ML-based Optimization for High Precision Wafer Dicing Process

APCSM Conference 2021

Mohamed Bahnas, Srividya Jayaram, Andres Torres - Siemens EDA
Dimitrios Velenis, Julien Bertheau - IMEC
Moritz Stoerring, Bruno Van de Vel, Maarten Liebens - KLA

SIEMENS

Agenda

- Overview on Smart Manufacturing Flow (MADEin4)
- Objective: Wafer Dicing Conditions Analysis and Optimization
- Approach: ML-based Feature Importance Analysis
- Summary

SIEMENS

MADEin4 (Metrology Advance for Digitized ECS Industry)

- **Industry 4.0** initiative by **ECSEL** (Electronic Components and Systems for European Leadership):
  - **MADEin4** Consortium (51 Partners)
  - Published work covers one project between **IMEC – KLA – Siemens EDA**

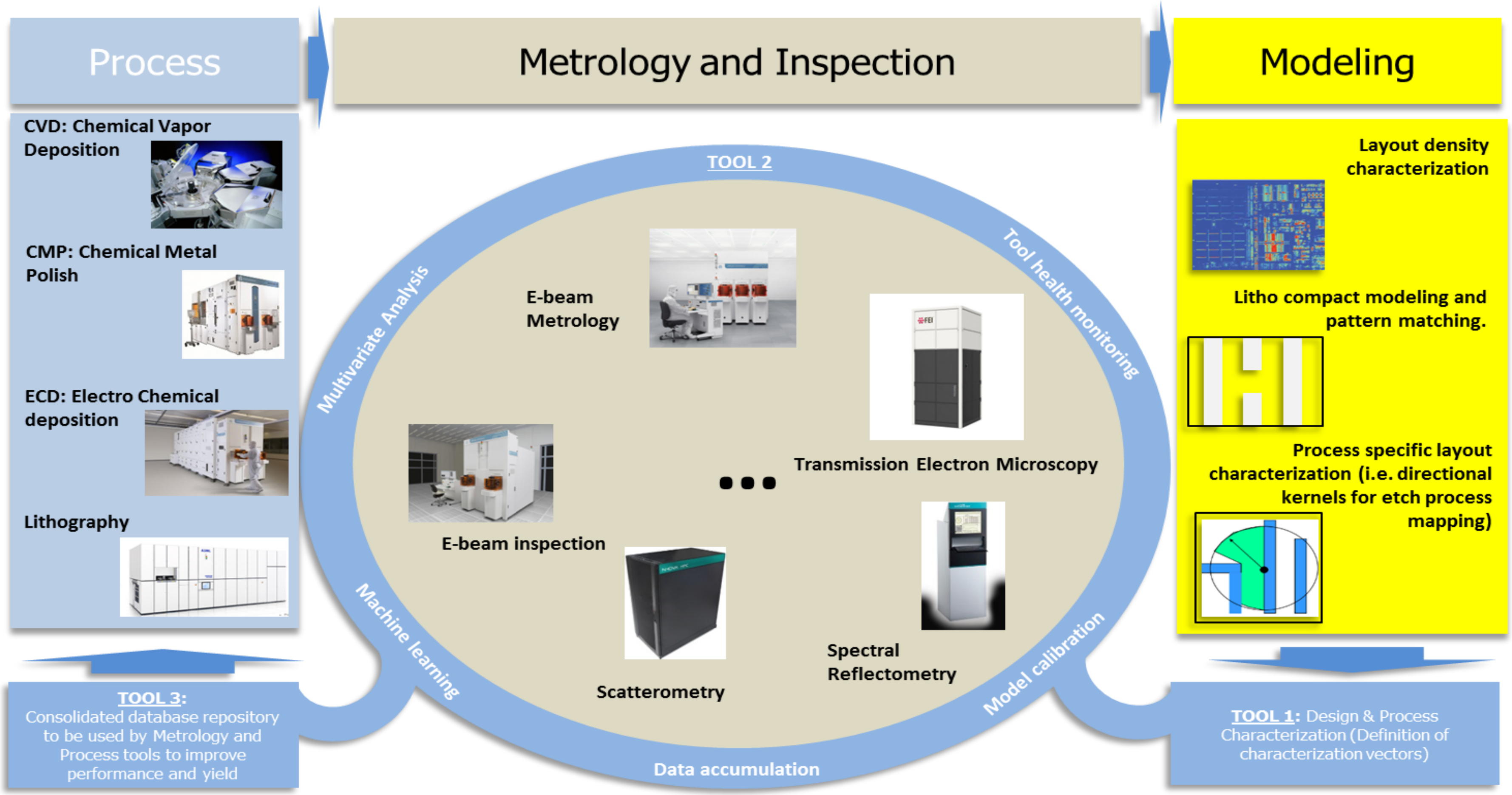
- Develop next gen metrology & inspection tools in support of Industry 4.0 HVM in the semiconductor manufacturing industry

