DATA BRIEF

Direct Emissions in Semiconductor Manufacturing Are Increasing Again – What Is Behind the Shift?

Julia Christina Hess, Maria Nowicka

January 20, 2026

Tech analysis and policy ideas for Europe

interface

Table of Contents

Abstract: After a sharp decline from 2021 to 2023, direct emissions from chip production are rising again. Could artificial intelligence be driving the increase? This brief explores one likely factor: surging demand for high-bandwidth memory (HBM). Vertically stacking more memory layers requires more frequent etching and cleaning, which results in higher fluorinated gas use. We unpack the technology behind HBM, the market forces shaping demand, and the varying emission trends among leading memory chip manufacturers.

Alongside this data brief, we have published an interactive microsite that visualises trends in semiconductor manufacturing emissions and energy consumption. It supports transparency by making complex sustainability data easier to explore and understand.

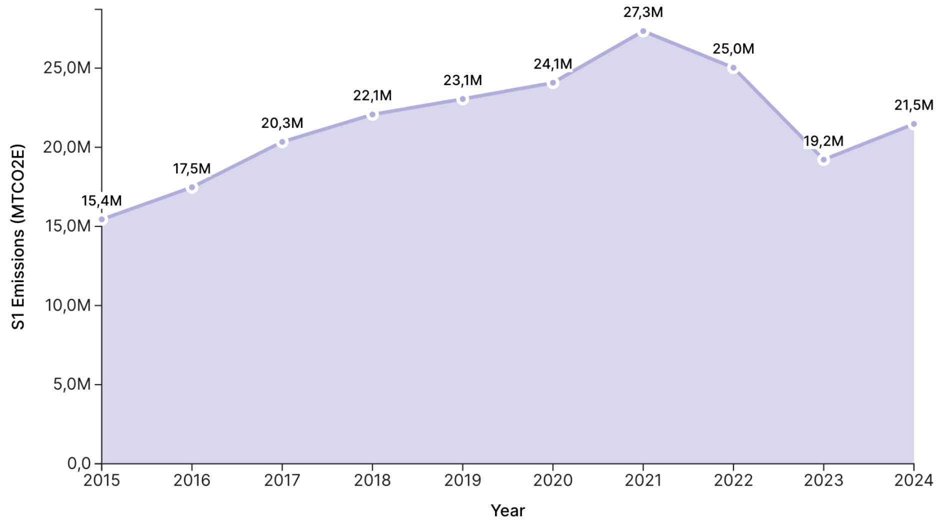
The microsite works alongside the 2025 Semiconductor Emissions Explorer study, which provides deeper analysis of reporting practices and data gaps across the sector.

Intro

This January, we launched a new microsite tracking emissions from chip production. Each year, we plan to update the site with the latest corporate social responsibility (CSR) data from semiconductor manufacturers, alongside a new annual data brief that examines emerging trends and probes potential explanations for notable or unexpected developments. This year, one shift stood out in particular: direct (scope 1) emissions reversed course between 2023 and 2024. In this brief, we explore whether the rapid growth in demand for artificial intelligence (AI) could be part of the story – and whether this change might signal the beginning of a longer-term shift.

Everyone is talking about the role graphic processing units (GPUs) play in powering AI – but that's only part of the story. Therefore, we are zooming into AI hardware beyond GPUs. Behind every advanced AI processor is high-bandwidth memory (HBM), the component that delivers the speed and capacity needed to keep data flowing into GPUs fast enough. HBM is a critical but often overlooked component of today's AI stack – and it may help explain some of the changes we are seeing in emissions.

Aggregated direct (scope 1) emissions from semiconductor manufacturing (2015-2024)



For a complete presentation of this graph, please see the online version of this publication.
https://www.interface-eu.org/publications/semiconductor-emissions-data-2026

First, let's take a step back and recap what direct (scope 1) emissions are. In chip production, these emissions mostly come from chemicals, especially fluorinated gases and wet chemicals, which make up 80%–90% of direct emissions. Short-term reductions usually come from abatement equipment, while changes to manufacturing processes and lower-GWP[1] alternatives take more time to implement.

