Out-of-training-range Synthetic FinFET and Inverter Data Generation using a Modified Generative Adversarial Network

Vasu Eranki, Nathan Yee, and Hiu Yung Wong, *Senior Member, IEEE*

***Abstract*—In this paper, a novel variation of Generative Adversarial Network (GAN) is proposed and used to predict device and circuit characteristics based on design parameters. Unlike regular GAN which takes white noise as inputs, this modified GAN uses device or circuit parameters as inputs. Unlike regular Physics-informed GAN (PI-GAN) which incorporates differential equations in the training process, this modified GAN learns physics through the inputs and has one extra step of supervised learning. FinFET is used as a device example and Technology Computer-Aided-Design (TCAD) is used to generate its current-voltage (I_DV_G, I_DV_D) and capacitance-voltage (C_GV_G) curves as the training data by varying the gate length (L_G), fin top width (W_{TOP}), and gate metal workfunction (WF). A CMOS inverter with source contact defects is used as a circuit example and a SPICE simulator is used to generate its Voltage Transfer Characteristics (VTC) by varying the source contact resistances. We show that 1) the GAN model is able to generate both the device and circuit electrical characteristics based on the input parameters, 2) it can predict the characteristics of the device and circuit out of the training range (in a testing volume 3.7x to 4.6x larger than the training volume), and 3) it is further verified on *experimentally measured* data in the inverter case that it does not overfit and has learned the underlying physics.**

***Index Terms*— FinFET, Generative Adversarial Networks (GANs), Inverter, Machine Learning, Simulation, Technology Computer-Aided Design (TCAD)**

I. Introduction

SIMULATION-augmented machine learning (SAML) [1]-[4] has been proposed to use simulation data to train better machines for defect analysis [1]-[5], device characteristic predictions [6][7], device/circuit manifold learning [8][9], inverse design [10], and surrogate model development [11][12]. Since simulations, such as Technology Computer-Aided-Design (TCAD) or spice circuit simulation, are very cost-effective [4][5] and well-controlled [1]-[14] (e.g. defects can be placed precisely in the simulation structure), using simulation data in SAML makes it very useful.

Manuscript received July 16, 2022. This material is based upon work supported by the NSF under Grant No. 2046220. V. Eranki and H. Y. Wong are affiliated with the Department of Electrical Engineering, San Jose State University, San Jose, CA 95192 USA. N. Yee is affiliated with the Department of Computer Science, San Jose State University, San Jose, CA 95192 USA. (Corresponding author: H. Wong; e-mail: hiuyung.wong@sjsu.edu).

To make SAML more useful and meaningful, we believe it needs to have *at least one of the following features*. 1) It uses a reasonably small set of training data so that it can save the simulation time when it is used to predict the device characteristics under unseen bias conditions [8]. 2) It obviates domain expertise and input feature extraction so that different devices and circuits can reuse the same framework without the need to go over another cycle of human testing [5][8] [13][14]. 3) It can be used for optimization and inverse design which cannot be done using brute force simulations [11][12]. 4) It can be used to predict the characteristics *out of the training range*. It is worth noting that ML with these features can also be used to develop novel devices/circuits by using limited experimental data as the training data [8]. However, to the best of our knowledge, 4) has not been demonstrated yet.

In this paper, Generative Adversarial Network (GAN) [16]-[17], which has been shown to produce higher-resolution and more accurate outputs in other applications, is studied. A modified GAN which takes the device/circuit parameters as the input (instead of white noise [16]-[19]) and has an extra supervised learning step is used. It is demonstrated that it can predict the IV and CV of a FinFET and VTC of an inverter out of the training range. It also fulfills the aforementioned merits that no domain expertise is required (applicable to device and circuit problems) and only a limited number of training data is required.

