Krasnikov, Dmitry V.; Khabushev, Eldar M.; Gaev, Andrey; Bogdanova, Alisa R.; Iakovlev, Vsevolod Ya.; Lantsberg, Anna; Kallio, Tanja; Nasibulin, Albert G.

Machine learning methods for aerosol synthesis of single-walled carbon nanotubes

Published in:
Carbon

DOI:
10.1016/j.carbon.2022.10.044

Published: 15/01/2023

Document Version
Publisher's PDF, also known as Version of record

Carbon 202 (2023) 76–82

Contents lists available at ScienceDirect

Carbon

journal homepage: www.elsevier.com/locate/carbon

Machine learning methods for aerosol synthesis of single-walled carbon nanotubes

Dmitry V. Krasnikov [a], Eldar M. Khabushev [b,**], Andrey Gaev [c], Alisa R. Bogdanova [b], Vsevolod Ya. Iakovlev [d], Anna Lantsberg [c], Tanja Kallio [b,*], Albert G. Nasibulin [a]

[a] Skolkovo Institute of Science and Technology, 3 Nobel Str., 121205, Moscow, Russia
[b] Aalto University, School of Chemical Engineering, Kemistintie 1, 02150, Espoo, Finland
[c] Bauman Moscow State Technical University, Baumanskaya 2-ya Str, 5/1, 105005, Moscow, Russia
[d] Kiev Institute of Business and Technology, Zoryanyi Lane, 1/5, 04078, Kyiv, Ukraine

ARTICLE INFO

Keywords:
Floating catalyst CVD
Single-walled carbon nanotube
Machine learning
Transparent conductive films

ABSTRACT

This work is devoted to the strategy towards the optimal development of multiparametric process of single-walled carbon nanotube (SWCNT) synthesis. Here, we examine the implementation of machine learning techniques and discuss features of the optimal dataset size and density for aerosol chemical vapor deposition method with a complex carbon source. We employ the dataset of 369 points, comprising synthesis parameters (catalyst amount, temperature, feed of carbon sources) and corresponding carbon nanotube characteristics (yield, quality, structure, optoelectrical figure of merit). Assessing the performance of six machine learning methods on the dataset, we demonstrate Artificial Neural Network to be the most suitable approach to predict the outcome of synthesis processes. We show that even a dataset of 250 points with the inhomogeneous distribution of input parameters is enough to reach an acceptable performance of the Artificial Neural Network, wherein the error is most likely to arise from experimental inaccuracy and hidden uncontrolled variables. We believe our work will contribute to the selection of an appropriate regression algorithm for the controlled carbon nanotube synthesis and further development of an autonomous synthesis system for an "on-demand" SWCNT production.

1. Introduction

Catalytic chemical vapor deposition (CVD) is nowadays one of the most well-established methods for a large-scale production of various nanomaterials such as low-dimensional carbon allotropes [1], semiconducting nanowires [2], 2D crystals [3], *etc*. The bottom-up synthesis ideology of CVD usually involves complex processes on different length scales including atomic level catalysis, nanoscale transport phenomena, catalyst particle-bed/particle interaction, as well as heat and mass transfer at macroscopic level [4]. The complex impact of synthesis parameters and reactor architecture on the CVD growth could be barely described by classical theoretical methods, while optimization of reaction outcome (a product) is typically a challenging and time-consuming task. Indeed, the typical CVD process employs a multidimensional and strongly interconnected parameter space making manual optimization with a trial-and-error approach "the worst form except all the others".

One of the best examples of conventional approaches falling short to facilitate the material development is the CVD growth of single-walled carbon nanotubes (SWCNTs). The SWCNT synthesis with tailored output has been the topic of numerous discussions driven by groundbreaking perspectives for transistor technology [5], optoelectronic applications [6,7], bioimaging [8], *etc*. Intensive three-decade-long research has resulted in the development of various routes to govern SWCNT synthesis including catalyst design [9–11], selective treatment [12,13], growth kinetics manipulation [14–18], and reactor tuning [19,20]. However, a lack of established theory of catalytic SWCNT growth challenges the efficient association of current advances and further development towards the desired goal of controlled output.

Machine learning (ML) was shown to be extremely helpful for tackling strongly nonlinear and multidimensional tasks [21] including the ones related to material science and physical chemistry [22–24]. ML methods have shown their applicability for experiment planning, optimization, and prediction of material properties in terms of thin-film technology [25] and nanomaterial development [26–28], as they

* Corresponding author.
** Corresponding author.
E-mail addresses: eldar.khabushev@aalto.fi (E.M. Khabushev), tanja.kallio@aalto.fi (T. Kallio).

https://doi.org/10.1016/j.carbon.2022.10.044
Received 6 August 2022; Received in revised form 11 October 2022; Accepted 18 October 2022
Available online 21 October 2022

efficiently restore the relationship between experimental conditions and the output of the reaction. Thus, for substrate-supported laboratory scale CVD of SWCNTs, a closed-loop automated system coupled with ML was developed [29] and successfully implemented for yield maximization [30] and diameter control [31], promoting autonomous reaction research.

The aerosol (floating-catalyst) CVD method is a special case when a catalyst is not fixed on a support but introduced into a flow reactor as a pre-formed aerosol [32,33] or as a volatile catalyst precursor [34]. The main benefits of the aerosol CVD method are continuous operation, excellent scalability, and technological effectiveness. However, aerosol CVD fails to comply with the classical optimization approach for a catalytic process as a single particle can be hardly immobilized and thoroughly examined. In combination with ML, aerosol CVD opens the way for a fully autonomous industrial-scale system for controlled nanotube production facilitating advanced nanotube applications. However, the development of an automated aerosol system requires more-sophisticated instrumentation for *in situ* characterization. Aerosol CVD brings to the field complex aerosol science and flow dynamics making ML implementation especially relevant. Previously, Artificial Neural Network (ANN) [35] and Support Vector Machine [36] served the simplest aerosol CVD based on the Boudouard reaction [37] to predict output SWCNT properties (diameter, yield, film sheet resistance, defectiveness, *etc.*) from experimental conditions. However, ML implementation to a hydrocarbon-based process looks more complicated and has not been yet performed. This type of more sophisticated aerosol CVD methods exhibits stronger nonlinearity of reaction parameter space and vast amount of possible reaction routes but may provide higher productivity [38] or even higher quality of carbon nanotubes [39]. Moreover, there is a lack of discussion on the optimal algorithm for CVD nanotube synthesis.