- **Predictive yield** and tools performance improvement by combining process, design & metrology data under one ML (machine learning) platform

- **Smart Manufacturing**: Develop data science methodologies that can be extended to other applications, by applying best practices in:
  - Smart sampling,
  - Feature engineering requirements, and
  - Deployed ML approaches

SIEMENS

Combining Process, Metrology and Design Data for Smart Manufacturing



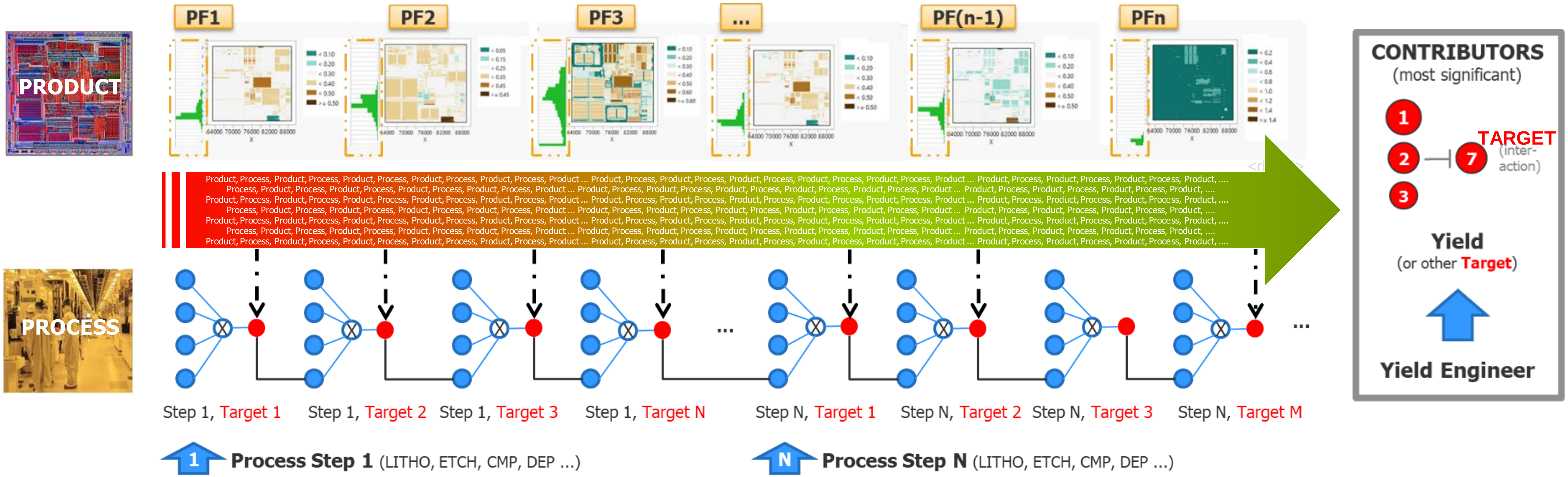
Open Questions

- What data?
- Which processes?
- What improvements are feasible?

