SEMICONDUCTOR FRONT-END · PART 3-13

Metrology, inspection and yield Finding small problems and learning where they began

A wafer passes through hundreds of processing steps before its circuits are ready to test. Manufacturers need to know whether dimensions and films are on target, where defects have appeared, and which changes will increase the fraction of usable chips. This lesson connects physical measurements, defect inspection, electrical testing and factory data to explain how that learning happens.

Semiconductor manufacturing series · Intermediate introduction

Suggested preparation: the manufacturing overview and basic photolithography · Updated 19 November 2026

What you will learn

- How defect size and location determine whether a circuit fails.
- How yield models account for die area and clustered defects.
- How optical inspection identifies candidates for detailed review.
- How dimensions, film thickness and overlay are measured.
- How physical and electrical data help engineers investigate a process.

Contents

1. Three questions a fab keeps asking
2. Yield and the defects that matter
3. How die area and clustering affect yield estimates
4. How inspection finds a candidate
5. Measuring dimensions, films and placement
6. Testing the circuit electrically
7. Turning measurements into yield learning
8. Suppliers and the markets they serve
9. Where materials knowledge is essential
10. Limits of inspection and the next challenges
11. Check your understanding
12. Terms used in this lesson
13. Sources
14. How the evidence is used

1. Three questions a fab keeps asking

Are the structures on target? Metrology measures quantities such as line width, film thickness and the relative placement of layers.

Where has something unusual appeared? Inspection searches for defects and other unexpected signals. Review and classification help determine what those signals represent.

Which circuits meet their requirements, and why? Electrical test measures device or circuit behaviour. Engineers combine those results with manufacturing history and physical measurements to investigate losses and improve yield.

2. Yield and the defects that matter

- **State what the yield figure counts**

Yield is a fraction of output meeting a defined condition. Different stages use different denominators. A wafer-level measure might count wafers reaching probe test. Functional die yield counts electrically working dies. Parametric yield concerns performance requirements such as speed, leakage or other specified limits. A circuit can function while falling outside a particular performance bin. [1]

The MIT teaching material discusses parametric failures from variations in quantities such as line width, random failures such as an unsuccessful via among many connections, and area-related losses associated with killer defects. These categories help organise the sources of yield loss. [1]

- **A particle's position matters as much as its size**

An added conductive particle can bridge neighbouring wires and create a short. Missing metal can interrupt a conductor and create an open. The same defect size may be harmless at one location and cause failure at another.

For a specified defect size and failure mechanism, the **critical area** is the set of positions at which the defect's centre would produce a failure. Critical area for a short differs from critical area for an open. This is a geometric way to connect a physical defect to an electrical consequence. [1]

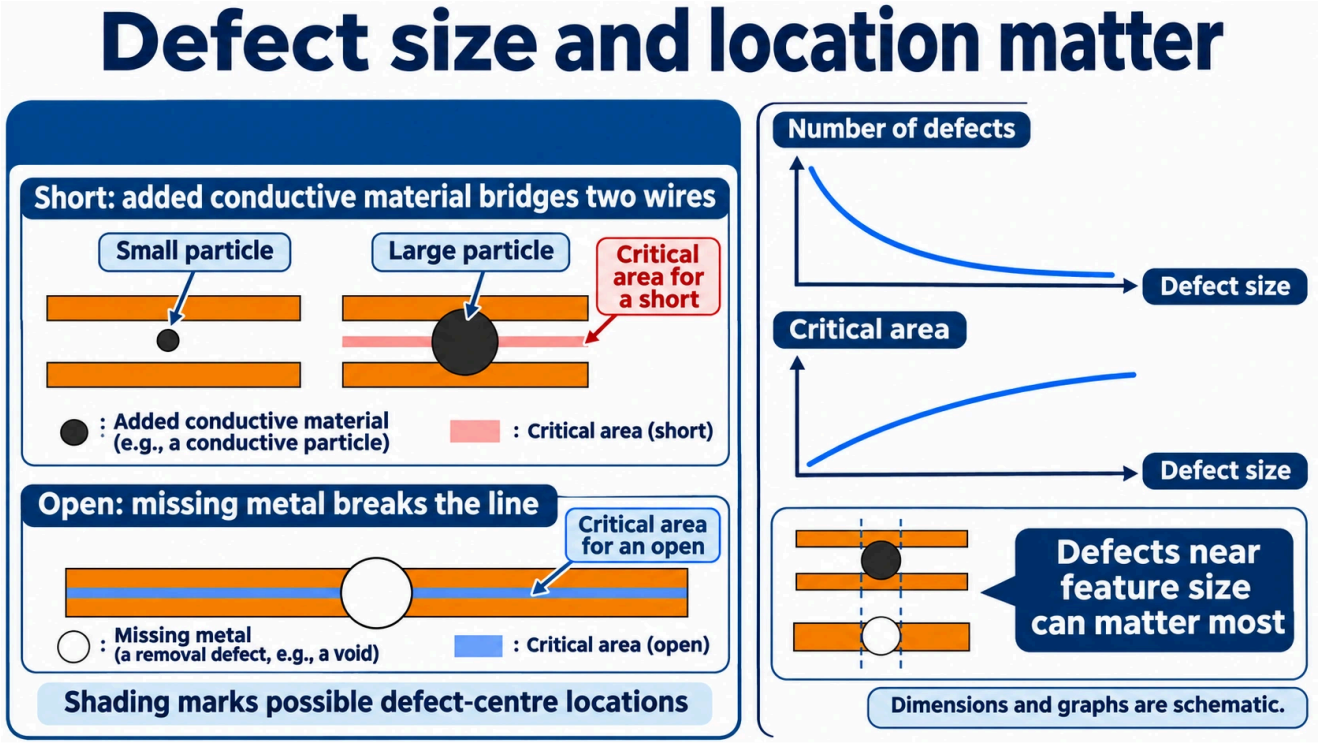


Figure 1. The shaded region identifies possible defect-centre positions. The short and open examples involve different kinds of material defect. Dimensions and trends are schematic.