In this work, we employ the hydrocarbon-based aerosol CVD method (implementing toluene and ethylene as carbon sources [40]), providing multiparametric data for the comparative study of different ML algorithms. We used sheet resistance at 90% transmittance–quality factor for transparent conductive films– as one of the output parameter for regression task. The acquired results allow us to describe the optimal dataset for the highest prediction accuracy. We believe our work to facilitate the successful ML algorithm selection for aerosol CVD, thus, guiding the path for the development of a fully autonomous system for controlled production of SWCNTs on an industrial scale.

2. Experimental

SWCNT films were produced using a vertical tubular reactor with inner diameter of 44 mm and length of 1.3 m. Briefly, ferrocene (catalyst precursor; Sigma-Aldrich, 98%) and thiophene (growth promoter; Sigma-Aldrich, 99%) were dissolved in toluene (Sigma-Aldrich, ≥99.9%); the solution was injected into the reactor with a syringe pump (NE-1000, New Era Instruments). Apart from toluene, ethylene (Linde Gas, 4.5 purity grade) was used as a second carbon source; it was premixed with a carrier gas – hydrogen (Linde Gas, 5.0 purity grade). High precision mass flow controllers (Alicat Scientific, MC-Series) were used to maintain the required gas flow rates. SWCNTs were collected at the outlet of the reactor on a filter (MF-Millipore, 0.45 μm effective pore size), and then the film was transferred to a quartz slab for characterization.

We utilized optical absorption spectroscopy (PerkinElmer, Lambda 1050) estimating the yield as the optical density of the film (absorbance; A_{550}) collected per unit time (t) and normalized by the collection area of the filter (S) and gas flow rate (Q):

$$Yield\ \left[cm^2 \bullet L^{-1}\right] = \frac{A_{550}}{t} \bullet \frac{S}{Q}.$$

The Raman spectroscopy of the samples was performed at 532 nm excitation wavelength using a confocal Raman microscope (Renishaw inVia). The spectra were fitted as a composition of Lorentzian peaks and the intensity ratio of graphitic (G) and disorder-induced (D) bands (I_G/I_D), describing the quality of nanotube structure [41], were calculated. Measurements for each sample were repeated at least three times, the results were averaged.

Four-probe unit (Jandel RM3000) was implemented for sheet resistance (R_S) measurements. For each film, the measurements were repeated five times at different probe arrangements (relative to the sample) and averaged. To evaluate the optoelectronic performance of the SWCNT films, we applied an equivalent sheet resistance at a transmittance of 90% in the middle of visible wavelength calculated according to the following equation [42]:

$$R_{90} = R_S \bullet \frac{A_{550}}{log_{10}(0.9)}$$

3. Results and discussion

3.1. Dataset characterization

Previously our research group published a systematic study of the joint ethylene and toluene effect on nanotube synthesis in terms of aerosol CVD exploring carbon feedstock effects on transparent conductive properties of SWCNT films with a focus on growth mechanism [43]. We expanded the previously obtained data with new experimental results combining them in a dataset of 369 points (Supplementary Information). Temperature, ethylene concentration, ferrocene and toluene feed rates formed a four-dimensional vector of features describing synthesis conditions. We kept other available variables (*e.g.*, gas flow rate, ratio of ferrocene and thiophene) constant to limit the dataset complexity, despite their apparent importance for the nanotube growth [18,19,44]. It is worth mentioning that those parameters might also be used as additional features for regression task and provide not only the better understanding of synthesis process but also improve the performance of certain regression models.

Each sample was studied by a set of characterization techniques including Raman and optical absorption spectroscopy and 4-probe resistivity measurements. The characteristics of produced SWCNT films, in particular, Yield (collection rate), I_G/I_D (as a measure of quality), R_{90} (a figure of merit for nanotube-based transparent electrodes), and SWCNT diameter (*d;* key feature for SWCNT-based optical applications) were considered as labels (output parameters) for the regression task (Fig. 1). The scatter plots depicting the mutual correlation between target variables and data features (Fig. S1) indicate the strongly non-linear nature of synthesis processes and resulting material properties that justifies the implementation of the ML approach.

The distributions of studied nanotube characteristics are depicted in Fig. 2. In general, both Yield and R_{90} labels range over 5 orders of magnitude with a close-to-log-normal distribution, while I_G/I_D values show close-to-normal distribution within a range from 1 to 300. Mean diameter (d) varies in a rather narrow range from 1 to 2.5 nm. Nevertheless, we used the logarithms of R_{90} and Yield as labels to reduce the effect of dataset inhomogeneity and a wide-range variation of labels. In addition, we carried out the data preprocessing of both features and labels prior to feeding the data to regression algorithms. For this, each feature and label were independently standardized as follows: $\overline{X} = \frac{X-\mu}{\sigma}$, where μ and σ are mean value and dispersion of certain feature/label, denoted here as X.

Measuring the film characteristics with several repetitions allowed us to estimate the mean square error of labels and use it further in data analysis (Fig. 2e). We found I_G/I_D to be the most fluctuating parameter with the widest distribution of relative deviation, reaching up to 90% for certain points, however, its third quartile (upper part of solid box) was below 18%. The mean values of relative deviation (hereinafter experimental error) of Yield, $R_{90,}$ and I_G/I_D were estimated to be 5.3%, 9.6%, and 15.2%, correspondingly. The deviation of mean diameter was

