

# Modeling of Schottky diode characteristic by machine learning techniques based on experimental data with wide temperature range

Yunis Torun <sup>a,b,\*</sup>, Hülya Doğan <sup>a</sup>

<sup>a</sup> Cumhuriyet University, Electrical and Electronics Engineering Department, Sivas, Turkey

<sup>b</sup> Cumhuriyet University, Artificial Intelligence Research Center, Sivas, Turkey



## ARTICLE INFO

### Keywords:

Machine learning  
Schottky diode  
Temperature based I-V characteristic

## ABSTRACT

In this study, 4 common machine learning methods have been used to model the I-V characteristic of the Au/Ni/n-GaN/undoped GaN Schottky diode. The current values of previously produced Au/Ni/n-GaN/undoped GaN Schottky diode against the voltages applied to the diode terminal starting from the temperature of 40K up to 400K with 20K steps were measured. Models were created using Adaptive Neuro Fuzzy System, Artificial Neural Network, Support Vector Regression, and Gaussian Process Regression techniques using experimental data containing 5192 samples in total. After determining the combinations and specifications for each one that provide the lowest model error of each model, the performances of the obtained models were compared with each other concerning the various performance indices. The performance of the ANFIS model was found to be much better than the others in both the learning and test phases with RMSE model errors as 6.231e-06 and 6.806e-06, respectively. Therefore, it was proposed as a powerful tool for modeling I-V characteristics at all temperature values between 40K and 400K.

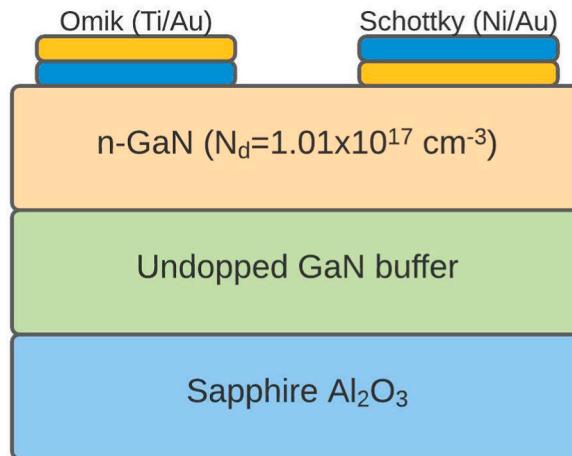
## 1. Introduction

Schottky diode is a metal-semiconductor rectifier contact diode [1]. Metal-semiconductor Schottky diodes are very important from a technological point of view. Schottky diodes, which have very fast switching capability and can operate at low forward voltages, have gained an important place in electronic technology [2,3]. Schottky rectifier contacts, which are formed by the tight contact of semiconductor and metal, are rapidly becoming widespread in semiconductor technology due to their fast transition to conduction and their ability to operate at high frequencies.

Many devices consisting of direct bandgap III-nitrides and their alloys have applications and have gained an important place in technology. Such devices can be examined in two categories as electronic and optoelectronic devices. Laser diodes (LDs), light-emitting diodes (LEDs), UV Schottky barrier (SB) photodetectors, and III-nitrides are particularly advantageous in high power and high-speed electronic applications due to their high-speed electron mobility and high breakdown field [4–8].

It is necessary to examine the electrical properties of devices obtained from GaN and III-nitride semiconductors under different temperature conditions. It is very important to determine the basic parameters such as Schottky barrier height, ideality factor, and series resistance obtained from the current-voltage properties of metal-semiconductor Schottky structures obtained at room

\* Corresponding author. Cumhuriyet University, Electrical and Electronics Engineering Department, Sivas, Turkey.  
E-mail address: [torun.yunus@gmail.com](mailto:torun.yunus@gmail.com) (Y. Torun).



**Fig. 1.** Schottky contact.

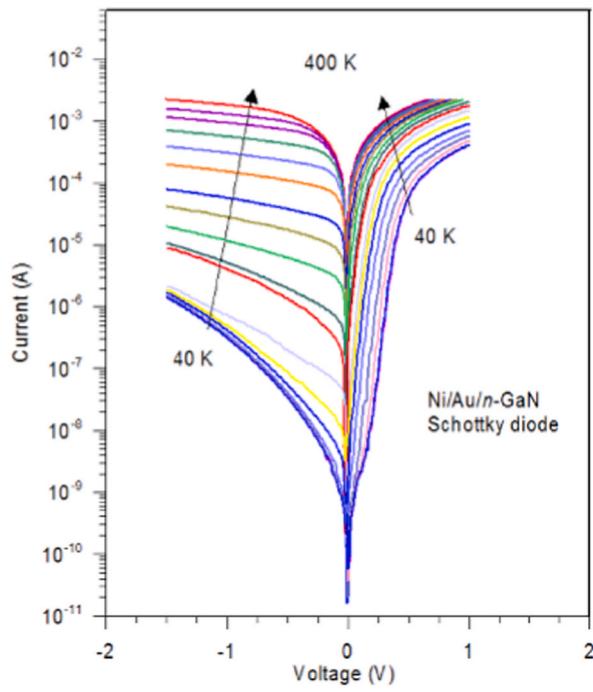
temperature [1–3]. However, obtaining reproducible rectifier contacts on GaN with a high and ideal Schottky barrier  $\Phi_B$  and a low leakage current remain as a serious technological concern, which is the physical origin of the subject of the research constantly. Especially with the determination of the diode parameters depending on the temperature, many studies have been carried out with regarding the fact that these parameters are an important function of the temperature [9–13]. Although machine learning techniques have been widely used in many real-world problems, there are a few studies that focus on the modeling of I–V characteristics of Schottky diodes. Güzel et al. [14] have developed a model to predict the current values of a polymer-interfaced 6H–SiC/MEH-PPV Schottky diode with an artificial intelligence approach based on the temperature and voltage within the temperature range between 100 and 250K with 244 experimental data. Similar work has been proposed by Çolak et al. for estimating the I–V characteristic of Schottky diode within the temperature range between 100 and 300K with 362 experimental data with ANN [15]. Rabehi A. et al. [16] have determined the Schottky barrier diode (SBD) parameters with the heuristic optimization techniques. Wang K. et al. [17] have presented the particle swarm optimization (PSO) method to solve the parameter estimation problem of the Schottky-barrier diode model and have showed that the proposed method has high parameter estimation accuracy. Rahmani et al. [18] have described a theoretical model to give an estimate of G/n-Si SJSC performance using artificial intelligence interpreted by the ANN model. They have stated that the simulation model they have designed was compatible with the experimental results. Karobaga et al. [19] have investigated a new method, which is modified artificial bee colony (MABC) algorithms to determine the basic parameters of the Schottky diodes such as barrier height, ideality factor and series resistance.

