University of Benghazi
Faculty of Education Almarj

جامعة بنغازي
كلية التربية – المرج
ISSN 2518-5845

Global Libyan Journal

المجلة الليبية العالمية

العدد السبعون / ديسمبر / 2023

An artificial neural network approach for the prediction of the growth rate of ZnO thin films

Dr. Maryam Mohammed Akhwater [1] , Dr. Seham Y. Dogharsham[2]

[1] Department of physics, Faculty of art and science – Almarj, University of Benghazi, Libya.
[2]Department of engineering science, Faculty of engineering – Ajdabiya, University of Benghazi, Libya.

marymuhammed14@gmail.com

An artificial neural network approach for the prediction of the growth rate of ZnO thin films

Abstract

In the present study, the artificial neural networks (ANNs) methods have been introduced based on two algorithms which were the scaled conjugate gradient (SCG) (known as multilayer perceptron networks) and the gradient descent (GD) with momentum and learning rate optimization coefficients, respectively. These algorithms have used to predict the growth rate in pulsed laser deposition (PLD)-grown ZnO thin films. After training and testing all ANNs models, a statistical analysis has carried out to assess the performance of each proposed model. A comparison between the two suggested ANNs algorithms have highlighted according to the statistical indicators, including the root mean square error (RMSE), the coefficient R^2, and the relative error (r). It has shown that model (1) of the SCG method had a high performance with RMSE of 0.483 and 0.325 for training and testing data, respectively, as well as better fitness with R^2 of 0.297. Furthermore, models (5) and (7) of the GD method have the highest values of R^2 of 0.374 and 0.391, respectively, which indicated higher forecasting performance than the other suggested models. In general, the GD algorithm models have found to be more efficient, with the highest value of the coefficient R^2 and lowest values of RMSE and r.

Keywords: artificial neural networks, growth rate, zinc oxide, coefficient of determination, relative error.

تطبيق الشبكات العصبية الاصطناعية للتنبؤ بمعدل النمو لأفلام أكسيد الزنك الرقيقة

د. مريم محمد عبد الكريم اخويطر & د. سهام دوغرشام

ملخص البحث

في هذا البحث تم استخدام الشبكات العصبية الاصطناعية للتنبؤ بمعدل النمو لأفلام أكسيد الزنك الرقيقة. وبصفة خاصة تم التركيز على تطبيق اثنان من اللوغاريتمات الأكثر شيوعا وهي scaled conjugate gradient (SCG) و gradient descent (GD). بعد عملية تدريب واختبار البيانات لنماذج الشبكات العصبية الاصطناعية تم اجراء تحليل احصائي لتقييم أداء وفعالية كل نموذج عل حدة. تمت المقارنة بين نتائج الشبكات العصبية الاصطناعية وفقا للمؤشرات الإحصائية التالية: الجذر التربيعي لمتوسط مربعات الخطأ (RMSE)، معامل التحديد (R2)، الخطأ النسبي (r). بناءا على النتائج المستخلصة يعتبر النموذج (1) لطريقة SCG الأكثر أداء حيث ان (RMSE=0.483) والأفضل تطابقا (R2 =0.297). بالإضافة الي

ان النموذجان (5) و (7) لطريقة GD كان لهما اعلي قيم لمعامل التحديد R2 والتي تدل على الأداء العالي للتنبؤ مقارنة بالنماذج الأخرى. بصورة عامة وجد ان نماذج الطريقة GD هي الأكثر فعالية مع معدلات عالية لمعامل التحديد R2 واخري منخفضة لكل من RMSE و r.

الكلمات مفتاحية: الشبكات العصبية الاصطناعية، معدل النمو، أكسيد الزنك، معامل التحديد، الخطأ النسبي.

Introduction

Among semiconducting oxide metal materials, ZnO has widely attracted attention in diverse technology and research area due to its electrical, piezoelectric, catalytic, and optical properties. In particular, it has been a promising candidate in different future nanoscale devices and applications such as, solar cells, photo-catalyst devices, biochemical sensors, optoelectronic devices, Li-ion battery, field emitter, etc. As an n-type semiconductor ZnO has a direct large band gap energy of 3.37 eV at room temperature which makes it translucent in the visible light region, and significantly photoactive in the ultraviolet radiation [1]. In addition, ZnO has large exciton binding energy (~60 m eV) which can boost (UV) luminescence at room temperature. In usual, ZnO nanostructures exhibit n-type conductivity resulted by intrinsic defects such as, zinc interstitials and oxygen vacancies [1]. More recently, it was reported that introducing imperfections such as oxygen and zinc vacancies into ZnO could efficiently enhance its functional characteristics or produce additional and surprising yet highly promising properties [2].

ANNs as a computational machine learning approach have been widely applied to forecast different product features and make influential decisions in optimizing physical and chemical manufacturing process parameters. The ANNs comprise the artificial neurons, which are a chain of connected units/nodes that can act in a similar way to a human neuron in the respect of performance and composition. Every node has the capability to send signals to other neurons. As time effective numerical tool, it can clarify linear and non-linear multivariate regression issues [3]. In addition, no explicit mathematical function is needed to define the considered case and the connection between the inputs and outputs of the process is established through a few numbers of designed experiments [4].