SIEMENS

ML-Flow Structure Dependency on Process, Metrology, Design Data

Product Feature (PF) – Extraction & Encoding



Characterizing process, metrology and design content sources of variation

SIEMENS

Published Work: Wafer Dicing Condition Analysis and Optimization

Motivation

- Identify the cause for sidewall and hairline cracks that occur during wafer dicing process
- Determine optimal conditions for dicing process
- Demonstrate the ML-based flow in software platform that fits with different datasets along the manufacturing process steps

Approach

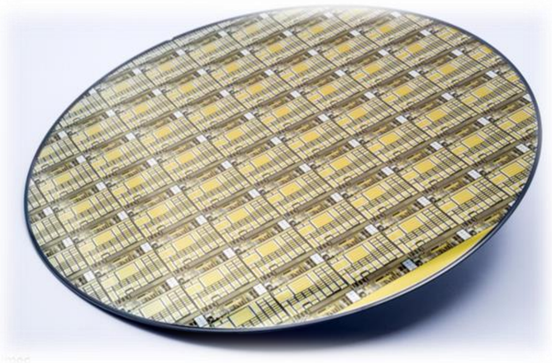
- Design of Experiment for applying different dicing conditions on several wafers of the same design
- Inspection with advanced technology for data collection of edge chipping depth (target)
- ML approach applied in input parameters during wafer dicing process conditions
- Model Analysis to focus on the individual contribution of input parameters on the model output

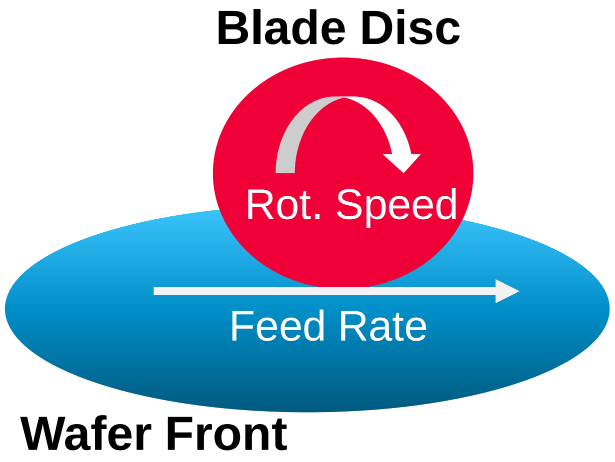
Outcome

- ML analysis ranked the feature based on its contribution on the model results (slower dicing speed and lower feed rate is the preferred direction), and the dicing speed has larger effect than feed rate.
- Analysis shows some dependence of a later dicing step on a former dicing step.
- There is slight effect of the radial location of the die on the dicing results.

Wafer Dicing Process Overview

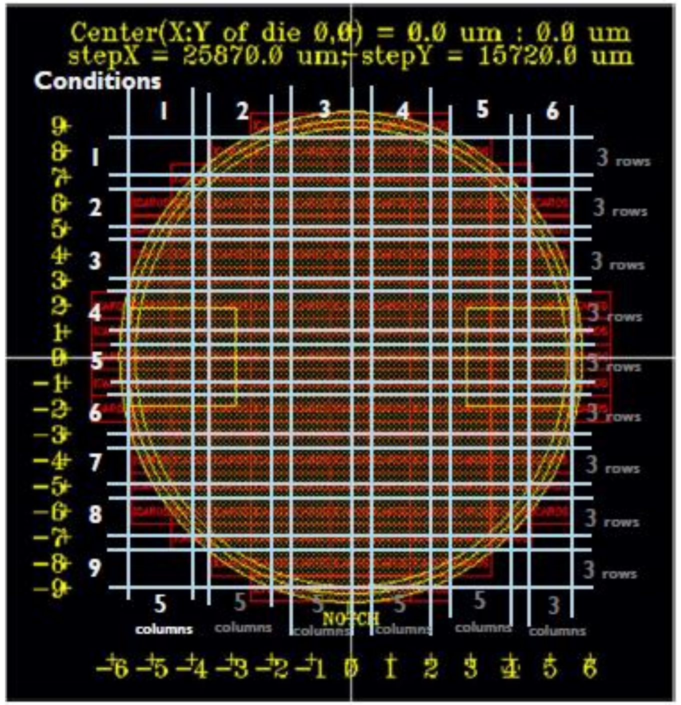
- Dicing is done at the end of the wafer processing to convert it into **individual dies**
- Dicing is done on the **finished wafer** after various cost intensive processing steps, it is crucial to avoid any defects introduced by dicing
- Dicing methods:
  - saw blade dicing, laser dicing, and plasma (etch) dicing
- **Blade dicing** is the original and still the most used and versatile method
- Blade dicing is a **mechanical** process that can result in defects on the die
  - **Chipping** on the side walls and cracks in the die
- To optimize the blade dicing process:
  - All parameters need to be **tuned** to the use case, and
  - constantly **monitored** during dicing
- Parameters are blade rotational speed and feed rate.