Although a few research papers exist in the literature, there is no research paper covers both the assessment of modeling capability of machine learning tools and the modeling within the high-temperature range. In this study, we have used the Au/Ni/n-GaN/undoped GaN Schottky barrier diode which was produced in the previous work for the physical system model [20]. The diode was epitaxially grown with GaN films by using metalorganic chemical vapor deposition (MOCVD) in the c-plane Al<sub>2</sub>O<sub>3</sub> substrate [20]. The electrical properties of (Au/Ni)/n-GaN Schottky barrier diode (SBDs) in a wide temperature range between 40 and 400 K with a temperature step of 20 K by using the reverse and forward bias I–V measurements have been acquired in experimental work. I–V measurement data samples were used to train and test of four machine learning tools; Adaptive Neuro Fuzzy Inference System (ANFIS), Artificial Neural Network (ANN), Support Vector Regression (SVR) and Gaussian Process Regression GPR. The four machine learning models have firstly been constructed with each of the best own combination for whole temperature range. Then, the performance of the modeling capability of the four machine learning tools has been compared with the various error performance indices with a 5-fold validation technique. The assessments have also been performed for both general approximation capability at the whole wide temperature range and the approximation capability at each local temperature step.

## 2. Material and methods

### 2.1. Experimental work

Experimental data has been acquired with preliminary produced GaN films which were epitaxially grown on sapphire substrate by using metal-organic chemical vapor deposition (MOCVD) method [20]. The growth stages of undoped and doped GaN films were performed as follows [20]; TMGa metalorganic source was used as a Ga source and NH<sub>3</sub> hydride source was used as nitrogen source to grow GaN on C-plane sapphire substrate. Sapphire substrate is widely used to grow the GaN epilayer. The undoped GaN layer is used to improve the quality of the n-type GaN epilayer. Water vapor etc. may be found on the substrate surface. In order to remove impurities from the surface, thermal surface cleaning was performed at 1047 °C for 10 minutes at the beginning of the growth stage. To grow the nucleation layer, the growth pressure was set to 200 mbar, at 477 °C, with the TMGa flow set to 15 sccm and the NH<sub>3</sub> flow to 2000 sccm. Then, for high-temperature GaN growth, 1250 nm thickness was grown at a reactor pressure of 200 mbar, at 1070 C, by



**Fig. 2.** Experimental measured current with respect to the applied Voltage and Temperature.

$$\begin{aligned}
 R_1 : & \text{ if } T \text{ is } A_1 \text{ AND } V \text{ is } B_1 \text{ then } f_1 = p_1 T + q_1 V \\
 \cdots \cdots \cdots \\
 R_k : & \text{ if } T \text{ is } A_k \text{ AND } V \text{ is } B_k \text{ then } f_k = p_k T + q_k V
 \end{aligned} \tag{1}$$

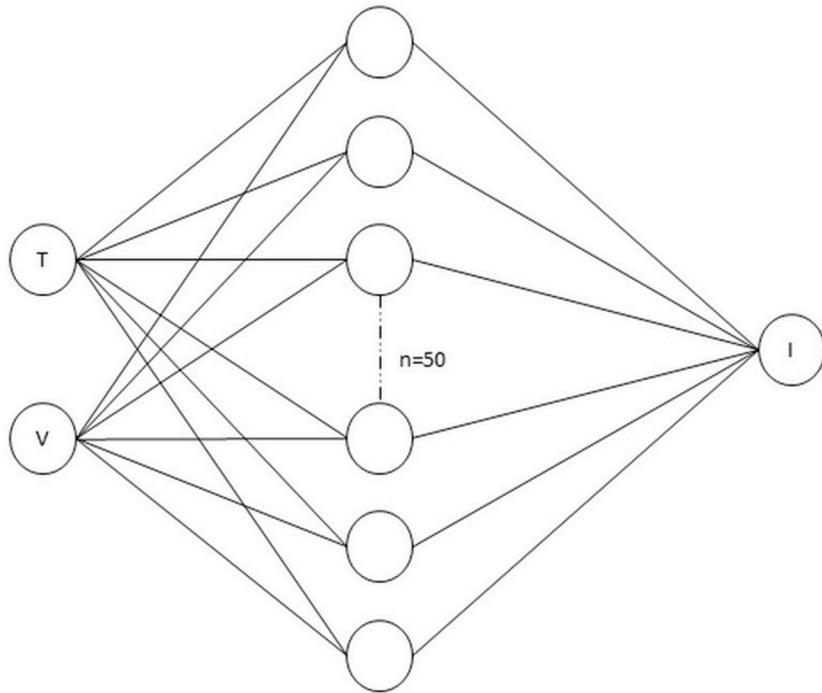
adjusting the TMGa flow to 20 sccm and the NH<sub>3</sub> flow to 1000 sccm. SiH<sub>4</sub> hydride source was used as Si source for doped GaN layer. n-GaN ( $N_d = 1.01 \times 10^{17} \text{ cm}^{-3}$ ) layer with the thickness of 1350 nm was grown by keeping the TMGa flow at 25 sccm, NH<sub>3</sub> flow at 1800 sccm and SiH<sub>4</sub> flow at  $2.83 \times 10^{-10} \text{ mol/min}$  by setting the reactor pressure at 150 mbar and the growth temperature at 1070 °C.

Before making ohmic and Schottky contacts, the samples were first cleaned by boiling in the trichloroethylene, methanol and rinsing in deionized water. The ohmic contacts were prepared by the evaporation deposition of Ti (90 nm)/Au (thermal evaporation) in a nitrogen gas environment. Ni (70 nm)/Au (thermal evaporation) metals were used for Schottky contacts. Schottky contacts are made using magnetron DC sputtering technique and Au. All these processes were carried out in a high vacuum system at a pressure of  $10^{-6}$  Torr. Schottky contacts were prepared as circular spots with a diameter of 1.5 mm (see Fig. 1) and the (Au–Ni)/n-GaN/undoped GaN structure was obtained. **Current-voltage (I–V) measurements of the Schottky diode were measured by using the Keithley 487 Picoammeter/Voltage Source with a temperature-controlled ARS HC-2 closed-loop helium cryostat in the 40–400 K temperature range.**

The acquired current values with respect to the temperature and voltage are shown in Fig. 2. Current transport in Schottky contacts is due to majority carriers and can be defined by thermionic emission through the interface barrier [2,21]. In Fig. 2, it can be said that the diode does not comply with the thermionic emission theory at high temperatures and other current mechanisms are dominant. Therefore, the diode deteriorated especially after 350 K and started to show ohmic properties. Ideality factor, barrier height, series resistance calculations were made in previous study [20]. The variation of the series resistance according to the temperature is also available in previous study [20,21].

## 2.2. Machine learning-based approximation of IV characteristic of Schottky diode

The I–V characteristic of Schottky diodes includes high nonlinearity with respect to temperature variations. Obtaining I–V characteristics is a process of measurement of current, voltage, and temperature via discrete sampling. To obtain any current value with respect to any certain voltage and temperature, machine learning models can be useful. These models can mimic the real physical system and are capable to predict any current value according to any voltage and temperature, without any extra experimental measurement. The four machine learning tools which have been used to model of I–V characterization of the Schottky diode in the current research are described in the following subsections.