Over the last decade, various ANN models have been applied for optimizing the nonlinear properties of nanostructured devices which can significantly influence the performance of those devices. Kim et al. [5] used ANN algorithms to build a forecasting model, in particular, they employed genetic algorithms (GAs) to optimize the initial weight distribution. Compared to conventional methods their prediction approach proved good efficiency as an optimization model. Cho et al. [6] investigated the characteristics of ITO/Al/ITO multilayer films at different Al film thicknesses and the annealing temperature. ANN algorithms were

utilized to determine the desired process parameters which were in good matching with the actual data. Recently, to achieve high product quality Kimaev and Ricardez-Sandoval [7] applied ANN to forecasting the optimum process parameters under uncertainty for thin film fabricated by chemical vapor deposition. To determine the influence of the thin film's two features which are roughness and growth rate, ANNs were used to anticipate the behaviour of the operation observables under uncertainty.

Besides, the optimum fit ANNs which can produce better forecasts of the real-time estimations generated from the operation were determined by using maximum likelihood estimation and mean square error. To investigate the plasma-enhanced chemical vapor deposition (PECVD) of SiO2 films fabricated under different settings, Han and May [8] applied neural network methods for the process modeling and optimized the varying condition on the predicted model via GAs. After predicting the recipes of the fabrication process using genetic algorithm models, they were experimentally confirmed. It was shown that the product quality was highly improved. Gherman et al. [9] built ANNs prediction models according to experimental measurements collected from the light-induced production of gold nanoparticles (AuNPs). It was shown that the powerful estimation efficiency of ANNs techniques was verified by the low relative errors and a high correlation factor between the real-time measurements and the ANN-driven data.

It is quite essential to determine the optimal and more efficient process in the analysis of a certain physical system under a particular condition [10]. In this study, the algorithms, including SCG, and GD were applied. The SCG approach is an algorithmic learning method for analysing the symmetrical, positively defined structure of linear equations and reducing unconstrained nonlinear functions [11]. It was reported that The SCG is a powerful, time-saving research tool, it can considerably be faster than the standard back-propagation algorithm [12]. In addition, the SCG method diverts the GD approach's direction by multiplying it by a positive multiple of the method's direction [10]. SCG method is efficient in coding and reducing functions of several variables as it needs less storage space [13]. In the SCG algorithm, a search is accomplished along conjugate directions, which mainly develops rapid convergence compared to gradient descent directions. On the other hand, the GD is another optimization algorithm that has been widely applied in machine learning due to

its simplicity in finding a local minimum of a differentiable function. In order to reach the local minimum, the learning rate should be identified at a suitable value, which is neither too small nor too large. The efficiency of this technique is highly influenced by the defined value of the learning rate. If it is set too large, the algorithm might fluctuate. Besides, low values of thelearning rate would make the algorithm consumes too much time to converge [14].

As mentioned earlier, ZnO nanomaterials have been the focus of interest for diverse industrial and technological applications due to their unique features. However, it is challenged to precisely control the inescapable nonlinear properties and the randomly changings through the fabrication process of these materials which limit their performance and reliability in nanoscale devices. Hence, in order to develop ZnO products of high quality, the correlation among the output and input parameters of the synthesise conditions should be investigated statistically before carrying out laboratory work. Herein, it is a try to evaluate how efficient ANNs models are in terms of predicting the most optimized fabrication parameters, and which ANNs model can give the high functionality and better performance.

Modelling methodology of ANNs

In general, due to the nature and the unavoidable random variations of the process. Therefore, the response factor must be characterized with respect to the varying process. The relationship between the process input factors and the performance metrics with the process fluctuations needs to be analysed statistically. Here, the aim was to build two different methods based on ANN algorithms using the experimental data sets obtained through the pulsed laser deposition of ZnO thin films (see Table. 1) [15]. These ANNs models were created and tested using the Neural Network Toolbox from IBM SPSS software version 26. Two input parameters namely, temperature and pressure were investigated with regard to the growth rate as the output parameter. The training and testing datasets were randomly selected, and the training partition was utilized to identify the weights and develop the model, whereas the testing data was applied to detect errors and avoid overtraining during the training stage. To check the model accuracy, the growth rates were predicted with the independent parameters in the test dataset and then the findings were compared with the actual values of the growth rate in the test dataset. In general, ANN design includes three layers that are the input layer which has obtained data, the output layer which generates

calculated data, and hidden layers to link the input and output layers. A neuron or node as a fundamental operating unit of a neural network can accomplish two important tasks which are gathering the input information and generating the output data. Each input is multiplied by a weight value and adding bias and then applying an activation function to generate an output. The weights are corresponded to the synaptic connections in human nervous system, each input value x_i is multiplied by a weight factor w_{ij} which connects different neurons in the input and hidden layers (See Fig. 1) [16]. The Back propagation ANNs are a simple and standard feed-forward network that were widely employed for training multi-layered ANNs. In the Back propagation network, the weight values are adjusted internally by approaching the non-linear correlation between the input and the output [16].