SIEMENS

Design Of Experiment – Dicing Wafer Map with Dicing Conditions



DICING CONDITIONS

- Parameters:
  - Feed rate: 5mm/sec; 10mm/sec; 25mm/sec
  - Spindle speed: 30000rpm; 40000rpm; 50000rpm

Conditions	1	2	3	4	5	6	7	8	9
Spindle speed (rpm)		30000			50000			40000	
Feed rate (mm/s)	5	10	25	25	10	5	5	10	25

SIEMENS

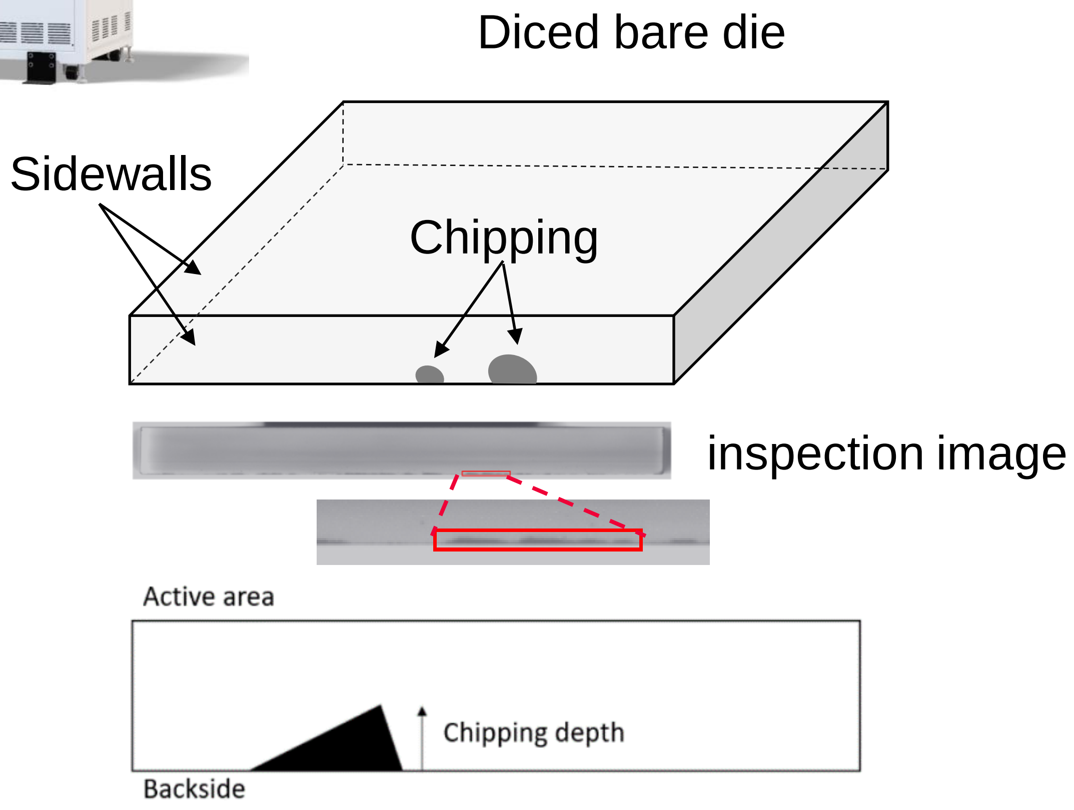
Data Collection – Die Sidewall IR Inspection – Chipping Depth as target

- KLA-ICOS F160XP die sorting and inspection system for Outgoing Quality Control (OQC), Known Good Die (KGD), Unit Level Traceability (ULT)

- Provides IR inspection for robust detection of invisible killer crack and chipping defects for fan-in wafer-level packages, memory chips and bare die