Very small defects can have little or no critical area for a particular failure. Very large defects can be less numerous. Integrating the defect-size distribution with size-dependent critical area identifies which sizes contribute most to loss. In the MIT examples, sizes near the minimum feature dimensions are especially important. The dominant range depends on the layout and failure mechanism. [1]

- **Why a roadmap discusses 3 nm particles**

The 2024 IRDS Yield Enhancement chapter describes electrically active particle sizes as small as **3.0 nm** in its most critical contamination discussion, using a definition related to half the critical logic dimension. This is a roadmap context for contamination control. It is not a single failure threshold for every material, layer or product. [2]

The same discussion identifies a metrology gap in ultrapure-water particle control. Contamination prevention, materials control and the ability to detect a particle must therefore be considered together.

3. How die area and clustering affect yield estimates

- **The Poisson starting point**

If yield-limiting defects are independently and uniformly distributed with mean density D_0, a simple zero-defect model gives $\mathbf{Y = \exp(-D_0 A)}$, where A is die area. A larger die samples more area and is more likely to contain a yield-limiting defect.

Real defect density can vary between locations, wafers or groups of dies. The MIT lecture introduces a negative-binomial model to account for such variation. At a fixed average defect density, clustering puts several defects into some dies while leaving more other dies free of them. [1]

A clustered-defect model

$$Y = (1 + D_0A/\alpha)^{-\alpha}$$

α is a clustering parameter. Larger α approaches the Poisson case. The lecture cites α values around 1.5–2 for typical memory and microprocessor examples. These values describe those teaching examples. [1]

Die area	Poisson	Negative binomial, α = 2	Negative binomial, α = 1.5
1 cm²	90.5%	90.7%	90.8%
6 cm²	54.9%	59.2%	60.4%

Calculation assumptions. D_0 = 0.1 defects/cm² in every column. For the larger die, the clustered models give roughly four to six percentage points more yield than the Poisson estimate. The models still predict lower yield as die area increases. Actual product yield also includes systematic and parametric losses.

Spatial patterns can reveal a different problem

Process variation can create a wafer-scale pattern, such as failures concentrated near an edge. The lecture treats these global losses separately from a simple critical-area defect calculation. IRDS also highlights the influence of wafer edges and backsides. Maps are useful because a spatial pattern can suggest which process mechanisms to investigate. [1][2]

4. How inspection finds a candidate

Comparing repeated patterns

A wafer often contains repeated copies of the same die. An inspection system can align images from equivalent positions and look for differences. A 1978 KLA Instruments patent describes electronic alignment and comparison of optical images from photomask dies. Its discussion also refers to reference-based and neighbouring-die comparisons. [8]

Consider three aligned copies, A, B and C. If B contains a particle absent from clean A and C, both the A–B and B–C comparisons highlight the same location. That makes the location in B a candidate for review. The image difference is a physical observation whose electrical importance still needs assessment.