We also note that there have been substantial works in applying physics to neural networks (NN) such as incorporating differential equations in Physics-informed GAN (PI-GAN) [19] and adding constraints [20], for initialization and pre-training

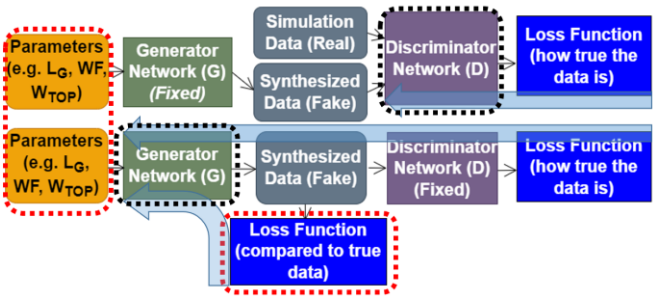

Fig. 1. The modified GAN used in this study (1 epoch shown). The red dotted boxes highlight the novelties of the proposed GAN. The first/top step is to train D (highlighted in black dotted box) while the bottom/second step is to train G (highlighted in black dotted box). 600 epochs are used. The blue semi-transparent arrows show the back propagation paths. Compared to a regular GaN, an extra training step on G is conducted by comparing to the true data.

[21], representing NN by Hamiltonian [22], for data generation [23], and hardwiring physics in NN [24]. They usually need significant domain expertise and are also not suitable if the underlying physics is unknown (as in a novel device).

II. Modified GAN

Fig. 1 shows the novel GAN. Like a regular GAN [17], it has a generator (G) and a discriminator (D). The purpose is to train G to capture the underlying physics of the true data (each data point is one simulated IV, CV, and VTC curve) so that it will generate synthetic (fake) data resembling the simulated data. D is used to distinguish if the data is true (simulated) or fake (generated by G). In each epoch, D is trained with the true data and the fake data through supervised learning with G fixed. Then D is fixed and G is trained so that it will generate data that makes D predict 1 (thought that it is simulated data). This is repeated for 600 epochs.

In this novel GAN, unlike regular GAN using white noise as the input, the device or circuit parameters are used (e.g. L_G, WF, W_{TOP}). The parameters are sampled uniformly and the sampling is described in detail in Section III and Fig. 3. This allows G's output to be correlated to device/circuit parameters. Moreover, in each epoch, G is trained one more time by comparing to the simulated curves (training data) through back-propagation. All training is applied sequentially. Both novelties help G to learn physics without incorporating differential equations as in [19].

To avoid mode collapse, W-GAN [16] is used with a learning rate of 10^{-3}. Weight clipping is also used during the training of D (-10^{-2} to 10^{-2} for VTC and -10^{-3} to 10^{-3} for I_DV_G, I_DV_D, and C_GV_G) to speed up the training and improve stability. Note that gradient penalty was not used. The training loss function and the structures of the neural networks are shown in Fig. 2. The NN structures used are obtained after a few trials of parameter optimization with hand. The optimizer used is the RMSprop optimizer with a momentum of 0, a smoothing factor of 0.99, a learning rate of 10^{-3}, and a parameters clamp of 10^{-2} after each training loop.

III. Data Generation and Preparation

3D TCAD simulation of FinFET, with calibrated TCAD models [25], is used to generate the $I_DV_G/I_DV_D/C_GV_G$ curves by varying the gate length (L_G), gate work function (WF), and the fin top width (W_{Top}). They are varied independently and uniformly in the range 15nm-25nm, 4.4eV-4.7eV, and 5nm-15nm, respectively. The details of the FinFET simulation can be found in [8].

