Achieving Sustainability in the Semiconductor Industry: The Impact of Simulation and AI

, , , , ,
and , *Senior Member,*

***Abstract*—Computational simulation has been used in the semiconductor industry since the 1950s to provide engineers and managers with a faster, more cost-effective method of designing semiconductors. With increased pressure in the semiconductor industry to move towards greener and more sustainable manufacturing, it is crucial to understand the impact of computational simulation and artificial intelligence on environmental sustainability, specifically reducing greenhouse gas (GHG) emissions. This paper quantifies the degree to which various types of simulation used for hardware, process, and device optimization can be adopted for different applications in wafer fabrication equipment research and development, along with the potential reduction in physical experimentation, saving silicon, gases, chemicals, and wafers. With this understanding and an estimation of the equivalent carbon cost impact of the computation itself, analyzed projects demonstrated a significant (>80%) decrease in emissions, primarily driven by the ability to use fewer patterned and blanket wafers whose carbon footprint appears to be orders of magnitude larger than that of used modeling resources. The paper concludes with an attempt to quantify the environmental savings from virtualization across our entire research organization and to illustrate the potential future impact of described activities.**

***Index Terms*—Artificial intelligence, carbon emissions, carbon footprint, carbon reduction, computational modeling, environmental factors manufacturing, low carbon economy, machine learning, net zero, numerical simulation, predictive models, semiconductor materials, simulation, sustainable development.**

I. Introduction

FOR DECADES, semiconductor manufacturing has powered increasingly large segments of the global economy. The drive for additional and less expensive computational power requires continuing advances in chip technologies, with semiconductors becoming more complex and intricate with each new generation [1]. Naturally, the increased sophistication of semiconductor devices cannot be achieved without an analogous growth in resources devoted to research and development (R&D). Each new generation of semiconductor chips calls for additional experimentation at every step of the R&D process, including developing more advanced wafer fabrication equipment (WFE) that can satisfy the tightening specifications required by chip manufacturers.

This technological challenge presents a conundrum for the semiconductor industry since it has committed itself to an environmentally sustainable path. Many major semiconductor companies have set a goal to achieve Net Zero emissions by 2050 – or earlier - to combat climate change and stop the continuous rise of greenhouse gas (GHG) concentration in the atmosphere [2]. This goal will require significantly curbing GHG emissions in each phase of the semiconductor manufacturing ecosystem [3], [4]. In our view, the R&D stage is a particularly promising area for decreasing the carbon footprint of the semiconductor industry despite increasing technical complexity. The reason for our optimism is that while fabricating actual devices invariably requires material resources, physical experimentation during the R&D phase can be at least partially replaced by simulation and virtualization. The end goal of R&D is technological insight, not the production of actual semiconductor devices. Simulation and virtual development can also lead to lower emission manufacturing pathways, further reducing GHG emissions.

Computational modeling and simulation have been indispensable for decades. These methods can save costs, accelerate development, and enable a greater understanding of the physical mechanisms behind complex processes in each manufacturing step (e.g., etching, deposition, lithography). Simulation can also predict overall device performance (e.g., response time, noise level, current-voltage characteristics) [5], [6], [7]. The rise of new computational semiconductor design techniques powered by artificial intelligence (AI) is accelerating the value provided by simulation and data modeling, so much so that it is expected to generate tens of billions of dollars in revenue for semiconductor companies [8]. However, it is important to study not only the financial and labor implications of these advanced design tools, but also their environmental footprint. In this paper, we analyze how several types of simulation used in the R&D environment can reduce the overall environmental impact of semiconductor projects. We will review the impact of simulation on R&D projects such as the development of new etch and deposition processes, prototyping of wafer fabrication equipment, and optimizing wafer integration processes. This work is representative of projects conducted at Lam Research during the design of equipment used to manufacture next-generation logic and memory devices.

Manuscript received 15 February 2024; revised 27 June 2024; accepted 31 July 2024. Date of publication 5 August 2024; date of current version 20 November 2024. *(Corresponding author: .)*

The authors are with the Office of the CTO and the Global Products Group, Lam Research Corporation, Fremont, CA 94538 USA (e-mail: wojtek.osowiecki@lamresearch.com).

This article has supplementary material provided by the authors and color versions of one or more figures available at https://doi.org/10.1109/TSM.2024.3438622.

Digital Object Identifier 10.1109/TSM.2024.3438622

The environmental impact of computer use has recently been studied to quantify GHG emissions and alert society to the costs of fast-growing computing usage. Large language models (LLMs) and the rise of generative AI are widely covered in the media for their concerningly high carbon footprint [9], [10], [11]. We have set out to study the emissions associated with computational modeling in the semiconductor industry, but we also want to understand the reduction in estimated emissions due to the decrease in physical experimentation enabled by computer-based modeling. To our knowledge, the trade-off between the environmental costs of computational modeling and the environmental savings from a reduction in physical wafer tests during semiconductor process development has never been examined. We wish to do so now to understand better how to make the semiconductor industry more environmentally sustainable.

II. Areas of Study

The areas we have selected for environmental study represent a variety of applications where computational modeling is used at Lam Research. Lam Research produces etch and deposition equipment used to manufacture semiconductor devices. One can broadly categorize WFE research into three application categories: 1) hardware development, 2) process and recipe development, and 3) integration flow and chip performance analysis. Hardware development involves testing various chamber subcomponent prototypes during the design of new chambers to achieve the best on-wafer performance. On-wafer performance is measured by the quality of etch and deposition control, manufacturing uniformity, particle control, and other monitored characteristics. Silicon wafer-based testing is used to quantify these performance measures. Simulation is often employed in place of silicon-based testing to limit the number of required prototypes and better understand fundamental equipment mechanisms. At this stage, simulation also provides insight into how the equipment might respond to process requirements or hardware modifications. Plasma environments used in semiconductor manufacturing equipment are particularly complex and sensitive to the shape and dimensions of hardware components, materials used, and surface finishes. Each parameter affects the local plasma density, ions' and radicals' flux and energy. Plasma-based wafer fabrication equipment can be designed more efficiently using computer modeling and simulation to understand the effect of hardware changes on the local plasma environment.

The second category of research, process and recipe development, has always been a significant R&D activity at WFE companies [6], [12]. Chipmakers purchase etch and deposition equipment only if the equipment meets their required specifications. For each new semiconductor device developed by the chipmakers, the right process steps need to be discovered, validated, and optimized to ensure that their device can be manufactured cost-effectively and at a high yield. Engineers test blanket and patterned wafers to prove the viability of a given process integration step. Blanket wafers are used to test etch and deposition rates and their uniformity. Patterned wafers provide much more information about process performance, but their manufacturing is significantly more expensive and often contains sensitive pattern information. Engineers often test small, patterned pieces (called "chips") mounted on a blanket wafer to increase the number of potential experiments. As with hardware prototyping, process simulation can provide crucial mechanistic insight while reducing the number of blanket and patterned wafers required to manufacture a product successfully.

