PRODUCTION-LEVEL ARTIFICIAL INTELLIGENCE APPLICATIONS IN SEMICONDUCTOR MANUFACTURING

Chen-Fu Chien
Department of Industrial Engineering and Engineering Management
National Tsing Hua University
No. 101, Section 2, Kuang-Fu Road
Hsinchu, 30013, TAIWAN

Hans Ehm
Infineon Supply Chain Innovation
Infineon Technologies AG
Am Campeon 1-15
Neubiberg, 85579, GERMANY

John Fowler
Department of Supply Chain Management
Arizona State University
300 Lemon Street
Tempe, AZ 85249, USA

Lars Mönch
Department of Mathematics and Computer Science
University of Hagen
Universitätsstraße 1
Hagen, 58097, GERMANY

Cheng-Hung Wu
College of Management
National Taiwan University
No. 1, Sec. 4, Roosevelt Rd., Da’an Dist
Taipei City 106, TAIWAN

ABSTRACT

This is a panel paper which discusses the use of Artificial Intelligence (AI) techniques to address production level problems in semiconductor manufacturing. We have gathered a group of expert semiconductor researchers and practitioners from around the world who have applied AI techniques to semiconductor problems and the paper provides their answers to an initial set of questions. These serve to provide a description of the AI work that has taken place already and to make suggestions for future directions in this arena.

1 INTRODUCTION

The increasing availability of data, advances in computational and storage capacities of IT systems, and algorithmic advances in Artificial Intelligence (AI), especially Machine Learning (ML), combine to make it possible for significant improvements in the efficiency, operations and throughput of manufacturing systems at the production level. The semiconductor industry is one of the most data-intensive industries and has seen increased use of several AI-based technologies over the last few years. In order to develop effective AI-based technologies in the semiconductor manufacturing industry several issues have to be taken into account including scalability, heterogeneity of data, and the need for interpretability.
The aim of the panel is to gather the thoughts of some leading researchers and practitioners regarding: 1) what semiconductor production problems could benefit from AI applications; 2) their experiences in applying AI-based techniques to production problems; 3) what production problems could be addressed using AI-based techniques; 4) what (if any) semiconductor production problems can be better solved by other techniques; and 5) what are some potential obstacles that need to be overcome to allow for the successful application of AI-based techniques. We have structured the paper around a set of questions on these topics. Each of the panelists provides their answer in turn. We use the panelists’ initials to identify who has written each response. The paper finishes with a short conclusion that aims to bring together the key points raised by the panelists.

2 QUESTIONS

2.1 What Makes a Semiconductor Manufacturing Problem on the Production-Level Amenable to the Application of Artificial Intelligence Techniques?

CFC: Semiconductor manufacturing on the production-level consists of lengthy complex processes that are reentrant competing for limited equipment capacity, subject to waiting time and precedence constraints and frequency-based setups. Meanwhile, as semiconductor manufacturing continuously migrates to advanced technology nodes for feature shrinkage, process challenges for yield enhancement have called for interrelated efforts involving advanced process control, advanced equipment control, and advanced quality control. With the increasing computational capabilities and the wide adoption of multimode sensors, internet of things, intelligent equipment, edge computing, and 5G, it is increasingly amenable to employ artificial intelligence techniques and big data analytics for semiconductor manufacturing on the production-level.

HE: The data available on the equipment level, on the process level, and on the yield level are only at a first level statistically distributed. If you have a more detailed look they follow certain patterns. AI and deep learning (DL) is very powerful in pattern recognition so it is amenable for production level, especially in the direction of predicting process parameters from machine parameters and yield from process parameters or in proposing equipment maintenance. The same is true for understanding customer order behavior, if there is an under-planning, an over-planning or another pattern. Again the power of pattern recognition can be used.