**Fig. 3.** Used 2-50-1 ANN architecture for the modeling of Schottky diode characteristic.

#### 2.2.1. Adaptive Neuro Fuzzy Inference System (ANFIS)

Adaptive Neuro Fuzzy inference system was developed by Jang with the cooperation of Fuzzy Logic and Artificial Neural Network training approach [22]. A fuzzy inference system is an alternative method for modeling the system instead of crisp sets if input-output relation can be represented by the linguistic fuzzy rules [23]. The neural network is capable of learning the nonlinear relation between input-output pairs with backpropagation of the modeling error through the synaptic weights of the network if a training data set exists. Conventional fuzzy inference systems could be constructed either as Mamdani [24] or Takagi-Sugeno (T-S) [25] types according to the rule base and output form of the fuzzy sets. The output of the T-S Fuzzy inference system is a linear combination of inputs while fuzzy sets were used for input fuzzification. In the lack of heuristic information about the rule base of the system, the T-S approach provides the construct of the input fuzzy sets by clustering or grid partitioning. In this work, for modeling of temperature-dependent IV characteristic of Schottky diode, the rule base of the T-S model has been constructed as; where  $A_{1..k}, B_{1..k}$ , fuzzy sets for input T (Temperature), and V (Voltage) respectively for the total k rule. The Gaussian type membership function is used as input fuzzy sets for temperature (T) and terminal voltage (V). The degree of membership for any instant can be calculated as;

$$\mu_{A_j}(T) = e^{\frac{1}{2} \left( \frac{T - c_j^k}{\sigma_j^k} \right)^2} \quad (2)$$

$$\mu_{B_j}(V) = e^{\frac{1}{2} \left( \frac{V - c_j^k}{\sigma_j^k} \right)^2} \quad (3)$$

where  $\mu_{A_j}(T)$  and  $\mu_{B_j}(V)$  are the degree of j th rule membership for crisp inputs T, and V respectively.  $c_j^k$  and  $\sigma_j^k$  correspond center and spread of Gaussian membership function. After fuzzification steps, ANFIS calculates each rule firing strength as;

$$w_j = \mu_{A_j}(T) \text{ AND } \mu_{B_j}(V) \quad (4)$$

where fuzzy AND operator can be selected as either minimum operator or product operator. The next step is the normalization of each firing strength with concerning all k rules as;

$$w'_j = \frac{W_j}{\sum_{l=1}^k W_l} \quad (5)$$

The fuzzy output should be defuzzified to convert fuzzy values to crisp ones as;

$$w'_j f_j = w'_j (p_j z + q_j n + r_j a) \quad (7)$$

Finally, the overall output, which is current ( $I$ ), is calculated as;

$$I = \sum_{j=1}^k w'_j f_j = \frac{\sum_{j=1}^k w'_j f_j}{\sum_{j=1}^k w_j} \quad (8)$$

The parameters of initial membership function,  $c_j^k$  and  $\sigma_j^k$ , and the total rule  $k$ , could be calculated by using either clustering or grid partitioning approach [26]. In this work, the subtractive clustering method is chosen. The least-square estimation is used to tune linear parameters,  $p_j$  and  $q_j$ , in rule consequent parts. A gradient descent based error-backpropagation is used to tune nonlinear parameters,  $c_j^k$  and  $\sigma_j^k$ , in rule antecedent part.

### 2.2.2. Artificial Neural Network (ANN)

Artificial neural network (ANN) is a strong nonlinear approximation tool that mimics the human brain [27]. The general architecture of ANN is composed of mainly three layers namely input, hidden, and output layer. Input data instants flow in one direction from input to outputs neuron for function approximation tasks while to reduce approximation error in training phase the error is backpropagated to the opposite direction as output to the input layer. The ANN architecture for modeling of temperature-dependent IV characteristic of Schottky is proposed in Fig. 3. A differentiable nonlinear activation function as the sigmoid, hyperbolic tangent, sine, Gaussian, etc. in each neuron, processes the data for approximation task. In this work the **hyperbolic tangent activation function** is used as;

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (9)$$

The training of the ANN is the finding the most appropriate set of weights between neurons for each connection. Although the error backpropagation method [28] is common for training in literature, Bayesian regulation was reported as more robust than standard gradient descent based algorithms [29]. Therefore, Bayesian regularization was chosen for training of the ANN model in the present study. Training is performed according to the measurement of the network performance by using the Mean Square Error function.

### 2.2.3. Gaussian Process Regression (GPR)

Gaussian Process Regression (GPR) is a non-parametric regression model which capable of estimating the value of unknown output variables as a function of known input variables by training the model using Gaussian Process (GP). The model is building on describing the covariance function which defines the changes of expected target value over the change of input variables. Covariance functions which are also called kernel functions encode all assumptions about the form of input-output modeling. A deep presentation on a variety of mean functions and kernel function can be read in Rasmussen and William [30]. The squared exponential kernel is appropriate for modeling very smooth functions for GP as described in Rasmussen [31];

$$k(x_i, x_j) = \sigma^2 e^{-\frac{1}{2}(x_i - x_j)^T \text{diag}(l)^{-2} (x_i - x_j)} \quad (10)$$

where  $\sigma^2$  Signal variance,  $l > 0$  is length scale that describes the smoothness,  $x_i$  and  $x_j$   $i$  th and  $j$  th input point corresponding  $i$  th and  $j$  th output,  $y_i$  and  $y_j$ .

The exponential kernel is defined as;

$$k(x_i, x_j) = \sigma^2 e^{-\left(\frac{r}{\sigma}\right)} \quad (11)$$

where  $r$  is the Euclidian distance between  $x_i$  and  $x_j$ .

Rational Quadratic Kernel is defined as;

$$k(x_i, x_j) = \sigma^2 \left(1 + r^2 / 2\alpha\sigma_l^2\right)^{-\alpha} \quad (12)$$

where  $\alpha > 0$  is the scale-mixture parameter.

A GP is defined by its mean function  $m(x)$  in addition to its kernel function. Hence training set has been normalized to zero mean, a zero mean function was selected for  $m(x)$  [32]. Considering the regression model as;

$$h(x)^T \beta + f(x) \quad (13)$$

where  $f(x)$  are from GP with  $m(x)$  and  $k(x_i, x_j)$  and the basis function  $h(x)^T$ . Any instant of  $i$  th output response  $y$  can be modeled by GPR as;

$$P(y_i | f(x_i), x_i) \sim N(y_i, h(x_i)^T \beta + f(x_i, \sigma^2)) \quad (14)$$

#### 2.2.4. Support Vector Regression (SVR)

Support Vector Regression (SVR) is one of the popular machine learning algorithms used for regression problems [33]. The probable lines are the smallest in linear regression. The objective function is the minimization of the least-squares of error. SVR allows defining the acceptable amount of error. The SVR finds a feasible line or hyperplane. Unlike linear regression, SVR deals with the minimization of the weight coefficient. It takes the maximum error less than or equal to the desired value to increase the model's accuracy. SVR considers data with a particular neighborhood distance ( $\epsilon$ ). In linear regression, the input and target data pairs are expressed as  $x_n$  and  $y_n$ . The linear regression function is presented in Eq. (15).

$$f(x) = xw + b \quad (15)$$

Since there is no function  $f(x)$  that satisfies the constraints for all points, slack variables  $\xi_i, \xi_i^*$  are defined for each point. The primal of the objective function is obtained by adding slack variables with the box constant ( $C$ ) and  $\epsilon$ .

$$\text{Min} \quad \frac{1}{2} \|w^2\| + c \sum_{i=1}^m (\xi_i + \xi_i^*) \quad (16)$$