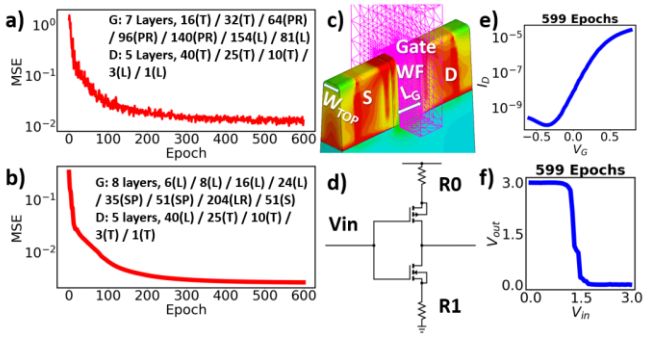

Fig. 2. Left: Training losses as a function of epoch for a) I_DV_G and VTC b). The neural network structures (hidden and output layers) are also shown with LR: Leaky Relu, PR: Parametric Relu, T: tanh, L: Linear, S: Sigmoid, SP: SoftPlus. Middle: c) FinFET structure and d) inverter circuit. Right: The corresponding predictions of I_DV_G (e) and VTC (f) for a FinFET with (L_G,WF, W_{TOP}) = (21nm, 13nm, 4.58eV) and for an inverter with (R_0, R_1) = (0.74kΩ, 0.74k kΩ), respectively.

Inverters formed by planar PMOS and NMOS are simulated for VTC curves with the PMOS and NMOS source contact resistances (R_0 and R_1) varied logarithmically uniformly and independently from 10Ω to 10MΩ. R_0 and R_1 are normalized to the NMOS resistance, R_n, at $V_{DS} = V_{GS} = V_{DD}/2 = 1.5V$ as $r_0 = \ln(R_0/R_n)$ and $r_1 = \ln(R_1/R_n)$. The SPICE simulation is calibrated to experiments. *Experimental circuits are also constructed and measured*. The details of the simulation and experimental setup can be found in [4]. Each I_DV_G/ I_DV_D/C_GV_G curve is discretized to 81 points and each VTC curve is discretized to 51 points. In data preprocessing, the VTC's are divided by 3 and used directly. For IV's, the logarithmic values are used followed by standard scaling. The I_DV_D (V_G=0.8V), saturation (V_D = 0.8V) and linear (V_D = 0.05V) I_DV_G, and C_GV_G are trained separately.

IV. Machine Learning Results

Firstly, we show the novel GAN network learns physics gradually as G and D are improved in each epoch. Fig. 2 shows the prediction errors as a function of epoch and the snapshots of the output of G for a given set of parameters.

Then, we show that the modified GAN network has better performance than the traditional manifold learning using an autoencoder to capture the underlying physics [8]. The machines are trained by *only 50 randomly chosen* simulation data and used to predict the rest of the *unseen* simulation data. Tables I and II show the R^2 of various quantities predicted by the machines and it can be seen that, overall, the novel GAN performs better than the autoencoder in the FinFET case. Note that the R^2 of G_{M2} is relatively low because it is the derivative of the IV curves.

TABLE I
FinFET Saturation I_D and G_M Prediction Accuracy (R^2)

	Saturation I_D @ V_G				G_{M2} @V_G[A]
Machine	0.8V	0.4V	-0.2V	-0.6V	0.8V
AE50 [8]	0.97	0.98	0.91	0.66	0.66
GAN	0.99	0.96	0.95	0.73	0.60

[A] G_{M2} is defined as the transconductance at $V_G = 0.8V$ in an 87.5mV interval.

TABLE II
Other FinFET Metrics Prediction Accuracy (R^2)

	CV Quantities		IV Quantities		I_DV_D
Machine	C_{high}[a]	C_{low}[b]	DIBL[d]	SS[c]	R_{ON}[e]
AE50[8]	0.84	0.93	0.84	0.87	N/A
GAN	0.98	0.97	0.94	0.98	0.98

[a]Gate capacitance at V_G=0.8V, [b]Gate capacitance at V_G=0V, [c]Subthreshold Slope defined in the region between 10^{-9}A to 10^{-6}A in saturation I_DV_G. [d]DIBL is calculated based on the difference between linear and saturation V_{TH}'s, where V_{TH} is defined as V_G@ I_D=10^{-7}A. [e]The ON-state resistance is the inverse of the slope of the I_DV_D curves calculated between V_D=0V and V_D=58mV.