LM: Planning and control problems in semiconductor supply chains are challenging due to the sheer size of the manufacturing facilities and the related supply chains. There is a high degree of uncertainty in the involved processes, but also for customer demand (Mönch et al. 2013). AI techniques are appropriate to deal with large-scale decision problems in dynamic and uncertain situations which require adaptive behavior of the decision-making procedures to the current situation (Russell and Norvig 2020). Despite the current AI hype, the application of AI techniques in semiconductor manufacturing has a long history which goes back to the late eighties and early nineties of the last century (Holman and Kempf 1988; Subbiah and Bodenstab 1992; Kempf 1993).

CHW: The manufacturing of 300mm semiconductor wafers requires thousands of processing steps. Enabled by the automated material handling systems (AMHS), manufacturing decisions could be executed fully automated in such very complex reentrant manufacturing systems. Currently, manufacturing decisions are made mainly by rule-based heuristics monitored by human decision-makers. Due to the limited complexity of decision rules and the limited rationality of human decision-makers, the current manufacturing decision framework fails to achieve fab-wide coordinated manufacturing decisions and fails to efficiently respond to dynamic events, such as machine failures. With artificial intelligence techniques, a higher level of coordination and seamless responses to manufacturing dynamics could be achieved with less human intervention and the potency of the automated manufacturing system could be fully unleashed.
2.2 Have You Applied Artificial Intelligence to any Semiconductor Manufacturing Problems on the Production-Level? If so, Please Briefly Describe.

CFC: Focusing on realistic needs, we have developed artificial intelligence solutions and big data analytics for semiconductor manufacturing as follows.

(1) Overlay error control: To address the increasing challenges for overlay error control due to nano technology nodes, we have integrated artificial intelligence technologies, big data analytics, statistical approaches, and in-situ decision support such as support vector regression (SVR) and a mixed-effect least-square support vector regression (LS-SVR) control system via a series of empirical studies sponsored by leading semiconductor companies in Taiwan (Chien et al. 2003; Chien and Hsu, 2009; Chien et al. 2014; Khakifirooz et al. 2018; Khakifirooz et al. 2019a; Khakifirooz et al. 2019b).

(2) Production status prediction and cycle time reduction: Due to the lengthy complex wafer fabrication process, we have integrated a back-propagation neural network (BPNN), a Gauss-Newton regression method, and simulation to determine control windows of WIP Levels of tool sets, forecast arrival rates, and determine the allocation of interchangeable tool sets for cycle time reduction (Kuo et al. 2011; Chien et al. 2012; Yu et al. 2017; Chien et al. 2020b).

(3) Demand forecast: Focusing on realistic needs, we constructed a UNISON framework that integrates deep reinforcement learning (RL) for dynamically selecting the optimal demand forecast model for the corresponding demand patterns to empower smart production and supply chain resilience (Chen and Chien, 2018; Fu and Chien, 2019; Chien et al. 2020c).

(4) Chiller configuration optimization for energy saving: To fulfill the cooling load demand and minimize electricity consumption, we integrate a cooling load forecasting model based on the SARIMAX model for optimizing the operational efficiency of the chillers and a chiller health evaluation model for prognostic and health management (PHM) for optimizing the combination of the operating chillers for chiller configuration strategy that have been effectively employed for semiconductor fabs and TFT-LCD manufacturing (Chien et al. 2020a).

HE: We have had activities for several years now in applying AI on the manufacturing level e.g. for predictive maintenance but also for recognizing customer ordering behavior (COB). On the example of COB we generate two dimensional heat maps and apply several methods like Gradient Boosting, Random Forest, Multilayer Perceptron or Convolutional Neural Networks which performed the best to classify the heat maps in several categories including Over-planning, Up-down, Constant, Up-Down-Up, Double Booking, Under-Planning, Random, Down-Up, Down-Up-Down and 10*booking.

LM: We apply multi-agent systems (MAS), a technique from distributed AI (DAI), to solve scheduling problems for wafer fabs. The FABMAS prototype implements a distributed hierarchical approach to solve complex job shop scheduling problems for the MIMAC I model (Fowler and Robinson 1995) in a rolling horizon setting (Mönch et al. 2006a). The FABMAS system was later extended from a single-fab setting to an entire semiconductor supply chain. The resulting prototype is called S2CMAS (Herding and Mönch 2016). Both the FABMAS and the S2CMAS prototypes are implemented based on the ManufAG framework (Mönch and Stehli 2006).

