

Surrogate-Assisted Calibration of a DRIE Process Model

Roman Kostal^{1*}, Manuel Kleinbichler³, Stefano Di Nicola⁴, Lado Filipovic^{1,2}

¹Institute for Microelectronics, TU Wien, 1040 Vienna, Austria

²CDL for Multi-Scale Process Modeling of Semiconductor Devices and Sensors at the Institute for Microelectronics

³KAI Kompetenzzentrum Automobil- und Industrieelektronik GmbH, 9524 Villach, Austria

⁴Infineon Technologies Austria AG, 9500 Villach, Austria

*Email: kostal@iue.tuwien.ac.at

Abstract—Accurate calibration of physics-based process simulation models for deep reactive ion etching (DRIE) requires exploring a high-dimensional parameter space with computationally expensive simulations. We present a profile-driven, multi-domain workflow using a machine-learning surrogate to calibrate an SF₆/C₄F₈ etching model to SEM cross-sections of trenches that we fabricated and characterized for this study across six domains (three cycle counts at two trench widths). Trained on 16,384 Sobol-sampled simulations, the multilayer perceptron (MLP) surrogate reaches test R² of 0.996 for trench depth and 0.966 for Chamfer distance. The surrogate enables optimization via a large-scale Sobol sweep with one billion candidates in about 15 minutes on a single GPU, followed by gradient-based refinement and full-physics validation. The calibrated surrogate is intended for the modeled process window and geometry domain considered in this study.

Index Terms—Process Simulation, DRIE, Surrogate Model, Machine Learning, Open Source Software, ViennaPS

I. INTRODUCTION AND MOTIVATION

Cyclic deep reactive ion etching (DRIE) enables high-aspect-ratio silicon etching with control over sidewall tapering, which is widely used in semiconductor device and microstructure fabrication. Each cycle comprises three steps: C₄F₈ passivation, a high-bias SF₆ breakthrough, and a bias-free SF₆ etch. The resulting trench geometry is notoriously sensitive to many interacting process and model parameters, making calibration to SEM cross-section targets challenging, as it requires many computationally expensive simulations across a high-dimensional parameter space.

To address this challenge, we generate a dataset of 2¹⁴ = 16 384 Sobol-sampled ViennaPS [1] simulations spanning the full parameter space of interest, and use

it to train a machine learning (ML) surrogate capable of predicting the trench cross-section geometry at a fraction of the cost of direct simulation. Surrogate-assisted optimization for expensive simulators is well established [2], and ML-assisted TCAD calibration has recently been explored for device-level models [3]. This surrogate enables efficient multi-domain calibration over the joint process and model parameter space, replacing the expensive repeated simulator evaluations. More importantly, it acts as a fast emulator of the calibration landscape, making the parameter-profile relationship explicit across process settings and geometries. Critically, cycle count and trench width are treated as explicit model inputs, allowing a *single* surrogate to generalize simultaneously across all six experimental domains extracted from the in-house fabricated test structures: 15, 30, and 45 etch cycles at two trench widths (Fig. 1).

II. SIMULATION SETUP

Each step of the DRIE cycle is implemented and simulated in ViennaPS [1], our open-source process simulator which represents the surface as a level set and handles particle transport via Monte Carlo ray tracing. The deposition and etch steps use a `SingleParticleProcess` with surface deposition/removal velocity $v = R\Phi$, where Φ is the normalized ray-traced flux. The breakthrough step uses the full SF₆/C₄F₈ plasma etch model loosely based on the feature-scale SF₆ model of Belen *et al.* [4]. Silicon is removed by chemical etching (fluorine), physical sputtering (ions), and ion-enhanced etching, consistent with the classical plasma-etch mechanism framework [7]. In this model, the pseudo-steady-state fluorine surface coverage