TABLE III
Prediction Accuracy (R^2) by GAN of Vin Giving Specific V_{OUT}

V_{OUT}	0.75V_{DD}	0.5V_{DD}	0.25V_{DD}	V_M (V_{in}=V_{out})
R^2	0.982	0.995	0.992	0.995

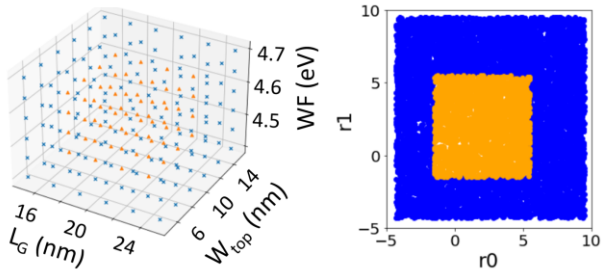

Fig. 3. The training (orange) and prediction (blue) data distribution of GAN for FinFET (left) and inverter (right).

To gauge the performance of VTC prediction by the GAN, the V_{IN}'s give $V_{OUT} = 0.75V_{DD}$, $0.5V_{DD}$, $0.25V_{DD}$, and $V_{IN} = V_{OUT} = V_M$ in the GAN's outputs are compared to the simulation one. Again, the GAN is only trained by 50 simulated VTC's and the R^2's of predicting unseen data are shown in Table III. It can be seen that the prediction of GAN is also very accurate.

Finally, to further verify that the novel GAN can capture the underlying physics, it is trained with data from a restricted parameter set and used to predict the curves *out of the training range*. We define a quantity called *parameter volume* as the product of the parameter ranges in the training. For the FinFET, the training parameter ranges are $L_G \in (17nm, 23nm)$, $WF \in (4.46eV, 4.64eV)$, $W_{TOP} \in (7nm, 13nm)$. Therefore, the parameter volume is $6.48\times10^{-18}m^2eV$ (64 data). The trained GAN is then used to predict the $I_DV_G/I_DV_D/C_GV_G$ of FinFET with $L_G \in (15nm, 25nm)$, $WF \in (4.4eV, 4.7eV)$, $W_{TOP} \in (5nm, 15nm)$ (parameter volume = $30\times10^{-18}m^2eV$ (151 data), which is 4.6x of the training volume). Similarly for the inverter, the training parameter range is $r_0 \in (-1.5, 5.5)$, $r_1 \in (-1.5, 5.5)$ (parameter volume = 49 (~2k data) and it is used to predict the inverter with $R_0 \in (-4.5, 9)$, $R_1 \in (-4.5, 9)$ (~10k data) which is 3.7x of the training volume. Fig. 3 shows the training and testing space for the FinFET and the inverter, respectively.

The FinFET GAN is used to predict the $I_DV_G/I_DV_D/C_GV_G$ curves of the FinFET with parameters out of the training range. Important characteristics are extracted from the generated curves and compared to TCAD simulations. Their prediction R^2 (V_D = 0.05V) are 0.93 for I_{on}, 0.89 for subthreshold slope, 0.98 for V_{TH} (defined as V_G when $I_D=10^{-7}A$), 0.97 for C_{high}, 0.97 for C_{low}, 0.91 for g_m, 0.89 for DIBL, 0.92 for R_{on}, 0.99 for I_D (V_G = 0V), 0.9 for I_D (V_G = -0.425V) and 0.34 for I_D (V_G = -0.6V). They all predict well except for I_D (V_G = -0.6V). Therefore, the machine can predict all curves well in the region $V_G \in (-0.425V, 0.8V)$. This is the same for the machine in Table I.