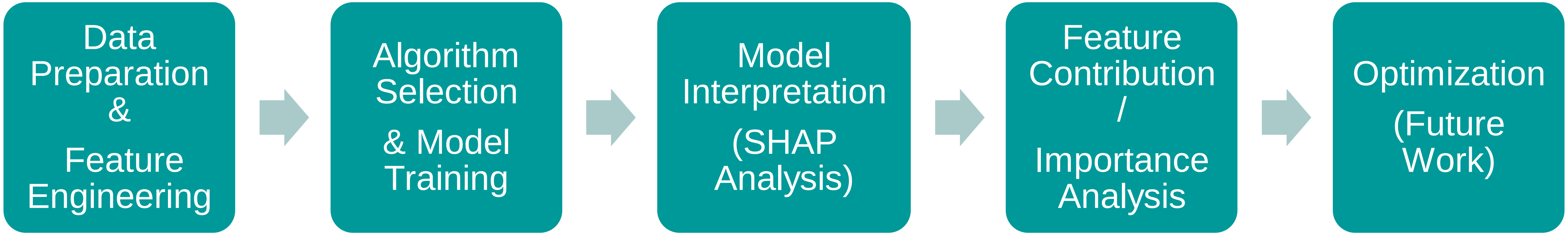
- IR2.0 module combines optical and IR inspection of sidewall providing high throughput and die sorting accuracy

- Die sidewalls were inspected for chipping defects and chipping depth was measured for use in ML



SIEMENS

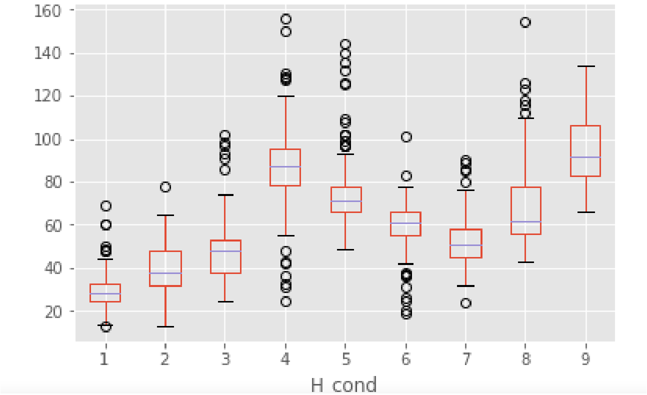
ML-based Flow Steps in ML Software Platform

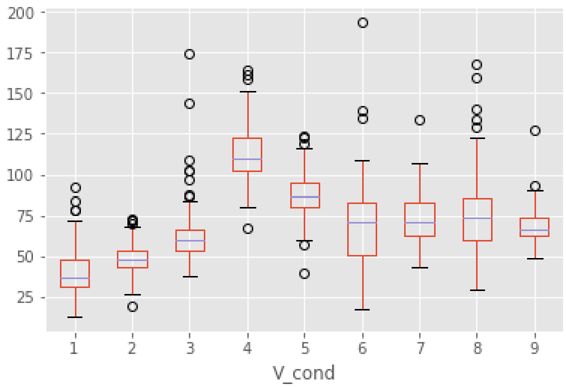


ML-based flow: Data Preparation and Feature Engineering

- Decomposing raw data (dicing condition and chipping depth) per die to be ready for ML model building
- Data enrichment with die spatial data (Radial distance from wafer center point) and diced edge orientation
- Split data to horizontal and vertical, while considering the process conditions of the orthogonal edge (i.e., hypothesis of some stress cross effect)

- Statistical analysis is applied on vertical/horizontal edges, by box plotting the target results (depth in um), split by the 9 distinct dicing conditions





Conditions	1	2	3	4	5	6	7	8	9
Spindle speed (rpm)		30000			50000			40000	
Feed rate (mm/s)	5	10	25	25	10	5	5	10	25

