Data Driven Optimisation of Spatial Dynamic Sampling Using FSCA

Seán McLoone, Federico Zocco, Queen's University Belfast

Introduction

Metrology plays an essential role in meeting process performance targets in semiconductor manufacturing; however it is an expensive, time consuming and non-value added activity. Consequently, monitoring of silicon wafers is performed via spatial (intra-wafer) and temporal (inter-wafer) sampling (Fig. 1). The challenge is to devise a dynamic sampling plan that minimises the number of locations measured on each wafer while retaining the capacity to: (1) generate an accurate reconstruction of the wafer profile, and; (2) detect previously unseen process behaviour in a finite time horizon.

Dynamic spatial sampling

We adopt a data driven approach [1, 2]:

Training phase:

- Historical wafer metrology data is collected into a data matrix (***X***) for a set of candidate measurement sites
- FSCA [3] is used to determine the most representative subset among this candidate set in the sense of the total variance explained metric
- Remaining sites are then clustered around the selected FSCA sites based on temporal correlation (Fig.2)

Operation phase:

- For each new wafer, one sample from each FSCA cluster is used to define the measurement plan (a different one for each wafer)
- Using the historical data, ***X***, VM models are estimated to reconstruct unmeasured sites using the measured sites as regressors
- Unmeasured sites are estimated using the VM models
- A 3-D wafer profile reconstruction is generated using the real and virtual measurements (Fig. 3)

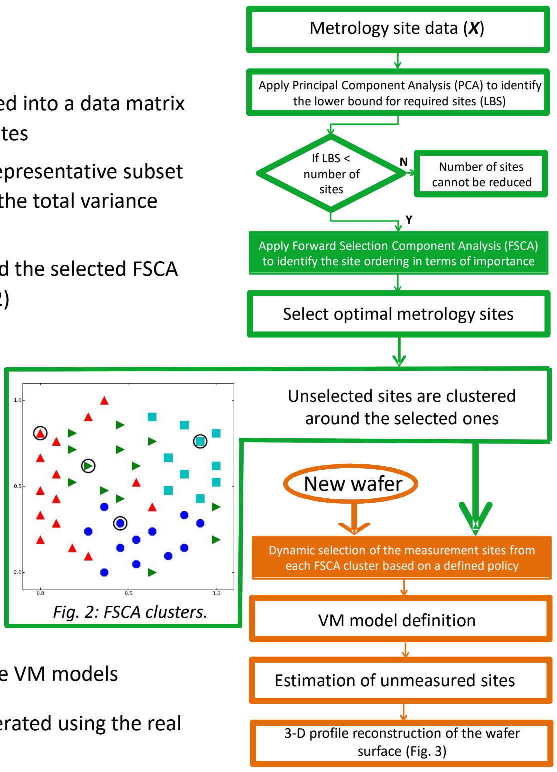

Fig. 2: FSCA clusters.

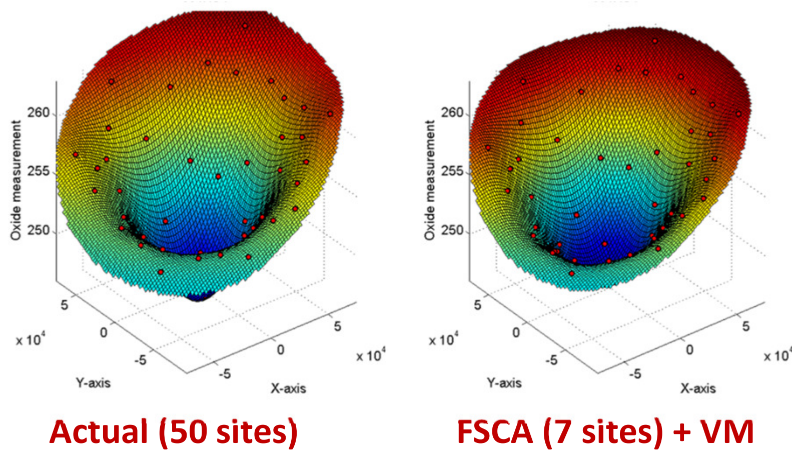

Fig. 3: Wafer profile reconstruction using the actual measurements (left) and the FSCA-based methodology (right).

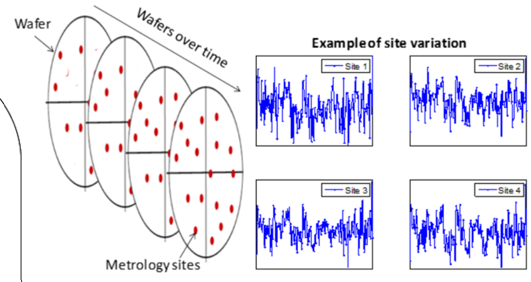

Fig. 1: Temporal and spatial sampling performed to monitor the wafer surface.

FSCA

FSCA is a greedy algorithm [3, 4] that estimates the subset of variables that best represents the information contained in a set of candidate variables ***X*** using an iterative procedure consisting of two steps:

(1) the variable with maximum contribution to the variation observed in ***X*** is identified;

(2) the contribution of the identified variable is removed from ***X***.

Case study

Monitoring of a PVD process where 50 sites were recorded historically.

- PCA shows that 5 PCs can explain 99% of the observed variance over the 50 sites
- FSCA shows that 7 FSCs (i.e. 7 sites out of 50) are sufficient to achieve 99% reconstruction accuracy (Fig. 4)

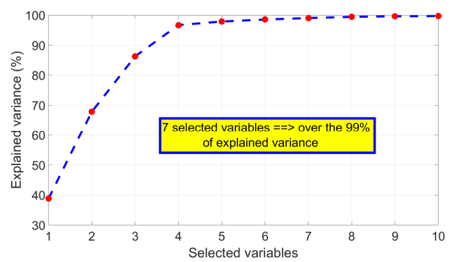

Fig. 4: Explained variance as function of the number of FSCs (selected variables).

References

[1] **S. McLoone**, A. Johnston, and G.A. Susto, A Methodology for Efficient Dynamic Spatial Sampling and Reconstruction of Wafer Profiles, *IEEE Transactions on Automation Science and Engineering*, (online January 2018, DOI: 10.1109/TASE.2017.2786213.

[2] **S. McLoone**, **F. Zocco**, M. Maggipinto, G.A. Susto, On Optimising Spatial Sampling Plans for Wafer Profile Reconstruction, *3rd IFAC Conference on Embedded Systems, Computational Intelligence and Telematics in Control, CESCIT 2018*, Faro, Portugal, 6-8 June 2018.

[3] L. Puggini, **S. McLoone**, Forward Selection Component Analysis: Algorithms and Applications, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39(12), pp. 2395-2408, December 2017, (online Jan 2017, DOI: 10.1109/TPAMI.2017.2648792).

[4] A.A. Bian, J.M. Buhmann, A. Krause, and S. Tschiatschek, Guarantees for greedy maximization of non-submodular functions with applications, *arXiv preprint arXiv:1703.02100*, 2017.

Email: {s.mcloone,fzocco01}@qub.ac.uk

QUEEN'S UNIVERSITY BELFAST

School of Electronics, Electrical Engineering and Computer Science

i-AMS
Centre for Intelligent Autonomous Manufacturing Systems