The final research category we will review is integrated process flow and related device (chip) performance. The customers of equipment manufacturers (i.e., chipmakers) have the greatest chip integration expertise. However, WFE companies like Lam Research still need to understand whether their equipment and suggested process recipes (equipment settings) can successfully manufacture the latest logic and memory devices produced by their customers. Without access to an entire semiconductor fabrication facility and the equipment required for every integration step, virtual process modeling is an efficient way for equipment manufacturers to study chip performance.

III. Carbon Footprint Reduction Enabled by Simulation

This paper addresses examples of carbon footprint reduction from all three categories of WFE research described in the previous section. In each of the six projects chosen, based on our completed analysis, the use of computational modeling led to a significant decrease (>84%) in carbon footprint compared to physical experiments alone (Figure 1). We collected the information required for carbon footprint calculations for each project in the same manner. The technical owner provided us with the details of the simulations used (duration, number, and types of computational cores) and the physical experiments (if any). The methodology for calculating carbon footprint for computing is well-established, with the biggest sensitivity coming from the electricity source and the percentage of renewable energy (which can be approximated by geographical location) [13]. For this work, all carbon footprint calculations assumed that the energy sources were based in the Western United States.

Estimating the carbon footprint expended during physical experimentation presents a larger challenge. These calculations are specific to wafer fabrication equipment, and the calculation methodology is still evolving. Fortunately, new tools and resources have recently become available to make these calculations due to increased attention by manufacturers and with the assistance of various academic partnerships, such as the imec.netzero application [1]. To establish emissions savings enabled by simulation, our technical experts also provided estimates of physical experiments that could be replaced by simulation for each analyzed project. While these estimates are hypothetical, we possess high confidence in the data we received due to our collective experience performing hundreds of similar experiments at Lam Research. The reported data are within the average resource range typically devoted to comparable activities when simulation is not used. This fact gives us further confidence that our comparison between

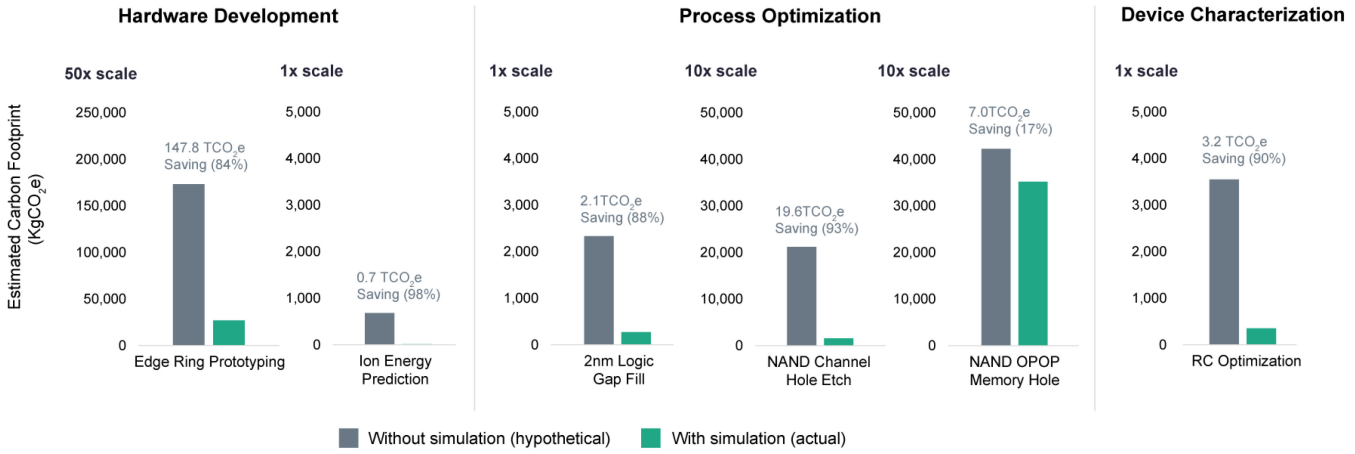

Fig. 1. Actual carbon footprint of successful simulation projects with simulation (green) and a hypothetical estimate without it (grey).

the actual carbon footprint of presented projects and their hypothetical "no-simulation" counterpart is meaningful.

The significant carbon footprint savings shown in Figure 1 accrue from a reduction in patterned and blanket wafer physical experiments and a decrease in rapid prototype hardware subcomponents. Naturally, computational modeling also has an associated environmental footprint of emissions from electricity consumption. However, in each case we studied, the carbon footprint of computational modeling remains far smaller than the equivalent footprint of using physical resources in the laboratory (Figure 2). Simulation represents 100% of 1D plasma simulation's emissions because it entirely eliminates the need for physical experimentation. In all other cases, simulation contributes less than 10% (for 2nm gap fill and edge ring prototyping) or even less than 1% (for RC optimization and NAND HAR etch) of the total carbon footprint.

The relatively small contribution of simulation compared to physical experimentation is due to the high carbon intensity of physical experimentation in the semiconductor industry. Manufacturing a single blanket silicon wafer, the simplest possible material required for on-tool testing (and a predecessor to any film integration), already requires emissions equal to at least 9 $kgCO_2e$ (kilogram CO_2 equivalent) [14]. Other reports have estimated its footprint above 20 $kgCO_2e$ [15], [16], but to remain conservative in the estimations of savings enabled by simulation, we chose to use the value of 9 $kgCO_2e$ per 300mm wafer. The inherent high emission intensity results from the high temperatures (greater than 1000 deg C) required to produce polysilicon and, subsequently, single-crystal silicon [17].

With every additional process step and integration layer, the environmental impact of processing physical silicon wafers grows [14], [18]. Researchers must often use realistic wafers consisting of partial (short-loop) or complete (full-loop) device integration to test new semiconductor processes and manufacture better wafer fabrication equipment. As stated before, engineers attempt to conserve material by cleaving the patterned wafers into several chips. However, the carbon footprint associated with producing one such wafer routinely exceeds 100 $kgCO_2e$ and can approach 2,000 $kgCO_2e$ [1], [19], so the patterned wafers often remain the biggest contributor to the overall emissions of WFE research projects. Any savings in the number of used patterned wafers bring a significant environmental benefit to WFE customers and partners manufacturing these wafers for research purposes.

Lastly, it is worth noting that simulation provides environmental benefits in addition to reducing carbon footprint. Physical experimentation consumes water, nitrogen, and compressed air while releasing pollutant by-products that can be air pollutants in addition to greenhouse gases [20]. Poly and perfluoroalkyl substances (PFAS) are also used in semiconductor manufacturing, but these substances are increasingly being regulated due to their potential biotoxicity [21]. Therefore, reducing the need for physical cleanroom experiments through simulation also contributes to decreasing the consumption of these resources and minimizing the impact of manufacturing on human and environmental health.