Table. 1. The experimental data sets of ZnO thin films.

T (^{0}c)	P (m Torr)	Growth rate (µm)
400	450	0.35
400	300	0.42
375	350	0.44
350	250	0.20
425	400	0.20
425	300	0.30
400	400	0.30
450	350	0.20
425	350	0.42
375	400	0.25
400	350	0.48
450	450	0.22
400	250	0.20
350	450	0.23
350	350	0.15
375	300	0.30
450	250	0.20

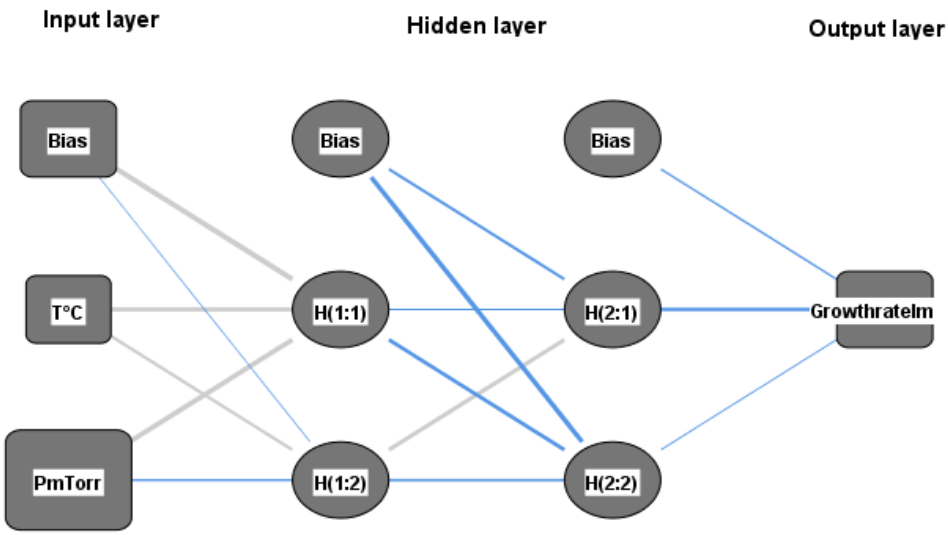

Fig. 1: A basic diagram of ANNS.

Using Eq. (1) the bias values b_j produce the net sum I_j and then pass through a nonlinear activation function (namely, sigmoid function and hyperbolic tangent function, respectively (see Eqs. (2 & 3)) to generate the output signal y_j as shown in Eq. (4):

$$I_j = \sum_{i=1}^{n} x_i w_{ij} + b_j \quad (1)$$

$$f(I) = {1 - e^{-2I}}/{1 + e^{-2I}} \quad (2)$$

$$f(I) = {e^{I} - e^{-I}}/{e^{I} + e^{-I}} \quad (3)$$

$$y_j = f(I_j) \quad (4)$$

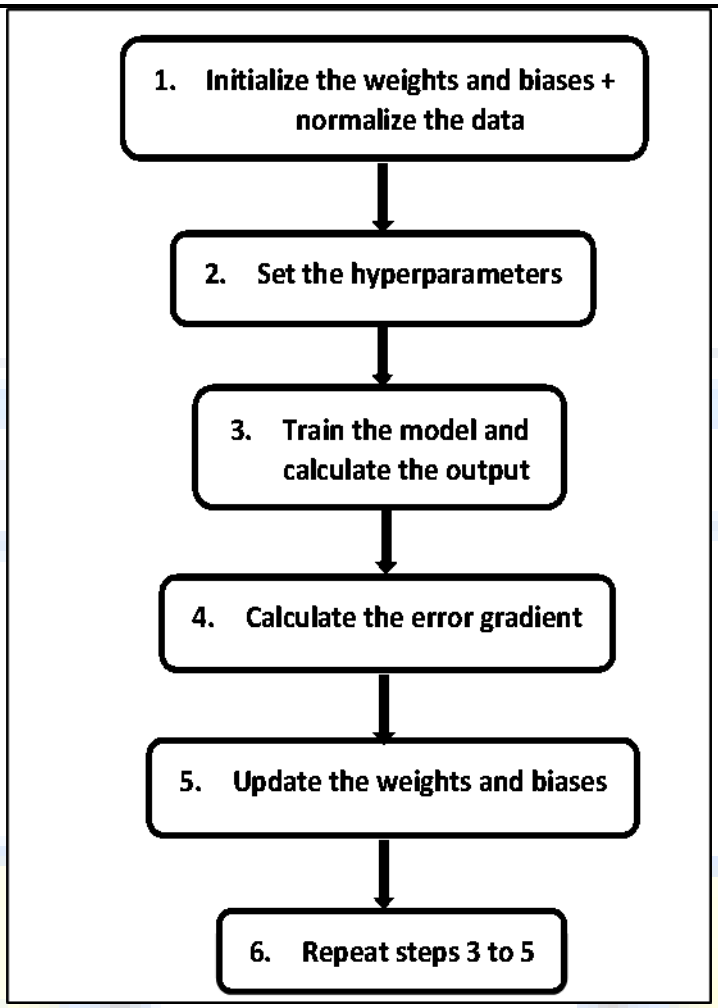

Fig. 2: A flow chart of the artificial neural network optimization process.