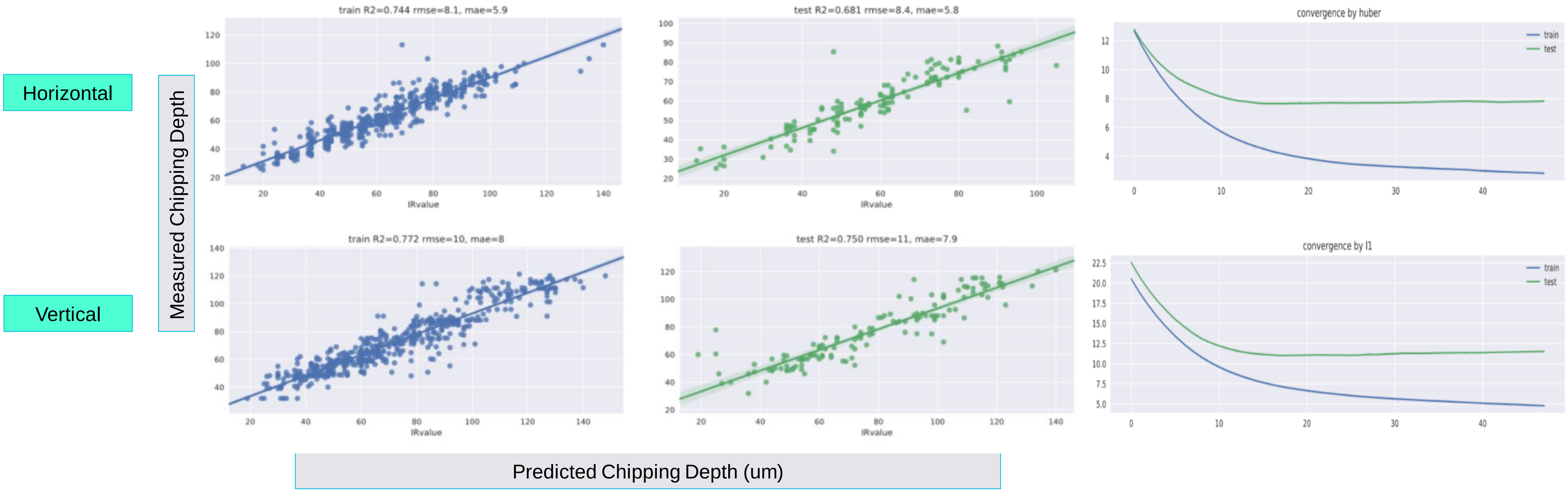
- Simple inference: condition # 1 is showing the least mean value of chipping depth ➔ preferrable direction
- Important pending question: which individual parameter is more effective, and how ranked compared to other design or process parameters? ➔ use ML-based feature importance ranking approach

ML-based flow: Selected Algorithm - Gradient Boosted Forests Advantages

- The structure of the trees carries interpretable information about the data
- Relatively fine control on the aspects of the trees
- Control the order of interaction between the features
- Explicit control on overfitting and accuracy
- Not sensitive to the scale of the input data
- Deal with mixture of categorical and numerical input data
- Deal with missing data (imputation)
- Control over subsampling
- Suitable for huge datasets as well as shallow data
- In hyper-parameter optimization, tree structure is relatable to performance