IV. Details of Physical Experiments and Simulations

A breakdown of carbon footprint for the six projects across the three already described R&D categories is shown in Figure 2 above. In each case, the carbon footprint of computational modeling is compared to the carbon footprint of physical experimentation (including the cost of blanket and patterned wafers, tool operations, and parts manufacturing). A description of each project, including experimental conditions and simulation costs, is listed below. See Electronic Supplement Table I for further details.

A. Edge Ring Prototyping for Tunable Edge Sheath Solutions

In all plasma etch reactors, the abrupt end of the wafer surface creates inherent electrical discontinuities at the edge region, forming voltage gradients that bend the plasma sheath. This, in turn, changes the direction of the plasma's components (ions and neutrals), which impacts etch results and causes unwanted variability. In the case of 3D NAND devices, the change in plasma properties at the wafer edge can cause tilted etch profiles or prevent features from being completely etched (Fig. 3a). In addition to affecting tilt angle, these edge effects

Fig. 2. Percentage breakdown of carbon footprint by activity for each project without (top) and with (bottom) simulation. Other emissions sources, such as metrology use, and wafer transportation were deemed relatively minor and are not included.

can result in non-uniform critical dimensions (CDs) or changes in local overlay metrics.

This project's objective was to design prototypes of edge rings required to deliver a tunable edge sheath solution for maximizing the yield of dies at the wafer edge. Up to ~10% of the dies are found at the last 20-30 mm of the wafer edge. Lam's tunable edge sheath solution, known as Corvus® technology, provides a novel capability to control extreme wafer edge performance by tuning the plasma sheath at the edge to produce a constant, user-defined etch rate and ion angle. With 3D NAND applications, Corvus technology has demonstrated the ability to minimize plasma sheath drift, preventing detrimental feature tilting at the wafer's edge and eliminating wafer-to-wafer yield variability.

The bulk of the conceptualization and design prototyping for Corvus was achieved through extensive plasma modeling and simulation. Several edge ring design prototype configurations were formulated and simulated in plasma modeling software packages such as HPEM [22] and Vizglow^{TM} [23]. The effect of the edge ring configurations relevant for a tunable sheath was assessed through simulation outputs such as on-wafer energy-dependent ion angular tilt and spread, radial uniformity of plasma density at sheath edge and on-wafer ion fluxes of positively charged species. Correlations between specific features in the edge ring shape, dimension, or layout and the assessed metrics were used to numerically optimize the required edge ring configurations for a range of process conditions. A short-listed set of best-case designs was then rapidly prototyped into actual hardware parts and used for testing in the etch chambers with both blanket and pattern wafers. In addition, by utilizing a design methodology based on the judicious use of plasma modeling and simulation in developing the successful Corvus technology, a five-fold cost reduction and three-fold acceleration of time to solution was realized."

B. 1D Simulation of Plasma for Ion Energy Prediction

A 1D simulation of plasma in a capacitively coupled plasma (CCP) reactor was used to understand the Ion Energy Distribution function (IEDF) at the wafer surface (1D plasma simulation).

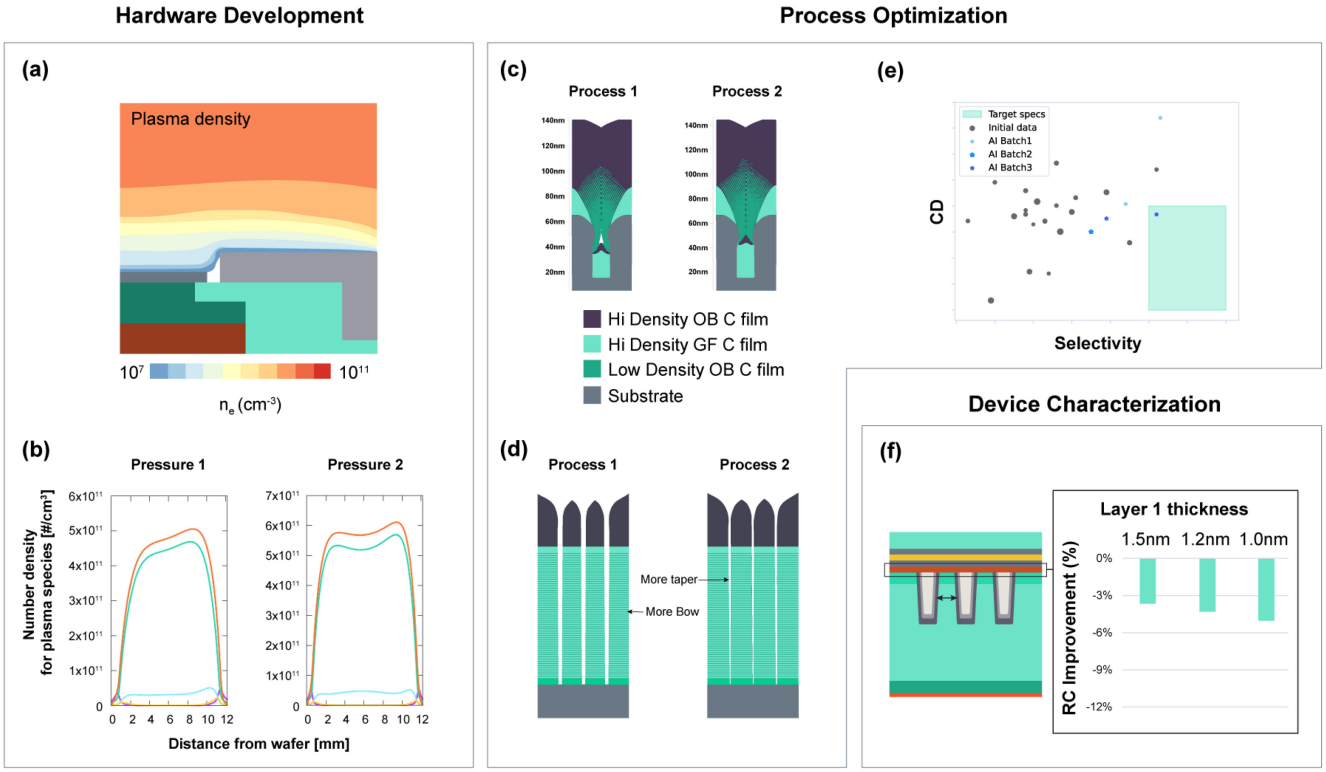

Fig. 3. Computational projects in WFE research used for 1) hardware development: a) effect of chamber edge ring geometry on plasma density and etch performance, b) 1D plasma simulation for understanding the Ion Energy Distribution function (IEDF) of various plasma species at the wafer surface for two different pressure conditions; 2) process development: c) simulation of 2nm logic gap fill deposition, d) simulation of 3D NAND etch, e) AI-guided recipe optimization to meet 3D NAND OPOP targets; 3) device characterization: f) RC optimization of 2nm logic device.