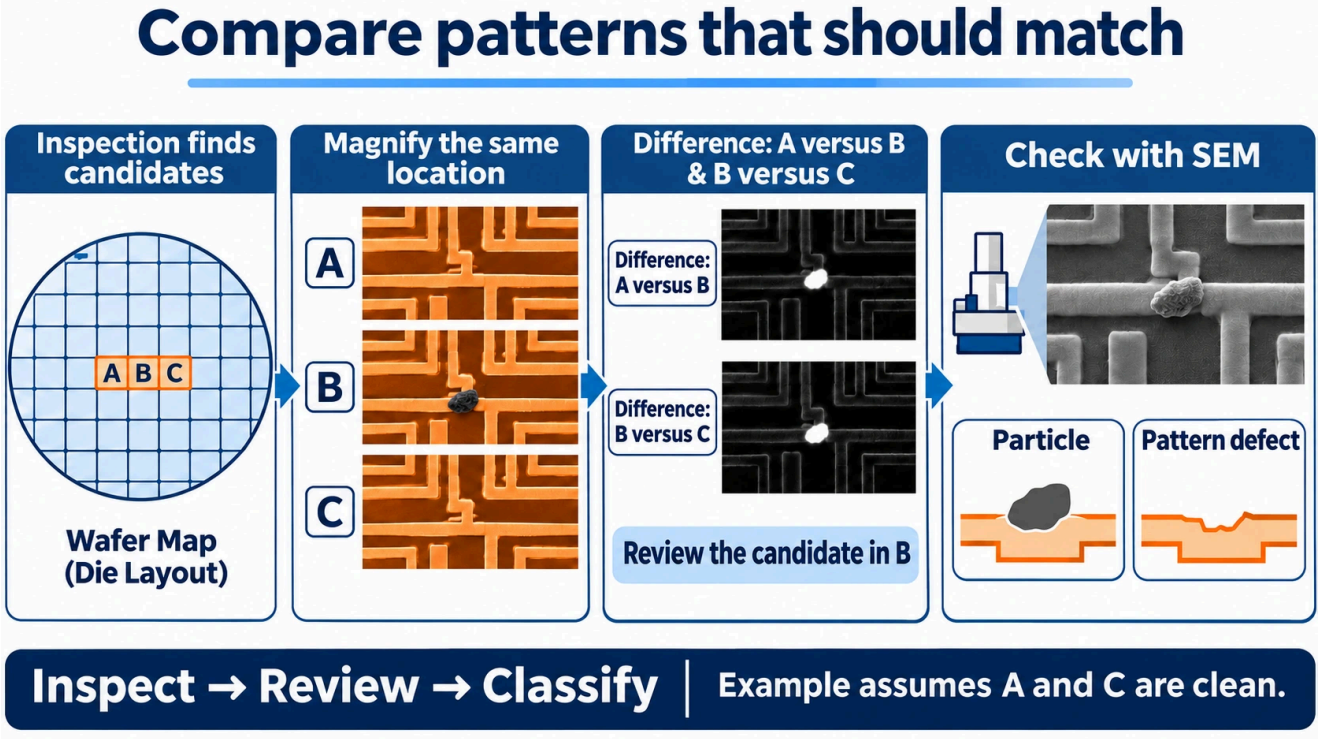


Figure 2. A simplified three-die example assuming A and C are clean. Inspection identifies a candidate; review examines it more closely and classification assigns a type.

Image alignment and the choice of reference are essential. A repeated error shared by the compared dies may require another reference or another measurement to identify it. Manufacturing inspection therefore uses several approaches suited to the structure and suspected defect.

Bright-field and dark-field collection

Bright-field inspection collects the reflected light that forms the normal pattern background. Changes in pattern shape or optical response can appear as contrast. Dark-field collection suppresses the main specular reflection and collects scattered light, allowing some particles or irregularities to stand out against a darker background. Illumination angle, collection angle, wavelength and polarisation all affect sensitivity.

KLA describes broadband optical inspection platforms and oblique laser-scanning inspection platforms used for production monitoring. Its history of broadband patterned-wafer inspection traces the source from tungsten-halogen illumination in the first generation to deep-ultraviolet operation in later systems. [6][7]

Detection and resolving a shape are different tasks

The supplied source gives a 190–260 nm wavelength range for the discussed KLA systems. Compared with a 3 nm particle, the wavelength is about 63–87 times larger. This ratio is a size comparison. It does not establish that a given optical inspection tool detects every 3 nm defect.

An optical instrument can detect a change in scattered or reflected light from an object whose detailed shape it cannot resolve. Whether that change is detectable depends on scattering strength, material contrast, pattern background, collection optics, noise and the inspection recipe. A roadmap's contamination-control requirement and a tool's demonstrated sensitivity therefore need to be read with their own conditions. [2][6]

Review then asks what the candidate actually is. A SEM can provide more detailed local information for many candidates, while buried or difficult structures may require additional methods. A physical candidate also needs to be related to electrical function.

Throughput and nuisance signals

KLA's historical account compares an image-data rate of 0.8 megapixels/s in its early platform with 24.8 gigapixels/s in the later series. Dividing those figures gives about 31,000. This is a comparison of the reported data rates; production throughput also depends on the inspection task. [6]

Higher sensitivity can produce many signals from harmless pattern variation, as well as false detections. Such **nuisance** signals must be separated from defects that matter to the process. Classification algorithms and machine learning help prioritise review. IRDS identifies the volume of nuisance signals as a continuing challenge. [2][7]

KLA describes automatic classification introduced in 2005 and later deep-learning-based approaches. It also links optical inspection with electron-beam review. Electron-beam inspection, including multibeam systems such as ASML's HMI products, provides another route to detailed pattern and voltage-contrast information. [6][7][12]

5. Measuring dimensions, films and placement

Metrology produces a quantitative estimate of a defined property. The appropriate method depends on the material, geometry, required uncertainty, sampling rate and whether the measurement itself can alter the structure. IRDS identifies critical dimension, overlay, defects and material properties as increasingly demanding parts of process control. [3]

Critical dimensions with electrons

A critical-dimension SEM, or CD-SEM, scans an electron beam and collects emitted electrons to form an image from which dimensions are estimated. Image formation, edge definition, charging and calibration affect the result.

The Semiconductor History Museum of Japan describes Hitachi's 1984 S-6000 as a landmark CD-SEM. A field-emission source and operation near 1 keV helped combine useful resolution with reduced charging and damage. Automated secondary-electron detection made rapid dimensional measurement practical. Those operating details describe the historical system. [5]

Different structures can require different energies. IRDS discusses higher-energy SEM measurements, above 10 keV, of bottom diameters in 3D NAND channel holes. The same discussion notes the possibility of electrical damage at measured locations. The best measurement recipe must consider the effect on the sample. [3]

Optical CD and model fitting

Scatterometry measures how a periodic structure changes incident light. Wavelength-dependent intensity and polarisation can be fitted to an optical model of line width, height and sidewall shape. This is often called optical critical dimension measurement, or OCD. The inferred profile depends on both the model and the information in the measurement. [3]

Film thickness and ellipsometry

Ellipsometry measures the change in polarisation when light reflects from a surface or film stack. A suitable model relates that response to film thickness and optical constants. Spectroscopic measurements add wavelength-dependent information.