The unlikeness between the two applied algorithm methods is in the weight updating process. In the GD algorithm with momentum and learning rate, the weights can be updated among the layers to minimize error gradient which is backward propagated from the output node to the input node as shown in Eq. (5).

$$w'_{jk}(t+1) = w'_{jk}(t) + \eta\delta_k y_j + \alpha(w'_{jk}(t) - w'_{jk}(t-1)) \quad (5)$$

where η is the learning rate and α is the momentum coefficient. In the SCG algorithm, the weight can be updated between the output layer and the hidden layer as displayed in Eq. (6). The neural networks flow chart is illustrated in Fig. 2.

$$w'_{jk}(t+1) = w'_{jk}(t) + \eta\delta_k y_j \qquad (6)$$

As can be seen from this figure, developing the best model requires setting up optimum hyperparameters including number of hidden layers, learning rate, and momentum coefficient, following by testing the model performance to assess its efficiency to predict the accurate values of the growth rate at any temperature and pressure parameters.

Results and discussions

The efficiency of ANNs is significantly affected by the kind of algorithm applied, the number of the hidden layers, learning rate and momentum. In addition, the activation functions can optimize the learning process of different data, therefore it is essential to choose the appropriate activation functions to enhance the neural network findings. In this work, two methods were applied. In the first one two types of mostly used nonlinear activation functions with different number of hidden layers were utilized, while the second technique including applying different momentum and learning rate values within the hyperbolic tangent function. This network was trained with the learning rates of 0.0025, 0.2, and 0.3 and the momentum parameters of 0.75 and 0.95. It is worth noting that the learning rate is the most significantly critical hyperparameter, as it commands the variation rate of the model in response to the gradient error. Therefor selecting the appropriate value of the learning rate is quite crucial as too small values could lead to an elongated training procedure, while too large values might cause a fluctuating training process. Here, to reach the local minimum and highly performed model suitable values of the learning rate, which are neither too small nor too large, were employed. Additionally, momentum is applied to speed up convergence and reduce oscillations in the weight update process. Hence, it helps avoiding the optimization process from getting stuck in local minimum and resulting in better network efficiency.

Moreover, to evaluate the efficiency of the SCG and GD algorithms models, the statistical indicators which are MSE, RMSE, and R^2 were applied using the following equations (see Eqs. (7)– (9)).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(X_t - X_t^i)^2 \quad (7)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_t - X_t^i)^2} \quad (8)$$

$$R^2 = 1 - \sum_{i=1}^{n}(X_t - X_t')^2 \Big/ \sum_{i=1}^{n}(X_t')^2 \quad (9)$$

Where X_t, and X'_t represent the observed and predicted growth rate values respectively, and n is the total number of datapoints. The statistical indicator error RMSE is utilized to evaluate the model performance; it has always a positive value and the optimum condition should be zero. The values of RMSE for each model developed via SCG algorithms (first method) are summarized in Table. 2.

Table. 2. The statistical data for training and testing of the first method.

Model	Hidden layers	Activation functions	MSE (training)	RMSE (training)	MSE (testing)	RMSE (testing)	R^2	r (training)	r (testing)
1	1	hyperbolic	0.233	0.483	0.105	0.325	0.297	0.661	0.99
2	1	sigmoid	0.353	0.594	0.166	0.408	0.002	1	3.18
3	2	hyperbolic	0.285	0.534	0.116	0.341	0.004	0.97	2.25
4	2	sigmoid	0.382	0.618	0.055	0.350	0.003	1	4.2

Smaller values of RMSE indicate optimum performance of the suggested models. As can be observed from Table. 2, when all models were trained using one and two hidden layers,

model (1) was the most optimum one with RMSE of 0.483 and 0.325 for training and testing data, respectively.

The R^2 coefficient explain how well the predicted data of model close to or fit the actual data, and ideally the biggest value means the model is well fit the data with close correlation between the simulated predicted data and actual data. According to R^2 values listed in Table 2, the model (1) showed a better fit than the other models. In addition, the efficiency of each model was identified with regard to the relative error (r) of ANNs, via the following equation:

$$r = \frac{Pre_{out} - Act_{out}}{Act_{out}} \times 100 \qquad (10)$$

Where Pre_{out} and Act_{out} are the predicted and the actual outputs, respectively.
Hence, the most creditably model has the lowest relative error for training and testing data, the model (1) emphasized significant performance of 0.661 and 0.99 for training and testing data sets. Moreover, GD algorithms models were trained with different learning rates and momentum parameters. The performances of these six models in terms of RMSE, r, and R^2 are illustrated in Table. 3.