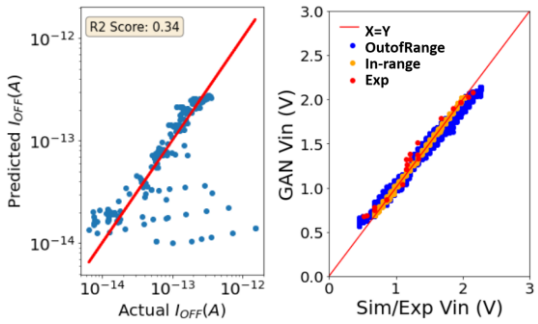

Fig. 4. FinFET $I_D(V_G$=-0.6V) prediction (left) and prediction of V_{in} that gives $V_{out} = V_{in} = V_M$ by the GAN within/out-of training range for simulation data and experimental data (right).

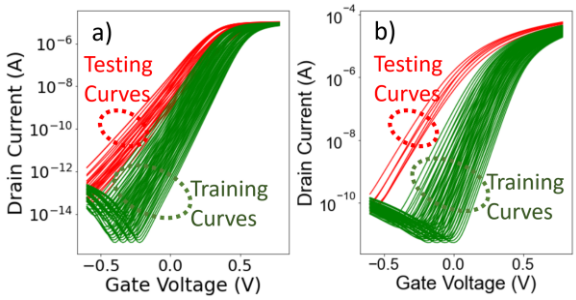

Fig. 5. Linear (a) and saturation (b) I_DV_G curves. The green lines are all the training data and the red lines are the testing data *which have bad prediction on* $I_D(V_G$=-0.6V).

Fig. 4 shows the scatter plot of the I_D (V_G = -0.6V) prediction. It is found that the leakages of 16 structures are not predicted well and they are all located at the corner of the training space (Fig. 3) with small L_G, low WF, and large W_{TOP}, thus with very low V_{TH} and Gate Induced Drain Leakage (GIDL) effect (i.e. I_D increases at very negative V_G) is *not* observed in the simulation range. Fig. 5a) plots the I_DV_G curves of the training dataset (green) and the outlier testing curves (red) from Fig. 4a). It can be seen that almost all of the training curves have a strong GIDL. Therefore, the machine "thinks" that every curve should have GIDL and predicts the outliers' I_D (V_G = -0.6V) wrong. This does not affect the prediction of the V_{TH} and SS because they are defined at a much higher current level (10^{-7}A for V_{TH} and 10^{-9}A to 10^{-6}A for SS). If the outliers are not included, the R^2 is 0.93.

To further confirm the understanding, the saturation curves are also plotted (V_D=0.8). It can be seen that, due to high V_D, GIDL occurs at a much higher current level (>10^4x) and only 4 of the testing curves are not showing GIDL. As a result, the R^2 of I_D (V_G = -0.6V) prediction is 0.72 even with the outliers. Therefore, the machine can learn the physics given in the training data set well. This also explains the relatively low R^2 of $I_D(V_G$ = -0.6V) in Table I. Fig. 4 also shows the prediction of V_{in} that gives V_{IN}= V_{OUT} =V_M. Prediction of experimental curves from [4] is also shown. The out-of-range prediction has $R^2 > 0.98$. Results for predicting V_{in}'s giving $V_{OUT} = 0.75V_{DD}$, $0.5V_{DD}$, and $0.25V_{DD}$ are similar. It should be noted that while ~2k low-cost training data are used to achieve the highest out-of-range accuracy (R^2=0.98), good prediction accuracy can still be achieved (R^2=0.96) even with 50 training data points.

V. Conclusions

A novel GAN for device and circuit curve generation is demonstrated. The modifications of the GAN allow the correlation of device/circuit parameters to the generated electrical outputs. It is believed to have learned physics *without* incorporating differential equations in the training because 1) it applies to different problems without using domain expertise, 2) it can be trained with as few as 50 data for with-in-training-range prediction, and 3) it can predict out-of-training-range structures' electrical characteristic well. We expect this can be used in novel device/circuit development in which the training curves can be replaced by the scarce experimental curves.

