Artificial Intelligence Driven Optimization of 1200V SiC DMOS

1st David Green
Silvaco Europe Ltd.
St Ives, United Kingdom
david.green@silvaco.com

2nd Thomas Jokinen
Silvaco Inc
Santa Clara, CA, USA
thomas.jokinen@silvaco.com

3rd Stefania Carapezzi
Silvaco France
Montbonnot-Saint-Martin, France
ORCID: 0000-0002-9271-1189

***Abstract*—We present an Artificial Intelligence (AI) flow to implement device optimization to upgrade traditional trial-and-error approach by technology computer aided design (TCAD) simulations. We combine advanced Design of Experiment (DOE) techniques with Machine Learning (ML) methods such as random forest algorithm and neural network (NN) modeling. Careful DOE design to sample the parameters space allows to build TCAD-generated datasets with few tens of elements but, despite that, suitable to build predictive virtual models connecting the relevant input factors to the figure of merits (FoMs) of the device. Finally, state-of-art optimization procedures are applied to this virtual model to retrieve the optimized device input parameters. We demonstrate such AI flow for a typical 1200V SiC DMOS, although the illustrated procedure is technology and target agnostic so can be easily extended.**

***Index Terms*—Machine learning, random forest algorithm, neural network, device optimization, TCAD, SiC DMOS.**

I. Introduction

A critical part of research and development in microelectronic industry consists in developing devices whose design meets target specification, to ensure maximum energy efficiency, optimum behavior in the target circuit application and the maximum number of devices per unit area. TCAD modeling plays a pivotal role in device design. Traditional TCAD approaches are mostly trial-and-error procedures by expert designers, with the time to achieve target specifications that can become prohibitively long. Thus, it is mandatory to develop innovative strategies to boost the time-to-market and to reduce development expenses. In this respect, ML and AI techniques are currently regarded as essential steps to integrate and upgrade TCAD strategies.

In this work, we present an AI device optimization flow that merges state-of-art ML techniques with TCAD simulations (Fig. 1). We test our approach over optimization of simulation of physical processing aligned to typical industrial processes for SiC DMOS rated at 25A, 1200V. We are able to retrieve optimized process parameters under target of realistic specifications of breakdown voltage (BV) of 1800V (derating to 1200V), threshold voltage (V_{th}) of 3V and the lowest drain-source on-resistance (R_{dsOn}) possible. Key aspects of our flow is easiness-to-set-up and the quite limited number (few tens) of TCAD simulations needed for building the datasets to implement ML techniques that yield accurate results. To-date, attempts to couple TCAD and ML require a much more significant number of simulations [1], [2].

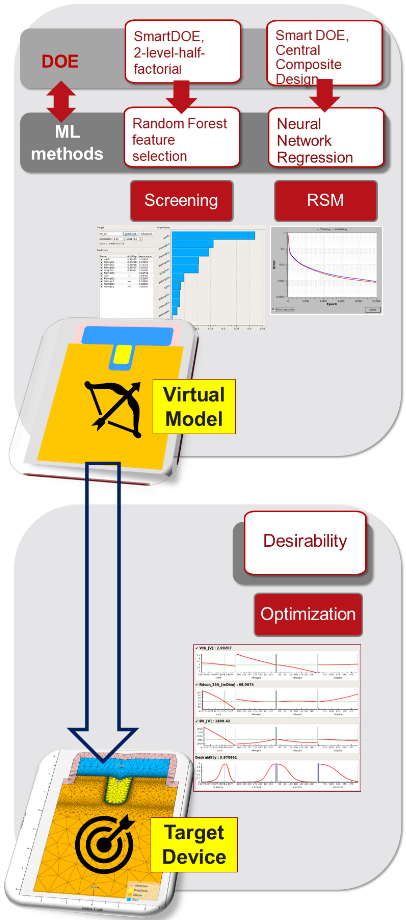

Fig. 1. AI device optimization flow made by the screening step and the response surface method (RSM) to produce the virtual model, and the optimization step to retrieve input factors for the targeted device FoMs. Both the screening and the RSM are based on ML techniques that build on DOE TCAD-generated datasets.

