296

[8] A.V. Redkov, Acta Mater., **287**, 120762 (2025). DOI: 10.1016/j.actamat.2025.120762

[9] V.N. Bessolov, D.V. Karpov, E.V. Konenkova, A.A. Lipovskii, A.V. Osipov, A.V. Redkov, I.P. Soshnikov, S.A. Kukushkin, Thin Solid Films, **606**, 74 (2016). DOI: 10.1016/j.tsf.2016.03.034

[10] A.V. Redkov, S.A. Kukushkin, Cryst. Growth Des., **20** (4), 2590 (2020). DOI: 10.1021/acs.cgd.9b01721

[11] A. Redkov, Front. Chem., **11**, 1189729 (2023). DOI: 10.3389/fchem.2023.1189729

[12] A. Redkov, S. Kukushkin, Faraday Discuss., **235**, 362 (2022). DOI: 10.1039/D1FD00083G

[13] A. Redkov, Crystals, **14** (1), 25 (2023). DOI: 10.3390/cryst14010025

[14] S.A. Kukushkin, A.V. Osipov, J. Phys. D, **47**, 313001 (2014). DOI: 10.1088/0022-3727/47/31/313001

[15] A.S. Grashchenko, S.A. Kukushkin, A.V. Osipov, A.V. Redkov, Catal. Today, **397**, 375 (2022). DOI: 10.1016/j.cattod.2021.08.012

[16] A.V. Redkov, A.S. Grashchenko, S.A. Kukushkin, A.V. Osipov, K.P. Kotlyar, A.I. Likhachev, A.V. Nashchekin, I.P. Soshnikov, Phys. Solid State, **61**, 299 (2019). DOI: 10.1134/S1063783419030272.

[17] A. Paszke, arXiv:1912.01703 (2019).

[18] T. Akiba, S. Sano, T. Yanase, T. Ohta, M. Koyama, in *Proc. of the 25th ACM SIGKDD Int.Conf. on Knowledge Discovery & Data Mining* (Anchorage, USA, 2019), p. 2623–2631. DOI: 10.1145/3292500.3330701

[19] M. Feurer, F. Hutter, in *Automated machine learning*, ed. by F. Hutter, L. Kotthoff, J. Vanschoren (Springer, Cham, 2019), p. 3–33. DOI: 10.1007/978-3-030-05318-5_1

[20] W.E. Marcílio, D.M. Eler, in *2020 33rd SIBGRAPI Conf. on Graphics, Patterns and Images (SIBGRAPI)* (IEEE, 2020), p. 340–347. DOI: 10.1109/SIBGRAPI51738.2020.00053

Translated by D.Safin