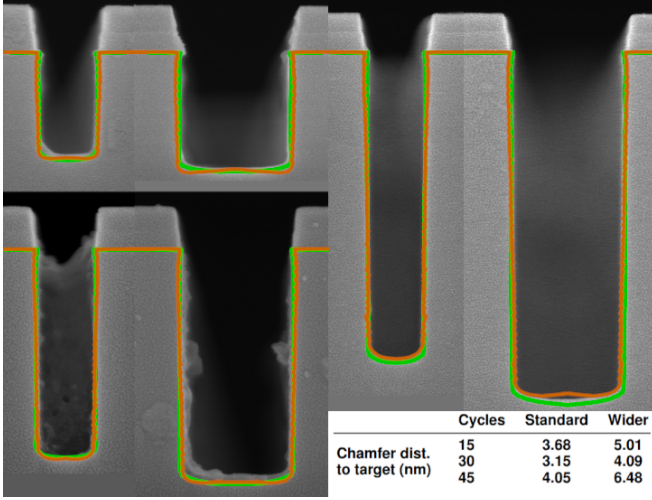


Fig. 1. SEM cross-sections of the trenches overlaid with extracted target profiles (green) and a candidate solution found by surrogate-guided calibration (orange): 15 cycles (top left), 30 cycles (bottom left), and 45 cycles (top right) for standard and wider trench widths. The mini-table (bottom right) shows the Chamfer distances of the solutions to the targets.

is

$$\theta_F = \left[1 + \frac{k\sigma_{\text{Si}} + 2Y_{ie}\Gamma_i}{\beta_F\Gamma_F} \right]^{-1}, \quad (1)$$

and the silicon removal rate is

$$v_{\text{Si}} = \frac{1}{\rho_{\text{Si}}} \left(\frac{k\sigma_{\text{Si}}\theta_F}{4} + Y_{sp}\Gamma_i + Y_{ie}\Gamma_i\theta_F \right), \quad (2)$$

where Γ_i , Γ_F are ion and etchant fluxes, β_F the fluorine sticking coefficient, and $k\sigma_{\text{Si}}$ the chemical etch rate coefficient. Each yield factorizes into energy and impact angle dependence as $Y(E, \theta) = \hat{Y}(E) \cdot f(\theta)$. Ion energies follow $\mathcal{N}(\bar{E}, \sigma_E)$ and directions a power-cosine distribution with exponent n . The C_4F_8 polymer is removed exclusively by sputtering:

$$v_{\text{poly}} = Y_{sp}\Gamma_i / \rho_{\text{poly}}. \quad (3)$$

The details of the models for each of the 3 steps are listed in Table I.

Self-consistent feature charging is too costly to include in every calibration run, since each update would require an electrostatic solve and ion tracing in the local electric field. Its layout-dependent contribution is therefore represented in the DRIE model by the effective ion-angle factor B_{sp} . To check the scale of this approximation, we performed a separate Poisson/ray-tracing calculation for the representative layouts. The dielectric charge is updated from the local ion-electron flux imbalance, and ions are traced through the resulting field [4], [5]. As

TABLE I. ViennaPS models used for the 3-step DRIE process.

	Deposition	Breakthrough	Etch
# Particles	1	2	1
Species	C_4F_8 radical	SFx^+ ion + F rad.	F radical
Rate on Si	$R_{\text{dep}} \Phi$	Eq. (2)	$R_{\text{etch}} \Phi$
Rate on poly	$R_{\text{dep}} \Phi$	Eq. (3)	0
$\hat{Y}_{sp}$ (Si+poly)	—	$A_{sp}(\sqrt{E} - \sqrt{E_{\text{th}}})$	—
f_{sp} on Si	—	$(1 + B_{sp}(1 - \cos^2\theta)) \cos\theta$	—
f_{sp} on poly	—	1	—
Y_{ie} on Si	—	$A_{ie}(\sqrt{E} - \sqrt{E_{\text{th}}})$	—
f_{ie} on Si	—	$\cos\theta \geq 0.5: 1$ else: $\max(3 - 6\theta/\pi, 0)$	—
Radical sticking	β_{dep} on both	β_F on Si β_{poly} on poly	β_{etch} on Si 0 on poly

TABLE II. Variable process parameters and search bounds.