II. TCAD Simulation of 25A, 1200V SiC DMOS

The 2D geometry of a realistic 25A, 1200V SiC DMOS is generated by a process simulation (Fig. 2) aligned to

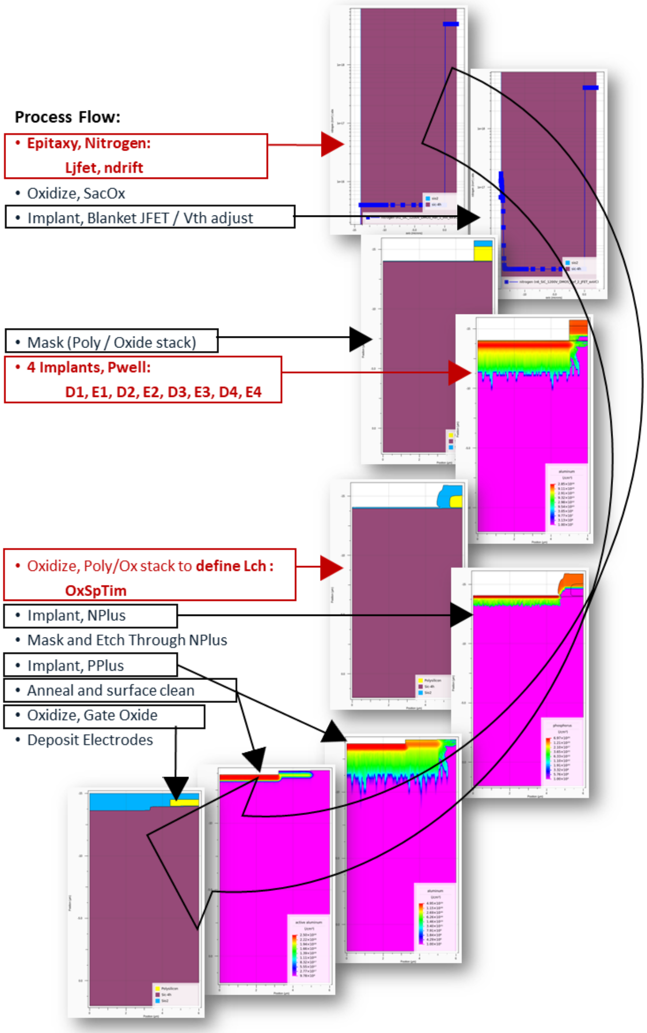

Fig. 2. Schematics of process flow. An epitaxial layer is grown on top of a highly doped substrate as drift region. Then, the JFET implant is made across the entire surface through a thermally-grown, sacrificial oxide. Next, the PWELL opening is defined by a deposited poly oxide mask stack, which is oxidized after the PWELL implanting. The thermal oxide now masks for the NPLUS source implant, thus the distance the oxide grows out laterally will define the gatet length. A heavily doped contact area is obtqined by etching through part of the NPLUS source implant and the PPLUS source implant is made. Finally, the gate oxide is grown and the electrodes are deposited. Highlighted in dark red are the sub-processes whose TCAD parameters have been used as input factors for the screening step.

typical industrial methods, using TCAD Process simulator [3]. In particular, JFET and PWell regions are each formed by four implants. The processed geometry is meshed using a delaunay meshing scheme, with implemented a number of refinements strategies such as a fine mesh around the location of the peak electric field for accurate breakdown voltage simulations. The models for incomplete ionization of impurities, bandgap narrowing, Fermi–Dirac carrier statistics, Shockley–Read–Hall and Trap-Assisted Auger recombinations are considered for device simulation with TCAD Device simulator [4]. Alternative CVT mobility model is used for low-field regime, combined with a parallel-field-dependent mobility model. Finally, an interface charge of 2×10^{12} cm^{-2} and distributions of density of interfacial defect at 4H-SiC/SiO2 interface are assumed. The area parameter *width* is chosen to be typical of a 25A device. The threshold voltage V_{th} is extracted at 5 mA on the drain after tiding it to the gate, and ramping to a compliance of 25 A. R_{dsOn} is extracted at a gate voltage of 15 volts and 25 A on the drain. The breakdown voltage BV is extracted at 0.1 pA while ramping the drain.

III. AI Device Optimization Flow

We apply an AI flow to optimize the design of 25A, 1200V SiC DMOS. This flow consists of three main steps: the screening, the response surface method (RSM), and the optimization (Fig. 1). The combination of screening and RSM generates a virtual model linking input process factors to device FoMs. Then, optimization of FoMs is applied to the virtual model and achieved quite rapidly since each optimization cycle does not require anymore TCAD simulations.