IRDS discusses the need for optical constants appropriate to the materials and wavelengths involved. The inverse problem can admit several plausible parameter combinations, which is why calibration, independent constraints and an appropriate film model are important. [3]



An ellipsometer at LAAS-CNRS in Toulouse, France, as identified by the source. The arms illuminate the sample obliquely and collect the reflected polarisation response. Image by Guillaume Paumier, Wikimedia Commons, CC BY-SA 3.0, resized and converted to WebP. Original and licence

Overlay and hybrid metrology

Overlay measures the relative placement of patterns on different layers. A dimension can be correct while its position relative to the layer below is wrong. Smaller alignment budgets therefore make overlay control an important part of yield. A common approach prints dedicated target marks in both layers and measures the offset between their centres. [3]

Hybrid metrology combines methods with complementary strengths. For example, AFM information about sidewall shape or surface profile can constrain a scatterometry model. Machine learning can assist fitting and analysis, while measurement quality still depends on suitable reference data and validation. [3]

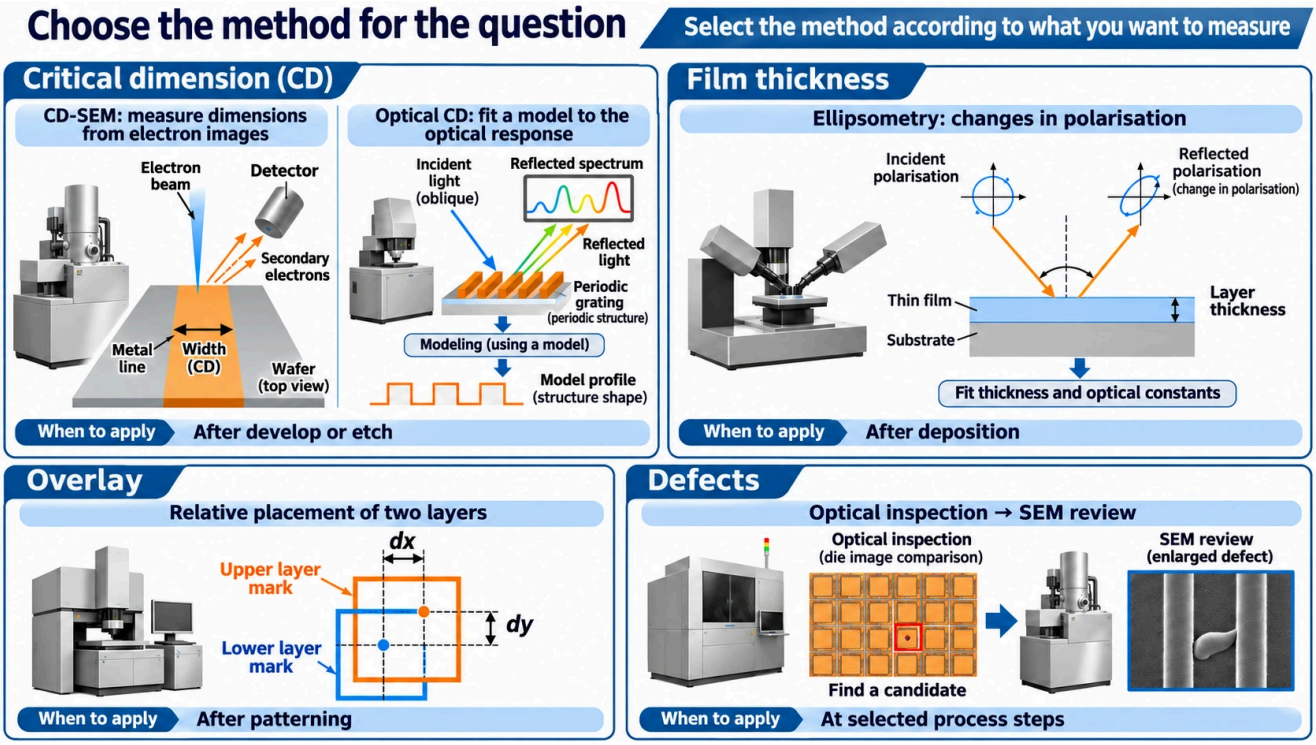
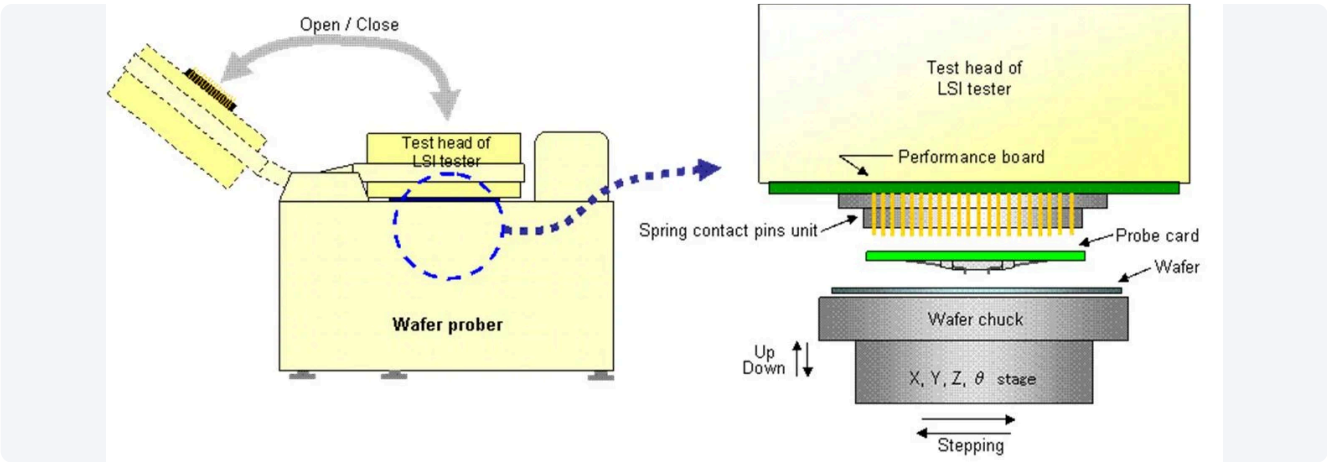