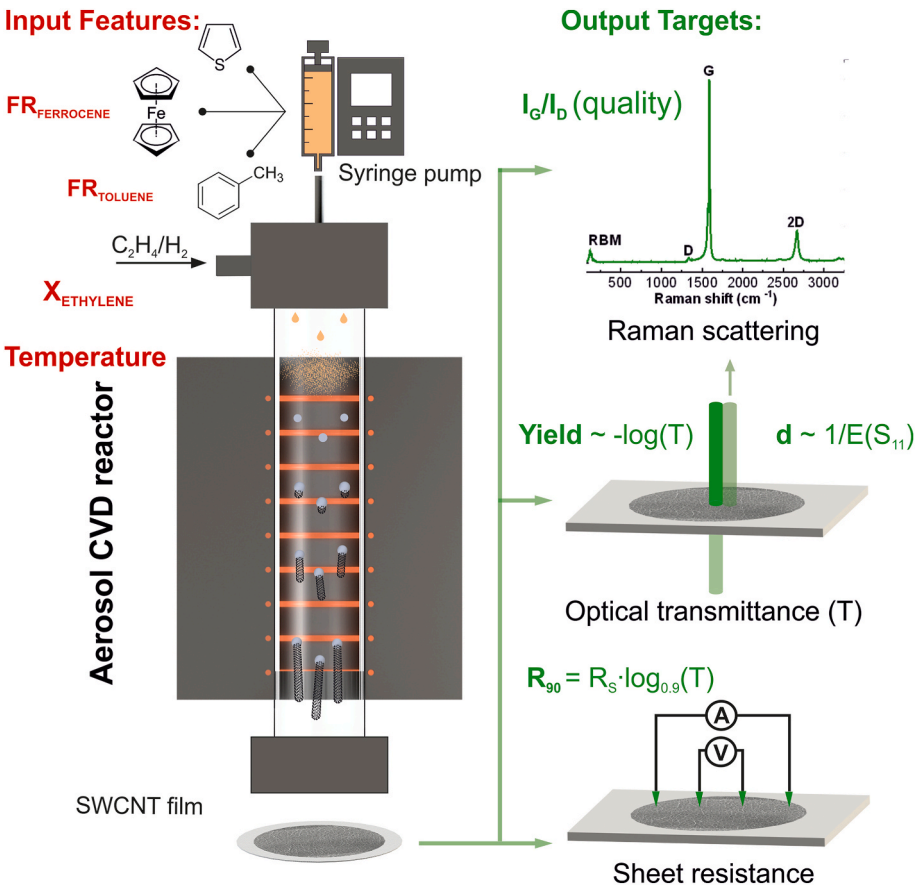

Fig. 1. Reactor set-up and description of input (reaction conditions) and output (nanotube characteristics) parameters. Visualization of collected dataset reflecting mutual dependence of nanotube film characteristics with reaction parameters. (A colour version of this figure can be viewed online.)

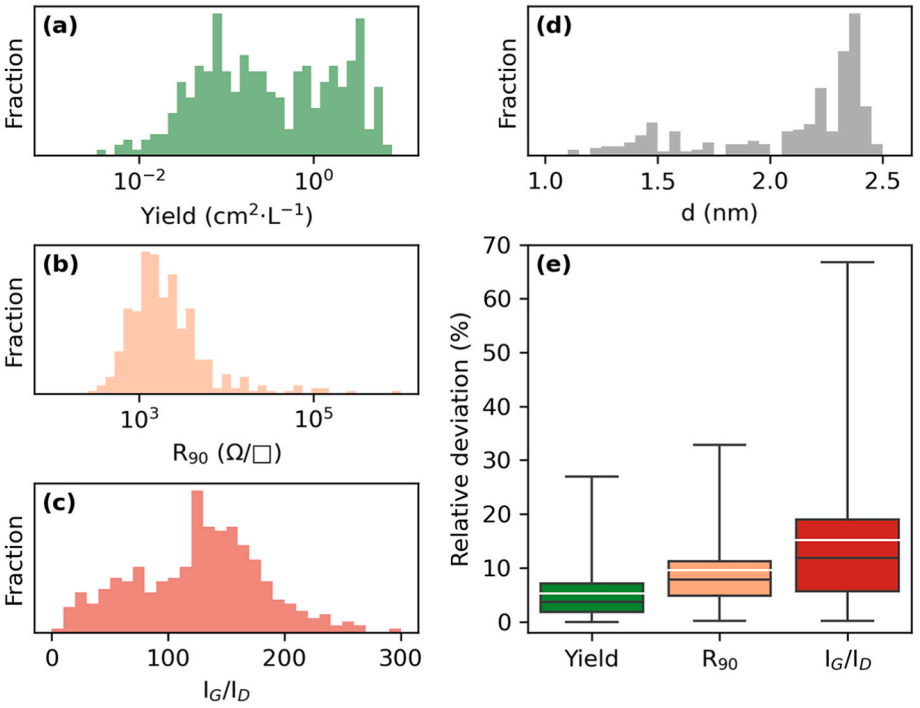

Fig. 2. The histograms of Yield, R_{90}, I_G/I_D and d distributions. Boxplot of relative deviation of Yield, R_{90}, and I_G/I_D. White and black lines in the box correspond to mean and median values of deviation. Whiskers on the box plot corresponds to the range four-times higher than interquartile one. (A colour version of this figure can be viewed online.)

estimated to be *ca.* 1.3% and it was considered the same for all the samples.

3.2. Machine learning methods

For the comparative analysis of the ML approaches on the experimental data we evaluated the performance of several classical regression models realized in Scikit-Learn Python package [45]. In particular, we employed simple Linear and Polynomial (linear model with polynomial features) models, as well as more advanced Support Vector Machine (SVM) and Random Forest algorithms successfully implemented in previous works [36,46]. Besides, in pursuit of achieving state-of-the-art performance, we tested XGBoost regressor – one of the most efficient gradient boosting decision tree method [47]. The experimental dataset was split into 5 folds and model performance was cross-validated on the data using mean absolute percentage error (MAPE) as a prediction score. Each fold was consequentially used as a testing subset for performance evaluation, while the other four were used for training, and the results were averaged. We performed the Grid Search hyperparameter optimization for each implemented model: the optimized hyperparameters (denoted according to Scikit-learn and XGBoost documentation) could be found in Supporting Information (Table S1).

