Engineering Applications of Artificial Intelligence 67 (2018) 126–135

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

An enhanced variable selection and Isolation Forest based methodology for anomaly detection with OES data

Luca Puggini [a,*], Seán McLoone [b]

[a] National University of Ireland Maynooth, Maynooth, Ireland
[b] Queen's University Belfast, Belfast, Northern Ireland, BT9 5AH, United Kingdom

ARTICLE INFO

Keywords:
Semiconductors
Fault detection
Dimensionality reduction
OES spectrum
Isolation Forest
Forward Selection Components Analysis

ABSTRACT

The development of efficient and interpretable anomaly detection systems is fundamental to keeping production costs low, and is an active area of research in semiconductor manufacturing, particularly in the context of using Optical Emission Spectroscopy (OES) data. The high dimension and correlated nature of OES data can limit the performance achievable with anomaly detection systems. In this paper we present a dimensionality reducing variable selection and isolation forest based anomaly detection and diagnosis methodology that addresses these issues. In particular, it takes account of isolated variables that can be overlooked when using conventional approaches such as PCA, and provides greater interpretability than afforded by PCA. The proposed methodology is illustrated with the aid of simulated and industrial plasma etch case studies.

1. Introduction

Semiconductor manufacturing is one of the largest industries in the world, employing almost 250,000 people in the USA alone (Yinug, 2015). It posted sales globally totalling 335.2 billion dollars in 2015 (SIA, 2016). It is a highly competitive sector with manufacturers continually delivering new devices that are smaller, faster and/or more energy efficient than previous generations. Keeping pace with these developments, which have largely followed Moore's law (Schaller, 1997), has resulted in the development of complex industrial processes, with product manufacture typically consisting of several hundred processing steps. Among these, plasma etching processes have been identified as critical to the production of semiconductor devices (Coburn and Winters, 1979). Fig. 1 shows the main characteristics of a plasma etch process. Gases are pumped into a chamber where they are excited by microwaves to generate a plasma. The plasma then interacts with the exposed surface of the wafer both chemically and mechanically to etch away the wafer surface in a controlled fashion (Abe et al., 2008).

Non-intrusive plasma monitoring can be achieved either by monitoring its electrical characteristics using a plasma impedance monitor (PIM) or by monitoring its optical output using optical emission spectroscopy (OES). OES monitoring is particularly attractive as it provides real-time information on the plasma chemical composition. This arises due to the unique wavelength signature that exists for each chemical species in the plasma. Data recorded from OES spectrometers consists of measurements of the intensity of the light emitted from the plasma as a function of a discrete set of wavelengths (channels) and time. As such, optical emission spectrometers are increasingly being deployed on plasma etch chambers to monitor plasmas, either directly through an optical window in the chamber, or indirectly through analysis of the exhaust gases from the chamber (as depicted in Fig. 1). Fig. 2 shows a typical spectrum generated by a plasma etch process. Several studies have shown OES to be an effective wafer processing monitoring signal (e.g. Chen et al. 1996, Puggini et al. 2014). It has also been employed for applications such as fault detection (Yue et al., 2000) and etch rate prediction (Puggini and McLoone, 2015; Zeng and Spanos, 2009).

Anomaly detection, in particular, is an active area of research in semiconductor manufacturing as the ability to detect faults early, and recognize anomalous behaviour in processes is key to improving product quality, overall process yield and throughput (He and Wang, 2007). Recent examples include Puggini et al. (2016) and Mahadevan and Shah (2009) who perform anomaly detection in OES time series data using unsupervised random forest and one class support vector machines (OC-SVM), respectively, and Ren and Lv (2014), He and Wang (2007), Verdier and Ferreira (2011) who employ clustering based methodologies to separate normal and anomaly samples.

Anomaly detection with OES data is a challenging problem, due to its high dimension and highly correlated variables (Prakash et al.,

* Corresponding author.
E-mail addresses: lpuggini@eeng.nuim.ie (L. Puggini), s.mcloone@qub.ac.uk (S. McLoone).

https://doi.org/10.1016/j.engappai.2017.09.021
Received 9 January 2017; Received in revised form 2 September 2017; Accepted 19 September 2017
0952-1976/