We analysed the direct emissions of 29 chip manufacturers by using data from their CSR reports. The chart illustrates that emissions nearly doubled from 15.4 MMTCO2E[2] in 2015 to 27.3 MMTCO2E in 2021, followed by a steep decline of 30%, reaching 19.2 MMTCO2E in 2023. But the 2024 data point to a potential trend reversal, with emissions ticking up 7% from the previous year. We want to dig into what is driving this latest shift.

Our hypothesis is that the high demand for memory chips is a key driver of rising

1 Global warming potential (GWP) is a measure of the climate impact of a greenhouse gas relative to one tonne of carbon dioxide (CO_2) over a specified time horizon, typically 100 years, with higher values indicating greater warming impact.

2 CSR reports typically measure GHG emissions in metric tons of carbon dioxide equivalents (MTCO2E). We are displaying the numbers in million metric tons of carbon dioxide equivalents (MMTCO2E). For more information on how emissions are measured according to the GHG Protocol, please refer to our 2025 publication: https://www.interface-eu.org/publications/semiconductor-emission-explorer.

direct emissions. To understand the potential connection, let's take a quick look at the technology behind AI – memory. Any digital device we use today relies on two main types of memory: dynamic random-access memory (DRAM), which provides short-term memory, and NAND flash, which provides long-term storage.[3] DRAM itself can be divided into DDR (double data rate) and HBM, which is increasingly important for AI applications. Unlike DDR (which is placed next to the central processing unit), HBM is placed directly next to or on top of the GPU, which allows for extremely fast data access. To deliver the required speed and capacity, DRAM is vertically stacked, which increases both memory density and bandwidth. This "cube" architecture enables the rapid transfer of very large amounts of data per second between the GPU and memory – one of the key requirements for modern AI workloads. Without HBM, it is a bit like trying to run a cutting-edge app on an older smartphone – the memory just cannot keep up with the software needs.

Growing demand for memory chips

The memory chip market is known for its boom-and-bust cycles. Since memory chips are treated as commodities, manufacturers compete fiercely on price, which makes the market especially volatile. But due to the crucial role of memory chips in AI innovation, memory manufacturers have seen a significant increase in revenue since 2024, particularly in the DRAM segment – and in HBM as a sub-category of DRAM, which we are focusing on. This surge can be linked to the strong pull dynamics created by the global race to build out GPU clusters. The days where memory chip manufacturers were struggling to get enough orders in are clearly over. While training large AI models especially during the reinforcement learning stage already requires substantial amounts of high-bandwidth DRAM, the deployment of AI, known as inference, demands even larger volumes of HBM. With more users sending ever-larger prompts to increasingly complex models, memory requires growing data storage and faster data movement at scale. According to Micron, producing one bit of HBM memory requires forgoing the production of three bits of more conventional memory for other devices. This drives a massive capacity allocation away from more traditional memory commodities to HBM for AI accelerators.

Recent advancements in AI hardware therefore focus heavily on expanding memory capacity to support both increasing model sizes and growing context lengths. A good example is NVIDIA's Blackwell Ultra GPU architecture (GB300), which includes 288 GB of HBM3e per GPU. This is about 1.5× more memory than the

3 The term "NAND" comes from the *NOT-AND* logic gate used in this type of flash memory.

previous GPU GB200 (192 GB). More memory enables GPUs to handle a larger number of data samples and longer input sequences in a single pass, which significantly improves throughput for complex models. The memory bandwidth has also increased dramatically: the Blackwell Ultra delivers around 8 TB/s per GPU, enabled by 12 stacks of HBM3e.

What the growing appetite for memory has to do with increasing scope 1 emissions

In front-end manufacturing, emissions are largely driven by the high global warming potential of seven fluorinated gases – namely tetrafluoromethane (CF4), octofluropropane (C3F8), octafluorobutane (C4F8), hexafluoroethane (C2F6), trifluoromethane (CHF3), nitrogen trifluoride (NF3), and hexafluoride (SF6), which together account for 80%–90% of scope 1 emissions in a fab. These gases are mainly used in etching and chamber cleaning, production steps that are particularly frequent in memory chip production.