The presented optimization model in Eq. (16) is easily solved with the Lagrange dual formulation. The optimal values of primal and dual problems do not need to be equal, and the difference is named as "duality gap". By adding the non-negative  $\alpha_i$  and  $\alpha_i^*$  multipliers for each  $x_i$  observation, the Linear SVR regression dual formula can be written as presented in Eq. (17). The parameter of  $\beta$  can be calculated by using Eq. (18). The new values are determined on the support vector with Eq. (19).

$$L(\alpha) = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (\alpha_i - \alpha_j^*) (\alpha_j - \alpha_j^*) x_i' x_j + \epsilon \sum_{i=1}^N (\alpha_i + \alpha_i^*) + \sum_{i=1}^N y_i (\alpha_i - \alpha_i^*) \quad (17)$$

$$\beta = \sum_{n=1}^N (\alpha_n - \alpha_n^*) x_n \quad (18)$$

$$f(x) = \sum_{n=1}^N (\alpha_n - \alpha_n^*) (x_n' x) + b \quad (19)$$

The most significant component of SVR is the kernel function which produces a non-linear decision surface that can transform a nonlinear problem into a linear problem [34]. A variety of kernel functions have different mapping capabilities for approximation tasks [35]. The most commonly used kernel functions are Linear, Quadratic, and Gaussian type kernel functions, which are defined below [36]; The linear kernel function is defined as;

$$k(x_i, x_j) = x_i \cdot x_j \quad (20)$$

The polynomial kernel function is defined as;

$$k(x_i, x_j) = ((x_i \cdot x_j) + c)^q \quad (21)$$

when  $q = 2$ , the polynomial kernel function is called as Quadratic Kernel function. Radial Base or Gaussian kernel function is defined as;

$$k(x_i, x_j) = e^{-\|x_i - x_j\|^2 / \sigma^2} \quad (22)$$

#### 2.3. Modeling environment

The four machine learning models, ANFIS, ANN, GPR, and SVR were programmed with the Matlab Fuzzy Logic Toolbox, Matlab Machine Learning, and Deep Learning Toolbox (Matlab R2018a). The model codes have been run on a desktop computer with a 4th generation i7 processor with 16 Gb RAM. The model performance not only has been measured by error-index but also the consumption time of each model on the training phase have been recorded. Data has been divided into train and test phases with 5-fold cross-validation to test each machine learning tool with random data. Therefore %80 of data was used for training while the remaining data was used to test the models for each fold. At the end of 5 fold, all models have been trained and tested with 5 different combinations of training and test sets, so the model performance was measured by averaging each fold. Performance assessment has been performed with three different indexes. Root mean square error (RMSE) has been calculated as;

$$RMSE = \left[ \frac{1}{n} \sum_{i=1}^n |I_i - \hat{I}_i|^2 \right]^{\frac{1}{2}} \quad (23)$$

Mean absolute error (MAE) has been calculated as;

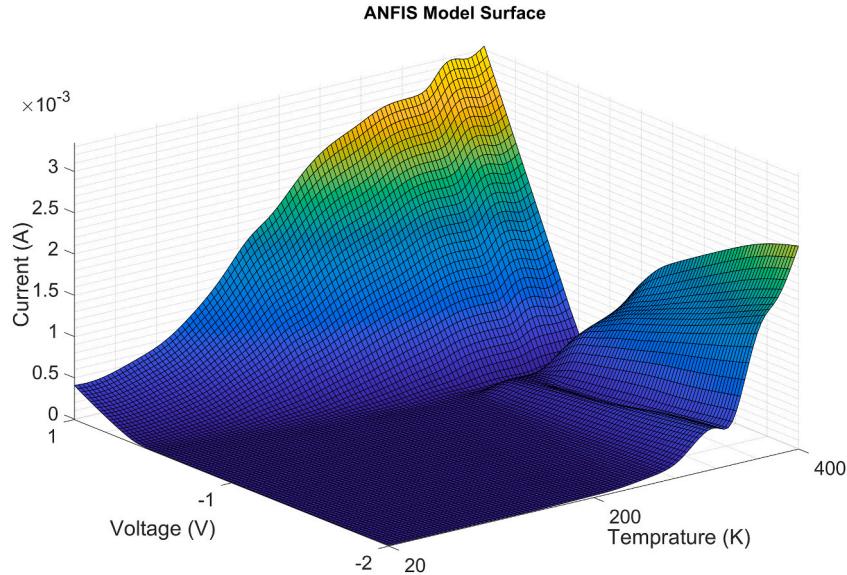


Fig. 4. Modeling surface for ANFIS.