Figure 3. Methods answer different process-control questions. Example measurement points vary with the product and process flow.

6. Testing the circuit electrically

A wafer prober positions a wafer so that a probe card can contact die pads. The connected tester then applies and measures electrical signals before the wafer is diced. Testing can identify functional failures and assign performance bins. [9]

Dedicated process-control monitor structures, or PCM structures, measure selected electrical properties of the process. MIT also discusses test patterns with different line spacings for investigating opens, shorts and defect-size distributions. These structures connect a physical process hypothesis to an electrical observation. [1]



A wafer-prober arrangement. The stage positions successive dies for contact with a probe card connected to a tester. Image by GUCCIN, Wikimedia Commons, CC BY-SA 3.0. Original and licence

7. Turning measurements into yield learning

- **Connect the observations to manufacturing history**

A defect map identifies where a signal occurred. A failed-die map identifies electrical outcomes. Tool traces and manufacturing history show which processes and equipment each wafer experienced. Combining these records helps engineers test a proposed explanation for loss.

The 2024 IRDS Yield Enhancement chapter describes both rapid feedback from inline inspection and more proactive approaches that connect data across the fab. Useful inputs include manufacturing execution system records, equipment traces, metrology, defect data, PCM, wafer probe, final test, material and facilities information, maintenance records, reliability results and failure analysis. [2]