We use ML tools from a statistical analysis software suite [5] to perform each of the above steps. Random forest algorithm, which builds several decision trees based on different random subsets of training TCAD data, is used to screen relevant input factors for a given target FoM. The output of random forest is determined by taking the majority vote of the individual trees, and it yields the ranking of importance among

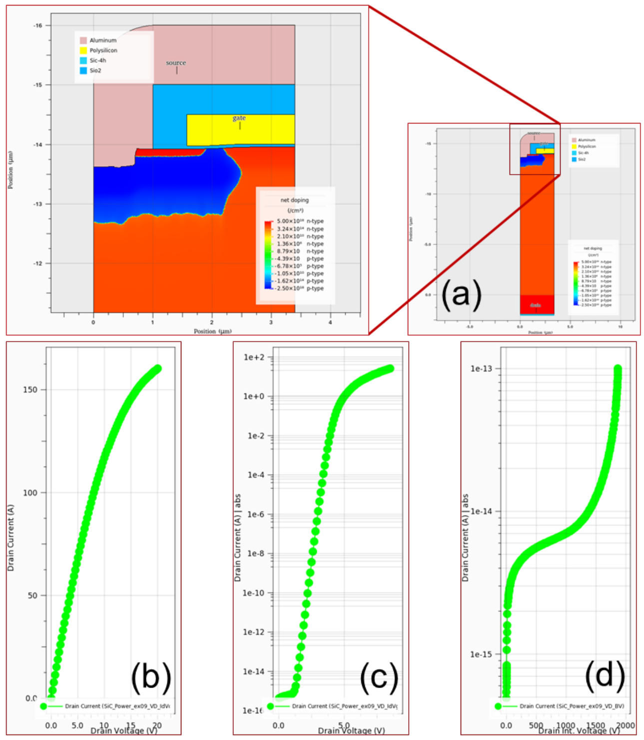

Fig. 3. (a) Device geometry and net doping, with inset showing Source / Gate region. Typical (b) output, (c) transfer and (d) Breakdown Voltage curves.