The averaged results of cross-validation are presented in Table 1. Firstly, all tested algorithms demonstrated good performance for predicting nanotube mean diameter with MAPE ranging from 3.6 to 7.5%. The main reason for the relatively good performance is narrow distribution of the label (diameter), as well as the prevailing effect of a single input parameter (ethylene concentration) on nanotube diameter [39]. This was not the case for other parameters (Yield, I_G/I_D, and R_{90}). As we expected, simple methods of classical machine learning (linear and polynomial regression) demonstrated poor performance on the rest of the dataset with MAPE above 50% for each investigated label. The reason for such a poor performance might lie in the highly nonlinear dependence of output targets (material characteristics) on input synthesis parameters, which cannot be properly described by simple models. The situation significantly improves with the use of SVM, demonstrating moderate performance on the dataset and reaching the lowest regression error of 28% for R_{90}. Decision tree-based methods such as Random Forest and XGBoost have slightly different scores handling well with I_G/I_D labels and being less effective for the R_{90} prediction. It is worth mentioning that XGBoost method noticeably outperforms its counterparts for I_G/I_D output parameter reaching MAPE score of 23%. Despite all three models (SVM, Random Forest and XGBoost) employ ensemble based approach, the best performance of XGBoost was anticipated because of its unique ability to adaptively insert new trees in the ensemble to reduce the residual loss from previous trees.

Despite classical algorithms of ML adequately coping with the experimental dataset, the error score achieved is significantly higher than the estimated experimental inaccuracy. In pursuit of prediction quality improvement, we referred to the deep learning approach – Artificial Neural Network.

Table 1
Performance (MAPE [%]) of classical ML algorithms on the experimental dataset; the accuracy of the experiment is also given as a reference value.

Model	R_{90}	I_G/I_D	Yield	d
Std. experimental error	9.6	15.2	5.3	1.3
Linear	61	52	115	7.5
Polynomial	61	52	63	5.9
SVM	**28**	30	35	**4.0**
Random Forest	33	27	33	4.1
XGBoost	33	**23**	**29**	4.3
ANN	**22**	**23**	**27**	**3.6**

3.3. Artificial neural network

Artificial neural networks (ANNs) have good potential for complex regression tasks since they possess a greater number of training parameters than classical methods. However, being a data-greedy algorithm, ANN has an increased risk of over-fitting. For comparison with classical ML algorithms, we selected a fully connected neural network, which was successfully implemented in our previous work for processing the data from a carbon monoxide-based synthesis system [35]; the architecture of the ANN is depicted in Fig. S2. Briefly, the sequential model of ANN, consisting of 3 hidden layers with a ReLu activation function, was developed using the open-source Keras library [48]. To counteract overfitting and boost training speed dropout layers followed the 1st and the 3rd hidden layers, and a batch normalization layer was placed after the 2nd hidden layer. Besides, in terms of ANN training, the following data augmentation procedure was used: we oversampled the training subsets generating 20 distorted copies of each data sample with the same features, but labels distributed normally around the actual value with the standard deviation equal to the experimental error. At the same time, the testing subset contained only real samples which were not oversampled. This approach allowed us not only to artificially enlarge the dataset facilitating ANN training but also provided an additional regularization scheme taking into account the accuracy of experimental labels.

In pursuit of optimizing ANN performance, we carried out the hyperparameter adjustment with a Bayesian Optimization Keras tuner, varying the number of nodes for each hidden layer, and drop-out parameters. Best hyperparameters could be found in Supplementary Materials (Table S2). The learning curves for ANN with optimized parameters are depicted in Fig. 3. In general, approximately 50–100 training epochs are sufficient for algorithms to reach convergence on the validation dataset for every output parameter. Surprisingly, despite a scarcity of the data used, we did not observe noticeable model overfitting (increase in error on validation dataset with epoch number) after 300 learning epochs. Based on the results of cross-validation, we can state the optimized ANN outperforms the methods of classical machine learning. Thus, the MAPE of 27%, 23%, 22%, and 3.6% were obtained after cross-validation for Yield, R_{90}, I_G/I_D, and diameter predictions correspondingly, demonstrating slight improvement over the previous results. It should be stressed that the ratio for mean average percentage errors and experimental errors (5, 2, 1.5, and *c.a.* 3 times higher for Yield, R_{90}, I_G/I_D, and d, respectively) corresponds to previous results of the ANN for the aerosol CVD based on the simple Boudouard reaction (3, 2.8, and 4 times for I_G/I_D, Yield, and d, respectively) [35]. Thus, the enhanced error is rather provided by the experiment design and method complexity.

ANN noticeably outperforms classical algorithms in the R_{90} parameter prediction demonstrating 6% percent improvement in score. This might be attributed to the unique ability of ANN to learn high-level features from data in an incremental manner and the complexity of the R_{90} parameter. Indeed, as it was demonstrated previously, the R_{90} is an intricate parameter strongly interconnected with the structural characteristics of SWCNTs [49] and it has a strongly nonlinear relationship with synthesis parameters [43].

It is worth mentioning that the best result on I_G/I_D prediction corresponds to the highest experimental error of 15.2% (Table 1), while the reverse result is obtained for the yield label. This fact might indicate the underestimated error values for yield, as follows, resulting in lower efficiency of implemented augmentation procedure in describing experimental inaccuracy.

3.4. Optimal dataset

Despite the improvement in prediction quality, achieved MAPE values are significantly higher than the estimated experimental accuracy: 5, 2, and 1.5 times higher for Yield, I_G/I_D, and R_{90}, respectively.