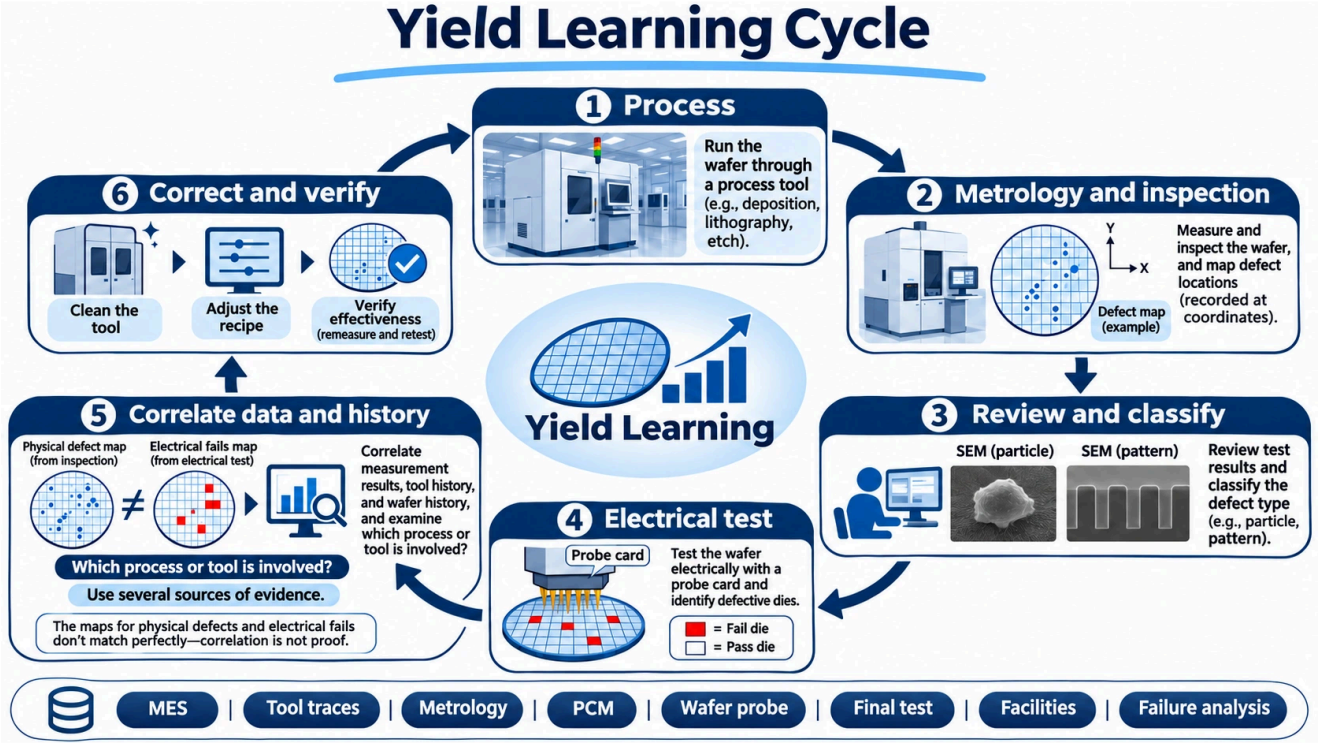


Figure 4. One explanation of the yield-learning loop. Several sources of evidence help narrow the cause, and the effect of a corrective action is checked.

Suppose failing dies and a particular defect class both become more common after wafers pass through one chamber. Engineers can examine timestamps, chamber history, maintenance and recipe data, then investigate the physical failure mechanism. A controlled correction and subsequent measurements help determine whether the change addresses the cause.

Virtual metrology and predictive analytics aim to extract earlier warnings from such data. Their practical value depends on reliable data association, suitable sampling and models that continue to work as the process changes. IRDS places automatic analysis and faster root-cause learning among the important requirements. [2]

Why learning speed has economic value

KLA's annual report describes advanced-fab investments well above US$10 billion and emphasises the importance of bringing yield up quickly. A slow feedback loop can allow many more wafers through an unstable process before its cause is understood. Faster learning helps manufacturers use their installed capacity more effectively. [10]

8. Suppliers and the markets they serve

Inspection, review, dimensional metrology and mask inspection are related activities with different requirements. Suppliers often participate in several of them. The following overview identifies roles, while the market figures use the category definitions of the source.

A historical market snapshot

The METI presentation cited in the supplied lesson reports **¥795.2 billion** for wafer surface inspection and **¥74.2 billion** for its wafer external-appearance inspection category. The measurement year is unspecified. The values below retain the source figures provided in the Japanese article. [4]

Source category	Reported selected shares
Wafer surface inspection	KLA 69%, ASML 10%, Applied Materials 6%, Hitachi High-Tech 4%, Onto Innovation 4%, others 6%.
Wafer external-appearance inspection	KLA 36%, Applied Materials 26%, Shanghai Precision 7%, others 12%, plus an unidentified Japanese-supplier band.

The first set adds to 99%, so it is presented without forcing a 100% total. In the second category, the supplied description does not give a name or numerical share for the Japanese band. An unexplained remainder is not assigned to a company. The source categories also need to be retained when comparing these figures with other market studies.

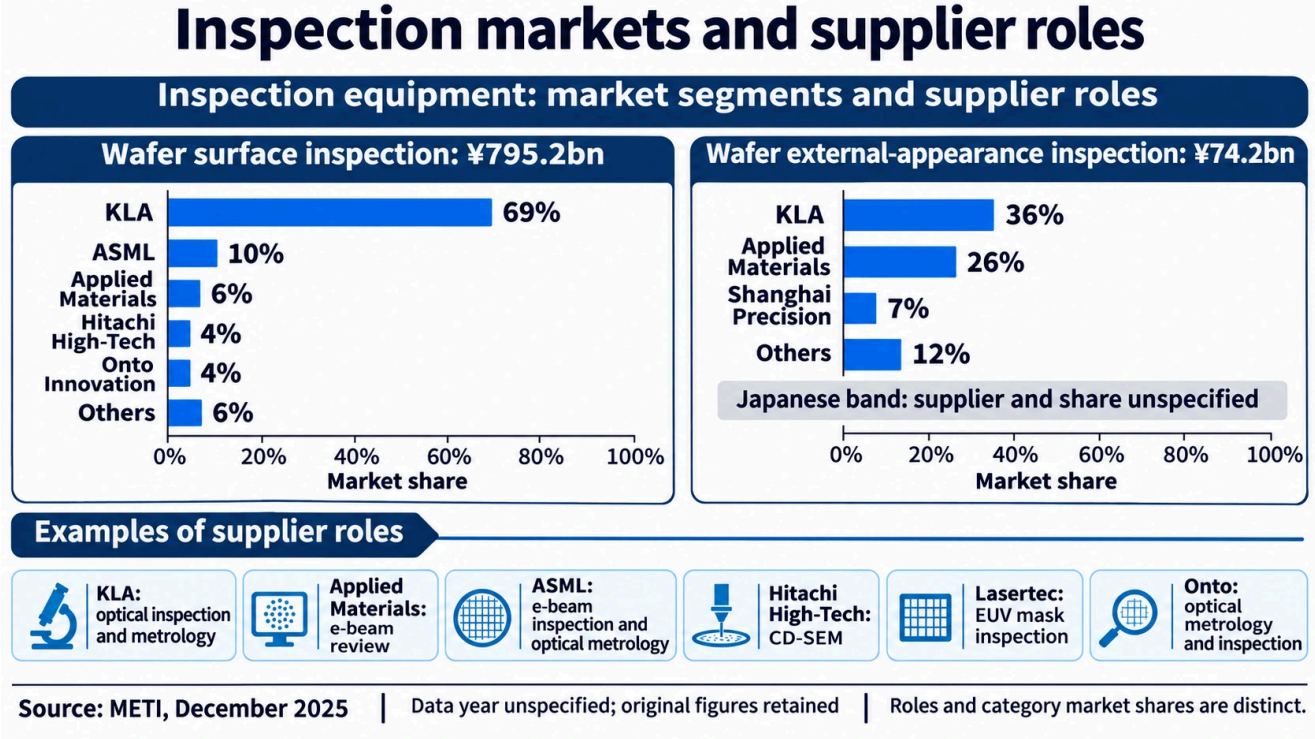


Figure 5. Original market figures and their limitations. Supplier roles describe product examples and have a different scope from the category shares. [4][7][11][12][13][14]