REFERENCES

[1] Y. S. Bankapalli and H. Y. Wong, "TCAD Augmented Machine Learning for Semiconductor Device Failure Troubleshooting and Reverse Engineering," *2019 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, Udine, Italy, 2019, pp. 1-4, doi: 10.1109/SISPAD.2019.8870467.

[2] C. Teo, K. L. Low, V. Narang and A. V. Thean, "TCAD-Enabled Machine Learning Defect Prediction to Accelerate Advanced Semiconductor Device Failure Analysis," *2019 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, Udine, Italy, 2019, pp. 1-4, doi: 10.1109/SISPAD.2019.8870440.

[3] J. Pan, K. L. Low, J. Ghosh, S. Jayavelu, M. M. Ferdaus, S. Y. Lim, E. Zamburg, Y. Li, B. Tang, X. Wang, J. F. Leong, S. Ramasamy, T. Buonassisi, C.-K. Tham, and A. Thean, "Transfer Learning-Based Artificial Intelligence-Integrated Physical Modeling to Enable Failure Analysis for 3 Nanometer and Smaller Silicon-Based CMOS Transistors," ACS Appl. Nano Mater. 2021, 4, 7, 6903–6915, doi.org/10.1021/acsanm.1c00960.

[4] T. Lu, V. Kanchi, K. Mehta, S. Oza, T. Ho and H. Y. Wong, "Rapid MOSFET Contact Resistance Extraction From Circuit Using SPICE-Augmented Machine Learning Without Feature Extraction," in IEEE Transactions on Electron Devices, vol. 68, no. 12, pp. 6026-6032, Dec. 2021, doi: 10.1109/TED.2021.3123092.

[5] H. Y. Wong, M. Xiao, B. Wang, Y. K. Chiu, X. Yan, J. Ma, K. Sasaki, H. Wang and Y. Zhang, "TCAD-Machine Learning Framework for Device Variation and Operating Temperature Analysis With Experimental Demonstration," in *IEEE Journal of the Electron Devices Society*, vol. 8, pp. 992-1000, 2020, doi: 10.1109/JEDS.2020.3024669.

[6] J. Chen, M. B. Alawieh, Y. Lin, M. Zhang, J. Zhang, Y. Guo and D. Z. Pan, "Powernet: SOI Lateral Power Device Breakdown Prediction With Deep Neural Networks," in *IEEE Access*, vol. 8, pp. 25372-25382, 2020, doi: 10.1109/ACCESS.2020.2970966.

[7] H. Carrillo-Nuñez, N. Dimitrova, A. Asenov and V. Georgiev, "Machine Learning Approach for Predicting the Effect of Statistical Variability in Si Junctionless Nanowire Transistors," in *IEEE Electron Device Letters*, vol. 40, no. 9, pp. 1366-1369, Sept. 2019. doi: 10.1109/LED.2019.2931839.

[8] K. Mehta and H. -Y. Wong, "Prediction of FinFET Current-Voltage and Capacitance-Voltage Curves Using Machine Learning With Autoencoder," in IEEE Electron Device Letters, vol. 42, no. 2, pp. 136-139, Feb. 2021, doi: 10.1109/LED.2020.3045064.

[9] V. Eranki, T. Lu, and H. Y. Wong, "Comparison of Manifold Learning Algorithms for Rapid Circuit Defect Extraction in SPICE-Augmented Machine Learning," 2022 IEEE 19th Annual Workshop on Microelectronics and Electron Devices (WMED), 2022, pp. 1-4, doi: 10.1109/WMED55302.2022.9758032.

[10] K. Mehta, S. S. Raju, M. Xiao, B. Wang, Y. Zhang and H. Y. Wong, "Improvement of TCAD Augmented Machine Learning Using Autoencoder for Semiconductor Variation Identification and Inverse Design," in *IEEE Access*, vol. 8, pp. 143519-143529, 2020, doi: 10.1109/ACCESS.2020.3014470.