Table. 3. The statistical data for training and testing of the second method.

Model	η	α	MSE (training)	RMSE (training)	MSE (testing)	RMSE (testing)	R^2	r (training)	r (testing)
5	0.0025	0.75	0.061	0.246	0.031	0.177	0.374	0.606	0.55
6	0.0025	0.95	0.151	0.388	0.047	0.218	0.003	1.03	3.48
7	0.2	0.75	0.076	0.276	0.035	0.187	0.391	0.630	0.604
8	0.2	0.95	0.111	0.333	0.036	0.189	0.015	1.02	1.88
9	0.3	0.75	0.091	0.301	0.015	0.121	0.278	0.711	3.60
10	0.3	0.75	0.128	0.358	0.016	0.128	0.001	1.02	1.05

Based on the obtained results, it seems that the RMSE values for models (5) and (7) were smaller compared to the other GD algorithms models. The RMSE training values of designed models (5) and (7) were 0.246 and 0.276, respectively, while their RMSE testing estimations had 0.177 and 0.187 values. In contrast, model (6) showed an unsatisfied performance according to its RMSE values of 0.388 and 0.218 for training and testing data sets, respectively. Furthermore, models (5) and (7) had the largest values of R^2 of 0.374 and 0.391, respectively, which showed that higher forecasting performance than the other models. On the other hand, the results of RMSE, R^2, and r for models (6), (8), (9), and (10) were not good in terms of the optimum prediction of data. Figure (3) represent the contrast between experimental values and predicted network output values for both training and test datapoints of models (5) and (7), respectively.