ML-based flow: Model Training and Results Analysis

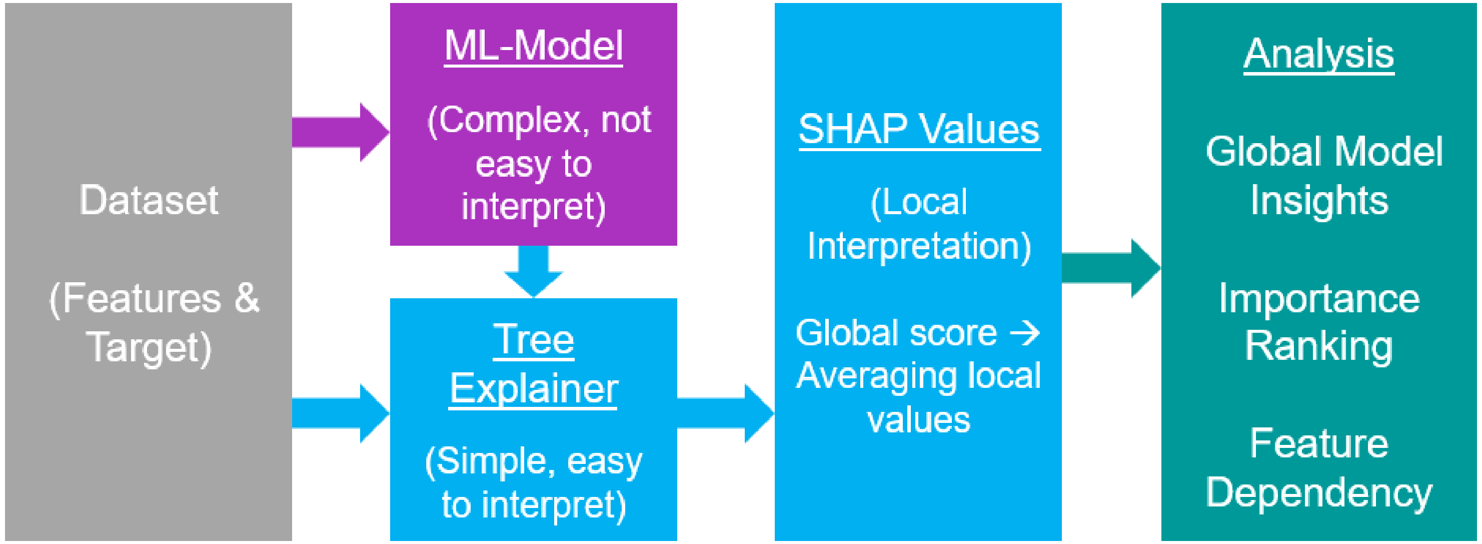
- Our ML model explains target variability from the input variability for a given **multivariate optimization** problem
- Rmse = measure of the **model prediction accuracy** on unseen data
- The closer test dataset (blue) to training dataset (green), the **better variance** the model is



ML-based flow: Model Interpretation

SHAP (SHapley Additive exPlanations)

- Uses simple explanation model to interpret the parameters contribution in the complex ML model
- Derives the global importance from local explanations of individual prediction samples
- Overall SHAP score of each input feature = Average of the Marginal contributions across all permutations



Lundberg, Scott M. and Su-In Lee. “A Unified Approach to Interpreting Model Predictions.” *NIPS* (2017)

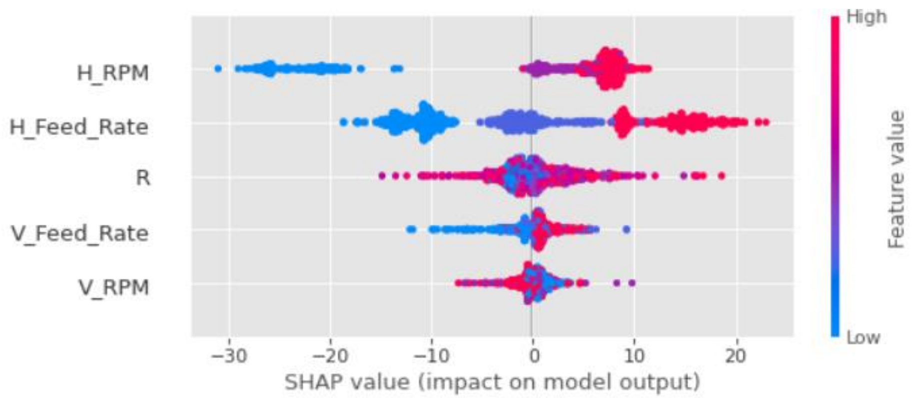
ML-based flow: Feature Importance Ranking

SHAP Analysis

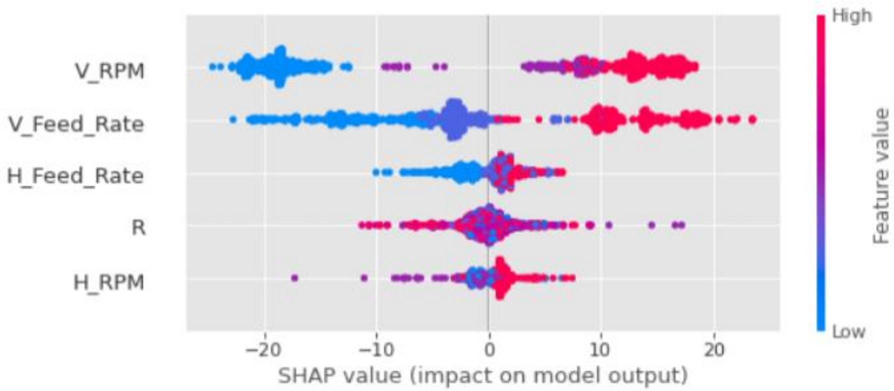
- **Ranking features** based on impact in the model's decision making
- **Quantify contribution** per each feature to the model's prediction value