Parameter	Min	Max
Deposition rate R_{dep} (nm/s)	1	30
Deposition sticking β_{dep}	0.01	0.99
Deposition source exponent n_{dep}	1	30
Ion flux Γ_i (nm ⁻² s ⁻¹)	10	400
Etchant flux Γ_F (nm ⁻² s ⁻¹)	100	1000
Fluorine sticking coeff. β_F	0.01	0.99
Polymer sputtering coeff. A_{sp}	0.02	0.2
Ion flux width factor B_{sp}	1.0	1.2
Ion angular exponent n	100	800
Number of cycles	15 / 30 / 45	
Trench width (relative)	standard / wider	
Remaining parameters fixed to TCP chamber-motivated values.		

expected, the strongest field build-up occurs near the mask corners in Fig. 3a and Fig. 3b, due to the different angular distributions of electrons and ions [6]. With the incoming plasma conditions fixed, the calculation predicts a normalized bottom ion flux ratio of 1.1259 for the 300 nm dense trench relative to the 165 nm isolated trench. The remaining difference to the experimentally inferred factor of approximately 1.15 can arise from differences in the impact-angle distribution generated by ion bending in the lateral electric field, as shown by the bottom-impact histograms in Fig. 3c and Fig. 3d. Together with the lateral-field maps, these distributions support using B_{sp} as a compact correction for charging-mediated ion-angle broadening and bottom-flux loading.

III. SURROGATE MODEL

A. Dataset

Using SALib [8], we draw a Sobol sequence of $2^{14} = 16384$ parameter combinations and run the full cyclic process simulation for each sample across all six domains, producing up to 98304 data points. Runs in which no etching was achieved were removed, leaving

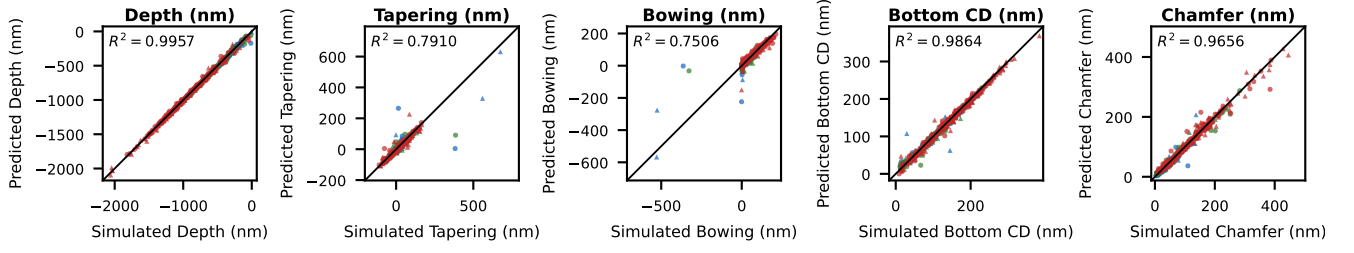


Fig. 2. Parity plots (MLP surrogate, test set) for five predicted targets, color-coded by cycle count (circle = standard, triangle = wider trench). 900 points shown (150 per domain); R^2 computed on the full test set of 17 232 samples.

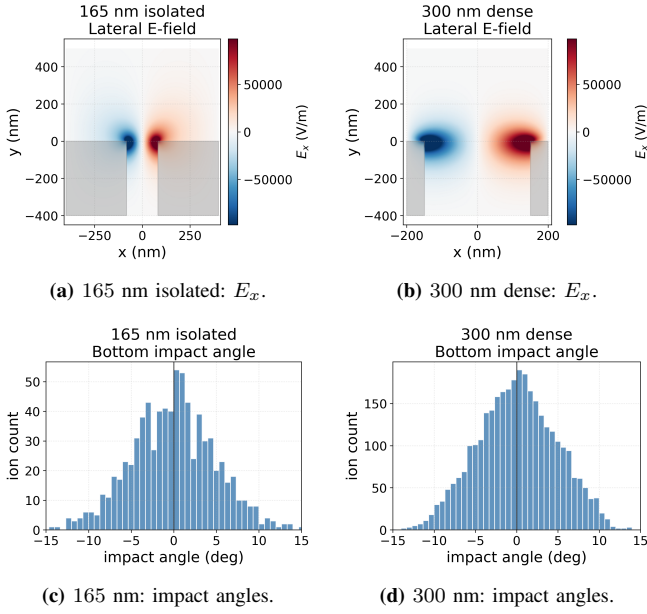


Fig. 3. Auxiliary charging calculation used to parameterize the layout-dependent correction. Top row: lateral electric field E_x ; bottom row: bottom ion impact-angle distributions.

