



Machine learning-based analyses for total ionizing dose effects in bipolar junction transistors

Bai-Chuan Wang^{1,2,3} · Meng-Tong Qiu³ · Wei Chen³ · Chen-Hui Wang^{3,4} · Chuan-Xiang Tang^{1,2}

Received: 16 May 2022 / Revised: 11 August 2022 / Accepted: 21 August 2022 / Published online: 12 October 2022
© The Author(s), under exclusive licence to China Science Publishing & Media Ltd. (Science Press), Shanghai Institute of Applied Physics, the Chinese Academy of Sciences, Chinese Nuclear Society 2022

Abstract Machine learning methods have proven to be powerful in various research fields. In this paper, we show that research on radiation effects could benefit from such methods and present a machine learning-based scientific discovery approach. The total ionizing dose (TID) effects usually cause gain degradation of bipolar junction transistors (BJTs), leading to functional failures of bipolar integrated circuits. Currently, many experiments of TID effects on BJTs have been conducted at different laboratories worldwide, producing a large amount of experimental data, which provides a wealth of information. However, it is difficult to utilize these data effectively. In this study, we proposed a new artificial neural network (ANN) approach to analyze the experimental data of TID effects on BJTs. An ANN model was built and trained using data collected from different experiments. The results indicate that the proposed ANN model has advantages in capturing non-linear correlations and predicting the data. The trained ANN model suggests that the TID hardness of a BJT tends to increase with base current I_{B0} . A possible cause for this

finding was analyzed and confirmed through irradiation experiments.

Keywords Total ionizing dose effects · Bipolar junction transistor · Artificial neural network · Machine learning · Radiation effects

1 Introduction

Bipolar junction transistors (BJTs) are widely used as analog components in electronic systems [1]. Unfortunately, BJTs exhibit total ionizing dose (TID) effects, which induce current gain degradation in radiation environments []. Such TID effects are complex and influenced by many factors. A large number of studies have been conducted on various dependences and underlying mechanics of TID effects, on bipolar devices 2–6, as well as on other types of devices [7–10]. Generally, to investigate the correlations between the TID hardness and the parameters of BJTs, a series of devices should be prepared for comparison. This method is effective in excluding the influence of other parameters, but is costly. Many laboratories worldwide have been conducting TID experiments for many years, producing a large amount of experimental data. Such data are growing rapidly and are readily available. The correlations between the TID hardness and the parameters of the BJTs may be extracted from these data. However, such highly nonlinear data cannot be accurately described using multiple linear regression. Therefore, a new generation of computational tools is needed to assist researchers in extracting valuable information from the growing volumes of data.

This work was supported by the National Natural Science Foundation of China (Nos. 11690040 and 11690043).

✉ Chuan-Xiang Tang
tang.xuh@tsinghua.edu.cn

- ¹ Department of Engineering Physics, Tsinghua University, Beijing 100084, China
- ² Key Laboratory of Particle & Radiation Imaging (Tsinghua University), Ministry of Education, Beijing 100084, China
- ³ State Key Laboratory of Intense Pulsed Radiation Simulation and Effect, Xi'an 710024, China
- ⁴ State Key Laboratory of Nuclear Physics and Technology, Peking University, Beijing 100871, China

Machine learning methods can capture correlations in data and make predictions, thereby providing an alternative approach for scientific investigations [11]. These methods, which exhibit huge potential for generalized suitability for scientific research, have been applied to different research areas for decades [12]. Examples include genetic research [11], antibiotic discoveries [13], material design [14–16], and quantum entanglement simulations [17, 18]. Recently, machine learning methods have also been successfully applied in an increasing number of areas in nuclear physics research [19], including nuclear theories [20–22], experimental methods [23–25], accelerator science [26, 27], and nuclear data processing [28, 29]. Neural networks are powerful methods that have proven to be effective for nonlinear function fitting and accurate prediction [30, 31].

In this paper, we demonstrate that research on radiation effects could benefit from machine learning methods and present a machine learning-based scientific discovery approach for the TID effects of BJTs. Specifically, an artificial neural network (ANN) model was built and trained using the data collected from different experiments. The results indicate that the trained neural network model has significant advantages over the traditional multiple linear regression in capturing nonlinear correlations and predicting data. The trained ANN model suggests that the TID hardness of a BJT tends to increase with base current I_{B0} . A possible mechanism for this phenomenon was analyzed and experimentally verified. Our work indicates that machine learning methods have advantages in discovering correlations and predicting experimental data. The proposed approach could be a powerful new tool to discover correlations from experimental datasets and make predictions for radiation effects.

2 Methods

2.1 Datasets

TID degradation is related to many parameters such as bias, layout, dose, dose rate, passivation layer, and hydrogen content [32–34]. A dataset containing all the related parameters is perfect for analysis. However, the collected historical experimental data do not contain all related parameters. Nevertheless, a dataset containing several parameters may still be helpful in assisting scientific discovery by providing useful information, as described herein.

The experimental dataset to be analyzed was collected from 10 articles [1, 35–43]. The Gummel-plot data of 12 bipolar devices (eight NPNs and four LPNPs) were obtained from the literature. The Gummel plot records the collector and base current values at different base-emitter

voltages. Experimental data of bipolar devices irradiated by cobalt-60 gamma sources with different doses and dose rates at room temperature were collected. The Gummel plots before and after irradiation were measured. Degradations of BJTs with different types, V_{BE} , Beta_0 , I_{B0} , doses, and dose rates were extracted to create the dataset depicted in Fig. 1. The dataset contains 565 radiation response data points for two types of BJTs: NPN and LPNP. V_{BE} is the base-emitter voltage, whereas Beta_0 and I_{B0} are the corresponding common emitter current gain and base current before irradiation, respectively. The maximum dose was 1,902 krad(Si). The dose rate was in the range of 0.0015–312 rad(Si)/s. The absorbed doses and dose rates of silicon were used in this study. The degradation of a BJT is represented by the change in the base current I_B/I_{B0} , where I_{B0} and I_B are the base currents before and after irradiation, respectively. The total ionizing dose effects mainly cause an increase in the base current I_B at a fixed base-emitter voltage V_{BE} , whereas the collector current I_C remains roughly constant [37, 39]. Therefore, I_B/I_{B0} represents gain degradation of the BJT.

In addition to the Gummel-plot data, the correlations between degradations and doses or dose rates, obtained via experimentation, have been exhibited in the literature [35, 40, 41]. These additional data were adopted as a test set to verify the generalization of the trained ANN model. None of the experimental data in the test set were used to train the ANN.

2.2 Artificial neural network model

We focused on three-layer ANNs, which have proven to be powerful in studies using relatively small datasets [44, 45]. The ANNs were implemented using a deep learning framework called Keras [46]. The inputs of the ANN are the type, $|V_{BE}|$, Beta_0 , I_{B0} , dose, and dose rate. The first neuron of the input layer represents the type of a BJT. Specifically, ‘0’ stands for the NPN type and ‘1’ stands for the LPNP type. Each input parameter is normalized to a mean of 0 and standard deviation of 1 over the training dataset. The output of the ANN was one neuron, corresponding to the change in the base current in the logarithm $\log(I_B/I_{B0})$. The rectified linear unit (ReLU) [47] was adopted as the activation function for the first two layers, while the linear activation function was adopted for the last layer. The dropout algorithm with the rate of 0.1 was employed for the first two layers. This algorithm is helpful in improving the generalization performance of ANNs [48]. The loss function is the mean squared error. Adam with the default learning rate of 0.001 was employed as the optimizer to update the weights during the training process [49]. The batch size was set to 32 for stable convergence during the training [50].