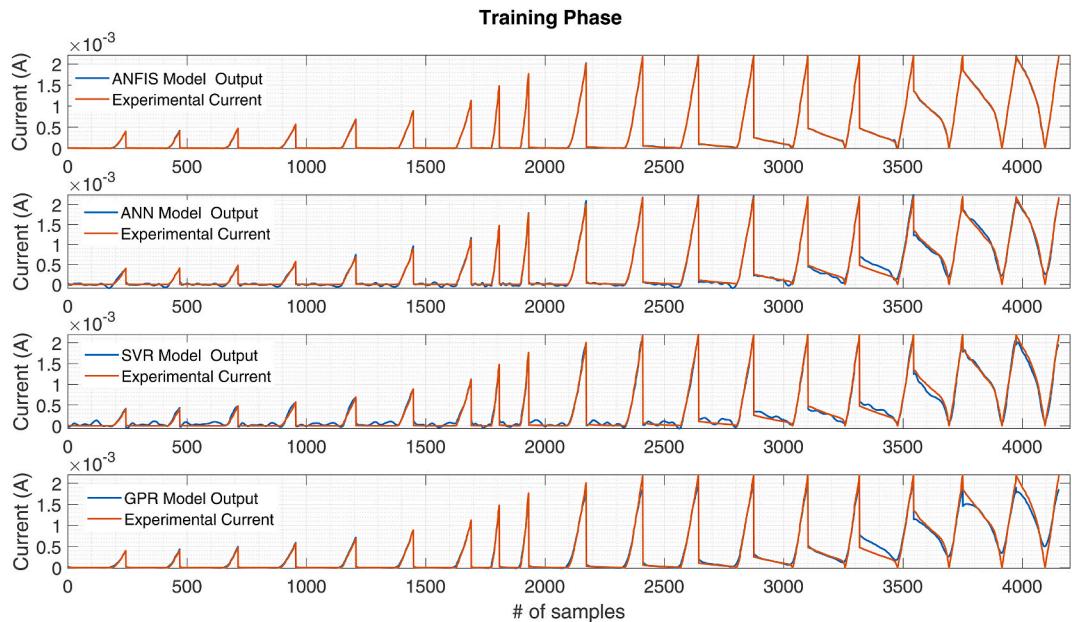


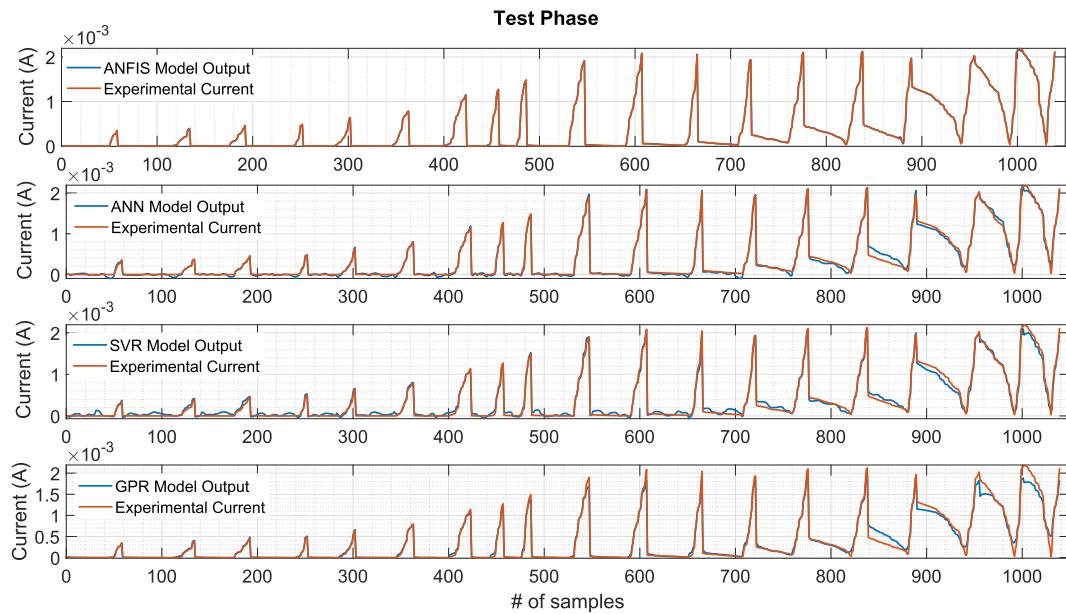
Fig. 5. Model versus actual output responses for the ANFIS, ANN, SVR, and GPR for training phase with all data sets.

$$MAE = \frac{1}{n} \sum_{i=1}^n |I_i - \hat{I}_i| \quad (24)$$

Mean square error (MSE) has been calculated as

$$MSE = \frac{1}{n} \sum_{i=1}^n (I_i - \hat{I}_i)^2 \quad (25)$$

where  $I_i$  is the acquired experiment value of current ( $I$ ),  $\hat{I}_i$  is the predicted current value for  $i$ th instant within the total  $n$  instant.



**Fig. 6.** Model versus actual output responses for the ANFIS, ANN, SVR, and GPR for test phase with all data sets.

**Table 1**  
Modeling performance.

	Model	RMSE	MSE	MAE	Consumption Time
<i>Test</i>	ANFIS	<b>6.806e-06</b>	<b>4.632e-11</b>	<b>3.914e-06</b>	62.50 msec
	ANN	4.698e-05	2.207e-09	3.297e-05	156.25 msec
	SVR	5.743e-05	3.298e-09	4.662e-05	<b>48.10 msec</b>
	GPR	7.772e-05	6.041e-09	3.935e-05	203.12 msec
<i>Train</i>	ANFIS	<b>6.231e-06</b>	<b>3.883e-11</b>	<b>3.782e-06</b>	98.593 s
	ANN	8.769 e-05	2.062e-09	3.328e-05	14.406 s
	SVR	5.818e-05	3.385e-09	4.741e-05	<b>2.437 s</b>
	GPR	7.778e-05	6.049e-09	3.947e-05	296.375 s

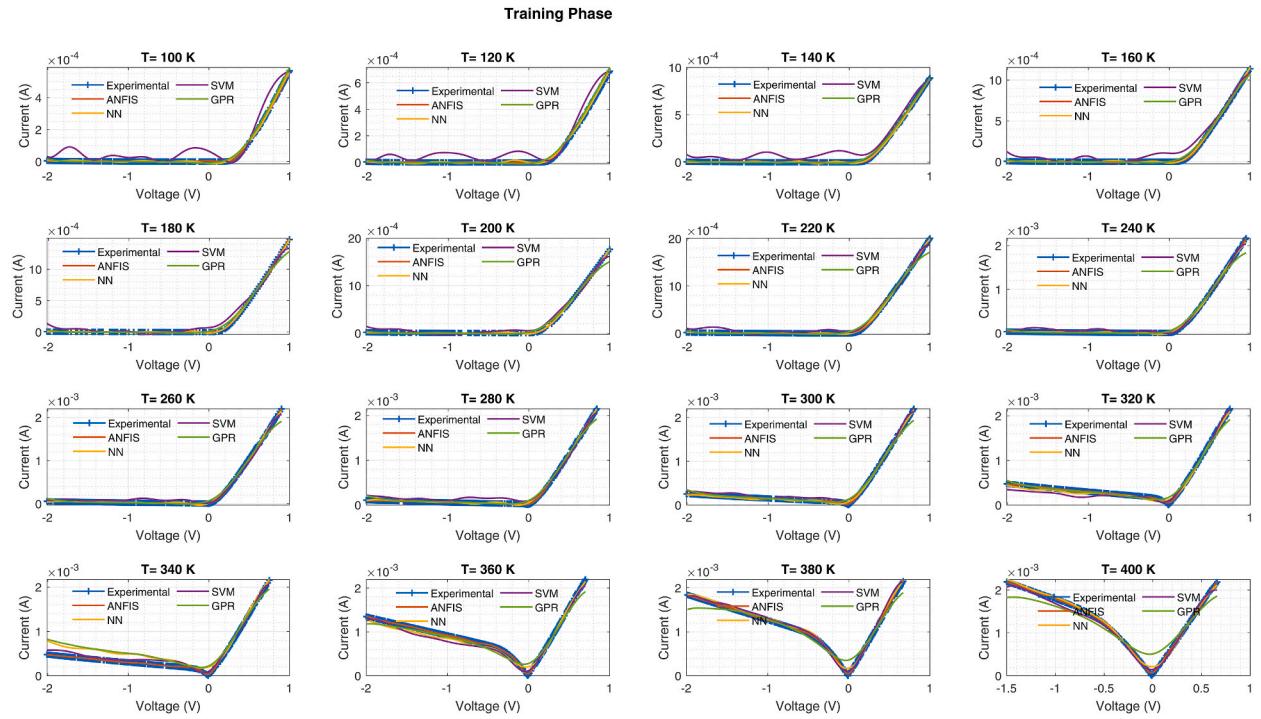
### 3. Results and discussion

Before comparison of modeling performance of algorithms, several algorithm-specific combinations have been tried for each algorithm. Cluster radii determines the number of membership functions and the number of rules in the ANFIS structure. Several clustering radii between 0.1 and 0.75 with 0.05 steps have been tried by assessment of RMSE error for both training and testing. Increasing the radii causes less number of rules, however, yields more RMSE error. On the other way, decreasing the radii causes a high number of rules, however it may cause of decreasing the RMSE error in training while increases RMSE error for testing. In order to avoid overfitting cases in training while subject to the lower RMSE error in testing, the best value for cluster radii has been chosen as 0.25 which yields 27 rules in the rule base. The model surface of the ANFIS model which was a demonstration of nonlinear relation of temperature-based I-V characteristic has been shown in Fig. 4.