86 158 valid samples. These failures occurred when the deposited polymer layer was too thick relative to the sputtering mechanism, blocking the mask opening entirely. The simulations were completed in approximately two days using four parallel workers across two NVIDIA L40S GPUs. Each sample is described by 11 inputs given in Table II (nine process parameters, the cycle count, and the trench width) and six output targets (trench depth, tapering angle, bowing, top and bottom critical dimensions (CDs), and Chamfer distance to the target profile). Sobol sensitivity analysis [8] on the dataset identifies Γ_i and A_{sp} as the dominant parameters for Chamfer distance, with R_{dep} contributing secondary effects. Feature profiles and critical dimensions (such as etch depth and sidewall tapering) were extracted

from SEM images and compared to simulations using our open-source tool, ViennaFit [9], which employs CD matching and the Chamfer distance as a shape-aware dissimilarity measure [10].

B. Architecture and training

We train a fully-connected PyTorch [11] MLP on a stratified 60/20/20 train/val/test split, with strata defined by the number of cycles and domain width, ensuring each domain is proportionally represented across all splits. Inputs are standardized via a `StandardScaler` fit on the training set only. The network maps the 11-dimensional input to 6 outputs through three hidden layers ($256 \rightarrow 256 \rightarrow 128$ neurons) with BatchNorm and Dropout(0.1) applied after each layer. Training uses the Adam optimizer ($\eta = 10^{-3}$) with a ReduceLROnPlateau scheduler and early stopping (patience 40) to prevent overfitting.

C. Results

Test-set R^2 scores together with parity plots are shown in Fig. 2. The surrogate predicts trench depth, Chamfer distance and bottom CD with high accuracy ($R^2 = 0.996$, 0.966 and 0.986 respectively), which are the quantities most directly relevant to calibration. Tapering and bowing are harder to predict ($R^2 = 0.791$ and 0.751), reflecting the sensitivity of sidewall shape to subtle parameter interactions. Overall, these errors are sufficiently low for fast candidate ranking before final full-physics verification.

IV. SURROGATE-GUIDED CALIBRATION

A single full-physics evaluation across all six domains requires approximately one minute of wall-clock time on an NVIDIA L40S GPU (ray tracing) coupled with 8 cores of an AMD EPYC 9355 CPU (level-set operations). The $2^{14} = 16\,384$ simulations used to train the surrogate represent the total simulation budget; spent instead as a direct search, they would yield

only 16 384 candidate evaluations in a nine-dimensional space, which is far too sparse for reliable optimization. The surrogate converts this fixed investment into the ability to score 10^9 candidates at microseconds each, exploring roughly $60\,000\times$ more of the parameter space at negligible marginal cost, and additionally exposes gradient information that is not directly available from a black-box simulator and would be costly to approximate via finite differences.

Stage 1 — Large-scale Sobol sweep. We draw 10^9 candidate parameter vectors from a scrambled Sobol sequence [8] and score each with the surrogate. The calibration objective sums predicted Chamfer distance and a depth penalty over all six domains:

$$\mathcal{L}(\theta) = \sum_{d=1}^6 \widehat{\text{CH}}_d(\theta) + \lambda \sum_{d=1}^6 |\hat{z}_d(\theta) - z_d^*|, \quad \lambda = 0.1, \quad (4)$$

where $\widehat{\text{CH}}_d$ and $\hat{z}_d$ are the surrogate-predicted Chamfer distance and trench depth for domain d , and z_d^* is the corresponding target depth. The complete sweep finishes in approximately 15 minutes on a single GPU and retains the top- $K = 10$ candidate sets by $\mathcal{L}$. With this wall-clock budget, direct full-physics search would permit only about 15 evaluations.