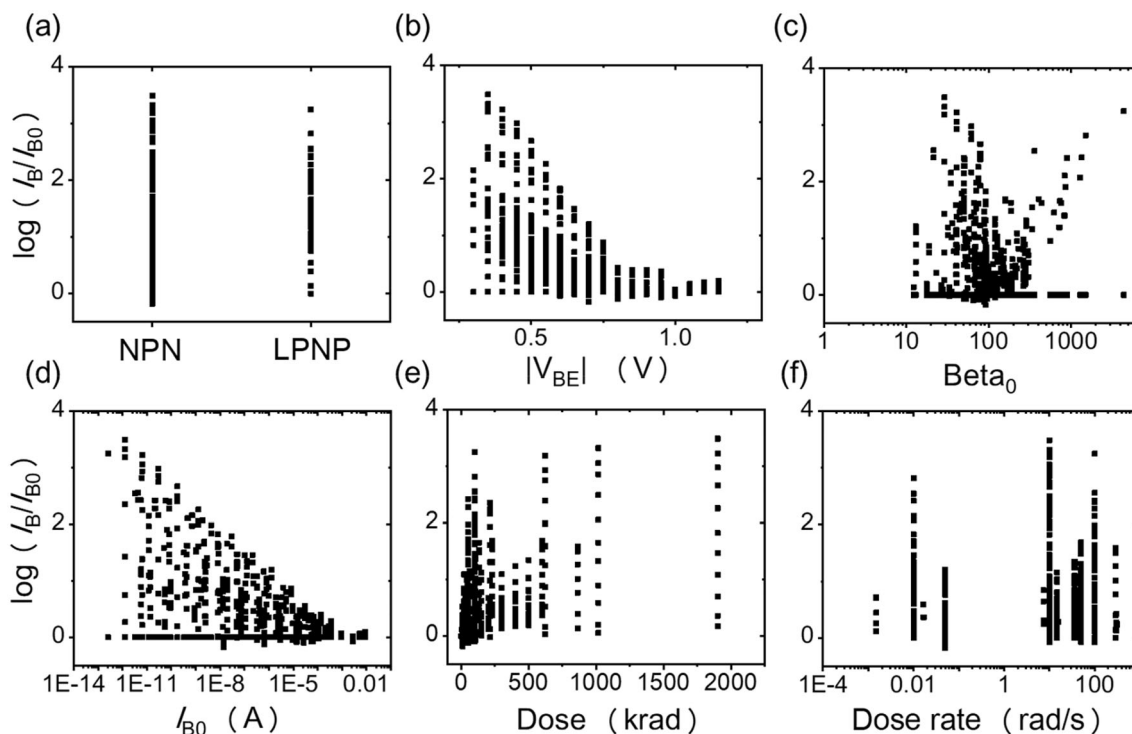


Fig. 1 Overview of the dataset: degradation $\log(I_B/I_{B0})$ versus (a) type, (b) bias condition V_{BE} , (c) common emitter current gain Beta_0 , (d) base current I_{B0} , (e) dose, and (f) dose rate

A suitable number of neurons and training epochs may vary significantly for different problems. Fivefold cross-validation (CV) was utilized to identify the suitable number of neurons in the two hidden layers and training epochs for fitting our dataset. In this technique, the dataset is randomly divided into five parts; five ANNs were trained and iteratively evaluated. Each ANN was trained using four parts of the dataset and validated using the remaining part. The CV technique was helpful in reducing the influence of training instability. The performance of the ANNs was evaluated using the mean absolute error (MAE), computed as:

$$\text{MAE} = \sum_{i=1}^N |y_i - \hat{y}_i|/N, \tag{1}$$

where y and \hat{y} represent the experimental and predicted values for $\log(I_B/I_{B0})$, respectively, and N represents the number of data points.

The performance of the ANNs with different training epochs and neuron numbers is shown in Fig. 2. The training and validation MAEs are the average values of all five-fold ANNs. The maximum number of neurons to be investigated was limited to 20×20 considering the small size of the dataset to be fitted. As can be seen in Fig. 2a, the MAE on the validation set converged after 1,000 epochs of training for the different sizes of ANNs evaluated in our study. Therefore, the number of epochs for training was set

to 1,000. Figure 2b presents the influence of the number of neurons in the two hidden layers on the valid MAE. The MAE on the validation set decreased with an increase in neurons. The distribution of the MAE is symmetrical about the diagonal. Therefore, it is a good choice to set the number of neurons in the two layers equal when the total number of neurons is fixed. Figure 2c shows the MAEs of the training and validation sets for ANNs with the same number of neurons in the two layers. As the number of neurons increased, the MAEs of the training and validation sets decreased. However, the MAE of the validation set decreased more slowly than that of the training set. As shown in Fig. 2d, the difference between the two increased. A large difference indicates that the model can fit the training set accurately, but cannot accurately predict the validation set. This indicates poor generalization performance owing to overfitting [51]. To accurately capture the correlations between data, a good model should have a small MAE and a good generalization performance. In our study, we set the number of neurons in each layer to 7. Figure 3 displays a schematic of the proposed ANN model. Applying more neurons could reduce the MAE slightly, but simultaneously reduces the generalization performance. We preferred fewer neurons to avoid capturing fake correlations caused by overfitting.

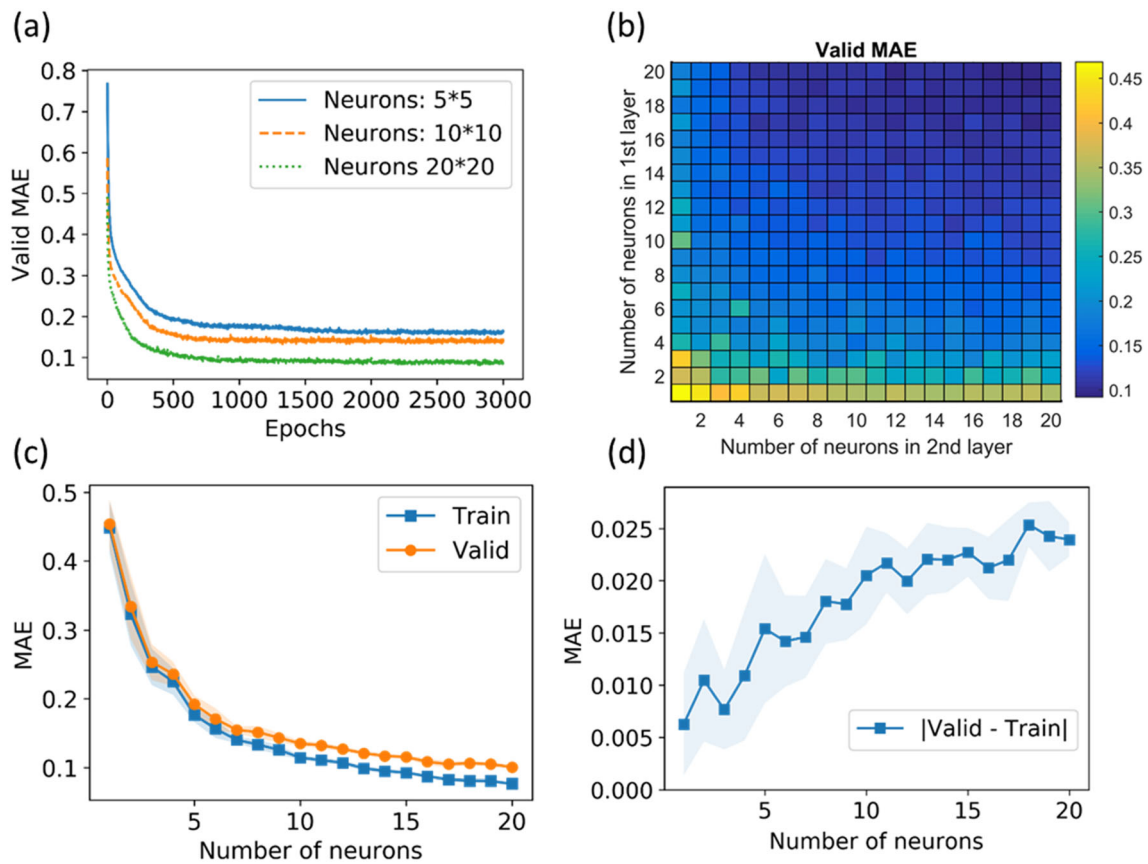


Fig. 2 Performance of ANNs with different training epochs and numbers of neurons. (a) Valid MAE for different training epochs. The valid MAE converges after 1,000 epochs for ANNs with different numbers of neurons. (b) Valid MAE for ANNs with different numbers of neurons in two hidden layers. (c) Training MAE and valid MAE for