- **Examples of supplier contributions**

KLA. Its portfolio combines inspection, metrology, review and software for process control. The company attributes a major historical change to its KLA 2020 patterned-wafer inspection platform introduced in 1984. Its fiscal-2026 filing reports revenue of approximately US$13.58 billion, net income of US$4.83 billion and service revenue around 23% of total revenue. It identifies Applied Materials, ASML, Hitachi High-Tech, Lasertec and Onto Innovation among competitors. [6][10]

Applied Materials. SEMVision is an electron-beam defect-review family. Its 2025 H20 announcement describes electron imaging combined with image analysis to examine and classify small defects. [13]

ASML. Alongside lithography, ASML offers HMI electron-beam inspection and optical metrology. Its eScan systems include multibeam inspection. These product offerings can be verified directly, while the precise composition of ASML's share in the METI category remains a separate market-definition question. [12]

Hitachi High-Tech. CD-SEM is a longstanding part of its contribution, with the historical 1984 S-6000 described earlier. A company's share in a broad inspection category is a different measure from its position in a particular metrology segment. [4][5]

Lasertec. Its ACTIS family addresses actinic EUV patterned-mask inspection, using EUV-wavelength information relevant to printing. The company describes A150 as the first EUV patterned-mask inspection system and A300 as supporting high-NA masks. Its ABICS products address EUV mask blanks. Mask inspection is important because a printable mask defect can be repeated across many exposure fields. [11]

Onto Innovation. Its product range includes optical critical-dimension, film and overlay metrology, inspection and process-control software. Those capabilities contribute at different stages of wafer and advanced-packaging production. [14]

9. Where materials knowledge is essential

- **Optical constants.** Model-based measurements need appropriate refractive-index and absorption data. Composition, structure, phase and wavelength can change the response. [3]
- **Reference samples.** Dimensional and thickness measurements need calibration and stable references. AFM and other methods can support dimensional-metrology calibration. [3]
- **Beam response.** Resist shrinkage, charging and damage must be understood when using electrons to measure a structure. [3][5]
- **Contamination control.** Water, chemicals and tool components can introduce particles whose detection is difficult. Source control and suitable materials are part of the solution. [2]

10. Limits of inspection and the next challenges

Five limits of inspection today

The methods in this lesson also run into limits that are hard to remove.

- **Particles below detection.** IRDS states that the critical particle size is already below what leading-edge metrology can detect, and that this has been true for ultrapure water for nearly a decade. [2]
- **Light cannot image the shape.** At 190–260 nm, the inspection wavelength is about 63–87 times larger than a 3 nm particle (section 4). Optical inspection can flag a candidate from a small change in scattering, but electron review is needed to see what it is.
- **Electrons can affect the sample and are slow over large areas.** Electron beams resolve fine detail, but charging and damage must be controlled. [3][5] Scanning a large area at high resolution also takes time.
- **Optical metrology depends on models.** Scatterometry and ellipsometry infer shape or thickness by fitting a model, and the inverse problem can have more than one plausible answer. [3]
- **Buried defects.** Voids and defects deep inside high-aspect-ratio structures cannot be seen from the surface. [2]

Inspection is always a balance. Higher sensitivity captures more nuisance signals, while faster or wider coverage leaves less time to examine each site, so more defects may be missed. This balance is an explanatory summary rather than a quoted finding. Particles too small to detect therefore have to be prevented rather than found (section 9).

Challenges ahead

Smaller critical contamination. Detection capability and contamination-control requirements must be developed together, including control upstream of the wafer. [2]

Three-dimensional and buried structures. Deep holes, voids and interfaces create questions that a simple surface image cannot answer. [2][3]

Stochastic patterning variation. EUV-related randomness and rare defects need suitable sampling and measurement strategies. [3]

Predictive learning. Connected factory data can help identify process changes earlier, provided the relationships are validated and the models are maintained. [2]

This completes Part 3, which followed the fabrication steps individually. Part 4 turns to the device structures assembled from them, including FinFET, gate-all-around and future stacked transistors, DRAM, 3D NAND and power devices.

11. Check your understanding

▼ What does a yield percentage need to specify?

The population being counted, the stage of manufacture and the condition for passing, such as functionality or a performance limit.

▼ What is critical area?

For a given defect size and failure mechanism, it is the area in which the defect's centre would cause a circuit failure.

▼ Why can clustered-defect models predict a higher zero-defect die fraction?

At the same mean density, several defects can be concentrated in some dies while more other dies remain free of them.

▼ What assumption underlies the three-die comparison example?