AI techniques from the machine learning area, namely artificial neural networks and inductive decision trees, are used to choose the look-ahead parameters in the Batched Apparent Tardiness Cost (BATC) dispatching rule (Mönch et al. 2006b). The BATC rule is appropriate to solve scheduling problems for batch processing machines.
Genetic algorithms (GAs) are nature-inspired search strategies. They are used to solve various scheduling problems for wafer fabs, very often scheduling problems for batch processing machines (cf., for instance, Sobeyko and Mönch 2016; Mönch and Roob 2018; Rocholl et al. 2020 amongst others). But we also apply GAs to solve NP-hard planning problems for wafer fabs (cf. Ponsignon and Mönch 2012 for a GA application in master planning). GAs are appropriate to deal with conflicting objectives typical for many scheduling problems (Rocholl et al. 2020).

Constraint programming (CP) is another AI technique that is used to solve scheduling problems for flexible flow shops with time constraints between consecutive process steps (Cailloux and Mönch 2019). CP techniques are appropriate when the scheduling problems are highly constrained and even computing feasible schedules is not straightforward.

CHW: When the queue time constraints (time window constraints) become stricter and consecutive processes steps suffers serial time constraints, the level of coordination between manufacturing decision is critical. Under time constraints, coordinating unreliable upstream and downstream workstations requires stochastic optimization models that could not be solved efficiently. DL and reinforcement learning are effective tools for such coordination that requires risk awareness across multiple workstations.

2.3 What are some of the Semiconductor Manufacturing Problems on the Production-Level that could be Addressed using Artificial Intelligence Techniques?

CFC: Semiconductor manufacturing production-level problems for (1) yield enhancement including engineering data analysis, FDC, R2R, PHM, APC/AEC/AQC, and optimal parameter setting for multiple factors, (2) fab-wise production planning, resource allocation, and production control for smart manufacturing for both productivity and yield, (3) back end operations including IC assembly and testing, and (4) supply chain effectiveness and virtual vertical integration.

HE: Wherever there are many data available, we can generate patterns and then use the patterns for classification or for forecasting. This includes equipment machine data for better maintenance planning and predicting process data, process data to predict SPC (Statistical process data) from Sample data for virtual metrology but also for yield forecasting; Yield data for yield patterns (including wafer maps) and use this for supply prediction.

LM: Dispatching is still an important production control technique for wafer fabs. Genetic programming seems to be a promising technique to automatically discover dispatching rules for complex manufacturing systems. Only initial results are available in this direction (Hildebrandt et al. 2014).

AI applications seem to be possible for production scheduling. For instance, it is interesting to use machine learning techniques to design triggers for rescheduling activities. A situation-dependent parameterization of scheduling heuristics seems to be another application area for AI techniques. Only initial results are available (Osisek and Aytuk 2004). It is also expected that GA-type heuristics can be used to design robust scheduling procedures that can cope with data uncertainty.

Another promising area is order release and more general production planning for wafer fabs. Some initial work is reported by Schneckenreither et al. (2020). It seems also possible to apply genetic programming to discover order release rules.

Different AI techniques, mainly machine learning, can be used to predict cycle times in wafer fabs. The resulting estimates, known as lead times, might be, for instance, important for master and production planning.

Another promising area is demand planning. For instance, it might be possible to capture the customer behavior by machine learning.
Automated parameter tuning for various planning and control algorithms is another task that can be supported by machine learning. Fully-automated parameter tuning is important for future automation efforts of decision-making processes in semiconductor supply chains.

It is also expected that next-generation planning and control systems for semiconductor manufacturing are more software agent-based.

**CHW:** In addition to the dynamic dispatching problems under time window constraints, the operation efficiency of material handling systems could also be significantly enhanced by artificial intelligence techniques and the design of multi-agent systems.

### 2.4 What, if any, Semiconductor Manufacturing Problems on the Production-Level would not Benefit from Artificial Intelligence Applications?