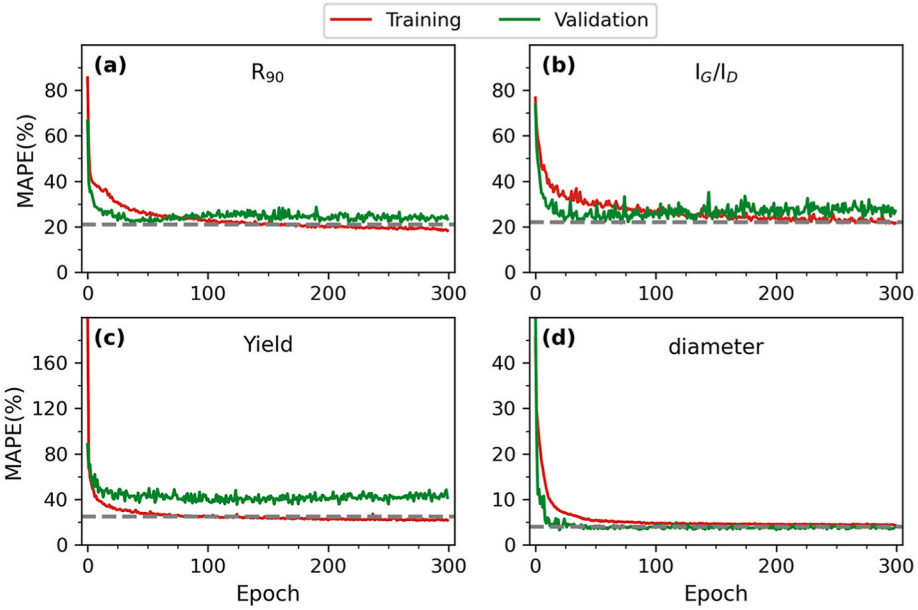

Fig. 3. ANN learning curves for different target parameters: R_{90} (a), I_G/I_D (b), Yield (c) and diameter (d). The grey dashed lines correspond to achieved prediction errors. (A colour version of this figure can be viewed online.)

Taking into account ANN being a data greedy method, firstly, we intuitively assumed that a dataset of a bigger size should improve the prediction performance. Since dataset enlargement is an expensive and laborious process, we approached the issue of appropriate dataset size from another angle instead – we randomly sliced our dataset into subsets of different sizes (training size was four-time bigger than testing) and evaluated the algorithm performance (implementing previously discussed augmentation procedure). The number of generated training subsets of each size was reversely proportional to the size (See Table S3).

The results of the analysis for MAPE of R_{90} (the most complex parameter) are depicted in Fig. 4. In general, a prediction score shows a downward trend below the training size of *ca.* 200 points and no noticeable change above. Small size of training subsets also results in a huge deviation of MAPE, though the situation improves further, reaching a standard deviation value of 4% at the training subset sizes above 250 points. Thus, we can assume the dataset of 250 points to be sufficient to effectively describe the space of synthesis parameters used in our study and no further improvement in the prediction score is expected to be achieved by dataset enlargement. In contrast, the prediction score is rather limited by the quality of the points or accuracy of the data collection.

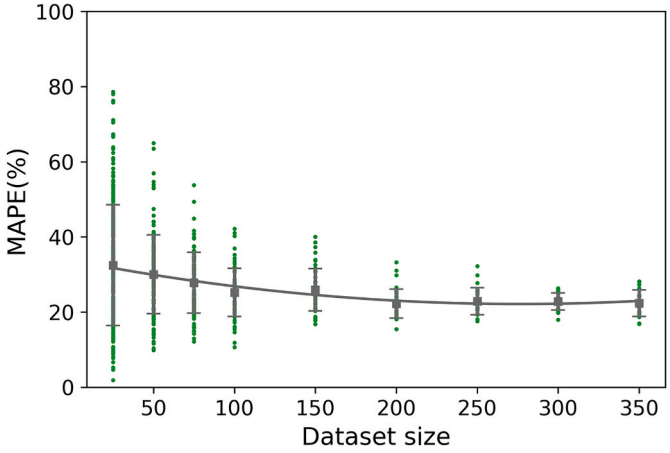

Fig. 4. The effect of training dataset size on ANN performance for R_{90} label. Ratio of training to testing sizes was kept constant at 4:1. Green dots indicate each repetition. The Grey error bars describe mean and standard deviation for each training size. (A colour version of this figure can be viewed online.)

Another hypothesis regarding the high ratio between MAPE and experimental error is a non-homogeneous distribution of the input parameters over the chemical space. Indeed, a significant error might be brought by outlying points of extreme parameter variation (*e.g.* too high ethylene concentration). To assess this hypothesis, the procedure for data discarding from a four-dimensional space might be elaborated. To reduce the dimensionality of the variables to one, we employed the principal component analysis (PCA), that allowed us to "trim" (Fig. 5), and better visualize the dataset (see Fig. S4). The dataset trimming has almost no effect on R_{90} prediction accuracy; though for other parameters (Fig. S3) the downward trend of MAPE with trimmed fraction is observed justifying to some extent the need for an even space of input parameters. Nevertheless, the two-fold difference between MAPE and experimental error implies the role of hidden parameters of the employed hydrocarbon aerosol CVD method (*e.g.* memory effect of the reactor walls [50] or SWCNT doping with ambient [51]).

Thus, the achieved results are comparable with experimental error and we believe them to be fruitful for the development of fully-autonomous research and advanced optimization.

4. Conclusions

In conclusion, we evaluated the performance of several ML algorithms for the regression task on the data from a complex hydrocarbon-based aerosol CVD process for the SWCNT synthesis. Having acquired the dataset of 369 points comprising 4 input and 3 target parameters, we showed the implementation of a fully connected artificial neural network to outperform other tested algorithms, especially, coping better with complex R_{90} parameters presumably due to its ability to learn high-level features. We reached mean average percentage errors comparable to experimental errors (5, 1.5, and 2 times higher for Yield, I_G/I_D, and R_{90}, respectively) and examined possible factors limiting further progress. We showed that even a dataset of 250 points (out of 369) is sufficient to provide the same accuracy refuting thereby the amount of data being a key barrier. As for the quality of the data, using principal component analysis, we were able to reduce the number of variables and show that even distribution of input features might improve ANN