Vertically stacking numerous layers of DRAM increases bandwidth but also demands highly selective etching to define nanoscale features within each layer and effective cleaning to remove residue and prevent contamination between layers. Particularly in the 3D architectures for HBM (such as stacking DRAM dies in a 3DIC assembly), bit density is made greater by increasing the layer counts of memory cells. This leads to a taller memory stack, which in turn requires more deposition and etching materials and therefore increased use on fluorinated gas. The etching process involves extremely deep and narrow trench structures, which require a durable hard mask, typically an amorphous carbon layer (ACL), to prevent sidewall collapse during deep etch steps. During etching, fluorinated gases precisely remove unwanted material to enable accurate patterning of stacked layers. During cleaning, they ensure ultrapure surfaces by eliminating residues from previous fabrication steps, including carbon deposits left by repeated ACL deposition.

Consequently, the focus of innovation in memory chip manufacturing has shifted from lithography to deposition and etching processes due to the growing complexity and material demands of advanced memory technologies. Finding viable alternatives for the range of fluorinated gases used in dry etching is especially challenging and could take another decade or even longer. On top of that, searching for alternatives goes hand-in-hand with major changes in the manufacturing process, which means that existing fabs cannot easily adopt them.

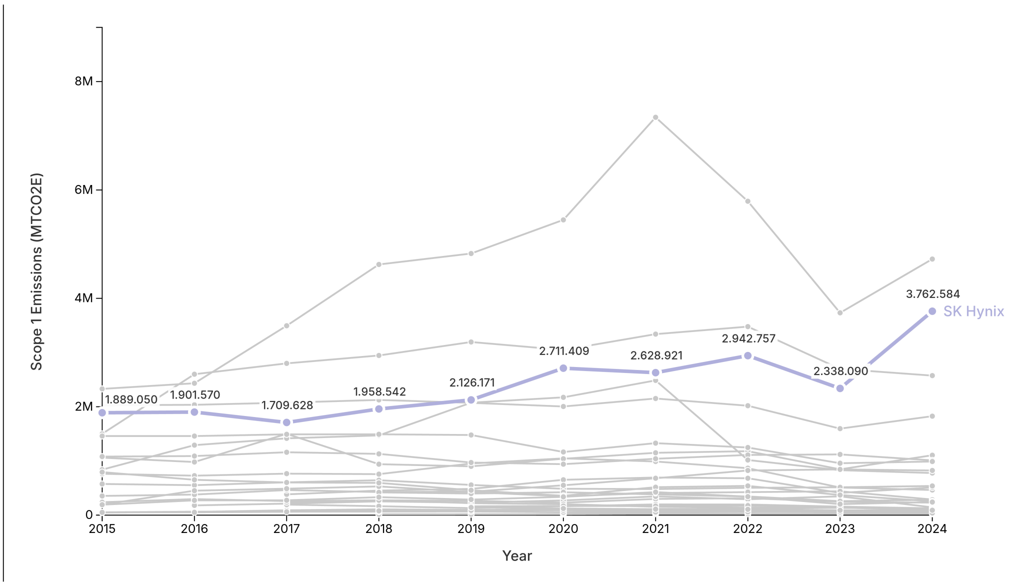
Putting it all together, the recent rise in direct emissions becomes clearer: booming demand for advanced memory chips, the push to stack more DRAM layers for AI, and the more frequent etching and cleaning required in front-end production likely contribute to the uptick.

Does the individual company tracker support our hypothesis?

On our microsite, we display a tracker that explores emissions trends for individual companies. We have consulted the tracker as an additional tool to understand the recent uptick in scope 1 emissions.

Since the most recent publicly available CSR data cover emissions from chip production only up to 2024, we focus on the DRAM and HBM market landscape at the end of 2024. The market is highly concentrated. In DRAM, three companies dominated global production: Samsung (38%), SK Hynix (35%), and Micron (22%). For HBM – representing roughly 5% of the global DRAM market in 2024 but 10% in 2025 – the distribution looks slightly different. SK Hynix held a clear lead with 51% of the global HBM market, followed by Samsung (40%) and Micron (9%).[4]

SK Hynix scope 1 emissions (2015–2024)