[11] Z. Zhang, R. Wang, C. Chen, Q. Huang, Y. Wang, C. Hu, D. Wu, J. Wang and R. Huang, "New-Generation Design-Technology Co-Optimization (DTCO): Machine-Learning Assisted Modeling Framework," 2019 Silicon Nanoelectronics Workshop (SNW), Kyoto, Japan, 2019, pp. 1-2, doi: 10.23919/SNW.2019.8782897.

[12] A. Lu, J. Marshall, Y. Wang, M. Xiao, Y. Zhang, and H. Y. Wong, "Vertical GaN Diode BV Maximization through Rapid TCAD Simulation and ML-enabled Surrogate Model," Accepted to 2022 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), 2022.

[13] S. S. Raju, B. Wang, K. Mehta, M. Xiao, Y. Zhang and H. -Y. Wong, "Application of Noise to Avoid Overfitting in TCAD Augmented Machine Learning," 2020 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), Kobe, Japan, 2020, pp. 351-354, doi: 10.23919/SISPAD49475.2020.9241654.

[14] H. Dhillon, K. Mehta, M. Xiao, B. Wang, Y. Zhang and H. Y. Wong, "TCAD-Augmented Machine Learning With and Without Domain Expertise," in IEEE TED, vol. 68, no. 11, pp. 5498-5503, Nov. 2021, doi: 10.1109/TED.2021.3073378.

[15] Sentaurus™ Device User Guide Version Q-2019.12, Dec. 2019.

[16] M. Arjovsky, S. Chintala and L. Bottou, "Wasserstein generative adversarial networks," in Proceedings of the 34th International Conference on Machine Learning, PMLR 70:214-223, 2017.

[17] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville and Y. Bengio, "Generative adversarial nets," Advances in Neural Information Processing Systems, vol. 27, 2014.

[18] L. de Oliveira, M. Paganini and B. Nachman, "Learning Particle Physics by Example: Location-Aware Generative Adversarial Networks for Physics Synthesis," Comput Softw Big Sci 1, 4 (2017). https://doi.org/10.1007/s41781-017-0004-6

[19] L. Yang, D. Zhang and Em Karniadakis, "Physics-informed generative adversarial networks for stochastic differential equations," SIAM J. Sci. Comput. 42 (2020), pp. A292 -- A317, https://doi.org/10.1137/18M1225409.

[20] M. Uzun, M. U. Demirezen and G. Inalhan, "Physics guided deep learning for data-driven aircraft fuel consumption modeling," Aerospace, vol. 8, no. 2, p. 44, 2021.

[21] X. Jia, J. Zwart, J. Sadler, A. Appling, S. Oliver, S. Markstrom, J. Willard, S. Xu, M. Steinbach, J. Read and V. Kumar, "Physics-guided recurrent graph networks for predicting flow and temperature in river networks," arXiv preprint arXiv:200912575, 2020.

[22] A. Tanaka, A. Tomiya and K. Hashimoto, Deep learning and physics. Springer, 2021.

[23] S. Shah, D. Dey, C. Lovett and A. Kapoor, "Airsim: High-fidelity visual and physical simulation for autonomous vehicles," in Field and Service Robotics, 2018: Springer, pp. 621-635.

[24] A. Daw, R. Q. Thomas, C. C. Carey, J. S. Read, A. P. Appling and A. Karpatne, "Physics-guided architecture (PGA) of neural networks for quantifying uncertainty in lake temperature modeling," in Proceedings of the 2020 SIAM International Conference on Data Mining, 2020: SIAM, pp. 532-540.

[25] Application Note: Three-Dimensional Simulation of 14/16 nm FinFETs With Round Fin Corners and Tapered Fin Shape, Synopsys Inc., 2019.