This project aimed to understand the directional behavior of plasma species at a wafer surface, while varying the reactor process parameters such as pressure, temperature, frequency, and supplied power. A one-dimensional plasma simulation provided the plasma species concentration and temperature across the electrode gap, while the ion energy and angle distribution functions were obtained on the wafer surface. The parameters were varied in a multivariate simulation experiment using 24 different conditions to understand this behavior.

The method used in this project was a theoretical simulation and did not use or require any silicon wafers for calibration or validation purposes. It is extremely challenging to test or validate which exact ion etches the surface of a wafer. In this case, 1D simulations explained etch behavior consistent with the physics at those operating conditions [24], [25]. This type of simulation provides a fundamental explanation of etch processes, which an expert engineer could use to test a new etch process recipe. In the absence of 1D simulation capabilities, an expert etch engineer would keep running wafer-based experiments (in a highly non-linear space) and would likely settle for the best-observed performance after multiple experiments. Figure 3b shows the concentration profiles of various species (represented by different colors) at different pressures. From these pictures, it is noted that certain species are generated near the plasma sheath regions and are significant for the fundamental understanding of the etch process on the wafer. Additionally, the effect of pressure on the concentration profile provides guidance regarding optimal recipe conditions. In short, 1D simulations are informed by a general understanding of the process (etch rate vs. temperature, etch depth vs. pressure) and plasma physics (a stronger electric field means higher ion energy). If these basic rules are not violated, some data (ion species concentration, impact energy, etc.) can be used to validate the simulation.

C. Gapfill Optimization (2nm Gap Fill, Process Simulation 1)

This project aimed to develop process conditions and recipe settings that optimized the planarity of a deposited film stack on top of an array of structures. Planarization is used to smooth the surface of a semiconductor wafer and is important in ensuring the electrical characteristics of devices produced on the wafer. This project simulated aspect ratio variations ranging from narrow trench to wide trench structures, with "mesa" structures built in between the trenches. A film stack was then virtually deposited on the structure. This simulation aimed to develop process conditions and recipe settings that ensured conformity of the deposition process with no voids or gaps in the deposited film, moving from an unoptimized Process 1 to the optimized Process 2 (Fig. 3c) [26]. The starting process contained three steps: two deposition steps followed by an etch back step. Using SEMulator3D®, we simulated the deposition and etch processes across various aspect ratio structures. This allowed us to understand the process behavior on these structures, instead of experimentally determining process windows and boundary conditions.

D. Cyclical Etch for High Aspect Ratio Applications (NAND Channel Hole HAR Etch, Process Simulation 2)

This project aimed to develop an understanding of process conditions and recipe settings that optimized the profile of a high aspect ratio etch for a NAND device. The project aimed to optimize and fine-tune plasma etch gas flows to reduce tapering and bowing of the etch profile that leads to poor contact and device performance. By developing a calibrated computational model of two different etch types, we optimized the timing to combine two etch processes and minimize tapering and bowing during the high aspect ratio etch process (Fig. 3d) [26]. Two separate etch recipe conditions were independently calibrated using CDSEM & TEM data. Once calibrated, these two recipe conditions were computationally combined to understand the co-optimization of the two conditions to achieve the desired etch profile.

Achieving the same level of understanding without simulation would require between 233 and 466 hours of processing time on a semiconductor etch system.

E. Dielectric Etch of 3DNAND OPOP Memory Hole (AI-Aided Recipe Development)

The objective of this project was to simultaneously optimize the diameter and the straightness of memory holes and to maximize the etch selectivity of the OPOP stacks over the hard mask. The increasing aspect ratio of the holes made the etch selectivity difficult to improve, and the smaller critical dimension requirement demanded fine-tuning of the process recipe to generate potential performance improvements. The process engineer experimentally reached process improvement limits when a strong process trade-off between selectivity and etch profile straightness was discovered. After running 20 experimental chips to explore the process regime, the team adopted an AI-aided recipe optimization tool developed at Lam Research [6] to determine the next experimental recipe. The AI tool trains a multivariate machine learning model to learn the relationship between recipe settings and metrology results using experimental data obtained from wafer tests. The model provides guidance for experiment design based on important process trends learned from data. While the model was less accurate at the beginning in predicting metrology because of lack of data, it improves its accuracy as more data are collected.

Moreover, it gauges the uncertainty of the model given the dataset and recommends new experimental process conditions considering model uncertainty and optimization objectives. Using a trained model, the tool recommended two recipes to test experimentally in the first round of experimentation (Fig. 3e). The first round of the software recipe predictions resulted in only a marginal improvement in the etch profile straightness and selectivity. The second round also did not show differentiated improvement. After updating the model by including the data from the two previous rounds, a third-round prediction provided a recipe that simultaneously improved all the targeted specs. We found that while the AI model was incapable of predicting metrology outcomes with a nanometer scale accuracy for an arbitrary recipe, it predicted process trends and trade-offs around the baseline recipe sufficiently and led to useful improvements.

F. RC Process Stack Sensitivity Analysis (RC Optimization)

This project aimed to understand the potential resistance/capacitance (RC) improvement that could be obtained by varying the thickness of four layers within a baseline semiconductor structure. After calibrating the structure's RC performance using a single chip, the team built a computational model with pre-specified process steps. They subsequently used the simulation model to investigate RC performance of the virtual device with multiple variations in the thickness of the layers. The image in Figure 3(f) represents a sample of the data generated in this experimental simulation and shows the simulated variation of the resistance/capacitance with different thicknesses of the first layer. Without simulation, a process engineering team must physically manufacture and electrically test these different geometries. The experimental simulation allowed the team to quickly understand the electrical characteristics of these thickness variations without performing extensive silicon-based testing.

V. Modeling Methodology and Analysis Assumptions

A. CO_2 Equivalent Calculation Methodology for Physical Experimentation

We obtained the carbon footprint of patterned wafers from imec.netzero v2.0, the IC chip footprint calculator developed by imec [1], using a "Scope 1+2 – Total Emissions" setting. Emissions connected to operating Lam Research etch and deposition tools were calculated using internal data sources, tool SEMI S23 reports, and the IPCC Tier 2c methodology for process gases [27]. For simplicity and to remain conservative, we assumed that all etch rate measurements consist of the emissions from the etch or deposition tool plus the emissions required to manufacture a single blanket Si wafer (9 $kgCO_2e$), regardless of the wafer top material (which may only add further footprint). For chip tests, we added the tool emissions, one blanket Si wafer carbon footprint, and counted how many patterned wafers are required for the whole project based on the number of available chips per given patterned wafer. Other inputs, such as the carbon footprint of manufacturing hardware prototypes (specifically the manufacturing of etch chamber edge rings), were estimated using publicly available emission factors (see the Electronic Supplement Table II for further details).