ANNs with the same number of neurons in the two layers. (d) Differences between training MAE and valid MAE. For (c) and (d), the data points in the figures are the average values of 10 random runs, while their standard deviations are marked with shaded areas. (Colour figure online)

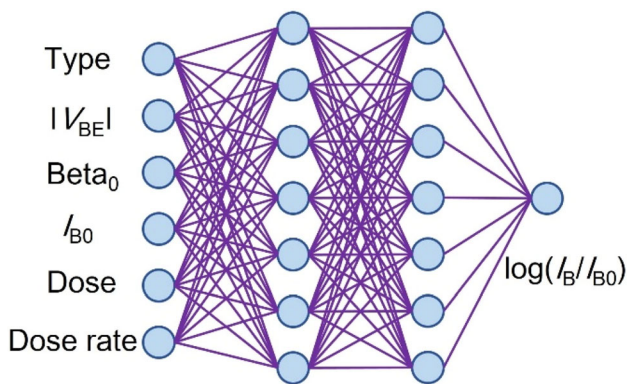


Fig. 3 Schematic of the proposed ANN model for our dataset. (Colour figure online)

3 Results

3.1 Predictions

With the help of CV technology, five ANNs were trained using different parts of the dataset. The predictions

made by these ANNs were slightly different. We used the average value as the prediction result. The performance of the trained ANNs regarding the test set is expressed in Fig. 4. Figure 4a–g shows that the degradations at different doses and dose rates predicted by the trained ANN model agree with the experimental results for the test set. Figure 4h presents a summary of the predictions for the test set. The MAE between the predictions and experimental data was 0.101.

Figure 5 presents a comparison of the learning curves for the three different models. The performance of our proposed ANN model was compared with those of the average and multiple linear regression models. The average model considers the average value of the training set samples as the prediction results. As the number of training samples increased, the MAE on the test set decreased significantly for the ANN and multiple linear regression models. When trained with 550 samples, the average MAE for 10 random runs of the ANN model was the smallest, which was nearly half of that of the multiple linear regression model. The results show that the proposed ANN

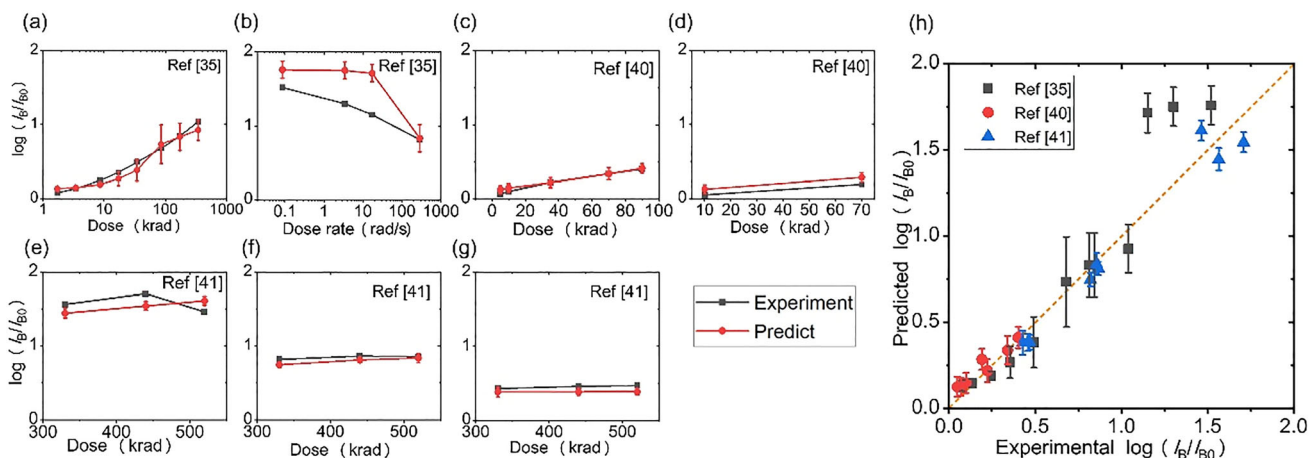


Fig. 4 Performance of the trained ANN model on the test set: (a)–(g) Experimental results and predictions of degradations at different doses and dose rates. (h) Summary of the predictions for the test set.

Error bars represent the standard deviations of different CV parts. (Colour figure online)

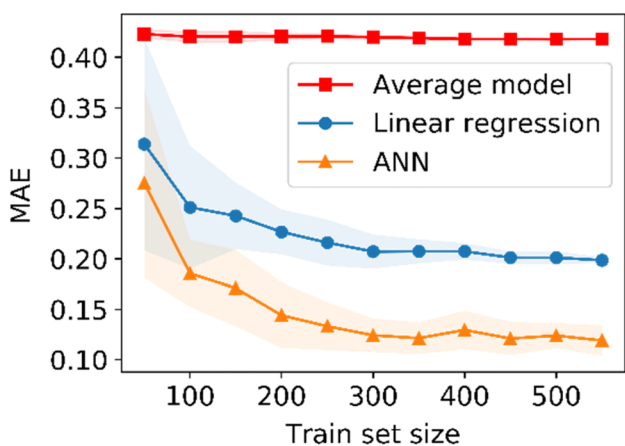


Fig. 5 Comparison of learning curves for average model, multiple linear regression, and ANN. The data points in the figures are the average values of 10 random runs, while their standard deviations are marked with shaded areas. (Colour figure online)

model performs better than the traditional multiple linear regression model in predicting the experimental results.

3.2 Correlations

The trained ANN model fits the experimental data well, indicating that the trained ANN model successfully learned the correlations between the degradation and input parameters in the dataset. These correlations were investigated by feeding the trained ANN model with specific inputs. Figure 6 displays the correlations between the degradation and each parameter generated by the trained ANN model. Specifically, dose = 100 krad(Si), dose rate = 100 rad(Si)/s, $|V_{BE}| = 0.6$ V, $\text{Beta}_0 = 100$, and $I_{B0} = 1 \times 10^{-8}$ A were selected as the reference parameter values. The data points in the figures are the average values

of the different CV parts, and the shaded areas represent their standard deviations. The correlations captured via multiple linear regression are shown in Fig. 7 for comparison.