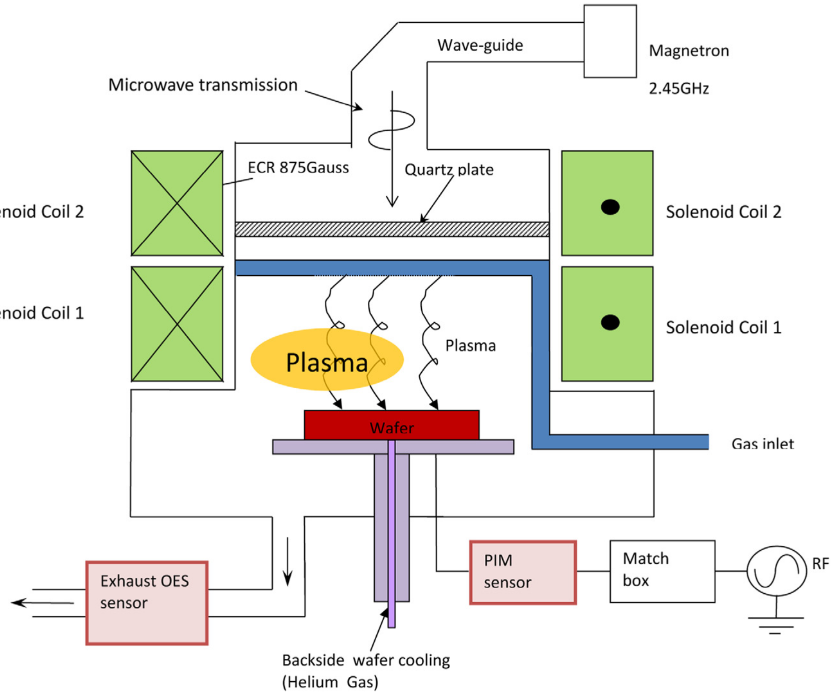

Fig. 1. A plasma etching chamber.

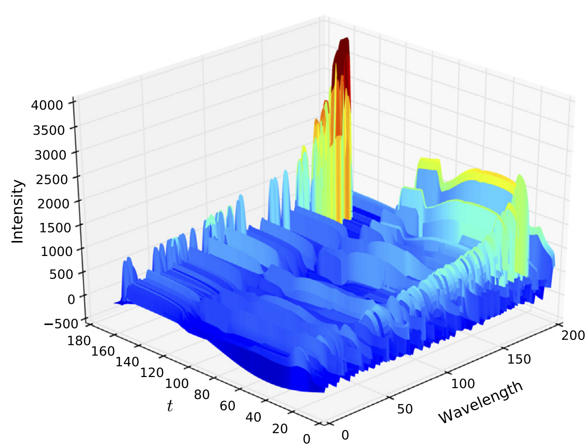

Fig. 2. A typical OES spectrum from the case study presented in Section 5.

2012). Both of these characteristics pose problems for anomaly detection algorithms. Most anomaly detection algorithms are based on a distance measure and it is known that such measures are unreliable in high dimensional spaces due to the so-called curse of dimensionality (Kriegel et al., 2008). As will be illustrated later in the paper, high levels of correlation among variables can degrade the performance of anomaly detection algorithms, as a small shift outside of the normal values in a group of correlated variables may generate a more anomalous result than a large shift in an isolated variable.

In this paper we propose a methodology for unsupervised anomaly detection and diagnosis using historical OES data that addresses both the dimensionality and correlation challenges. This consists of a dimensionality reduction pre-processing step, anomaly detection using the Isolation Forest algorithm (Liu et al., 2008), and a novel anomaly diagnosis procedure based on interrogation of the Isolation Forest (IF) model. In particular, building on our preliminary work in Puggini and McLoone (2016), we propose variable selection based dimensionality reduction techniques as a means of enhancing the interpretability of the IF model and improve its performance in the context of anomaly detection. In Puggini and McLoone (2016) we proposed a Forward Selection Independent Variables (FSIV) algorithm as an unsupervised variable selection technique specifically designed for anomaly detection. Here, we extend this work with more comprehensive industrial case studies, the introduction of a new variant of the algorithm that performs variable selection based on minimizing the maximum reconstruction error (and referred to as FSMM), and the development of a novel Isolation Forest based anomaly detection and diagnosis procedure.

The remainder of the paper is organized as follows. Section 2 introduces FSIV and FSMM and also briefly describes the underpinning Forward Selection Component Analysis (FSCA) (Puggini and McLoone, 2017) algorithm. Section 3 provides an overview of IF and introduces the novel IF based fault diagnosis procedure. It also discusses the limitations of IF based anomaly detection with regard to correlated variables. A simulated example is then presented in Section 4 to illustrate this point, and to highlight the differences between FSIV, FSCA, and FSMM in the context of anomaly detection. For comparison purposes results are also presented for Principal Component Analysis (PCA) (Jolliffe, 2002). The algorithms are then evaluated on two plasma etch process industrial cases studies in Sections 5 and 6. In the first case study OES time series data is available for a plasma etch process where a chamber seasoning effect is known to cause significant performance issues. In the second case study OES time series summary statistics are available for each wavelength for a plasma etch chamber that exhibits faulty behaviour, as evidenced by ground truth etch rate metrology data that is also available. Through these case studies the performance of the various unsupervised variable selection methods with IF based anomaly detection are compared and the application of the anomaly diagnosis procedure demonstrated. Finally, conclusions are presented in Section 7.

2. Dimensionality reduction in anomaly detection

Dimensionality reduction techniques such as PCA and FSCA seek to obtain lower dimensional approximations of datasets that retain the majority of the information in the original high dimensional datasets, usually defined in terms of the percentage of explained variance. PCA