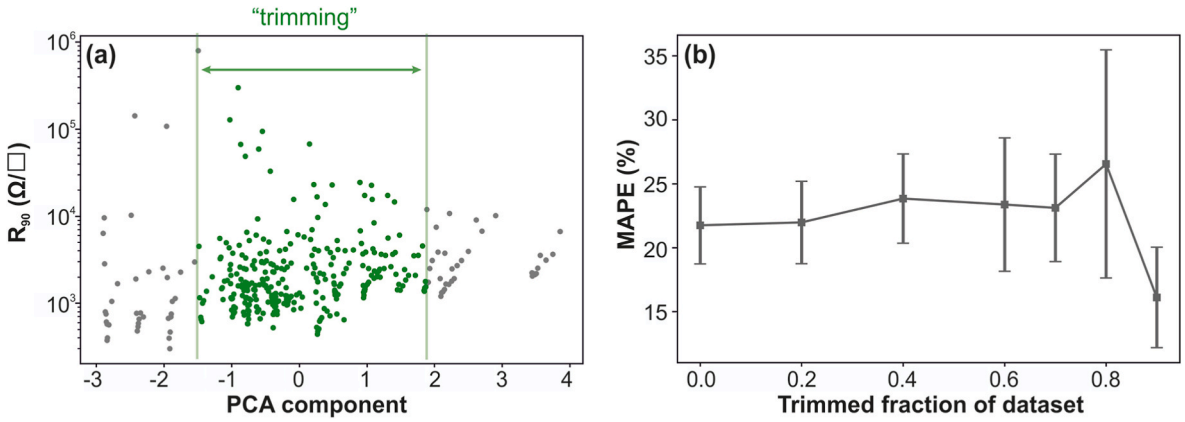

Fig. 5. (a) PCA analysis of the data and illustration of data reduction ("trimming"). (b) MAPE of R_{90} prediction with ANN as a function of trimmed dataset fraction. (A colour version of this figure can be viewed online.)

performance to MAPE (e.g. 25% for I_G/I_D). We believe ANN to be an effective approach for the controlled aerosol CVD synthesis task, in particular for autonomous aerosol CVD operation speeding up the "on-demand" synthesis of SWCNTs.

CRediT authorship contribution statement

Dmitry V. Krasnikov: Conceptualization, Funding acquisition, Writing – review & editing. **Eldar M. Khabushev:** Methodology, Investigation, Data curation, Visualization, Writing – original draft. **Andrey Gaev:** Formal analysis, Data curation. **Alisa R. Bogdanova:** Investigation, Formal analysis. **Vsevolod Ya. Iakovlev:** Data curation, Validation. **Anna Lantsberg:** Validation, Supervision. **Tanja Kallio:** Resources, Funding acquisition, Supervision, Writing – review & editing. **Albert G. Nasibulin:** Conceptualization, Funding acquisition, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank Prof. Esko I. Kauppinen and Dr. Fedor Fedorov for fruitful discussions. E.M.Kh. and T.K. acknowledge Academy of Finland (Profi 5 Project) and Finnish National Agency for Education (the EDUFI fellowship grant). D.V.K. acknowledges RSF (grant No. 20-73-10256; ANN, optimal dataset). D.V.K and A.G.N. acknowledge Council on grants of RF (grant No НШ-1330.2022.1.3).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.carbon.2022.10.044.

References

[1] R. Rao, C.L. Pint, A.E. Islam, R.S. Weatherup, S. Hofmann, E.R. Meshot, et al., Carbon nanotubes and related nanomaterials: critical advances and challenges for synthesis toward mainstream commercial applications, ACS Nano 12 (2018) 11756–11784, https://doi.org/10.1021/acsnano.8b06511.

[2] L. Güniat, P. Caroff, A. Fontcuberta I Morral, Vapor phase growth of semiconductor nanowires: key developments and open questions, Chem. Rev. 119 (2019) 8958–8971, https://doi.org/10.1021/ACS.CHEMREV.8B00649/ASSET/IMAGES/LARGE/CR-2018-00649N_0006.JPEG.

[3] J. Zhang, F. Wang, V.B. Shenoy, M. Tang, J. Lou, Towards controlled synthesis of 2D crystals by chemical vapor deposition (CVD), Mater. Today 40 (2020) 132–139, https://doi.org/10.1016/J.MATTOD.2020.06.012.

[4] I. Chorkendorff, J.W. Niemantsverdriet, Concepts of Modern Catalysis and Kinetics, first ed., Wiley, 2003 https://doi.org/10.1002/3527602658.

[5] J. Lefebvre, J. Ding, Z. Li, P. Finnie, G. Lopinski, P.R.L. Malenfant, High-purity semiconducting single-walled carbon nanotubes: a key enabling material in emerging electronics, Acc. Chem. Res. 50 (2017) 2479–2486, https://doi.org/10.1021/acs.accounts.7b00234.

[6] X. Wei, S. Li, W. Wang, X. Zhang, W. Zhou, S. Xie, et al., Recent advances in structure separation of single-wall carbon nanotubes and their application in optics, electronics, and optoelectronics, Adv. Sci. (2022), https://doi.org/10.1002/ADVS.202200054.

[7] D.S. Kopylova, N.Y. Boldyrev, V.Y. Iakovlev, Y.G. Gladush, A.G. Nasibulin, A bolometer based on single-walled carbon nanotubes and hybrid materials, Quant. Electron. 46 (2016) 1163–1169, https://doi.org/10.1070/QEL16146/XML.

[8] J. Pan, F. Li, J.H. Choi, Single-walled carbon nanotubes as optical probes for bio-sensing and imaging, J. Mater. Chem. B. 5 (2017) 6511–6522, https://doi.org/10.1039/C7TB00748E.

[9] M. He, S. Zhang, Q. Wu, H. Xue, B. Xin, D. Wang, et al., Designing catalysts for chirality-selective synthesis of single-walled carbon nanotubes: past success and future opportunity, Adv. Mater. 31 (2019) 1–25, https://doi.org/10.1002/adma.201800805.

[10] S. Forel, A. Castan, H. Amara, I. Florea, F. Fossard, L. Catala, et al., Tuning bimetallic catalysts for a selective growth of SWCNTs, Nanoscale 11 (2019) 3952–3957, https://doi.org/10.1039/c8nr09589b.

[11] S. Ahmad, Y. Liao, A. Hussain, Q. Zhang, E.X. Ding, H. Jiang, et al., Systematic investigation of the catalyst composition effects on single-walled carbon nanotubes synthesis in floating-catalyst CVD, Carbon N. Y. 149 (2019) 318–327, https://doi.org/10.1016/j.carbon.2019.04.026.