Most correlations captured by the trained ANN model were consistent with classical theories, whereas a novel correlation was found. These correlations are nonlinear and cannot be accurately described using multiple linear regression model. In Fig. 6a, the degradation trend with dose is shown. It is apparent that the degradation of the BJT increases with increasing dose before saturation. In particular, degradation of the LPNP-type BJT was more severe and saturated earlier. These phenomena are consistent with those reported in previous studies [35]. As depicted in Fig. 6b, the degradation at a low dose rate is more severe, implying that enhanced low dose rate sensitivity (ELDRS) effects are captured in the ANN model. Figure 6c indicates the consistency with previous research results that degradation is more severe at lower bias voltages [1]. In Fig. 6d, the captured correlation between degradation and Beta_0 is unclear. It seems that for the NPN-type device, the degradation increased with Beta_0 . However, the degradation changed little with Beta_0 for the LPNP-type. In Fig. 6e, it is noteworthy that the trained ANN model captured the correlation that degradation of a BJT will lessen as the pristine base current I_{B0} increases, that is, the TID hardness of a BJT tends to increase with base current I_{B0} . To the best of our knowledge, this hypothesis has rarely been reported in the literature.

The proposed hypothesis shown in Fig. 6e agrees with the statistical analyses. The pristine base current I_{B0} and V_{BE} are related. To exclude the influence of V_{BE} , the experimental data of the same bias ($V_{BE} = 0.6$ V) were selected to analyze the correlation between I_{B0} and the

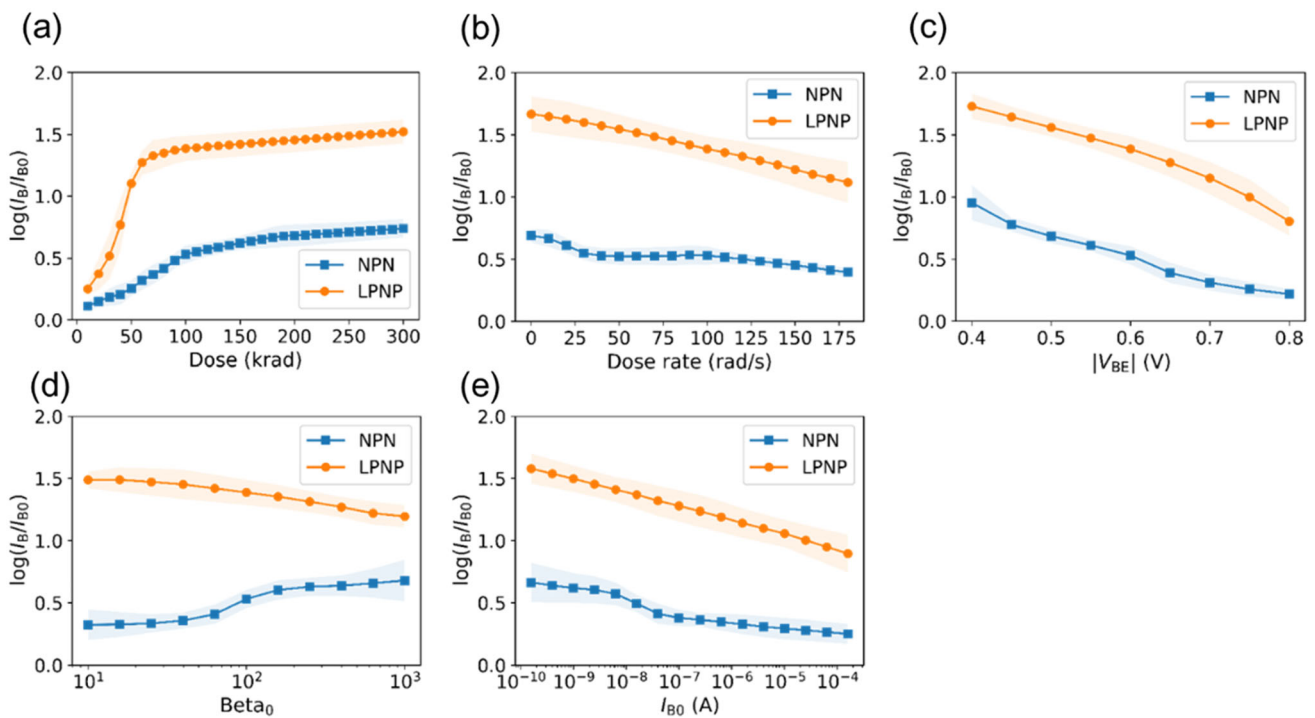


Fig. 6 Correlations captured by the trained ANN model between the degradation and each parameter. (a)–(e) results from the trained ANN model with the reference parameters: dose = 100 krad(Si), dose rate = 100 rad(Si)/s, $|V_{BE}| = 0.6$ V, $\text{Beta}_0 = 100$, and $I_{B0} = 1 \times 10^{-8}$ A. Shaded areas represent the standard deviations of different CV parts. (Colour figure online)

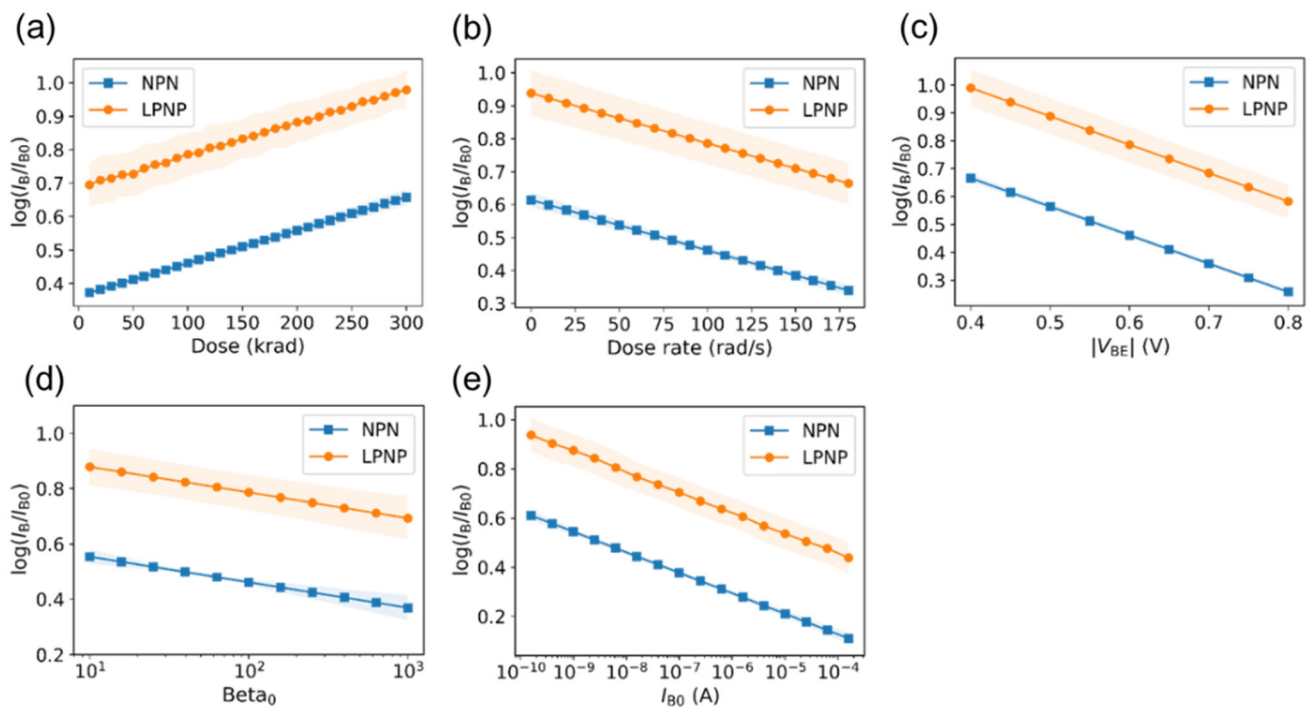


Fig. 7 Correlations captured by the multiple linear regression model between the degradation and each parameter. Shaded areas represent the standard deviations of different CV parts. (Colour figure online)

degradation $\log(I_B/I_{B0})$. The correlation coefficient r was calculated from the following equation as -0.344 :

$$r = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sqrt{\sum(x - \bar{x})^2} \sqrt{\sum(y - \bar{y})^2}}, \quad (2)$$

where x and y represent I_{B0} and $\log(I_B/I_{B0})$, respectively; and \bar{x} and \bar{y} represent the average values. The above value indicates a weak correlation. The two-tailed p-value obtained from the Student's t-test was 0.010, indicating that the negative correlation coefficient was statistically significant.