**CFC:** Indeed, various AI techniques should be employed to empower the solutions for different semiconductor manufacturing problems on the production-level. Nevertheless, effective solutions shall integrate AI technologies and other approaches such as soft computing, big data analytics, and digital decision.

**HE:** Wherever there is a high knowledge domain of operators, technicians and engineers available and human interaction based on experience drive the main decision, AI has problems to surpass human decisions. We have had several experiences where the AI result was quite good (from an AI perspective) but not good enough to surpass human decisions.

**LM:** A key aspect of any successful decision-making activity is the analysis and understanding of the problem at hand. Afterwards, an appropriate solution technique can be chosen. Since in the opinion of the panelist most of the successful industrial decision-making procedures are hybrids of different techniques the panelist sees potential of AI techniques for the majority of the semiconductor manufacturing problems on the production level. This is especially true for automated parameter tuning and for a meaningful automation of decision-making processes.

**CHW:** It depends on the definition of AI. When the narrower definition of artificial intelligence is limited to ML techniques, the limitation of AI is obvious. For example, when the production level problems lack the smoothness required by AI methods, traditional rule-based or optimization-based methods will be more efficient. (e.g. mixed integer linear programming (MILP) problems could not be replaced by AI in general, but Markov decision process (MDP) problems with smooth optimal policy structure could be).

### 2.5 What Hurdles must be Overcome in Order to Effectively Apply Artificial Intelligence Techniques to Semiconductor Manufacturing Problems on the Production-Level?

**CFC:** Most AI techniques require large amounts of data for training, while advanced technology ramping requires quick response under limited samples. To enhance solution effectiveness and quality, it will be more customized for specific problem setting that make AI solution transferrable to other problems in different contexts, not to mention potential issues for explainability, transparency, and traceability. Furthermore, as more technologies can be employed to address semiconductor manufacturing production-level problems with increasing scope and scale, limitations of the existing studies can be traced in part to the lack of a framework within which different technologies can be integrated, while domain knowledge can be incorporated for systematic decision analysis under uncertainty (Chien et al. 2020d; Fu et al. 2020).

**HE:** The primary problem is to understand the interconnection of data. We call it the semantics of the data. For that we built up a semantic web, an ontology of manufacturing data beyond the pure data lake where
the data are. The semantic web should also allow to add a priori know how from the high knowledge of
the people. First results on improving AI results using this are promising. Having good results and applying
the good results are two sides of the coin. When humans like Operators, Technicians, Engineers or Planners
are involved they usually demand a reason for the decision. So explainable AI (xAI) would be beneficial.
However, applying xAI often reduces the overall quality of the result. An alternative is to use the AI results
without explanation but this requires an alternative way of working together between the human and AI.
We featured this by setting up a game, which we call the HAI (human AI game) where people learn to trust
AI when it is in its area of competence but also recognize when it is not and bravely take over.

LM: Given the history of AI techniques in semiconductor manufacturing despite the current hype there is
little hope in the opinion of the panelist that the application of AI techniques in a black box manner will
revolutionize decision-making in semiconductor planning and control. The long lasting problems in
planning and control remain. Without a deep understanding of the problems it seem hard to obtain AI-based
solutions that go beyond the state-of-the-art. The panelist also believes that the tool box of the future for
decision-making will include different highly tailored methods that need to be combined in a problem-
specific manner. AI is only one of these tools.

Another hurdle is the fact that AI applications often require advanced computer science skills which
also hinders to some extent a widespread penetration of AI into real-world applications. For instance,
designing, engineering, and maintaining a MAS for a wafer fab is a highly non-trivial task.

Although data availability is better than 30 years ago and will be further improved due to additional
sensors and better hard- and software infrastructures, AI applications must be integrated into the existing
information system landscape of semiconductor companies which is challenging. When new paradigms
such as service-oriented architectures and software agents replace existing information systems in
semiconductor supply chains, the integration of AI techniques will be easier.