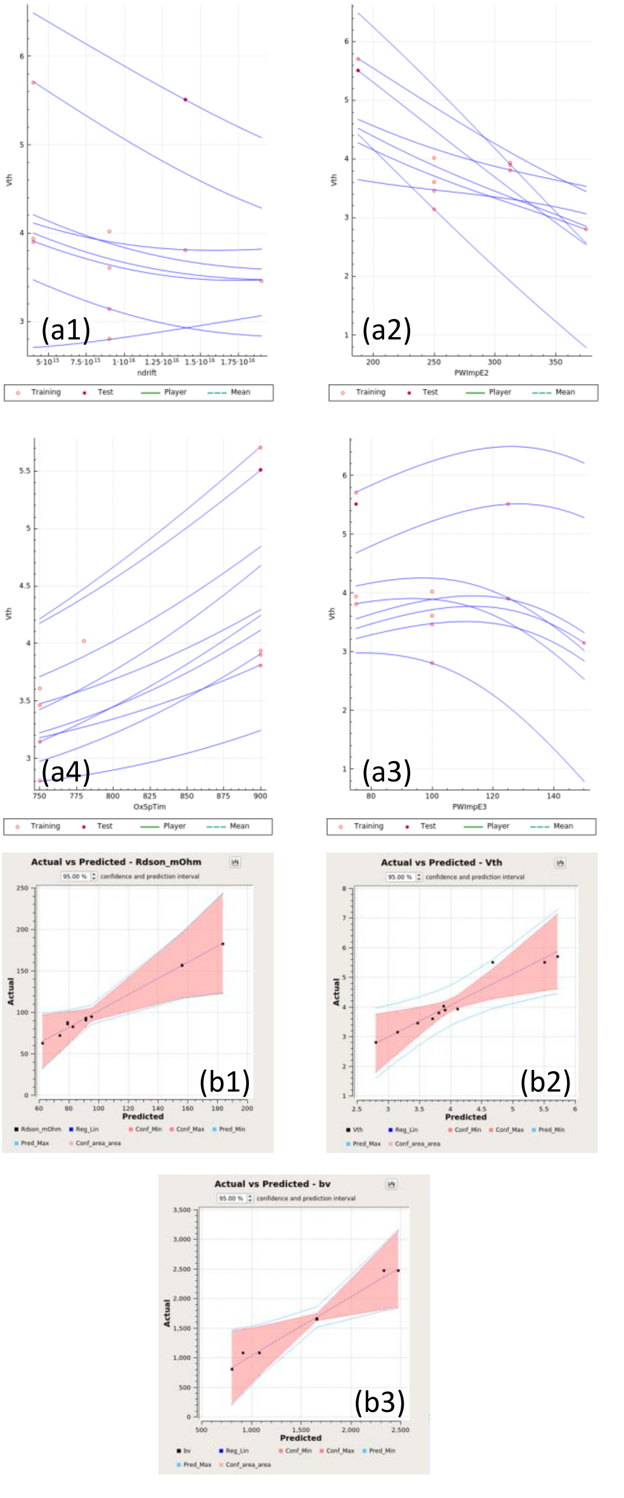

Fig. 4. (a1 - a4) Plots of V_{th} model fit (blue) curves vs. ndrift (a1), PWImpE2 (a2), PWImpE3 (a3) and OxSpTimdata (a4). Red dots are the training dataset and black dots are the test dataset. (b1 - b3) Plots of actual vs. predicted values for R_{dsOn} (b1), V_{th} (b2) and BV (b3). It is apparent the strong correlation achieved for all three modeled FoMs.

input factors. Thus, for any target it is possible to evaluate the most influential input factors by thresholding their importance above a given value.

The next step is to use randomly selected training/validating/testing TCAD data from the sampled simulations for the screened factors to build a ML multi-output regression model of the device FoMs. We use the default subsetting partition 80%:10%:10% $\leftrightarrow$ {Training: Validating: Testing} of all TCAD dataset, with Test dataset not included in the Training dataset. The ML model is created using NN with Feed-Forward architecture, with all nodes fully connected, and activation flows from input layer to output one. Optimization techniques are applied to find the NN parameters to fit the data. We use default NN hyperparameters: one hidden layer, two neurons for each screened input factor, and SoftPlus activation function.

Finally, optimized input factors are retrieved from targeted FoMs by means of Desirability function approach applied to NN regression (NNR) model. The Desirability function approach is widely applied to optimize multiple-output models, for example of industrial processes. It works as follows: each output Y_i is associated to a desirability function $d_i(Y_i)$ that maps the possible values of $Y_i(x)$ to real numbers comprised in the interval $[0, 1]$. $d_i(Y_i) = 0$ is a completely undesirable value of Y_i, while $d_i(Y_i) = 1$ is a completely desirable value. The overall desirability D is calculated as the geometric mean of all desirabilities $\{d_i(Y_i)\}_{i=0,...k}$. The overall desirability is then maximized depending on each output Y_i being maximized, minimized, or assigned to a target value [6].

Both the screening and the RSM steps build on datasets generated by TCAD simulations whose values of input factors are samples of the full TCAD parameters space. A typical challenge is the number of input parameters involved into a realistic simulation, which correlates to the amount of elements of datasets necessary to derive accurate analysis. In this work, we consider doses (D_i) and energies (E_i) of the four PWell implants, JFET width, gate length which is formed through a self-aligned process and NDrift doping, for a total of 11 parameters. We set up the DOE project and run the TCAD simulations [7]. We do not use full factorial (FF) DOE design since such approach is sometimes redundant and scales steeply with the number of inputs. We use advanced DOE design features [7] to build datasets with limited number (few tens) of TCAD simulations while still getting accurate results when applying ML techniques to such datasets.

We screen four input factors that rank higher importance for BV, V_{th} and R_{dsOn}: NDrift doping (ndrift), PWell energies of 2nd (PWImpE2) and 3rd (PWImpE3) implant, and the oxide growth time (OxSpTim). Then a second DOE is set up over screened ndrift, PWImpE2, PWImpE3 and OxSpTim, whose TCAD-generated dataset is used to build the NNR model. We then review the performance of the model with different methods. First we check how it fits against actual data by the combined plot of model fit curves and data used to build the model. Fig. 4 (a1 - a4) show representative plots for V_{th}. The plots provide visual confirmation that the generated model for

BV, V_{th} and R_{dsOn} is not affected by the issues of underfitting or overfitting. We also analyze the correlation between actual vs. predicted values to get a direct measure of the alignment of model's predictions with the actual values, to detect systematic biases or errors. Fig. 4 (b1 - b3) show the strong correlation between actual vs. predicted values for R_{dsOn} (panel b1), V_{th} (panel b2) and BV (panel b3). Such strong correlation is a further validation of the model's predictive power, ruling out any overfitting.

The NNR model can be now used to predict the FoMs for combinations of parameters values for which no TCAD simulated results are available. For example, Fig. 5 show colormaps projected on the various parameters planes of sliced 3D cross-sections of generated 5D V_{th} response surface. They provide a visual assessment of the various trends of V_{th}. Also, Desirability function approach can be applied to the NNR model to retrieve optimized input values for targeted BV, V_{th} and R_{dsOn}. Predicted values of input parameters are showcased in Table I. For final benchmark, we perform TCAD simulations with optimized input parameters and extract the corresponding FoMs. We find that TCAD results (Table II) well compare with targeted values.