We did not include emissions sources such as metrology used at Lam Research for the projects or transportation required to move material between Lam Research and its customers and vendors, as the operational changes from individual projects in emissions from these factors are considerably minor.

B. CO_2 Equivalent Calculation Methodology for Simulation

We reviewed methodologies to calculate the carbon footprint of a computational simulation and have adopted the

methodology developed by Lannelongue et al. [13] The carbon footprint of running a computational simulation is dependent on two factors: 1) the energy required to run the algorithm and 2) the pollutants emitted when producing such energy. We calculated the CO2 equivalent from simulation as follows: Kg CO2 Equivalent = Power Consumption × Carbon Produced Based on the Local Grid. See Electronic Supplement Table III for details on runtime, number of cores, and memory available for each analyzed project.

For all the algorithms used to calculate the CO2 equivalent, we consistently relied on Microsoft Azure, a cloud computing platform, assuming the use of Intel Xeon, E7-8880-v4 CPU Server chips, located in North America, west U.S. region [28]. The carbon intensity of the server is a significant factor in the CO_2 equivalent calculation. There is significant variation from country to country, driven by the method of electricity generation. For example, a kWh from Switzerland generated by hydropower will have a significantly lower CO_2 equivalent than a kWh generated in Australia from coal and gas. For future use and applications of this work, it is, of course, favorable to use servers in a location with a lower carbon footprint. The geographical variation can be seen in the paper by Lannelongue et al. [13].

We did not include the emissions required to manufacture the computers and servers that performed simulations, assessing that the analyzed projects were not sufficiently computationally heavy to drive the demand for additional hardware.

C. Sources of Uncertainty and Their Impact on Our Conclusions

Because most of the carbon footprint comes from patterned wafers (Fig. 2), their estimation carries the largest potential source of error in the final assessment. While we remain confident in the number of wafers used, the exact footprint required to manufacture them remains dependent on several variables outside this paper's scope. To our knowledge, the present best source of information on patterned wafers is the imec.netzero database, but it is still a relatively nascent initiative, and without a doubt, the manufacturing assessments will continue to change. Furthermore, we relied on the available averaged values for each device's generation and narrowed the integration flow only to the level of the relevant process zones (e.g., FEOL, BEOL, array). It is prudent to assume that the produced values are burdened by uncertainty of up to 50%.

Naturally, all the other data sources used in our analysis, such as tool consumption estimates, S23 reports, and the IPCC methodology, are burdened with many assumptions, simplifications, and averages, as well. Due to the existing uncertainty, we aimed to choose a conservative set of assumptions, such as the lowest number for blanket Si wafer carbon footprint found in the literature, and we requested equally conservative estimates of saved physical experiments from the technical owners. It is impossible to guarantee bias elimination entirely, but fortunately, the studied emission sources vary by several orders of magnitude (from gCO_2e for computation to 1,000 $kgCO_2e$ for patterned wafers). Therefore, the results demonstrated in this paper hold even if, despite our best efforts, we significantly (i.e., up to an entire order of magnitude) overestimated the emissions from physical experimentation.

VI. General Comparison Framework of Carbon Footprint of Running Silicon Against Different Level of Compute Capabilities

One of the fundamental questions we aimed to answer in this paper was how one can directly compare the carbon footprint associated with running silicon to running large-scale computational algorithms on servers. In addition to answering this question by addressing the specific projects and applications, we also found it helpful to set out a carbon footprint comparison of different silicon processes with equivalent compute capabilities. This framework is simplified in Figure 4, which shows that a single modern full-loop silicon wafer can be approximated to the carbon footprint of 1,500kg CO_2e. The comparable amount of computational resources required to generate a similar carbon footprint would require a 32 core/32GB system, Azure cloud system in the western U.S. region (the same compute used for our 2nm logic gapfill application) to run for 2,7000 hours or 3.1 years. A similar comparison can be made for a partial-loop old technology wafer with a carbon footprint of 150kg CO_2e. The equivalent computational resources needed to generate a similar carbon footprint, again assuming the 32 Core/32GB system, would have to run for 2,700 hours or 3.7 months. The carbon footprint associated with creating an etch rate test on a blanket wafer is approximately 10kg CO_2e, equivalent to 180 hours of runtime.

VII. Impact on Lam Research and the Wafer Fabrication Equipment Industry

This work demonstrates a range of potential (CO2e) emission reduction from 17% to 98% for the applications we investigated. Our selection of applications, however, contains a natural bias. Engineering teams always want to showcase their best work. To understand how we could scale these examples up to create a fair representation of the potential impact of simulation, we reviewed a range of scenarios on how simulation can be further adopted in WFE research. We set a range for the following two parameters:

- The simulation adoption percentage: The percentage of physical experiments that could be replaced by simulation and AI in each application domain. Our conservative estimate is similar to the present-day adoption rate at Lam Research, and the best-case scenario assumes a 90% overall adoption rate, reflecting our development roadmap.
- The CO_2e savings rate: Based on the projects and applications we have measured, we assumed a set of reasonable best-case scenarios ranging from 90-99%. The exact values for each application can be seen in Electronic