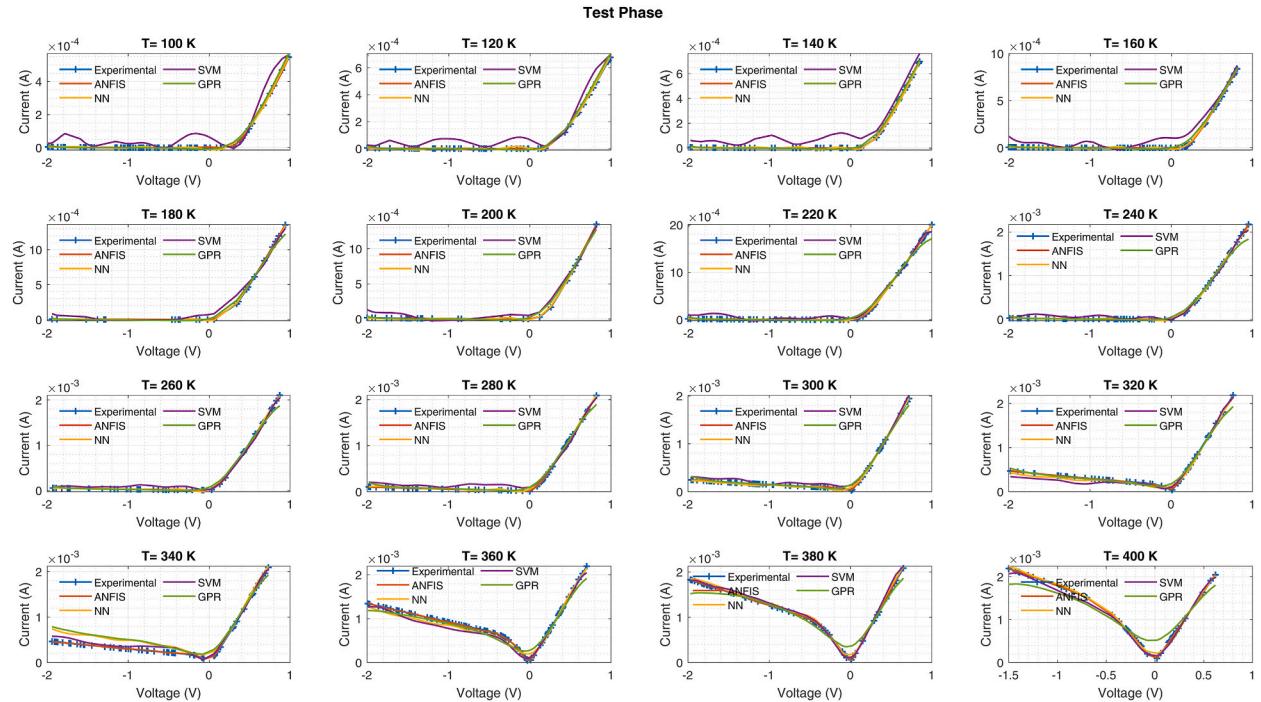
The number of neurons in the hidden layer of ANN affects the network performance as a high number, which has caused of overfitting and high consumption time while a lower number caused of increasing the RMSE in both training and testing phases. Therefore, 10 to 100 neurons with 10 steps have been performed to achieve local optimum neuron size. The best MSE has been achieved with 50 neurons in the hidden layer with the Bayesian Regulation learning in ANN.

The type of kernel functions of SVR affects the modeling performance. Linear kernel, Quadratic kernel, Cubic kernel, and Gaussian kernel function have been tried in SVR modeling. The best RMSE and the minimum time consumption have been obtained with Gaussian kernel function-based SVR. It has been also observed that the type of kernel function in GPR affects the modeling performance in the same manner. Rational Quadratic, Exponential, Square Exponential kernel functions have been tried to an assessment of each kernel function on the performance of the GPR model. The best performance has been obtained with the Square Exponential kernel function.

The predicted and actual value of Schottky current with all four models have been shown in Fig. 5 and Fig. 6 for the training and test phase respectively. The training dataset contains 80% of the total data, and the test phase contains 20% of the total data, which corresponds to 4161 samples in training data and 1030 samples in test data set in whole temperature ranges between 40K–400K.



**Fig. 7.** Models versus experimental current values for temperature steps within the range 100K–400K in the training phase.



**Fig. 8.** Models versus experimental current values for temperature steps within the range 100K–400K in the testing phase.

Partitioning of data set has been performed by using 5-fold cross-validation techniques, therefore each model has been trained and tested with 5 different train and data set combinations. Although all algorithms have been run 5 times with 5-fold cross-validation, the only first-fold response has been shown in figures for simplicity. The comparison mean values of 5-fold cross-validation for RMSE, MSE,

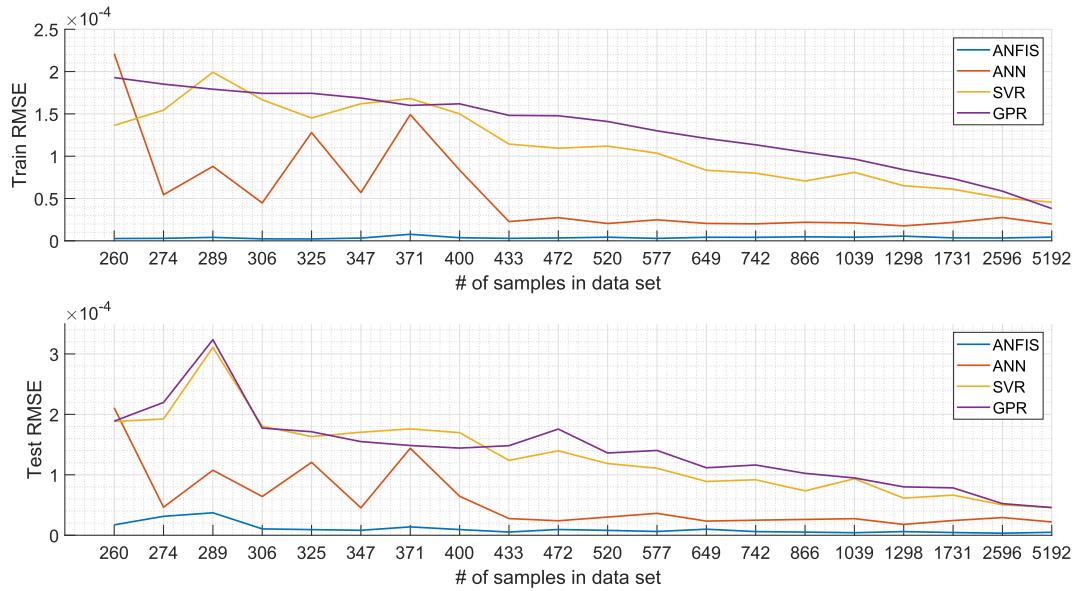


Fig. 9. Effect of number of samples in data set on training and test performances.

and MAE performance indices of four models with the best combination have been shown in Table 1. The best performances both training and test have been achieved with ANFIS with the lowest mean value of RMSE, MSE, and MAE for 5-fold average. Although ANN has the second-best performance for training, the performance gets worse for the testing phase for RMSE. It means that there was a bit of over-training for ANN. While SVR has moderate performance, GPR performance has been worse than others for both the training and test phases. In Table 1, the mean performance within the range of 40K–400K has been summarized. The sample number versus experimental current and model output current have been shown in Figs. 5 and 6 for training and testing, respectively. It is clear that the sample numbers have affected the performance as shown in Figs. 5 and 6. Although ANFIS has good approximation capability almost within the all data range, the other 3 methods have moderate approximation capability within some certain sample number range as 0–3000 for training phase as shown in Fig. 5. The same manner can be observed for between 0 and 800 samples for the test phase as shown in Fig. 6. Especially, after these ranges which correspond to the high-temperature region, the performances of the three models gets worst for both training and test phase.