CHW: The primary problem is to understand the interconnection of data. We call it the semantics of the
data. For that we built up a semantic web, an ontology of manufacturing data beyond the pure data lake
where the data are. The semantic web should us also allow to add a priori know how from the high
knowledge of the people. First results on improving AI results using this are promising. Having good
results and applying the good results are two sides of the coin. When humans like Operators, Technicians,
Engineers or Planners are involved they usually demand a reason for the decision. So xAI would be
beneficial. However, applying xAI often reduces the overall quality of the result. An alternative is to use
the AI results without explanation but this requires an alternative way of working together between the
human and AI. We featured this by setting up a game, which we call the HAI (human AI game) where
people learn to trust AI when it is in its area of competence but also recognize when it is not and bravely
take over.

3 CONCLUSION

Our group of expert semiconductor researchers and practitioners from around the world provided their
insights on the application of Artificial Intelligence techniques to semiconductor problems. This included
their own experiences in applying AI and identifying some additional problems that could benefit from
the use of AI. They also discuss some limitations of the use of AI for production level semiconductor
manufacturing problems and potential obstacles that might hinder their use.

REFERENCES


Chien, Ehm, Fowler, Mönch, and Wu


AUTHOR BIOGRAPHIES

CHEN-FU CHIEN is Tsinghua Chair Professor & Micron Chair Professor, in the Department of Industrial Engineering & Engineering Management, National Tsing Hua University (NTHU), Hsinchu, Taiwan. He is the Director of Artificial Intelligence for Intelligent Manufacturing Systems (AIMS) Research Center that is one of four national AI centers sponsored by Ministry of Science & Technology (MOST), Taiwan. He is the PI and Director for the NTHU-TSMC Center for Manufacturing Excellence and the Zhen Ding-NTHU Joint Research Center in NTHU, Taiwan. He has received 12 USA invention patents, and published 12 case studies in Harvard Business School and more than 190 journal papers including the 2011 Best Paper Award of IEEE Transactions on Automation Science and Engineering and the 2015 Best Paper Award of IEEE Transactions on Semiconductor Manufacturing. His email address is cfchien@mx.nthu.edu.tw.

HANS EHM is a Master of Science in Mecahnical Engineering from Oregoin State University and holds a diploma degree in Applied Physics from HS Munich. He is Lead Principal Supply Chain heading the Supply Chain Business Innovation department at Infineon Technologies. His email address is hans.ehm@infineon.com.

JOHN W. FOWLER is the Motorola Professor of Supply Chain Management and recently served as Chair of the Supply Chain Management department in the W.P. Carey School of Business at Arizona State University. His research interests include discrete event simulation, deterministic scheduling, multi-criteria decision making, and applied operations research with applications in semiconductor manufacturing and healthcare. He has published over 130 journal articles and over 100 conference papers. He was the Program Chair for the 2002 and 2008 Industrial Engineering Research Conferences, Program Chair for the 2008 Winter Simulation Conference (WSC), and Program Co-Chair for the 2012 INFORMS National Meeting. He is an Editor of the Journal of Simulation and an Associate Editor of IEEE Transactions on Semiconductor Manufacturing and the Journal of Scheduling. His email address is john.fowler@asu.edu.

LARS MÖNCH is Professor in the Department of Mathematics and Computer Science at the University of Hagen, Germany. He received a master’s degree in Applied Mathematics and a Ph.D. in the same subject from the University of Göttingen, Germany. His current research interests are in simulation-based production control of semiconductor wafer fabrication facilities, applied optimization and artificial intelligence applications in manufacturing, logistics, and service operations. His email address is lars.moench@fernuni-hagen.de.

CHENG-HUNG WU is an Associate Professor at the Institute of Industrial Engineering, National Taiwan University, Taipei, TAIWAN. He received his Ph.D. and M.S. degree in industrial and operations engineering from the University of Michigan – Ann Arbor in 2005 and 2006. His research interests include: (1) decisions under uncertainties, particularly in applications such as production_supply-chain system analysis and dynamic control; (2) information and decision support systems (especially operations management and supply chain management); and (3) theoretical work in Markov decision processes and stochastic programming. His email address is wuchn@ntu.edu.tw.