Stage 2 — Gradient refinement. The surrogate is end-to-end differentiable, so each of the K seeds is refined with Adam ($\eta = 2 \times 10^{-3}$, 5 000 steps), minimizing Eq. (4) while clamping θ to the bounds in Table II. The refined set with the lowest $\mathcal{L}$ is then validated against the full ViennaPS simulator. An example candidate solution is shown in Fig. 1, showing good agreement with the target profiles.

V. CONCLUSION

We have demonstrated a surrogate-assisted, multi-domain, profile-driven calibration workflow for cyclic DRIE across nine parameters and six experimental domains. The MLP surrogate achieves $R^2 = 0.996$ for trench depth and supports a large-scale 10^9 -candidate Sobol sweep (~ 15 min on a single GPU), followed by gradient-based Adam refinement. Beyond acceleration, the surrogate acts as a practical emulator of the calibration landscape, helping to reveal links between process parameters and observed profile evolution across domains. The one-time training cost is amortized across any number of calibration targets. The calibrated surrogate is valid within the modeled process window and geometry domains defined in Table II; extrapolation beyond these bounds is not claimed. Future work will

investigate physics-informed input features to improve sidewall-shape prediction and extend the workflow to broader process conditions.

ACKNOWLEDGMENT

Financial support by the Federal Ministry of Labour and Economy, the National Foundation for Research, Technology and Development, and the Christian Doppler Research Association is gratefully acknowledged.

REFERENCES

- [1] T. Reiter and L. Filipovic, “ViennaPS: A flexible framework for semiconductor process simulation,” *SoftwareX*, vol. 30, Art. no. 102453, 2025.
- [2] A. I. J. Forrester and A. J. Keane, “Recent advances in surrogate-based optimization,” *Progress in Aerospace Sciences*, vol. 45, no. 1–3, pp. 50–79, 2009.
- [3] B.-C. Wang, C.-X. Tang, M.-T. Qiu, W. Chen, T. Wang, J.-Y. Xu and L.-L. Ding, “A machine learning approach to TCAD model calibration for MOSFET,” *Nuclear Science and Techniques*, vol. 34, no. 12, Art. no. 192, 2023.
- [4] R. J. Belen, S. Gomez, M. Kiehlbauch, D. Cooperberg and E. S. Aydil, “Feature-scale model of Si etching in SF_6 plasma and comparison with experiments,” *Journal of Vacuum Science & Technology A*, vol. 23, no. 1, pp. 99–113, 2005.
- [5] G. S. Hwang and K. P. Giapis, “Pattern-dependent charging in plasmas: Electron temperature effects,” *Physical Review Letters*, vol. 79, no. 5, pp. 845–848, 1997.
- [6] Y. Zhai, R. Ge, Z. Hu, J. Li, H. Shao, J. Cheng, L. Filipovic and R. Chen, “Modeling the charging effect of the hardmask and silicon substrate during plasma etching in advanced nodes,” *Journal of Applied Physics*, vol. 137, Art. no. 063302, 2025.
- [7] J. W. Coburn and H. F. Winters, “Plasma etching—A discussion of mechanisms,” *Journal of Vacuum Science and Technology*, vol. 16, no. 2, pp. 391–403, 1979.
- [8] J. Herman and W. Usher, “SALib: An open-source Python library for sensitivity analysis,” *Journal of Open Source Software*, vol. 2, no. 9, Art. no. 97, 2017.
- [9] R. Kostal, T. Reiter and L. Filipovic, “A new module for automated optimization of process TCAD model parameters,” in *Proc. International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, 2025.
- [10] R. Kostal, T. Reiter, S. Di Nicola, M. Kleinbichler, P. Manstetten and L. Filipovic, “Shape-aware dissimilarity measures for process TCAD model calibration,” *J. Comput. Electron.*, vol. 25, no. 4, p. 132, Jun. 2026.
- [11] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai and S. Chintala, “PyTorch: An imperative style, high-performance deep learning library,” in *Advances in Neural Information Processing Systems* 32, 2019.