- The figures show feature ranking formulated by the density distribution of all the individual prediction samples
- Features with distributions of longer tails imply the strong impact on ML model outcome and accordingly its importance
- The red/blue colouring is an indication on the overall correlation between each feature and ML model output

Features Ranked by Impact high-to-low



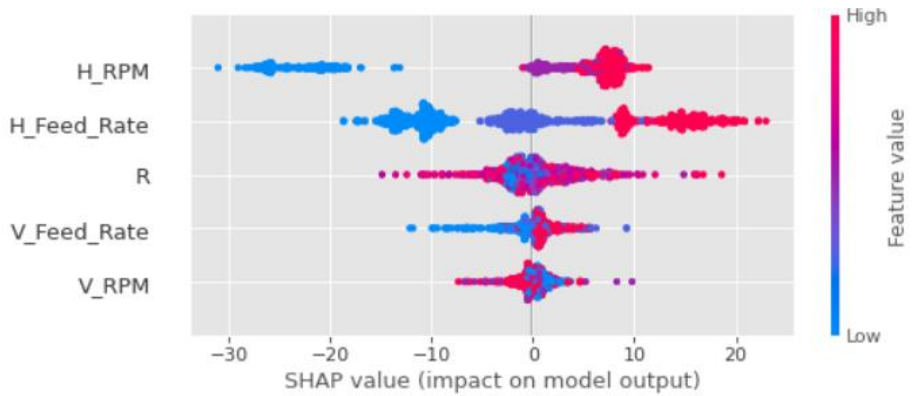
Horizontal Edges



Vertical Edges

ML-based flow: Feature Importance Outcome

Features Ranked by Impact high-to-low



Horizontal Edges

Vertical Edges

- **Highest ranked feature** is the horizontal spindle speed (**H_RPM**) followed by the horizontal feed rate (**H_Feed_Rate**) for the horizontal cut process and similar equivalent conclusion for the vertical cut process
- The colours plot confirms that the **lower RPM** and feed rate will lead to **smaller chipping depth**.
- The impact of the **orthogonal direction conditions** is mild. However, given that horizontal cuts are done first, so horizontal feed rate comes 3rd in ranking of feature importance for vertical cut results, again with positive correlation direction.
- For the **radial distance (R)** the impact was medium on the experimented data, and there is a mixture of red and blue dots, so maybe adding results from more wafers can help in revealing some impact here.

ML-based flow: Further Analysis Capabilities not covered here

Dependency Plot

- Simultaneous analysis of 2 features to understand their interactions before initiating process change.

Highlighting worst/best datapoints

- Feature Ranking for model top/bottom 10 predicted datapoints
- With an optimization goal of focusing on reducing the worst-case conditions. Sort of zoom-in to local impacts of features on few data points instead of the global ranking

Optimization

- Provides “recommended values” for Top Ranking Features
- Optimized values calculated based on operating conditions provided by the user
- Up to 10 features can be concurrently optimized

Summary

- Deploying the state-of-the-art technologies in wafer dicing and chipping inspection
- Applying ML-based contemporary methodology to analyze the impact of individual process features, and to give insight on the direction for process optimization
- Demonstrating the concept of smart manufacturing flow by combining process, metrology and design data under one model, with an end goal to improve yield and quality

MADEIN4

Part of the work presented in this project has received funding from the ECSEL Joint Undertaking (JU) under grant agreement No 826589. The JU receives support from the European Union's Horizon 2020 research and innovation program and Netherlands, Belgium, Germany, France, Italy, Austria, Hungary, Romania, Sweden and Israel

SIEMENS

THANK YOU !

Questions ?

SIEMENS

Contact

Mohamed Bahnas
Siemens Digital Industries Software

E-mail Mohamed.bahnas@siemens.com

SIEMENS