[12] C.M. Yang, K.H. An, J.S. Park, K.A. Park, S.C. Lim, S.H. Cho, et al., Preferential etching of metallic single-walled carbon nanotubes with small diameter by fluorine gas, Phys. Rev. B Condens. Matter 73 (2006) 1–7, https://doi.org/10.1103/PhysRevB.73.075419.

[13] Y. Tian, A. Nasibulin, B. Aitchison, T. Nikitin, J. Pfaler, H.J. Hua Jiang, et al., Controlled Synthesis of Single-Walled Carbon Nanotubes in an Aerosol Reactor, J. Phys. Chem. C 115 (n.d.) 7309–7318. https://doi.org/10.1021/jp112291f.

[14] J.S. Barnard, C. Paukner, K.K. Koziol, The role of carbon precursor on carbon nanotube chirality in floating catalyst chemical vapour deposition, Nanoscale 8 (2016) 17262–17270, https://doi.org/10.1039/C6NR03895F.

[15] C. Hoecker, F. Smail, M. Pick, A. Boies, The influence of carbon source and catalyst nanoparticles on CVD synthesis of CNT aerogel, Chem. Eng. J. 314 (2017) 388–395, https://doi.org/10.1016/j.cej.2016.11.157.

[16] I.V. Anoshkin, A.G. Nasibulin, Y. Tian, B. Liu, H. Jiang, E.I. Kauppinen, Hybrid carbon source for single-walled carbon nanotube synthesis by aerosol CVD method, Carbon N. Y. 78 (2014) 130–136, https://doi.org/10.1016/j.carbon.2014.06.057.

[17] A. Hussain, Y. Liao, Q. Zhang, E. Ding, P. Laiho, S. Ahmad, et al., Floating catalyst CVD synthesis of single walled carbon nanotubes from ethylene for high performance transparent electrodes, Nanoscale 10 (2018) 9752–9759, https://doi.org/10.1039/C8NR00716K.

[18] L. Weller, F.R. Smail, J.A. Elliott, A.H. Windle, A.M. Boies, S. Hochgreb, Mapping the parameter space for direct-spun carbon nanotube aerogels, Carbon N. Y. 146 (2019) 789–812, https://doi.org/10.1016/j.carbon.2019.01.091.

[19] I.V. Novikov, E.M. Khabushev, D.V. Krasnikov, A.V. Bubis, A.E. Goldt, S. D. Shandakov, et al., Residence time effect on single-walled carbon nanotube synthesis in an aerosol CVD reactor, Chem. Eng. J. 420 (2021), 129869, https://doi.org/10.1016/j.cej.2021.129869.

[20] E.M. Khabushev, J.V. Kolodiazhnaia, D.V. Krasnikov, A.G. Nasibulin, Activation of catalyst particles for single-walled carbon nanotube synthesis, Chem. Eng. J. 413 (2021), 127475, https://doi.org/10.1016/j.cej.2020.127475.