3.3 Mechanism analyses

The possible causes of the proposed hypothesis could be analyzed using classical TID mechanism theories. For an ideal NPN BJT, the pristine base current can typically be approximated as:

$$I_{B0} = Aq \frac{D_p n_i^2}{N_E W_E} \exp\left(\frac{V_{BE} - I_E R_s}{V_T}\right), \quad (3)$$

where A is the active emitter area, q is the magnitude of the electronic charge, D_p is the diffusion constant, N_E is the emitter doping, V_T is the thermal voltage, W_E is the effective emitter width, I_E is the emitter current, and R_s is the series resistance between the emitter and base [52, 53].

Ionizing radiation increases the base current owing to the fixed positive oxide-trapped charge (N_{OT}) and interface traps (N_{IT}) [53]. The base current after gamma irradiation is often modeled as [34, 53, 54]:

$$I_B = I_{B0} + \Delta I_R. \quad (4)$$

The increment in the base current mainly contains two parts of the electron-hole recombination current:

$$\Delta I_R = \Delta I_{R-SCR} + \Delta I_{R-NBS}. \quad (5)$$

One part was located above the emitter-base space-charge region (SCR). The other was located above the neutral-base surface (NBS). The excess base current in the SCR can be expressed as:

$$\Delta I_{R-SCR} = \frac{P_E \Delta s q V_T \pi n_i}{2E_m} \exp\left(\frac{V_{BE} - I_E R_s}{2V_T}\right), \quad (6)$$

where P_E is the emitter perimeter, Δs is the surface recombination velocity, n_i is the intrinsic carrier concentration of silicon, and E_m is the maximum electric field in the SCR [53]. The excess base current in the NBS is expressed as:

$$\Delta I_{R-NBS} = \frac{P_E \Delta s q W_B n_i^2}{2n_s} \left[\exp\left(\frac{V_{BE} - I_E R_s}{V_T}\right) - 1 \right], \quad (7)$$

where n_s is the majority carrier concentration at the surface and W_B is the width from the emitter to collector [53, 55]. The degradation can be written as:

$$I_B/I_{B0} = 1 + \Delta I_R/I_{B0}. \quad (8)$$

Notably, the total excess base current ΔI_R , which is the sum of ΔI_{R-SCR} and ΔI_{R-NBS} , is proportional to the perimeter of the emitter P_E . The pristine base current I_{B0} is proportional to the area of the emitter A . When other parameters of the BJTs are approximately the same, the degradation should be proportional to the perimeter-to-area ratio, that is:

$$\Delta I_R/I_{B0} \propto P_E/A. \quad (9)$$

The perimeter increased as the area of emitter A increased, but the perimeter-to-area ratio (P_E/A) tended to decrease. Consequently, I_B/I_{B0} decreased. Therefore, the difference in the perimeter-to-area ratio P_E/A may be one of the possible mechanisms leading to the phenomenon that a BJT with a larger base current I_{B0} tends to have a smaller degradation I_B/I_{B0} .

3.4 Irradiation experiments

Irradiation experiments were conducted to verify the influence of the perimeter-to-area ratio on degradation. The irradiation experiments were performed using a cobalt-60 gamma source at room temperature. Five kinds of NPN BJTs with different emitter sizes were irradiated, as shown in Fig. 8a. The other parameters of the BJTs are approximately the same. The devices were manufactured by Analog Foundries, based on a 6-inch bipolar process platform. Accounting for the uncertainties caused by manufacturing process fluctuations, three devices of each kind were used for irradiation. The total doses were 40, 80, 120, and 160 krad(Si), and the dose rate was 0.685 rad(Si)/s in the experiments. All BJTs were grounded during irradiation. The Gummel-plot data were measured using a semiconductor analyzer at room temperature before and after irradiation. V_E was swept from -0.4 to -1.2 V, maintaining $V_B = V_C = 0$ V. The delay between irradiation and every measurement was within 2 h. Typical Gummel-plot data at different doses are presented in Fig. 8b.

As shown in Fig. 8c, the degradation I_B/I_{B0} increases linearly with the perimeter-to-area ratio P_E/A , which agrees with Eqs. (8) and (9). The measured correlations between the degradation and pristine base current at different doses, shown in Fig. 8d, indicate that a BJT with a larger base current I_{B0} tends to have a smaller degradation I_B/I_{B0} , which is consistent with the ANN result. Experiments confirmed that the perimeter-to-area ratio could be one of the causes of this phenomenon.

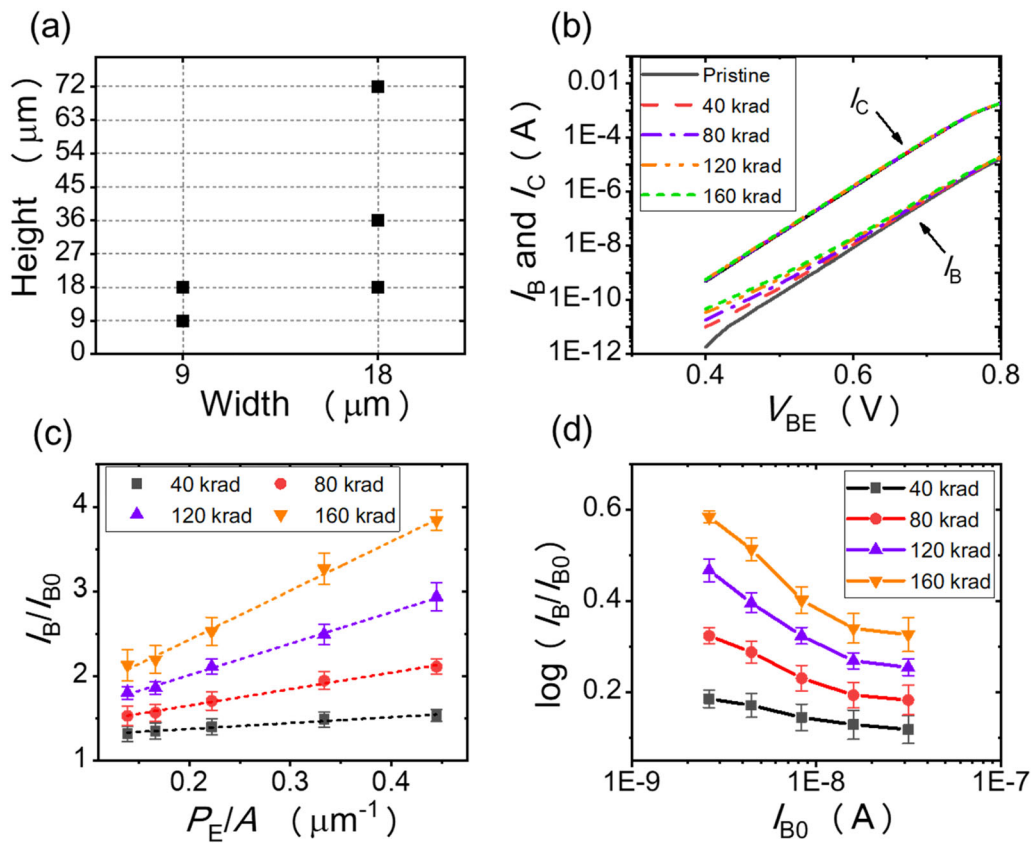


Fig. 8 Results of the irradiation experiments. **(a)** Emitter sizes of BJTs used in experiments. **(b)** Gummel plots for one of the $18 \times 18 \mu\text{m}^2$ emitter BJTs at different doses. **(c)** Degradation versus perimeter-to-area ratio at different doses ($V_{BE} = 0.6$ V). Error bars

are the standard deviations of the measurement results. Dashed lines are the linear fitting results of the mean values. **(d)** Degradation versus pristine base current at different doses ($V_{BE} = 0.6$ V). (Colour figure online)