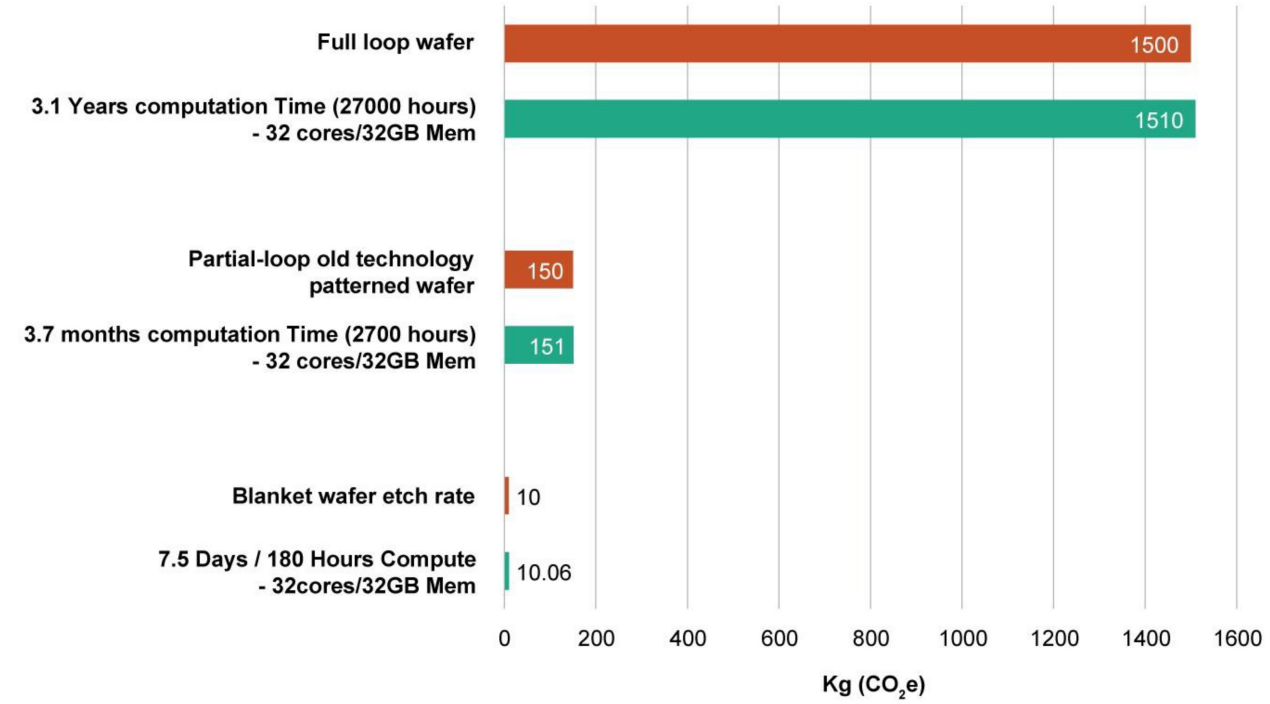

Fig. 4. Comparison of carbon footprints of physical silicon experimentation (red) against different levels of computing resources (green). The figure intends to establish an equivalency framework between physical and virtual experimentation.

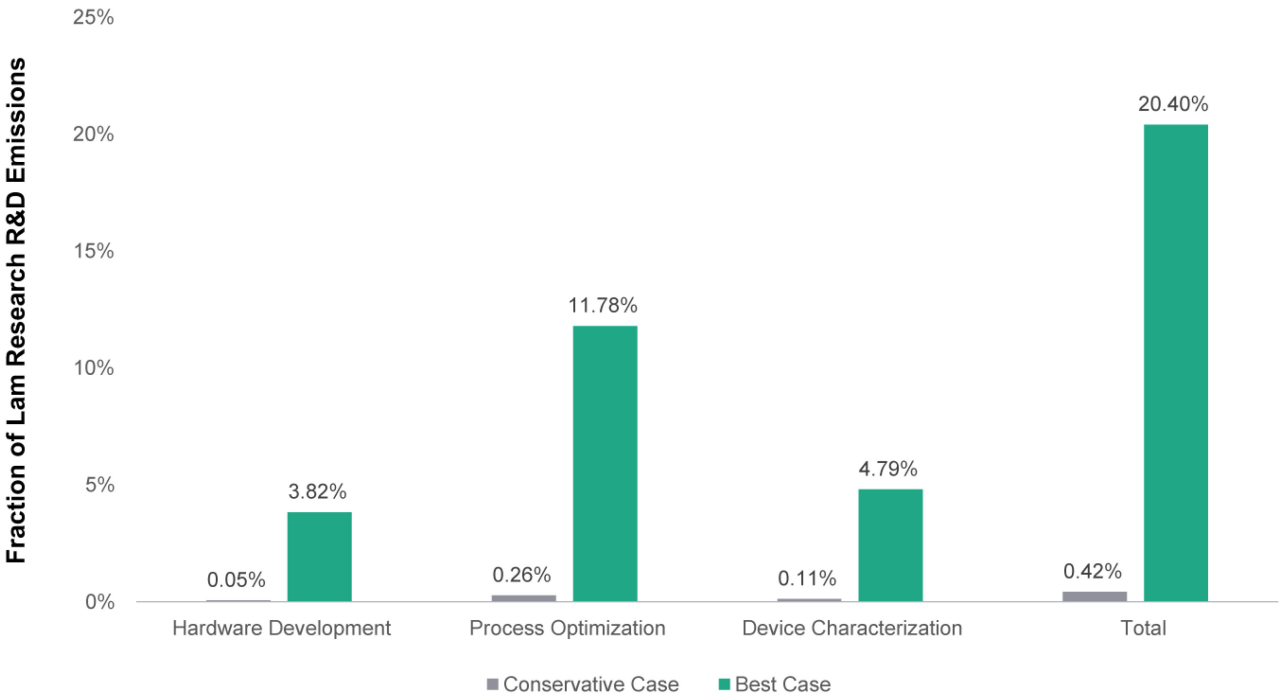

Fig. 5. Conservative (grey) and best-case (green) estimates of annualized CO_2e emission reduction in each application space as a percentage of the total scope 1 and 2 emissions from Lam R&D.

Supplement Table IV. For the conservative scenario, the carbon footprint rate of savings was set to 20%.

Combining these two ranges can estimate an upper and lower bound of CO2e for each application area investigated. The upper and lower bounds are then compared to the total Scope 1 and Scope 2 emissions of Lam R&D to better understand the context of the potential impact of simulation in R&D (Fig. 5). In this context, Scope 1 emissions are emissions direct from Lam sources. Scope 2 emissions are indirect emissions from the generation of purchased energy.

We chose to display the two distinct scenarios in Figure 5 to display the sense of potential associated with simulation, but also to demonstrate the uncertainty we are facing in predicting overall impact. Not all modeling projects are successful which we reflected our calculations, but over time, we observe increasingly higher success rates in our R&D, so we possess reasonable confidence that the future Lam Research trajectory will be closer to the best-case scenario than the conservative one. As new types of simulations are invented and implemented by our researchers, it will be possible to exceed the 20% fraction of Lam R&D emissions and achieve even larger savings in physical experimentation.

The magnitude of potential emissions changes primarily scales with the opportunity for limiting wafer processing, as high-purity silicon is by far the most carbon-intensive material of the ones studied in this work (Table II). Therefore, hardware development and device characterization appear to offer relatively less impact than process optimization, based on the current operations in our company.

Nevertheless, while exploring the best-case scenario for simulation adoption and the savings associated with minimizing physical experiments, it is essential to remember what trade-offs drive the decision-making in our industry. In almost all cases, simulation will be cheaper and carry a smaller carbon footprint than the comparable tests involving silicon wafers. However, the researchers in charge of solving a hardware or process problem must include other vital considerations such as simulation accuracy and time to solution. If the results from computation are not trustworthy enough to drive subsequent development or it takes significantly longer to establish the right modeling setup in comparison to testing wafers in the lab, we will not be able to see an increase in simulation adoption. With the growing adoption of virtualization in WFE R&D, the remaining problems are continuously harder to resolve and may bring smaller associated savings. Fortunately, the quality and capacity of simulation techniques continuously improve, as we cover in the next section, leaving us optimistic that the presented best-case scenario is not out of reach in the upcoming decade.