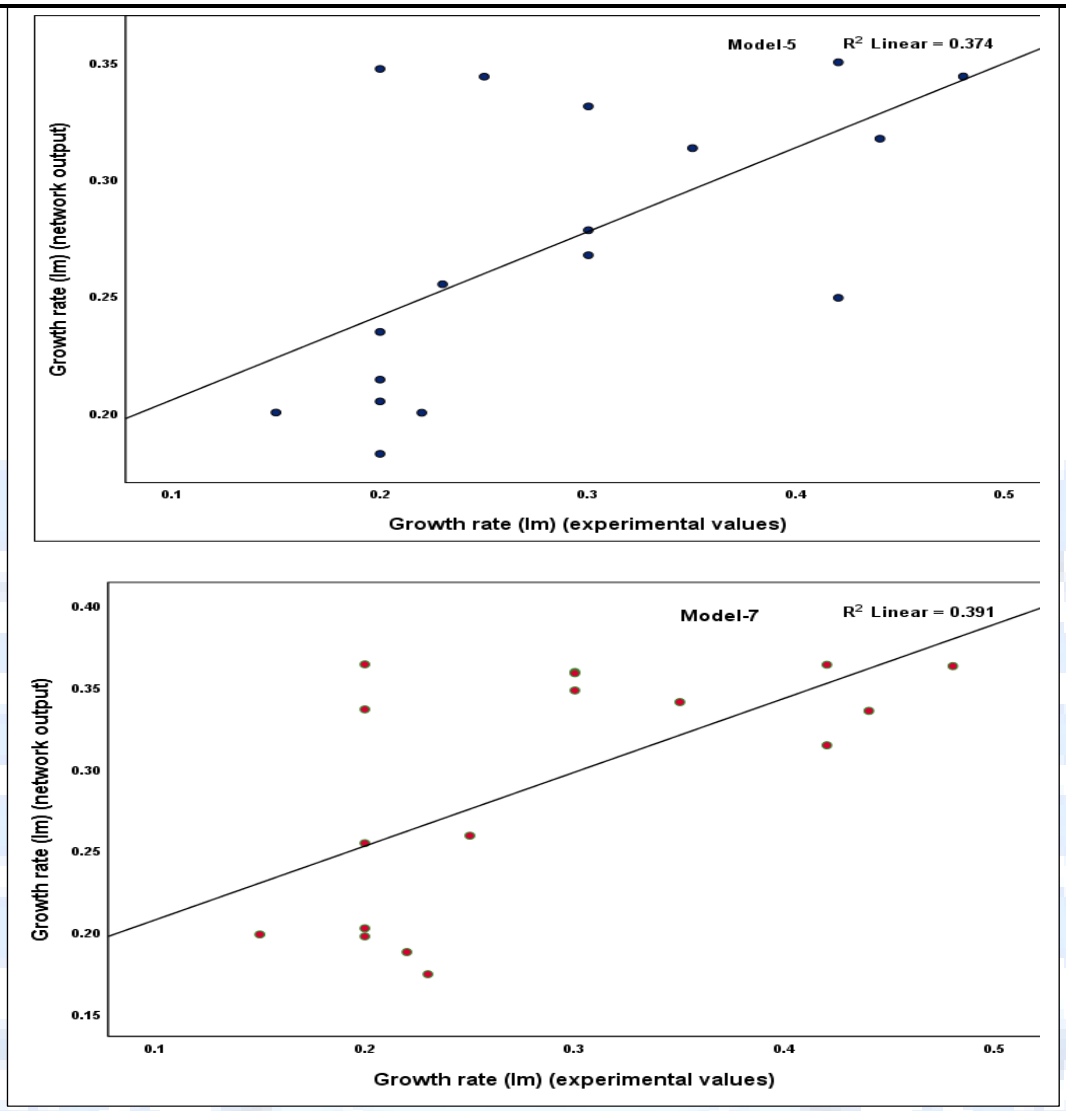

Fig. 3. Experimental values versus predicted network output of models (5) and (7).

Datapoints that locate in nearness area to the line are more precise compared to values place a little away to the line in both figures. It can be noticed there is no deviation between predicted datapoints and experimental values and both was in superior conformance. Those regression plots displayed a 'ideal fit' implying that the output of proposed models was usually identical to the actual output. The comparison between the growth rate generated by the proposed ANN algorithm methods and the actual growth rate of ZnO thin films is shown in Fig. 4. It is

proved that the linear correlation between the predicted network output datasets and the experimental datasets.

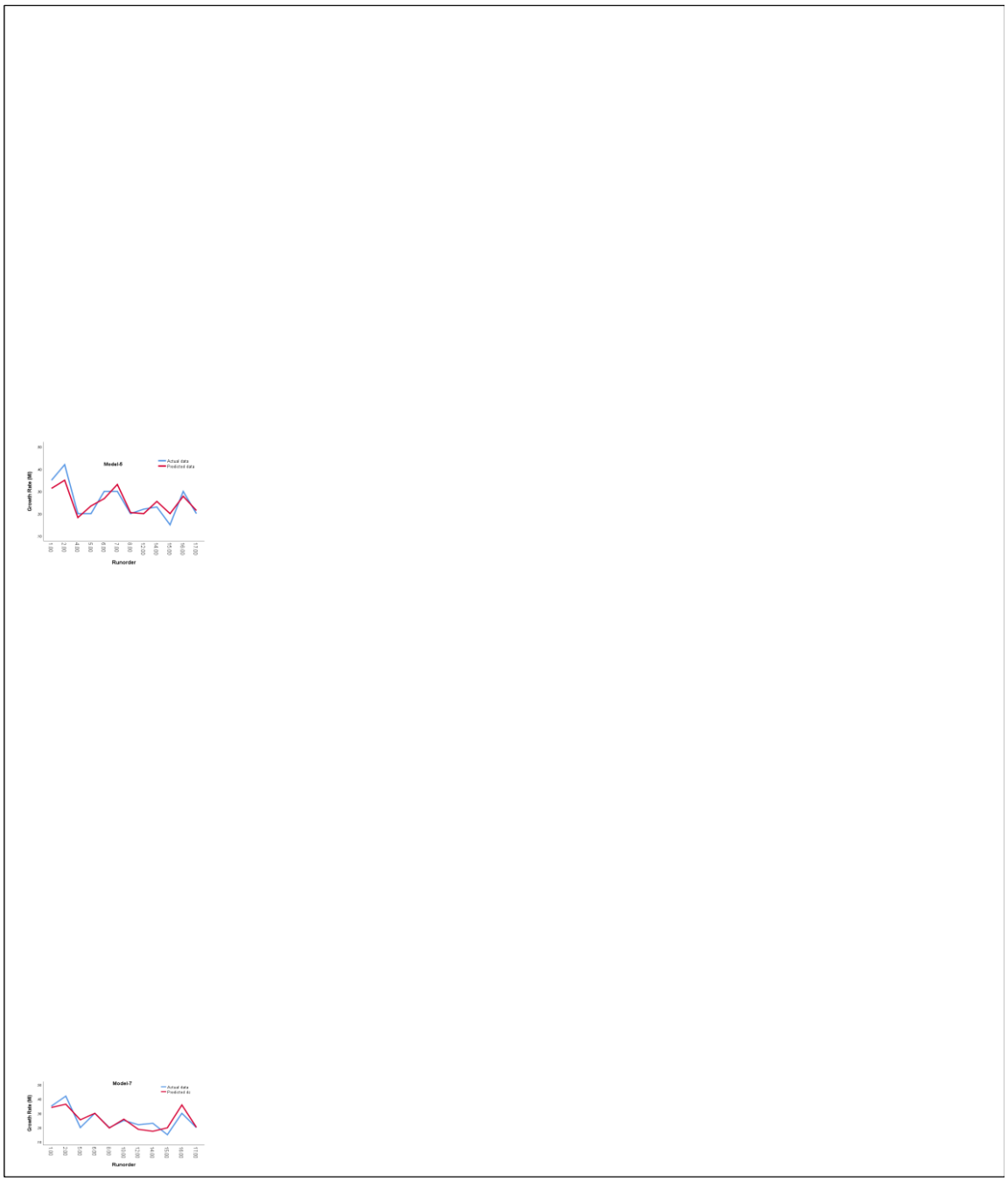

Fig. 4. An actual and predicted growth rate of models (5) and (7).