4 Discussions

4.1 Predictive ability

It should be noted that the trained ANN cannot guarantee the degradation prediction for a device that does not have any data in the training set. This is because of the systematic deviations between different devices and experiments. An ANN may overestimate or underestimate a device that has never been observed. The trained ANN is more suitable for predicting missing values, such as predicting degradations at other doses or dose rates in Sect. 3. More specifically, the ANN was trained to predict missing values in the training process. The dataset included experimental data from different devices. During training, the dataset was shuffled and divided into training and validation sets. The validation set can be considered the missing value of the training set. The ANN model was trained using the training set and verified by predicting the validation set. Therefore, the performance of the validation set correlates with ANN's ability to predict missing values.

As shown in Fig. 9a, the ANN failed to predict the data from the irradiation experiments in Fig. 8. The degradation was overestimated and the MAE was as large as 0.45. However, after the 130 krad Gummel-plot data of the $18 \times 18 \mu\text{m}^2$ NPN were included in the dataset, the predictions of the newly trained ANN model on irradiation experiments could be significantly improved. As shown in Fig. 9b, the MAE decreased to 0.14. It is noteworthy that only the degradations at 130 krad of one device were included in the dataset, but the predictions of 9×9 , 9×18 , 18×18 , 18×36 , and $18 \times 72 \mu\text{m}^2$ NPNs at 40, 80, 120, and 160 krad were improved. This implies that the ANN model learned the correlations from the other devices in the dataset. When more data are included, predictions can be further improved. Figure 9c shows the predictions after 50 krad Gummel-plot data of the $18 \times 72 \mu\text{m}^2$ NPN were added to the dataset. MAE reduced to 0.09. Note that all of the experimental data to be predicted in Fig. 9 have never been used for training.

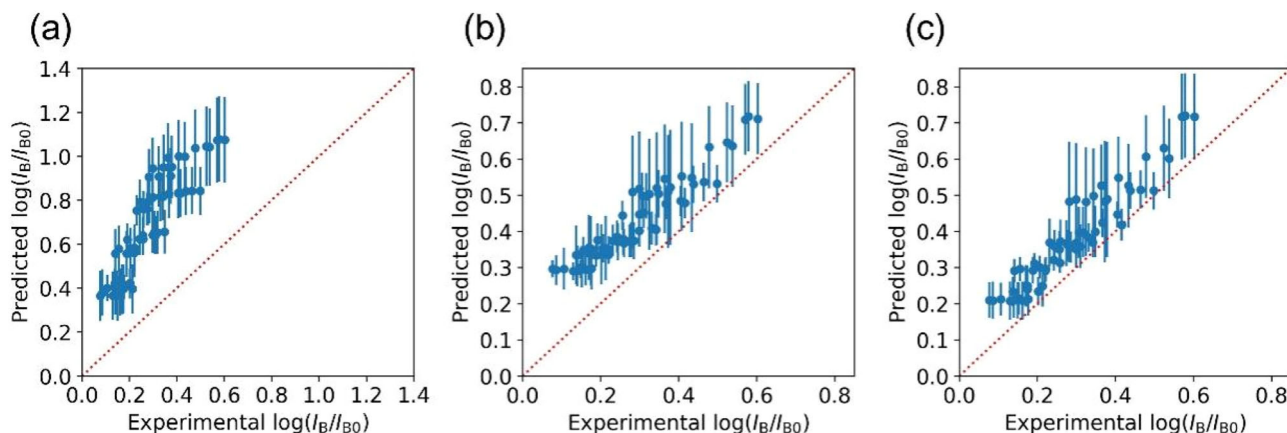


Fig. 9 Performance of predictions on experimental data of Fig. 8. (a) ANN trained with original dataset. (b) ANN trained with dataset adding 130 krad Gummel-plot data ($18 \times 18 \mu\text{m}^2$ NPN). (c) ANN trained with dataset adding 130 krad Gummel-plot data ($18 \times 18 \mu\text{m}^2$ NPN) and 50 krad Gummel-plot data ($18 \times 72 \mu\text{m}^2$ NPN)

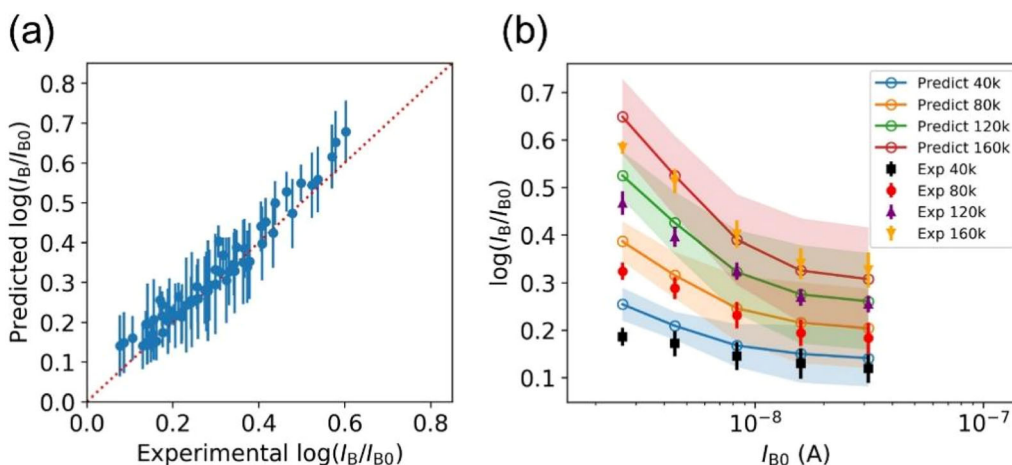


Fig. 10 Predictions on irradiation experiments of ANN trained with a new feature factor F . (a) Predictions versus experimental results in Fig. 8. (b) Degradations versus pristine base current at different doses ($V_{BE} = 0.6$ V). Shaded areas represent the standard deviations of different CV parts. (Colour figure online)

4.2 Systematic deviations between different devices

There may be systematic deviations between different experiments, such as differences between the radiation sources and measurement instruments. Moreover, devices from different manufacturers may also exhibit systematic deviations. Systematic deviations limit the ability to predict new devices. If we could account for systematic deviations, the predictions of the new devices could be improved.

We propose a simple method for approximately characterizing the system deviations between devices. We introduced a factor F to represent the deviation of one device from the average value of the other devices. To evaluate the factor F , a multiple linear regression model was trained using the dataset, excluding the device to be evaluated. The trained model represents the average of the

other devices. The predictions of the device to be evaluated are denoted by \hat{y} , and the actual degradations are denoted by y . Factor F was calculated by linear fitting as follows:

$$\hat{y} = F \cdot y. \tag{10}$$

Specifically, the fitting coefficient F can be calculated by:

$$F = \frac{\sum_{i=1}^N (y_i \cdot \hat{y}_i)}{\sum_{i=1}^N y_i^2}. \tag{11}$$

If the factor F of a device is larger than 1, the device is likely to have less degradation than the other devices in the dataset. The calculated F for the 12 devices in the dataset ranged from 0.36 to 1.48. The calculated F value was set as a new feature of the dataset to account for systematic deviations. A new ANN model with the same structure as

the previous model was trained, and the predictions of our irradiation experiments are shown in Fig. 10. It is clear that this model can accurately predict the experimental results. However, the factor F of the device is required when predicting. This implies that one should have some device data to compute factor F in advance. In our case, the factor F was calculated with degradations at 130 krad.