The repeated patterns are aligned, and the reference dies A and C are clean at that location. A common difference points to a candidate in B for review.

▼ Does a wavelength-to-particle-size ratio establish optical detection sensitivity?

No. Sensitivity also depends on material, contrast, scattering, background, noise and recipe conditions. Detection of a signal change and resolving a detailed shape are different measurement tasks.

▼ How do inspection and review fit together?

Inspection locates candidates. Review examines selected locations in more detail, and classification describes the type of feature or defect.

▼ How does ellipsometry estimate film properties?

It measures changes in reflected polarisation and fits an appropriate model for thickness and optical properties, often using measurements at several wavelengths.

▼ Why combine physical inspection with electrical test?

They provide complementary evidence about a physical feature and the circuit's measured behaviour, helping engineers assess relevance and investigate a cause.

▼ Which records help a root-cause investigation?

Wafer history, equipment traces, metrology, defect maps, electrical tests, materials and facilities data, maintenance records and failure analysis.

12. Terms used in this lesson

Term	Meaning
Metrology / inspection	Quantitative measurement of a defined property / searching for defects or unexpected signals.
Yield / parametric yield	The fraction meeting a stated pass condition / the fraction meeting specified electrical or performance limits.
Defect density D_0	A mean number of defects per unit area under the stated model and defect definition.
Critical area	The region of defect-centre locations that causes a specified failure for a specified defect size.
Killer defect	A defect capable of making a circuit fail under the relevant conditions.
Negative-binomial yield model	A model incorporating variation or clustering in defect density.
Die-to-die comparison	Comparing aligned images of repeated die patterns.
Bright field / dark field	Collection involving the main reflected background / collection that suppresses specular background and uses scattered light.
Review / nuisance	Detailed examination of a candidate / an inspection signal of little relevance to the intended yield-control task.
CD-SEM	Scanning electron microscope used for critical-dimension measurements.
Scatterometry / OCD	Inferring dimensions or profiles by fitting a model to a structure's optical response.

Term	Meaning
Ellipsometry	Using reflected polarisation changes to estimate film properties.
Overlay	The relative placement of patterns in different layers.
Hybrid metrology	Combining complementary measurement methods and constraints.
Wafer prober / PCM	A positioning and contacting system for wafer electrical test / process-control monitor structures or measurements.
Yield learning	Investigating causes of loss, applying corrections and checking the result.

13. Sources

Historical accounts, roadmap requirements, worked calculations, company statements and market-category figures are identified in the text.

1. MIT OpenCourseWare, Control of Manufacturing Processes, Lecture 10, Yield Modeling, 2008. Source document
2. International Roadmap for Devices and Systems, 2024 Edition, Yield Enhancement. Includes contamination-control requirements and yield-learning data. Source document
3. International Roadmap for Devices and Systems, 2024 Edition, Metrology. Source document
4. Ministry of Economy, Trade and Industry, Japan, Future Direction of the Semiconductor and Digital Industry Strategy, 23 December 2025, equipment chart p.5. Measurement year unspecified. Source document
5. Semiconductor History Museum of Japan, 1984 Critical Dimension-Scanning Electron Microscope (CD-SEM), English account, 2022 revision. Source document
6. KLA, BBP 40th Anniversary, company history of broadband patterned-wafer inspection. Source document
7. KLA, Defect Inspection and Review, product information. Source document
8. K. Levy and P. Sandland, KLA Instruments, US Patent 4,247,203, Automatic photomask inspection system and apparatus. Filed 3 April 1978, granted 27 January 1981. Source document
9. Advantest, Wafer Inspection—A Crucial Step in Semiconductor Manufacturing. Source document
10. KLA, Form 10-K for the fiscal year ended 30 June 2026, Item 1 and financial statements. Source document
11. Lasertec, semiconductor product range, including ACTIS and ABICS. Source document
12. ASML, HMI eScan 1100, electron-beam inspection product information. Source document
13. Applied Materials, SEMVision H20, electron-beam defect-review product information. Source document
14. Onto Innovation, Metrology, product information covering optical CD, films and overlay. Source document

14. How the evidence is used

Topic	Basis	Interpretation
Yield definitions, critical area and clustering	S1	Teaching models. The percentage table is calculated from stated parameters.
3 nm contamination and data integration	S2	Roadmap discussion with a specific contamination-control context.
Dimensions, films, overlay and hybrid methods	S3, S5	Metrology roadmap and historical technical account.
Die comparison	S8	Patent principle. The three-die teaching example assumes two clean references.
Wavelength ratio and data-rate ratio	S2, S6 and calculations	Arithmetic comparisons with explicit limits on what can be inferred.
Electrical wafer test	S9	Process description combined with the teaching material.
Company history, financials and products	S5–S7, S10–S14	Attributed historical accounts and primary company disclosures.
Inspection market shares	S4	Source figures retained from the supplied article, with category and missing-data caveats.

Illustrations created with ImageGen. Geometry, instruments, image differences and data maps are explanatory schematics. The market figures retain the supplied article's values and source categories.

© 2026 Troy Technical

Unauthorised reproduction, redistribution or re-uploading of copyright-protected text, illustrations and video is prohibited, subject to exceptions permitted by law. Rights in cited material and trademarks belong to their respective owners. Links to this article are welcome. Please contact Troy Technical about reuse.