It can be noticed that the simulated and targeted data were in a good agreement. Furthermore, it was obvious that the GD algorithm models produced better prediction performance and good efficiency than the SCG models. In general, the performance of all GD models was considered to be more efficient with the largest value of the coefficient R^2 and lowest values of RMSE and r.

As mentioned earlier, models (5) and (7) were shown to have the higher prediction performance and better fitness compared to the other suggested models through values of the coefficient R^2 of 0.374 and 0.391, respectively. These stimulating results verify that ANN is a powerful methodology to figure out the connections between linear and nonlinear properties especially, in a complex system such as the fabrication process of the nanostructured materials where different physical and chemical parameters are involved.

As residuals are a measure to the variations between an actual and a predicted data of the assumed model so, it essential to plot residuals against fitted data. It was found that the residuals were approximately normally distributed around the line of fit (zero) and no evident patterns or points were observed in residuals plots (see Fig. 5).

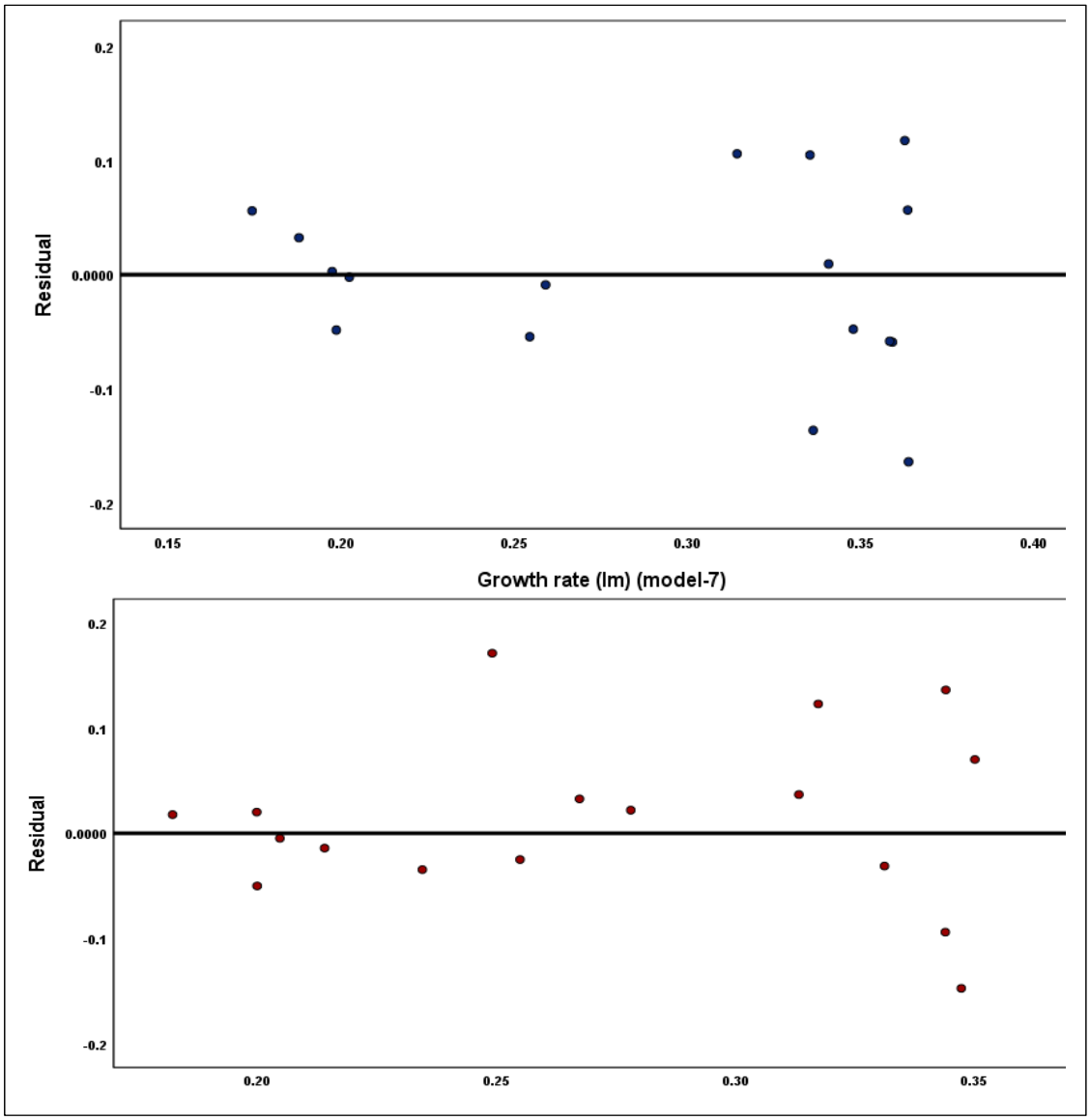

Fig. 5. The residuals versus run order for the growth rate of models (5) and (7).