VIII. Proposed Best Practices and Future Trends

Based on the findings of our analysis, we want to propose a few practical points to WFE researchers interested in lowering the environmental footprint of their work:

- Simulation appears to be almost universally less resource-intensive than physical experimentation, so we recommend that researchers look for opportunities to minimize experiments and attempt to solve their problems using computation methods.
- At the same time, WFE R&D is often conducted on tight deadlines, so modeling experts should collaborate closely with lab researchers to ensure that their virtualizations are accurate and fast enough to avoid the need for running wafer tests in the cleanroom.
- We recommend prioritizing techniques that deliver the best result, regardless of their relative computational intensity, as minimizing the overall carbon footprint will mostly depend on omitting physical experiments.
- Full-loop leading-edge patterned wafers are the most carbon-intensive, so it is best to work with small chips instead of whole wafers wherever possible.

Fundamentally, we believe that the capabilities of experimental simulation are still in their infancy relative to where they will eventually evolve. For the past 70 years, there has always been a need for the next level of accuracy in building effective simulation capability that can describe sub-atomic semiconductor processes and hardware interactions. We do not expect that to change in the near future, and this tradeoff between required complexity and computational speed and accuracy will need to be addressed over the coming years by increased research and development spending in artificial intelligence, machine learning, and data science. These algorithmic advances, coupled with the increased availability of high-power computing architectures, will accelerate the timeframe required for computational modeling and semiconductor process optimization. There is already a clear commercial advantage for companies that leverage these technologies to expedite their R&D. We see these commercial advantages expanding from R&D into the manufacturing domain as the accuracy and speed of simulation improve over time. While manufacturing physical chips will always have the associated carbon footprint, simulation can be critical in yield and efficiency improvement and optimizing for the most sustainable production methods. The significant finding in this paper is that organizations can not only leverage simulation across various semiconductor applications for commercial benefits but can also reap the added benefits of improved sustainability and reduced carbon emissions.

With the growing number of simulation projects, it is reasonable to expect a rise in demand for compute resources, resulting in additional carbon emissions associated with manufacturing that hardware. Our work has only covered the impact on operations involving computing and physical experimentation. Therefore, a valuable extension would model the capital changes driven by the growing demand for computational hardware on the one hand and the opportunity to limit the number of R&D tools and shrink cleanroom space on the other. It is worth noting that cleanroom air maintenance requires approximately the same energy consumption as operating the tools inside that space [29]. The semiconductor equipment is very specialized and resource-intensive while the hardware required for simulation is relatively standard, so the growing replacement of physical R&D tools by virtualization should likely further limit carbon emissions.

IX. Conclusion

We analyzed the impact of simulation on the carbon footprint of wafer fab equipment R&D using a variety of representative use cases: hardware prototyping, process optimization, and device characterization. In all cases, the implementation of modeling techniques led to a decrease in emission intensity due to the reduction in physical experimentation. As shown in Figs. 2 and 4, emissions associated with

computing are relatively minor compared to the environmental cost of producing semiconductor-grade Si wafers and patterning them into short-loop and full-loop structures. While the exact quantification of the impact remains highly dependent on the individual application, we remain confident that the increase in simulation and modeling techniques in R&D will positively change our industry and contribute to the transition towards sustainable operations.

We hope to see further studies detailing the environmental considerations of semiconductor development, not only in the context of GHG emissions but also in water consumption, air pollution, PFAS, and other topics. While our paper has focused on the operational emissions savings from simulation over physical experimentation, the associated emissions with the production of virtualization hardware will likely be smaller than that required to manufacture semiconductor test equipment and cleanroom spaces. We expect similar dynamics to occur in manufacturing, where physical resources will continue to play a vital role as the industry makes physical goods. There are many opportunities to improve yield, decrease the need for metrology, and simplify integration flows with simulation.

The extent to which simulation methods successfully transform the semiconductor industry will depend on their accuracy and the time to solution versus traditional physical experimentation. Therefore, continuing research to improve and expand simulation capabilities while considering environmental implications is critical. Our industry continuously expands and adapts to pursue increasingly more challenging technological objectives. We are used to balancing complex trade-offs such as cost, performance, and speed. We hope to see the environmental impact become a fixed consideration for all upcoming transitions, as, without a doubt, simulation techniques have a lot to offer for our industry.

ACKNOWLEDGMENT

We thank S. Covell, J. Lapidas, and L. Carver for helpful conversations and comments on the manuscript; S. Ramadurgam, D. Brown, J. Carrel, and H. Haruff for the information on edge ring manufacturing; J. Marsalla for input on Lam Research emissions; Y. Chuang for process information; D. Belanger for the graphics; A. Parker for editing; J. Ervin, J. Huang, Y. D. Chen, V. Rai, J. Lucas, V. Vijayakumar, H. Meena, and A. Faucet for the valuable input on process modeling and simulation applications; and Y. W. Lin for input on compute infrastructure.

REFERENCES

[1] G. Bardon et al., "DTCO including sustainability: Power-performance-area-cost-environmental score (PPACE) analysis for logic technologies," in *Int. Electron Devices Meeting (IEDM) Tech. Dig.*, 2020, pp. 41.4.1–41.4.4, doi: 10.1109/IEDM13553.2020.9372004.

[2] S. Fankhauser et al., "The meaning of net zero and how to get it right," *Nat. Cli Change*, vol. 12, no. 1, pp. 15–21, 2022, doi: 10.1038/s41558-021-01245-w.

[3] Bhat, Vervoorn, and C. Jones, "Transparency, ambition, and collaboration: Advancing the climate agenda of the semiconductor value Chain." 2003. Accessed: Jan. 22, 2024. [Online]. Available: https://discover.semi.org/transparency-ambition-and-collaboration-white-paper-download-registration.html

[4] S. Raoux, "Fluorinated greenhouse gas and net-zero emissions from the electronics industry: the proof is in the pudding," *Carbon Manag.*, vol. 14, no. 1, 2023, Art. no. 2179941, doi: 10.1080/17583004.2023.2179941.

[5] D. C. Cole et al., "The use of simulation in semiconductor technology development," *Solid State Electron*, vol. 33, no. 6, pp. 591–623, Jun. 1990, doi: 10.1016/0038-1101(90)90173-C.

[6] K. J. Kanarik et al., "Human–machine collaboration for improving semiconductor process development," *Nature*, vol. 616, no. 7958, pp. 707–711, Apr. 2023, doi: 10.1038/s41586-023-05773-7.

[7] J. A. Paulson and C. Tsay, "Bayesian optimization as a flexible and efficient design framework for sustainable process systems," 2024, *arXiv:2401.16373*.