As shown in Figs. 5 and 6, 19 regions correspond to the 19 temperature region between 40K–400K with 20K step. Especially when the number of samples gets high, it has been observed that the modeling performance gets worse. In order to assessment of each machine learning tool for each temperature step, the model performances for I–V characterization have also been observed for the training and test phase as shown in Figs. 6 and 7 respectively. Although the performances have been observed for 19 regions that correspond 40K–400K temperature range, it has only shown 16 regions that correspond to the temperature range between 100K–400K for graphical simplicity. ANFIS model almost perfectly fits with the experimentally acquired data at all temperature steps in both the training and test phase. While GPR has good approximation capability within the temperature range between 100K–300K, the output of GPR model doesn't completely fit the experimental current values at the temperature value greater than 300K. When the temperature rises, the modeling error also rises for GPR after 300K temperature. The same modeling performance behavior is also valid for ANN as shown in Figs. 7 and 8 for the training and test phase. Contrary to the ANN and GPR, it has been observed that SVR performance improves as the temperature increases from 100K to 300K. The performance decreasing is also less than ANN and GPR for above 300K.

One of the contributions of current work is to investigate the modeling capabilities of the different machine learning tools with a wide range of temperature regions. There are only a few research works for modeling I–V characteristics with machine learning tools to the best of our knowledge. The ANN performance has been given within the range 100–250K in the research work of Guzel et al. with 244 experimental data [14], and 100–300K in the research work of Çolak et al. with 362 experimental data [15]. MSE errors have been reported as 5.67046E-07 and 8.65054E-07 in Ref. [14] and 9.91E-09 and 2.13E-08 in Ref. [15] for the training and test phase respectively. While the ANN model performance of the current work is slightly better than the results of work in Refs. [14,15], the ANFIS model performance is precisely better than the related works. The number of data for training and testing in the current work is higher than the recent literature. In order to assess the size of the data set on the performance of the model, we have modified the experimental data set with downsampling. Downsampling is a method that decreases the size of the data by decreasing the sample rate of  $x$  by keeping the first sample and then every  $n$ th sample after the first. We have a total of 5192 samples in the experimental data set, then we applied downsampling to form 20 different data sets with downsampling 20 to 1. All four models have been trained and tested with the new 20 data sets. Training and test data were split randomly as 80% and 20% of each data set without any cross-validation. The calculated RMSE values for each data set were given in Fig. 9 for both training and testing. It is clearly observable that a noticeable decrease of RMSE when the number of samples in the data set increases for ANN, SVR, and GPR. However, the performance of ANFIS slightly increases when the decreasing the size of the data set as shown in Fig. 9. The results show that the ANFIS model was still a

powerful tool when a small number of experimental data exist.

#### 4. Conclusions

In this article, 4 different methods have been proposed for modeling the I-V characteristic of Au/Ni/n-GaN/doped GaN Schottky barrier diode obtained by making ohmic and Schottky contacts on n-GaN films grown epitaxially by using metal-organic chemical vapor deposition (MOCVD). Experimentally measured current values of the produced Schottky diode in the temperature that ranges between 40K–400K and voltage range between -2V and +1V were used in the modeling of I-V characteristic with ANFIS, ANN, SVR, and GPR. Unlike the current literature, the I-V characteristic has been modeled in the wider temperature region between 40K–400K, and also the performances of more than one machine learning tool on modeling have been evaluated. While the ANN and GPR model performances between 40K–300K were sufficient, it was observed that the model performances have decreased when the temperature increased. On the other hand, SVR has been observed to perform better at low high temperatures than low temperatures. It has been observed that ANFIS gives better results when compared to ANN, GPR, and SVR, as well as other studies in the literature, in the entire temperature range. Thus, it has been shown that ANFIS can be used in temperature-dependent I-V characteristic modeling.

#### Credit author statement

Yunis TORUN; Conceptualization, Methodology, Software, Formal analysis, Writing – original draft, Hülya DOĞAN; Validation, Investigation, Resources, Visualization, Supervision.