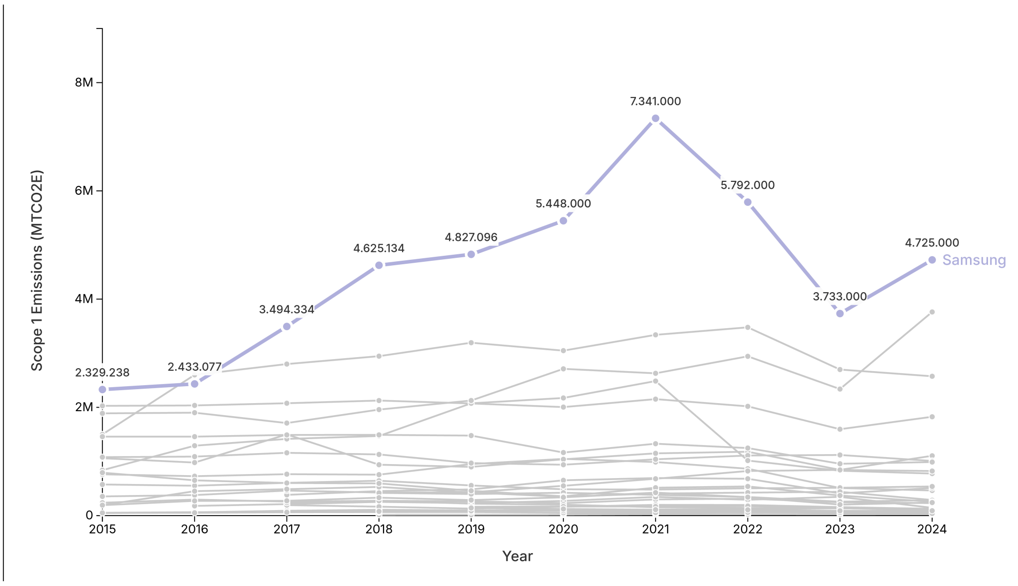


4 Some companies are starting to compete in HBM, primarily China's ChangXin Memory Technologies and some smaller players, such as Nanya Technology and Winbond, who have not yet entered the HBM market with competitive products.

In our dataset, SK Hynix shows the largest increase in scope 1 emissions, which rose from 2.3 MMTCO2E to 3.7 MMTCO2E. This supports our hypothesis that the current surge in demand for DRAM – particularly HBM used in AI applications – is already influencing emission trajectories. Today, SK Hynix is the leading global supplier of HBM, supported by its close partnership with NVIDIA.

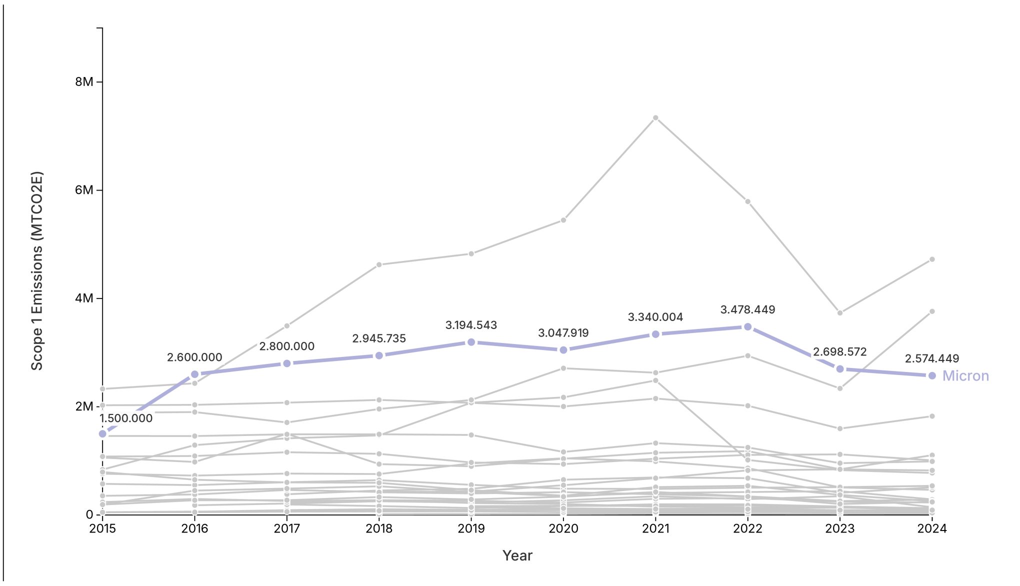
A key factor in this position is its first-mover advantage: SK Hynix began producing HBM well before the surge in AI demand. In addition, the company introduced an innovative packaging technique (MR-MUF), which strengthened its competitive position. In 2025, SK Hynix expanded its market lead and became the first company to supply 12-layer HBM4**,** reinforcing its technological edge in the rapidly growing AI memory market. For 2026, they already announced secured demand for its entire 2026 RAM production capacity.