[8] S. Göke, K. Staight, and R. Vrijen, (McKinsey Co., New York, NY, USA). *Scaling AI in the Sector that Enables it: Lessons for Semiconductor-Device Makers*. (2021). Accessed: Jan. 8, 2024. [Online]. Available: https://www.mckinsey.com/industries/semiconductors/our-insights/scaling-ai-in-the-sector-that-enables-it-lessons-for-semiconductor-device-makers#/

[9] N. Sundberg, "Tackling AI's climate change problem," *MIT Sloan Manag. Rev.*, vol. 65, no. 2, pp. 38–41, 2024.

[10] A. A. Chien, L. Lin, H. Nguyen, V. Rao, T. Sharma, and R. Wijayawardana, "Reducing the carbon impact of generative AI inference (today and in 2035)," in *Proc. 2nd Workshop Sustain. Comput. Syst.*, 2023, pp. 1–7.

[11] A. S. George, A. S. H. George, and A. S. G. Martin, "The environmental impact of AI: A case study of water consumption by chat GPT," *Partners Univ. Int.l Innov. J.*, vol. 1, no. 2, pp. 97–104, Apr. 2023, doi: 10.5281/zenodo.7855594.

[12] K. J. Kanarik, "Inside the mysterious world of plasma: A process engineer's perspective," *J. Vacuum Sci. Technol. A, Vacuum, Surfaces, Films*, vol. 38, no. 3, May 2020, Art. no. 31004, doi: 10.1116/1.5141863.

[13] L. Lannelongue, J. Grealey, and Inouye, "Green algorithms: Quantifying the carbon footprint of computation," *Adv. Sci.*, vol. 8, no. 12, Jun. 2021, Art. no. 2100707, doi: 10.1002/advs.202100707.

[14] Schmidt, H. Hottenroth, Schottler, G. Fetzer, and B. Schlüter, "Life cycle assessment of silicon wafer processing for microelectronic chips and solar cells," *Int. J. Life Cycle Assess.*, vol. 17, no. 2, pp. 126–144, Feb. 2012, doi: 10.1007/s11367-011-0351-1.

[15] (SUMCO Corp., Tokyo, Japan). *Sumco Sustainability Report 2023*. (2023). Accessed: Jan. 1, 2024. [Online]. Available: https://www.sumcosi.com/english/csr/pdf/csr_rep23en.pdf

[16] J.-H. Mohr, G. Tembey, K. Breidenbach, N. Sah, J. Jeschke, and T. Harder, (Boston Consul. Group Manag. Consul. Co., Boston, MA, USA). *For Chip Makers, the Decarbonization Challenge Lies Upstrea* Accessed: Jan. 19, 2024. [Online]. Available: https://www.bcg.com/publications/2023/why-chip-makers-need-to-focus-on-the-upcoming-decarbonization-challenges

[17] G. Fisher, R. Seacrist, and R. W. Standley, "Silicon crystal growth and wafer technologies," *Proc. IEEE*, vol. 100, pp. 1454–1474, May 2012, doi: 10.1109/JPROC.2012.2189786.

[18] S.-C. Hu, A. Shiue, H.-C. Chuang, and T. Xu, "Life cycle assessment of high-technology buildings: Energy consumption and associated environmental impacts of wafer fabrication plants," *Energy Build*, vol. 56, pp. 126–133, Jan. 2013, doi: 10.1016/j.enbuild.2012.09.023.

[19] T. Pirson, T. P. Delhaye, A. G. Pip, G. Le Brun, J.-P. Raskin, and D. Bol, "The environmental footprint of IC production: Review, analysis, and lessons from historical trends," *IEEE Trans. Semicond. Manuf.*, vol. 36, no. 1, pp. 56–67, Feb. 2023, doi: 10.1109/TS 2022.3228311.

[20] A. Villard, A. Lelah, and D. Brissaud, "Drawing a chip environmental profile: Environmental indicators for the semiconductor industry," *J. Clean. Prod.*, vol. 86, pp. 98–109, Jan. 2015, doi: 10.1016/j.jclepro.2014.08.061.

[21] R. Sharma, S. Shelke, B. Kashani, G. Morose, C. Christuk, and R. Nagarajan, "Safer and effective alternatives to perfluoroalkyl-based surfactants in etching solutions for the semiconductor industry," *J. Clean. Prod.*, vol. 415, Aug. 2023, Art. no. 137879, doi: 10.1016/j.jclepro.2023.137879.

[22] J. Kushner, "Hybrid modelling of low temperature plasmas for fundamental investigations and equipment design," *J. Phys. D, Appl. Phys.*, vol. 42, no. 19, 2009, Art. no. 194013, doi: 10.1088/0022-3727/42/19/194013.

[23] (Lam Res. Corp., Fremont, CA, USA). *OverViz—Plasma Modeling*. Accessed: Feb. 4, 2024. [Online]. Available: https://www.lamresearch.com/products/overviz/

[24] R. R. Upadhyay, K. Suzuki, L. L. Raja, P. L. G. Ventzek, and , "Experimentally validated computations of simultaneous ion and fast neutral energy and angular distributions in a capacitively coupled plasma reactor," *J. Phys. D, Appl. Phys.*, vol. 53, no. 43, Aug. 2020, Art. no. 435209, doi: 10.1088/1361-6463/aba068.

[25] D. Levko, R. R. Upadhyay, L. L. Raja, , and P. Ventzek, "Influence of electron energy distribution on fluid models of a low-pressure inductively coupled plasma discharge," *Phys. Plasmas*, vol. 29, no. 4, 2022, Art. no. 43510,

[26] (Lam Res. Corp., Fremont, CA, USA). *SEMulator3D: Semiconductor Process Modeling*. Accessed: Jan. 29, 2024. [Online]. Available: https://www.lamresearch.com/products/semulator3d/

[27] L. Beu, S. Raoux, and Y. C. Chang, "Refinement to the 2006 IPCC guidelines for national greenhouse gas inventories, volume 3: Industrial processes and product use, Chapter 6: electronics industry emissions." Accessed: Jan. 14, 2024. [Online]. Available: https://www.ipcc-nggip.iges.or.jp/public/2019rf/pdf/3_Volume3/19R_V3_Ch06_Electronics.pdf

[28] Russinovich, (Microsoft, Redmond, WA, USA). *Azure Confidential Computing on 4th Gen Intel Xeon Scalable Processors with Intel TDX*. Accessed: Feb. 6, 2024. [Online]. Available: https://azure.microsoft.com/en-us/blog/azure-confidential-computing-on-4th-gen-intel-xeon-scalable-processors-with-intel-tdx/

[29] W. Zhao, H. Li, and S. Wang, "Energy performance and energy conservation technologies for high-tech cleanrooms: State of the art and future perspectives," *Renew. Sustain. Energy Rev.*, vol. 183, Sep. 2023, Art. no. 113532, doi: https://doi.org/10.1016/j.rser.2023.113532.