5 Conclusion

We presented a machine learning-based scientific discovery approach for radiation effect research. It is shown that the machine learning method could be a powerful new tool to discover correlations from experimental datasets and make predictions. An ANN model was built and trained using the dataset collected from different experiments. The results indicate that the proposed ANN model has advantages over multiple linear regression in capturing most nonlinear correlations and predicting data. Most correlations captured by the trained ANN model were consistent with classical theories, whereas a new correlation was found. The trained ANN model suggests that the TID hardness of a BJT tends to increase with base current I_{B0} . Further mechanistic analyses and experiments confirmed that the differences in the emitter perimeter-to-area ratio of the BJTs could be one of the causes of this phenomenon. The ANN model presented in this study was trained using a relatively small and simple dataset. It is expected to be more powerful if a larger and more detailed dataset is provided. We plan to conduct ANN analyses on the historical data from several laboratories and build models for other kinds of devices.

Author contributions All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by Bai-Chuan Wang, Meng-Tong Qiu, and Chuan-Xiang Tang. The first draft of the manuscript was written by Bai-Chuan Wang and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

References

1. R. Li, C. Wang, W. Chen et al., Synergistic effects of TID and ATREE in vertical NPN bipolar transistor. *IEEE Trans. Nucl. Sci.* **66**, 1566–1573 (2019). <https://doi.org/10.1109/TNS.2019.2909690>
2. X. Li, J. Yang, D.M. Fleetwood et al., Hydrogen soaking, displacement damage effects, and charge yield in gated lateral bipolar junction transistors. *IEEE Trans. Nucl. Sci.* **65**, 1271–1276 (2018). <https://doi.org/10.1109/TNS.2018.2837032>
3. R.L. Pease, R.D. Schrimpf, D.M. Fleetwood, ELDRS in bipolar linear circuits: a review. *IEEE Trans. Nucl. Sci.* **56**, 1894–1908 (2009). <https://doi.org/10.1109/TNS.2008.2011485>
4. D.M. Fleetwood, Total ionizing dose effects in MOS and low-dose-rate-sensitive linear-bipolar devices. *IEEE Trans. Nucl. Sci.* **60**, 1706–1730 (2013). <https://doi.org/10.1109/TNS.2013.2259260>
5. G.I. Zebrev, A.S. Petrov, R.G. Useinov et al., Simulation of bipolar transistor degradation at various dose rates and electrical modes for high dose conditions. *IEEE Trans. Nucl. Sci.* **61**, 1785–1790 (2014). <https://doi.org/10.1109/TNS.2014.2315672>
6. L. Li, X. Chen, Y. Jian et al., Modeling the ionization damage on excess base current in pnp BJTs for circuit-level simulation. *IEEE Trans. Nucl. Sci.* **68**, 2220–2231 (2021)
7. S.Y. Chang, S.S. Yang, X. Wen et al., Experimental study on the total ionizing dose effects of nonvolatile RRAM. *Nucl. Tech.* **43**(12), 120502 (2020). <https://doi.org/10.11889/j.0253-3219.2020.hjs.43.120502>. (in Chinese)
8. P. Wu, L. Wen, Z. Xu et al., Synergistic effects of total ionizing dose and radiated electromagnetic interference on analog-to-digital converter. *Nucl. Sci. Tech.* **33**, 39 (2022). <https://doi.org/10.1007/s41365-022-01017-x>
9. X.Y. Zhang, Q. Guo, Y.D. Li et al., Total ionizing dose and synergistic effects of magnetoresistive random access memory. *Nucl. Sci. Tech.* **29**, 111 (2018). <https://doi.org/10.1007/s41365-018-0451-8>
10. Y.N. Liu, Y.P. Yang, F.G. Chen et al., Radiation effect sensitive peripheral of 180 nm CMOS microprocessor and probabilistic model analysis of its damage dose. *Nucl. Tech.* **44**(03), 030502 (2021). <https://doi.org/10.11889/j.0253-3219.2021.hjs.44.030502>. (in Chinese)
11. G. Eraslan, Z. Avsec, J. Gagneur et al., Deep learning: new computational modelling techniques for genomics. *Nat. Rev. Genet.* **20**, 389–403 (2019). <https://doi.org/10.1038/s41576-019-0122-6>
12. Y. Li, Y. Xu, M. Jiang et al., Self-learning perfect optical chirality via a deep neural network. *Phys. Rev. Lett.* **123**, 213902 (2019). <https://doi.org/10.1103/PhysRevLett.123.213902>
13. J.M. Stokes, K. Yang, K. Swanson et al., A deep learning approach to antibiotic discovery. *Cell* **180**, 688–702 (2020). <https://doi.org/10.1016/j.cell.2020.01.021>
14. D. Visaria, A. Jain, Machine-learning-assisted space-transformation accelerates discovery of high thermal conductivity alloys. *Appl. Phys. Lett.* **117**, 202107 (2020). <https://doi.org/10.1063/1.5028241>
15. K.T. Butler, D.W. Davies, H. Cartwright et al., Machine learning for molecular and materials science. *Nature* **559**, 547–555 (2018). <https://doi.org/10.1038/s41586-018-0337-2>
16. J. Li, H. Zhang, J.Z.Y. Chen, Structural prediction and inverse design by a strongly correlated neural network. *Phys. Rev. Lett.* **123**, 108002 (2019). <https://doi.org/10.1103/PhysRevLett.123.108002>
17. O. Sharir, Y. Levine, N. Wies et al., Deep autoregressive models for the efficient variational simulation of many-body quantum systems. *Phys. Rev. Lett.* **124**, 020503 (2020). <https://doi.org/10.1103/PhysRevLett.124.020503>
18. M.J. Hartmann, G. Carleo, Neural-network approach to dissipative quantum many-body dynamics. *Phys. Rev. Lett.* **122**, 250502 (2019). <https://doi.org/10.1103/PhysRevLett.122.250502>
19. A. Boehnlein, M. Diefenthaler, C. Fanelli et al., Machine learning in nuclear physics., *arXiv preprint arXiv:2112.02309* (2021).
20. S. Akkoyun, Estimation of fusion reaction cross-sections by artificial neural networks. *Nucl. Instrum. Meth. B* **462**, 51–54 (2020). <https://doi.org/10.1016/j.nimb.2019.11.014>
21. A.E. Lovell, A.T. Mohan, P. Talou, Quantifying uncertainties on fission fragment mass yields with mixture density networks. *J. Phys. G Nucl. Particle Phys.* **47**, 114001 (2020). <https://doi.org/10.1088/1361-6471/ab9f58>