Samsung scope 1 emissions (2015–2024)



Samsung's activities in chip production can be divided between logic and memory chip production, with the latter making up approximately 75% of their overall manufacturing capacity. Apart from the typical cyclical trends in the memory market, Samsung's memory chip segment has experienced a roller coaster dynamic in recent years. The company struggled to keep pace with competitors due to qualification challenges with its HBM chips for NVIDIA, but it maintained its role as a major DRAM supplier overall. Interestingly, despite the HBM innovation challenges, their 2024 emissions showed a notable increase, which may be explained partly by that being the first year they reported NF3 usage as part of their direct emissions. Looking ahead, Samsung appears to be catching up, achieving its strongest Q3 performance in the past three years in 2025.

Micron Scope 1 emissions (2015–2024)



Compared to SK Hynix and Samsung, Micron has held a smaller share of the DRAM market, and it did not have a significant commercial HBM presence until February 2024. These factors are reflected in Micron's emissions trend: unlike its competitors, their emissions remained largely stagnant from 2023 to 2024. Fast forward to 2025, however, and Micron was on track to catch up in the HBM segment. Their entire HBM supply was sold out in Q3 2025, and they successfully passed NVIDIA's quality verification for HMB3 even before Samsung. On top of that, they recently announced a discontinuation of their memory for consumer PC segment, signalling a rapid shift to AI chips and servers. This will in turn strengthening of their position in the HBM market – a development that will most likely lead to an increase in Micron's scope 1 emissions in the next reporting cycle.

Conclusion

Looking at the aggregated change of direct emissions from 2023 to 2024 and zooming in on the three largest memory chip manufacturers in our dataset supports our hypothesis: the observed beginning of a trend reversal may be driven by the current AI boom, which has translated into strong demand for GPUs and HBM. It is important to note that in 2024, the HBM market represented only 5% of the overall DRAM market. As a result, conventional DRAM was the primary demand driver in 2024, and the material impacts of the recent surge in HBM-specific demand are likely to become evident only in the yet-to-be-published 2025 and 2026 CSR data.

Thus, the trend reversal might be only the start of an increase in direct emissions due to the rapid increase in demand for AI. At the moment, the HBM market is on a steep growth trajectory with demand currently exceeding supply. Micron forecasts that the total addressable market will reach US$100 billion by 2028. In comparison, in 2024 it was only US$16 billion. It remains to be seen whether the projected rapid expansion will affect direct emissions over the longer term.

Acknowledgements

We would like to thank Sarah Budai for her patience and thoroughness during data collection; Pasha Semenova for their dedicated data science work; Alina Siebert for her indispensable design support; Maximilian Gottwald and Jack Walmsley for their help during the research and text edits; Luisa Seeling, Iana Pervazova and Sebastian Rieger for helping to spread the word about this publication.

Authors

Julia Christina Hess
Lead Global Chip Dynamics
jhess@interface-eu.org
+49 30 81 45 03 78 80

Maria Nowicka
Policy Researcher Global Chip Dynamics
mnowicka@interface-eu.org

Imprint

interface – Tech analysis and policy ideas for Europe
(formerly Stiftung Neue Verantwortung)

W www.interface-eu.org
E info@interface-eu.org
T +49 (0) 30 81 45 03 78 80
F +49 (0) 30 81 45 03 78 97

interface – Tech analysis and policy ideas for Europe e.V.
c/o Publix
Hermannstraße 90
D-12051 Berlin

Design by Make Studio
www.make.studio
Code by Convoy
www.convoyinteractive.com