#### 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.

#### References

- [1] S.M. Sze, D.C. Mattis, Physics of semiconductor devices, Phys. Today 23 (1970) 75, <https://doi.org/10.1063/1.3022205>, 75.
- [2] E.H. Rhoderick, Metal-semiconductor contacts, IEE Proc. I Solid State Electron Devices 129 (1982) 1–14, <https://doi.org/10.1049/ip-i-1.1982.0001>.
- [3] R. Tung, Recent advances in Schottky barrier concepts, Mater. Sci. Eng. R Rep. 35 (2001) 1–138, [https://doi.org/10.1016/S0927-796X\(01\)00037-7](https://doi.org/10.1016/S0927-796X(01)00037-7).
- [4] W. Mtangi, P. van Rensburg, M. Dialet, F.D. Auret, C. Nyamhere, J.M. Nel, A. Chawanda, Analysis of current-voltage measurements on Au/Ni/n-GaN Schottky contacts in a wide temperature range, Mater. Sci. Eng. B-Advanced Funct. Solid-State Mater. 171 (2010) 1–4, <https://doi.org/10.1016/j.mseb.2010.03.044>.
- [5] K. Çınar, N. Yıldırım, C. Coşkun, A. Turut, Temperature dependence of current-voltage characteristics in highly doped Ag/p-GaN/In Schottky diodes, J. Appl. Phys. 106 (2009), 073717, <https://doi.org/10.1063/1.3236647>.
- [6] K. Suzue, S. Mohammad, Z. Fan, W. Kim, O. Aktas, A.E. Botchkarev, H. Morkoç, Electrical conduction in platinum–gallium nitride Schottky diodes, J. Appl. Phys. 80 (1996) 4467–4478, <https://doi.org/10.1063/1.363408>.
- [7] H. Doğan, N. Yıldırım, İ. Orak, S. Elagoz, A. Turut, Capacitance-conductance-frequency characteristics of Au/Ni/n-GaN/undoped GaN Structures, Phys. B Condens. Matter 457 (2015) 48–53, <https://doi.org/10.1016/j.physb.2014.09.033>.
- [8] H. Dogan, N. Yıldırım, İ. Orak, S. Elagoz, A. Turut, Capacitance-conductance-frequency characteristics of Au/Ni/n-GaN/undoped GaN Structures, Phys. B Condens. Matter 457 (2015) 48–53, <https://doi.org/10.1016/j.physb.2014.09.033>.
- [9] D. Korucu, A. Turut, H. Efeoglu, Temperature dependent I-V characteristics of an Au/n-GaAs Schottky diode analyzed using Tung's model, Phys. B Condens. Matter 414 (2013) 35–41, <https://doi.org/10.1016/j.physb.2013.01.010>.
- [10] A. Li, Q. Feng, J. Zhang, Z. Hu, Z. Feng, K. Zhang, C. Zhang, H. Zhou, Y. Hao, Investigation of temperature dependent electrical characteristics on Au/Ni/β-Ga2O3 Schottky diodes, Superlattice. Microst. 119 (2018) 212–217, <https://doi.org/10.1016/j.spmi.2018.04.045>.
- [11] Y. Shen, Q. Feng, K. Zhang, Z. Hu, G. Yan, Y. Cai, W. Mu, Z. Jia, C. Zhang, H. Zhou, J. Zhang, X. Lian, Z. Lai, Y. Hao, The investigation of temperature dependent electrical characteristics of Au/Ni/β-InGa2O3 Schottky diode, Superlattice. Microst. 133 (2019) 106179, <https://doi.org/10.1016/j.spmi.2019.106179>.
- [12] W. Filali, N. Sengouga, S. Oussalah, R.H. Mari, D. Jameel, N.A. Al Saqr, M. Aziz, D. Taylor, M. Henini, Characterisation of temperature dependent parameters of multi-quantum well (MQW) Ti/Au/n-AlGaAs/n-GaAs/n-AlGaAs Schottky diodes, Superlattice. Microst. 111 (2017) 1010–1021, <https://doi.org/10.1016/j.spmi.2017.07.059>.
- [13] J. Jang, J. Song, S.S. Lee, S. Jeong, B.J. Lee, S. Kim, Analysis of temperature-dependent I-V characteristics of the Au/n-GaSb Schottky diode, Mater. Sci. Semicond. Process. 131 (2021) 105882, <https://doi.org/10.1016/j.mssp.2021.105882>.
- [14] T. Güzel, A.B. Çolak, Artificial intelligence approach on predicting current values of polymer interface Schottky diode based on temperature and voltage: an experimental study, Superlattice. Microst. 153 (2021) 106864, <https://doi.org/10.1016/j.spmi.2021.106864>.
- [15] A.B. Çolak, T. Güzel, O. Yıldız, M. Özer, An experimental study on determination of the shottky diode current-voltage characteristic depending on temperature with artificial neural network, Phys. B Condens. Matter 608 (2021), <https://doi.org/10.1016/j.physb.2021.412852>.
- [16] A. Rabehi, B. Nail, H. Helal, A. Douara, A. Ziane, M. Amrani, B. Akkal, Z. Benamara, Optimal estimation of Schottky diode parameters using advanced swarm intelligence algorithms, Semiconductors 54 (2020), <https://doi.org/10.1134/S1063782620110214>.
- [17] K. Wang, M. Ye, Parameter estimation of Schottky-barrier diode model by particle swarm optimization, Int. J. Mod. Phys. C 20 (2009), <https://doi.org/10.1142/S029181390139911>.
- [18] M. Rahmani, Z. Meziani, Z. Dibi, Modelling graphene/n-Si Schottky junction solar cells by artificial neural networks, in: Proc. - 2019 1st Int. Conf. Sustain. Renew. Energy Syst. Appl. ICSRESA 2019, IEEE, 2019, pp. 1–6, <https://doi.org/10.1109/ICRESA49121.2019.9182697>.
- [19] N. Karaboga, S. Kockanat, H. Dogan, The parameter extraction of the thermally annealed Schottky barrier diode using the modified artificial bee colony, Appl. Intell. 38 (2013), <https://doi.org/10.1007/s10489-012-0372-x>.
- [20] H. Doğan, S. Elagoz, Temperature-dependent electrical transport properties of (Au/Ni)/n-GaN Schottky barrier diodes, Phys. E Low-Dimensional Syst. Nanostructures. 63 (2014), <https://doi.org/10.1016/j.physe.2014.04.019>.
- [21] A. Türüt, On current-voltage and capacitance-voltage characteristics of metal-semiconductor contacts, Turk. J. Phys. 44 (2021), <https://doi.org/10.3906/fiz-2007-11>.
- [22] J.-S.R. Jang, ANFIS: adaptive-network-based fuzzy inference system, IEEE Trans. Syst. Man. Cybern. 23 (1993) 665–685, <https://doi.org/10.1109/21.256541>.
- [23] I. Eker, Y. Torun, Fuzzy logic control to be conventional method, Energy Convers. Manag. 47 (2006), <https://doi.org/10.1016/j.enconman.2005.05.008>.
- [24] E.H. Mamdani, S. Assilian, An experiment in linguistic synthesis with a fuzzy logic controller, Int. J. Man Mach. Stud. 7 (1975) 1–13, [https://doi.org/10.1016/S0020-7373\(75\)80002-2](https://doi.org/10.1016/S0020-7373(75)80002-2).

- [25] M. Sugeno, K. Tanaka, Successive identification of a fuzzy model and its applications to prediction of a complex system, *Fuzzy Set Syst.* 42 (1991) 315–334, [https://doi.org/10.1016/0165-0114\(91\)90110-C](https://doi.org/10.1016/0165-0114(91)90110-C).
- [26] Y. Torun, G. Tohumolu, Designing simulated annealing and subtractive clustering based fuzzy classifier, *Appl. Soft Comput. J.* 11 (2011), <https://doi.org/10.1016/j.asoc.2010.07.020>.
- [27] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, *Neural Network.* 2 (1989) 359–366, [https://doi.org/10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8).
- [28] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back-propagating errors, *Nature* 323 (1986), <https://doi.org/10.1038/323533a0>.
- [29] F. Burden, D. Winkler, Bayesian regularization of neural networks, *Methods Mol. Biol.* 458 (2008), <https://doi.org/10.1007/978-1-60327-101-3>.
- [30] M. Seeger, Gaussian processes for machine learning, *Int. J. Neural Syst.* 14 (2004) 69–106, <https://doi.org/10.1142/S0129065704001899>.
- [31] C.E. Rasmussen, H. Nickisch, *Gaussian processes for machine learning (GPML) toolbox*, *J. Mach. Learn. Res.* 11 (2010) 3011–3015.
- [32] T.O. Zander, V.R. De Sa, N. Langer, S.-H. Hsu, J.B. Colombe, M.S. Caywood, D.M. Roberts, H.S. Greenwald, M.Z. Weiland, Gaussian Process Regression for Predictive but Interpretable Machine Learning Models: An Example of Predicting Mental Workload across Tasks, 2017, <https://doi.org/10.3389/fnhum.2016.00647>.
- [33] F. Zhang, L.J. O'Donnell, Support vector regression, in: *Mach. Learn. Methods Appl. To Brain Disord.*, Elsevier, 2019, pp. 123–140, <https://doi.org/10.1016/B978-0-12-815739-8.00007-9>.
- [34] Z. Zhong, T.R. Carr, Application of mixed kernels function (MKF) based support vector regression model (SVR) for CO<sub>2</sub> – reservoir oil minimum miscibility pressure prediction, *Fuel* 184 (2016), <https://doi.org/10.1016/j.fuel.2016.07.030>.
- [35] K. Cheng, Z. Lu, Y. Wei, Y. Shi, Y. Zhou, Mixed kernel function support vector regression for global sensitivity analysis, *Mech. Syst. Signal Process.* 96 (2017), <https://doi.org/10.1016/j.ymssp.2017.04.014>.
- [36] Y. Zhang, Y. Wang, G. Zhou, J. Jin, B. Wang, X. Wang, A. Cichocki, Multi-kernel extreme learning machine for EEG classification in brain-computer interfaces, *Expert Syst. Appl.* 96 (2018), <https://doi.org/10.1016/j.eswa.2017.12.015>.