TABLE I
OPTIMIZED INPUT FACTORS FROM NNR MODEL

Input Factor	NNR Model
ndrift	7.83E15 cm^{-3}
PWImpE2	373.6 keV
PWImpE3	88.81 keV
OxSpTim	750 s

TABLE II
FOMS DERIVED FROM TCAD SIMULATIONS AFTER OPTIMIZED INPUT FACTORS

FoM	NNR Model	Targets
BV	1847.2 V	1800 V
V_{th}	3.1 V	3 V
R_{dsOn}	91.3 mΩ	lowest

IV. CONCLUSION

In this work we have illustrated an AI flow for device optimization based on TCAD simulations. We have used advanced DOE techniques and ML methods such as random forest algorithm and NN regression modeling. By sampling the parameters space with well-designed DOE we demonstrate that TCAD-generated datasets with just few tens of elements are nonetheless suitable to build predictive virtual models connecting the relevant input factors to the FoMs of the device. Finally, state-of-art optimization procedures are applied to this virtual model to retrieve the optimized device input parameters. We find that predicted FoMs well compare against FoMs from TCAD simulation with optimized input parameters, validating our AI flow for device optimization.

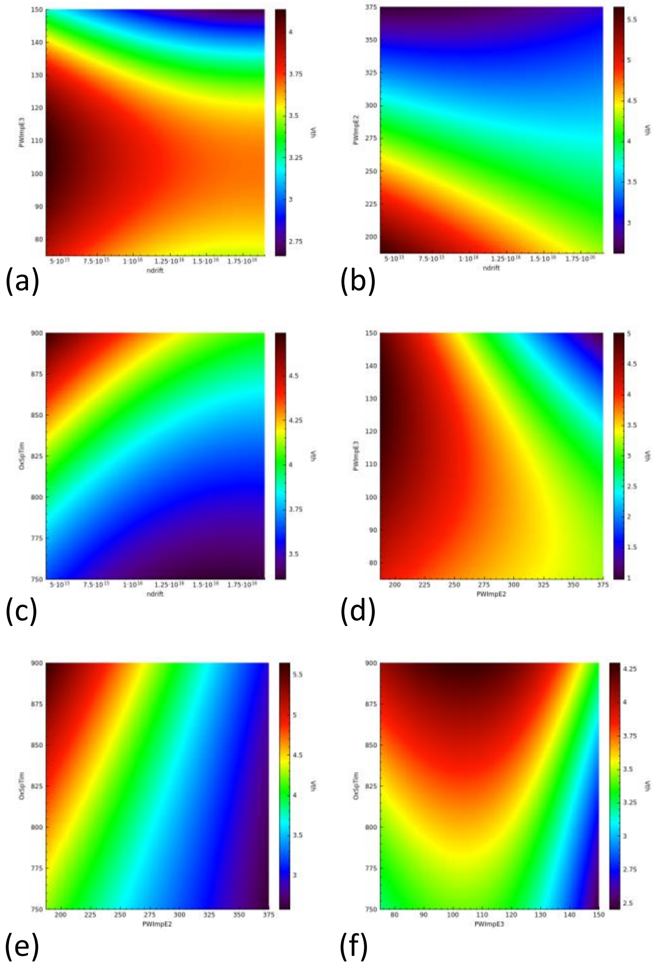

Fig. 5. (a - f) Colormaps of sliced 3D cross-sections of 5D V_{th} response surface projected on the various parameters planes: ndrift - PWImpE2 (a), ndrift - PWImpE3 (b), ndrift - OxSpTim (c), PWImpE2 - PWImpE3 (d), PWImpE2 - OxSpTim (e), OxSpTim - PWImpE3 (f).

We would like to highlight that the proposed AI device optimization flow is technology and target agnostic and then it can be easily extended.

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] H. Y. Wong et al., "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.

[3] "2024 Victory Process User Manual version 8.42.5.R", Silvaco Inc.

[4] "2024 Victory Device User Manual version 1.24.0.R", Silvaco Inc.

[5] "2024 Victory Analytics User Manual version 5.2.3", Silvaco Inc.

[6] "NIST/SEMATECH e-Handbook of Statistical Methods", 2025, subsection "5.5.3.2.2.Multiple responses: The desirability approach". [Online]. Available: https://www.itl.nist.gov/div898/handbook/pri/section5/pri5322.htm

[7] "2024 Victory DOE User Manual version 1.1.8.R", Silvaco Inc.