Conclusion

Since it is essential to develop a mathematical model based on ANNs algorithm that can precisely estimate the related control parameters before the product manufacturing process. In this work proposed ANNs methods were applied to comparably forecast the growth rate of PLD-deposited ZnO thin films. Based on the statistical estimations, the second algorithm technique was more capable than the first ANN algorithm. It indicated that applying an

optimizer parameter such as momentum and learning rate into ANN algorithm can significantly improve the network output. In according with the statistical indicators, the GD models (5) and (7) have the lowest RMSE values and the highest R^2 which mean that those models had a very small error on their performance. However, hence only two synthesis parameters were included in this study, it would be highly effective if further processing parameters were taken into consideration to generate more an accurate and efficient prediction system.

Abbreviations

Artificial neural networks	ANNs
Scaled conjugate gradient	SCG
Gradient descent	GD
Pulsed laser deposition	PLD
Mean square error	MSE
Root mean square error	RMSE
Genetic algorithms	Gas
Plasma enhanced chemical vapor deposition	PECVD
Gold nanoparticles	AuNPs

Reference

[1] Lakshmi Prasanna, V. and Vijayaraghavan, R., 2015. Insight into the mechanism of antibacterial activity of ZnO: surface defects mediated reactive oxygen species even in the dark. Langmuir, 31(33), pp.9155-9162.

[2] Gurylev, V. and Perng, T.P., 2021. Defect engineering of ZnO: Review on oxygen and zinc vacancies. Journal of the European Ceramic Society, 41(10), pp.4977-4996.

[3] Hakeem, M.A. and Kamil, M., 2017. Analysis of artificial neural network in prediction of circulation rate for a natural circulation vertical thermosiphon reboiler. Applied Thermal Engineering, 112, pp.1057-1069.

[4] Suzuki, K. ed., 2011. Artificial neural networks: methodological advances and biomedical applications. BoD–Books on Demand.

[5] Kim, B., Lee, H. and Kim, D., 2009. Modelling of thin film process data using a genetic algorithm-optimized initial weight of backpropagation neural network. Applied artificial intelligence, 23(2), pp.168-178.

[6] Cho, E.N., Moon, P., Kim, C.E. and Yun, I., 2012. Modelling and optimization of ITO/Al/ITO multilayer films characteristics using neural network and genetic algorithm. Expert Systems with Applications, 39(10), pp.8885-8889.

[7] Kimaev, G. and Ricardez-Sandoval, L.A., 2020. Artificial neural network discrimination for parameter estimation and optimal product design of thin films manufactured by chemical vapor deposition. The Journal of Physical Chemistry C, 124(34), pp.18615-18627.

[8] Han, S.S. and May, G.S., 1997. Using neural network process models to perform PECVD silicon dioxide recipe synthesis via genetic algorithms. IEEE transactions on semiconductor manufacturing, 10(2), pp.279-287.

[9] Gherman, A.M.M., Tosa, N., Cristea, M.V., Tosa, V., Porav, S. and Agachi, P.S., 2018. Artificial neural networks modeling of the parameterized gold nanoparticles generation through photo-induced process. Materials Research Express, 5(8), p.085011.

[10] Djordjevic, S.S., 2019. Unconstrained Optimization Methods: Conjugate Gradient Methods and Trust-Region Methods. In Applied Mathematics. IntechOpen.

[11] Faraj, B.M. and Taha, D., 2021. Comparison Between Steepest Descent Method and Conjugate Gradient Method by Using Matlab. Salih, DTM, & Faraj, BM (2021). Comparison

Between Steepest Descent Method and Conjugate Gradient Method by Using Matlab. Journal of Studies in Science and Engineering, 1(1), pp.20-31.

[12] Babani, L., Jadhav, S. and Chaudhari, B., 2016, September. Scaled conjugate gradient based adaptive ANN control for SVM-DTC induction motor drive. In IFIP International Conference on Artificial Intelligence Applications and Innovations (pp. 384-395). Springer, Cham.

[13] Powell, M.J.D., 1977. Restart procedures for the conjugate gradient method. Mathematical programming, 12(1), pp.241-254.

[14] Ruder, S., 2016. An overview of gradient descent optimization algorithms. arXiv preprint arXiv:1609.04747.

[15] Ko, Y.D., Moon, P., Kim, C.E., Ham, M.H., Myoung, J.M. and Yun, I., 2009. Modeling and optimization of the growth rate for ZnO thin films using neural networks and genetic algorithms. Expert Systems with Applications, 36(2), pp.4061-4066.

[16] Krenker, A., Bešter, J. and Kos, A., 2011. Introduction to the artificial neural networks. Artificial Neural Networks: Methodological Advances and Biomedical Applications. InTech, pp.1-18.