22. Z. Gao, Y. Wang, H. Lü et al., Machine learning the nuclear mass. *Nucl. Sci. Tech.* **32**, 109 (2021). <https://doi.org/10.1007/s41365-021-00956-1>
23. E. Doucet, T. Brown, P. Chowdhury et al., Machine learning n/γ discrimination in CLYC scintillators. *Nucl. Instrum. Meth. A* **954**, 161201 (2020). <https://doi.org/10.1016/j.nima.2018.09.036>
24. Z. Qian, V. Belavin, V. Bokov et al., Vertex and energy reconstruction in JUNO with machine learning methods. *Nucl. Instrum. Meth. A* **1010**, 165527 (2021). <https://doi.org/10.1016/j.nima.2021.165527>
25. Z.H. Wu, J.J. Bai, D.D. Zhang et al., Statistical analysis of helium bubbles in transmission electron microscopy images based on machine learning method. *Nucl. Sci. Tech.* **32**, 54 (2021). <https://doi.org/10.1007/s41365-021-00886-y>
26. Y. Zou, Q. Xing, B. Wang et al., Application of the asynchronous advantage actor-critic machine learning algorithm to real-time accelerator tuning. *Nucl. Sci. Tech.* **30**, 158 (2019). <https://doi.org/10.1007/s41365-019-0668-1>
27. Y. Yu, G. Liu, W. Xu et al., Research on tune feedback of the Hefei Light Source II based on machine learning. *Nucl. Sci. Tech.* **33**, 28 (2022). <https://doi.org/10.1007/s41365-022-01018-w>
28. C.E. Romano, L.A. Bernstein, T. Bailey et al., Proceedings of the Workshop for Applied Nuclear Data: WANDA2020. Oak Ridge National Lab.(ORNL), Oak Ridge, TN (United States) (2020).
29. D. Neudecker, O. Cabellos, A.R. Clark et al., Informing nuclear physics via machine learning methods with differential and integral experiments. *Phys. Rev. C* **104**, 34611 (2021)
30. S.C. Leemann, S. Liu, A. Hexemer et al., Demonstration of machine learning-based model-independent stabilization of source properties in synchrotron light sources. *Phys. Rev. Lett.* **123**, 194801 (2019). <https://doi.org/10.1103/PhysRevLett.123.194801>
31. Y. LeCun, Y. Bengio, G. Hinton, Deep learning. *Nature* **521**, 436–444 (2015). <https://doi.org/10.1038/nature14539>
32. K.F. Galloway, R.L. Pease, R.D. Schrimpf et al., From displacement damage to ELDRS: fifty years of bipolar transistor radiation effects at the NSREC. *IEEE Trans. Nucl. Sci.* **60**, 1731–1739 (2013). <https://doi.org/10.1109/TNS.2013.2244615>
33. J. Boch, F. Saigne, A.D. Touboul et al., Dose rate effects in bipolar oxides: competition between trap filling and recombination. *Appl. Phys. Lett.* **88**, 232113 (2006). <https://doi.org/10.1063/1.2210293>
34. C. Wang, W. Chen, X. Jin et al., Dependence on base width and doping concentration of current degradation in gate-controlled lateral PNP bipolar transistors exposed to reactor neutrons and gamma rays. *Energy Procedia* **127**, 110–119 (2017). <https://doi.org/10.1016/j.egypro.2017.08.119>
35. D.M. Schmidt, D.M. Fleetwood, R.D. Schrimpf et al., Comparison of ionizing-radiation-induced gain degradation in lateral, substrate, and vertical PNP BJTs. *IEEE Trans. Nucl. Sci.* **42**, 1541–1549 (1995). <https://doi.org/10.1109/23.488748>
36. J. Boch, F. Saigne, T. Maurel et al., Dose and dose rate effects on NPN bipolar junction transistors irradiated at high temperature. *RADECS* **2001**, 357–362 (2001)
37. S.L. Kosier, R.D. Schrimpf, R.N. Nowlin et al., Charge separation for bipolar transistors. *IEEE Trans. Nucl. Sci.* **40**, 1276–1285 (1993). <https://doi.org/10.1109/23.273541>
38. S.R. Kulkarni, R. Damle, ^{60}Co Gamma-ray induced gain degradation in bipolar junction transistors. *Indian J. Phys.* **85**, 391–400 (2011)
39. P. Zhang, X. Wu, Q. Yi et al., A comparison of the effects of cobalt-60 γ ray irradiation on DPSA bipolar transistors at high and low injection levels. *Microelectron. Reliab.* **71**, 86–90 (2017). <https://doi.org/10.1016/j.microrel.2017.02.015>
40. J.Y. Zhao, J.Q. Yang, L. Dong et al., Hydrogen soaking irradiation acceleration method: application to and damage mechanism analysis on 3DG111 transistors. *Acta Phys. Sin.* **68**, 068501 (2019). [https://doi.org/10.7498/aps.68.20181992\(inChinese\)](https://doi.org/10.7498/aps.68.20181992(inChinese))
41. O.M. Lawal, S. Liu, Z. Li et al., Experimental studies of collector-emitter voltage bias influence on the total ionization dose effects in NPN Si BJTs. *Superlattices Microst.* **122**, 194–202 (2018). <https://doi.org/10.1016/j.spmi.2018.08.008>
42. J.Q. Yang, L. Dong, C.M. Liu et al., Impact of nitride passivation layer on ionizing irradiation damage on LPNP bipolar transistors. *Acta Phys. Sin.* **67**, 168501 (2018). <https://doi.org/10.7498/aps.67.20172215>. (in Chinese)
43. X. Li, L. Dong, J. Yang et al., Impact of passivation layers on irradiation response of PNP transistors under different dose rates. *IEEE Access* **5**, 22194–22198 (2017). <https://doi.org/10.1109/ACCESS.2017.2756701>
44. Y. Pan, X. Nie, Z. Li et al., Data-driven vehicle modeling of longitudinal dynamics based on a multibody model and deep neural networks. *Measurement* **180**, 109541 (2021). <https://doi.org/10.1016/j.measurement.2021.109541>
45. J. Ma, S. Dong, G. Chen et al., A data-driven normal contact force model based on artificial neural network for complex contacting surfaces. *Mech. Syst. Signal Pr.* **156**, 107612 (2021). <https://doi.org/10.1016/j.ymsp.2021.107612>
46. Keras Documentation. <https://keras.io> Accessed 22 September 2022.
47. V. Nair, G. Hinton, Rectified linear units improve restricted boltzmann machines. In: *27th International Conference on Machine Learning (ICML-10)* (2010).
48. G.E. Hinton, N. Srivastava, A. Krizhevsky et al., Improving neural networks by preventing co-adaptation of feature detectors. *arXiv:1207.0580* (2012) <https://doi.org/10.48550/arXiv.1207.0580>
49. D. Kingma, J. Ba, Adam: A Method for Stochastic Optimization. 3rd International Conference on Learning Representations (ICLR 2015) (2015).
50. D. Masters, C. Luschi, Revisiting small batch training for deep neural networks. *arXiv:1804.07612* (2018).
51. W.G. Jiang, G. Hagen, T. Papenbrock, Extrapolation of nuclear structure observables with artificial neural networks. *Phys. Rev. C.* **100**, 54326 (2019). <https://doi.org/10.1103/PhysRevC.100.054326>
52. R.S. Müller, T.I. Kamins, *Device Electronics for Integrated Circuits* (Wiley, New York, 2003), pp.281–286
53. B.S. Tolleson, P.C. Adell, B. Rax et al., Improved model for excess base current in irradiated lateral p-n-p bipolar junction transistors. *IEEE Trans. Nucl. Sci.* **65**, 1488–1495 (2018). <https://doi.org/10.1109/TNS.2018.2829110>
54. H.J. Barnaby, B. Vermeire, M.J. Campola, Improved model for increased surface recombination current in irradiated bipolar junction transistors. *IEEE Trans. Nucl. Sci.* **62**, 1658–1664 (2015). <https://doi.org/10.1109/TNS.2015.2452229>
55. H.J. Barnaby, S.K. Smith, R.D. Schrimpf et al., Analytical model for proton radiation effects in bipolar devices. *IEEE Trans. Nucl. Sci.* **49**, 2643–2649 (2002). <https://doi.org/10.1109/TNS.2002.805410>

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.