[21] A. Scheinker, F. Cropp, S. Paiagua, D. Filippetto, An adaptive approach to machine learning for compact particle accelerators, Sci. Rep. 111 (11) (2021) 1–11, https://doi.org/10.1038/s41598-021-98785-0, 2021.
[22] O.V. Prezhdo, Advancing physical chemistry with machine learning, J. Phys. Chem. Lett. 11 (2020) 9656–9658, https://doi.org/10.1021/acs.jpclett.0c03130.
[23] D. Ranke, C.A.R. Perini, J.P. Correa-Baena, In data science we trust: machine learning for stable halide perovskites, Matter 4 (2021) 1092–1094, https://doi.org/10.1016/J.MATT.2021.03.007.
[24] Y. Nonoguchi, T. Miyao, C. Goto, T. Kawai, K. Funatsu, Y. Nonoguchi, et al., Governing factors for carbon nanotube dispersion in organic solvents estimated by machine learning, Adv. Mater. Interfac. 9 (2022), 2101723, https://doi.org/10.1002/ADMI.202101723.
[25] B.P. MacLeod, F.G.L. Parlane, T.D. Morrissey, F. Häse, L.M. Roch, K.E. Dettelbach, et al., Self-driving laboratory for accelerated discovery of thin-film materials, Sci. Adv. 6 (2020) eaaz8867, https://doi.org/10.1126/sciadv.aaz8867.
[26] H. Tao, T. Wu, M. Aldeghi, T.C. Wu, A. Aspuru-Guzik, E. Kumacheva, Nanoparticle synthesis assisted by machine learning, Nat. Rev. Mater. 6 (2021) 701–716, https://doi.org/10.1038/s41578-021-00337-5.
[27] S. Bhowmik, A. Govind Rajan, Chemical vapor deposition of 2D materials: a review of modeling, simulation, and machine learning studies, iScience 25 (2022), 103832, https://doi.org/10.1016/J.ISCI.2022.103832.
[28] L.E. Vivanco-Benavides, C.L. Martínez-González, C. Mercado-Zúñiga, C. Torres-Torres, Machine learning and materials informatics approaches in the analysis of physical properties of carbon nanotubes: a review, Comput. Mater. Sci. 201 (2022), 110939, https://doi.org/10.1016/J.COMMATSCI.2021.110939.
[29] P. Nikolaev, D. Hooper, F. Webber, R. Rao, K. Decker, M. Krein, et al., Autonomy in materials research: a case study in carbon nanotube growth, Npj Comput. Mater. 2 (2016), https://doi.org/10.1038/npjcompumats.2016.31.
[30] J. Chang, P. Nikolaev, J. Carpena-Núñez, R. Rao, K. Decker, A.E. Islam, et al., Efficient closed-loop maximization of carbon nanotube growth rate using bayesian optimization, Sci. Rep. 10 (2020) 9040, https://doi.org/10.1038/s41598-020-64397-3.
[31] R. Rao, J. Carpena-Núñez, P. Nikolaev, M.A. Susner, K.G. Reyes, B. Maruyama, Advanced machine learning decision policies for diameter control of carbon nanotubes, Npj Comput. Mater. 7 (2021) 157, https://doi.org/10.1038/s41524-021-00629-y.
[32] S. Ahmad, P. Laiho, Q. Zhang, H. Jiang, A. Hussain, Y. Liao, et al., Gas phase synthesis of metallic and bimetallic catalyst nanoparticles by rod-to-tube type spark discharge generator, J. Aerosol Sci. 123 (2018) 208–218, https://doi.org/10.1016/J.JAEROSCI.2018.05.011.
[33] D.V. Krasnikov, B.Y. Zabelich, V.Y. Iakovlev, A.P. Tsapenko, S.A. Romanov, A. A. Alekseeva, et al., A spark discharge generator for scalable aerosol CVD synthesis of single-walled carbon nanotubes with tailored characteristics, Chem. Eng. J. 372 (2019) 462–470, https://doi.org/10.1016/j.cej.2019.04.173.
[34] A.G. Nasibulin, A. Moisala, D.P. Brown, H. Jiang, E.I. Kauppinen, A novel aerosol method for single walled carbon nanotube synthesis, Chem. Phys. Lett. 402 (2005) 227–232, https://doi.org/10.1016/j.cplett.2004.12.040.
[35] V.Y. Iakovlev, D.V. Krasnikov, E.M. Khabushev, J.V. Kolodiazhnaia, A. G. Nasibulin, Artificial neural network for predictive synthesis of single-walled carbon nanotubes by aerosol CVD method, Carbon N. Y. 153 (2019), https://doi.org/10.1016/j.carbon.2019.07.013.
[36] E.M. Khabushev, D.V. Krasnikov, O.T. Zaremba, A.P. Tsapenko, A.E. Goldt, A. G. Nasibulin, Machine learning for tailoring optoelectronic properties of single-walled carbon nanotube films, J. Phys. Chem. Lett. (2019) 6962–6966, https://doi.org/10.1021/acs.jpclett.9b02777.
[37] A. Moisala, A.G. Nasibulin, D.P. Brown, H. Jiang, L. Khriachtchev, E.I. Kauppinen, Single-walled carbon nanotube synthesis using ferrocene and iron pentacarbonyl in a laminar flow reactor, Chem. Eng. Sci. 61 (2006) 4393–4402, https://doi.org/10.1016/j.ces.2006.02.020.
[38] Q. Zhang, J.S. Nam, J. Han, S. Datta, N. Wei, E.X. Ding, et al., Large-diameter carbon nanotube transparent conductor overcoming performance–yield tradeoff, Adv. Funct. Mater. 32 (2022), 2103397, https://doi.org/10.1002/ADFM.202103397.
[39] S. Jiang, P.X. Hou, M.L. Chen, B.W. Wang, D.M. Sun, D.M. Tang, et al., Ultrahigh-performance transparent conductive films of carbon-welded isolated single-wall carbon nanotubes, Sci. Adv. 4 (2018), https://doi.org/10.1126/sciadv.aap9264.
[40] T. Saito, S. Ohshima, T. Okazaki, S. Ohmori, M. Yumura, S. Iijima, Selective Diameter Control of Single-Walled Carbon Nanotubes in the Gas-Phase Synthesis, 8, 2008, pp. 6153–6157, https://doi.org/10.1166/jnn.2008.SW23.
[41] A. Jorio, A.G.S. Filho, Raman Studies of Carbon Nanostructures, (n.d.). https://doi.org/10.1146/annurev-matsci-070115-032140.
[42] D.A. Ilatovskii, E.P. Gilshtein, O.E. Glukhova, A.G. Nasibulin, Transparent conducting films based on carbon nanotubes: rational design toward the theoretical limit, Adv. Sci. (2022), 2201673, https://doi.org/10.1002/ADVS.202201673.
[43] E.M. Khabushev, D.V. Krasnikov, A.E. Goldt, E.O. Fedorovskaya, A.P. Tsapenko, Q. Zhang, et al., Joint effect of ethylene and toluene on carbon nanotube growth, Carbon N. Y. 189 (2022) 474–483, https://doi.org/10.1016/J.CARBON.2021.12.052.
[44] Z. Zhang, H. Dong, Y. Liao, X. Xiong, J. Yan, H. Li, et al., Controlled synthesis of single-walled carbon nanotubes by floating catalyst CVD for transparent conducting films: a critical role of loops, Diam. Relat. Mater. 124 (2022), 108942, https://doi.org/10.1016/J.DIAMOND.2022.108942.
[45] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, et al., Scikit-learn: machine learning in Python, J. Mach. Learn. Res. 12 (2012) 2825–2830. http://arxiv.org/abs/1201.0490.
[46] Z.H. Ji, L. Zhang, D.M. Tang, C.M. Chen, T.E.M. Nordling, Z. De Zhang, et al., High-throughput screening and machine learning for the efficient growth of high-quality single-wall carbon nanotubes, Nano Res. 14 (2021) 4610–4615, https://doi.org/10.1007/s12274-021-3387-y.
[47] T. Chen, C. Guestrin, XGBoost: A Scalable Tree Boosting System, (n.d.). https://doi.org/10.1145/2939672.2939785.
[48] A. Gulli, S. Pal, Deep Learning with Keras, PackT Publishing, 2017. https://www.packtpub.com/big-data-and-business-intelligence/deep-learning-keras. (Accessed 14 May 2022).
[49] E.M. Khabushev, D.V. Krasnikov, J.V. Kolodiazhnaia, A.V. Bubis, A.G. Nasibulin, Structure-dependent performance of single-walled carbon nanotube films in transparent and conductive applications, Carbon N. Y. 161 (2020) 712–717, https://doi.org/10.1016/j.carbon.2020.01.068.
[50] A.G. Nasibulin, D.P. Brown, P. Queipo, D. Gonzalez, H. Jiang, E.I. Kauppinen, An essential role of CO2 and H2O during single-walled CNT synthesis from carbon monoxide, Chem. Phys. Lett. 417 (2006) 179–184, https://doi.org/10.1016/j.cplett.2005.10.022.
[51] S.D. Shandakov, M.V. Lomakin, A.G. Nasibulin, The effect of the environment on the electronic properties of single-walled carbon nanotubes, Tech. Phys. Lett. 42 (2016) 1071–1075, https://doi.org/10.1134/S1063785016110080, 2016 4211.