

CHAMBER MATCHING IN SEMICONDUCTOR MANUFACTURING USING  
STATISTICAL ANALYSIS AND RUN-TO-RUN CONTROL

by  
Feyza Haskaraman

Bachelor of Science in Mechanical Engineering  
Massachusetts Institute of Technology, January 2016

Submitted to the Department of Mechanical Engineering  
in partial fulfillment of the requirements for the degree of



MASTERS OF ENGINEERING IN ADVANCED MANUFACTURING AND DESIGN  
at the  
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2016

© 2016 Feyza Haskaraman. All rights reserved.

Signature redacted

Signature of Author: \_\_\_\_\_

Feyza Haskaraman  
Department of Mechanical Engineering  
August 19, 2016

Signature redacted

Certified by: \_\_\_\_\_

Prof. Duane Boning  
Clarence J. Lebel Professor of Electrical Engineering  
Thesis Supervisor

Signature redacted

Accepted by: \_\_\_\_\_

Prof. Rowan Abeyaratne  
Quentin Berg Professor of Mechanics  
Chair, Committee of Graduate Students

The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.

**CHAMBER MATCHING IN SEMICONDUCTOR MANUFACTURING USING  
STATISTICAL ANALYSIS AND RUN-TO-RUN CONTROL**

by

Feyza Haskaraman

Bachelor of Science in Mechanical Engineering  
Massachusetts Institute of Technology, January 2016

Submitted to the Department of Mechanical Engineering  
in partial fulfillment of the requirements for the degree of

**MASTERS OF ENGINEERING IN ADVANCED MANUFACTURING AND DESIGN  
at the  
MASSACHUSETTS INSTITUTE OF TECHNOLOGY**

September 2016

© 2016 Feyza Haskaraman. All rights reserved.

Signature of Author: \_\_\_\_\_  
Feyza Haskaraman  
Department of Mechanical Engineering  
August 19, 2016

Certified by: \_\_\_\_\_  
Prof. Duane Boning  
Clarence J. Lebel Professor of Electrical Engineering  
Thesis Supervisor

Accepted by: \_\_\_\_\_  
Prof. Rohan Abeyaratne  
Quentin Berg Professor of Mechanics  
Chair, Committee of Graduate Students

The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.



77 Massachusetts Avenue  
Cambridge, MA 02139  
<http://libraries.mit.edu/ask>

## **DISCLAIMER NOTICE**

Due to the condition of the original material, there are unavoidable flaws in this reproduction. We have made every effort possible to provide you with the best copy available.

Thank you.

**The images contained in this document are of the best quality available.**

*This page is intentionally left blank.*



# **Chamber Matching in Semiconductor Manufacturing Using Statistical Analysis and Run-to-Run Control**

by

Feyza Haskaraman

Bachelor of Science in Mechanical Engineering  
Massachusetts Institute of Technology, January 2016

Submitted to the Department of Mechanical Engineering  
In partial fulfillment of the requirements for the degree of  
Masters of Engineering in Manufacturing

## **Abstract**

This thesis focuses on a chamber matching methodology for semiconductor manufacturing in Analog Devices Inc.'s fabrication sites. As ADI extends its efforts to implement Internet of Things and predictive maintenance (PdM) to its fabrication facilities, it is also seeking to increase their overall yield by implementing better monitoring and control of processes and matching the performance of chambers. This thesis project was conducted by F. Haskaraman, T. Nilgianskul and T. Nerurkar as a team to make a series of recommendations to improve process yields using statistical control and to show the benefits of chamber matching in particular. Nilgianskul's thesis focuses on the statistical process control and Nerurkar's thesis focuses on Design on Experiments (DOEs).

A chamber matching methodology is created and applied to chambers that run the plasma-ashing process. Using design of experiments, the machines are modeled individually and globally. While individual models reveal the mismatch, a global model is proposed as a step to optimize the process recipes for matching. The root cause of the differences is diagnosed with instrumented wafers and in-situ sensor monitoring. Recommendations are made to standardize the hardware and software along with calibration methods. First batch of streamed raw data from an in-situ thermocouple is analyzed and found to be another tool to monitor the chamber performance differences.

The process is simulated using an EWMA controller and is found to achieve lower mismatch by keeping outputs of the machine closer to the strip thickness in the case of a process drift. At the end of the project, a chamber matching methodology was recommended to the Analog Devices to complement its Internet of Things efforts. By increasing the routing flexibility and decreasing yield variability and tool qualification, this strategy is expected to save significant amount of costs and increase the quality of its products.

**Thesis Supervisors:** Duane Boning, Clarence J. Lebel Professor of Electrical Engineering and Computer Science

Rohan Abeyaratne, Quentin Berg Professor of Mechanics

*This page is intentionally left blank.*

## **Acknowledgements**

I would like to thank to Professor David E. Hardt for teaching us the concepts and the approaches to the technical questions throughout the year. I would like to thank Prof. Duane Boning for not only teaching us but also coaching us and encouraging us to come up with different approaches and providing us with ultimate guidance when we felt discouraged in technical difficulties. Thank you Jack Dillon and Ken Flanders for creating such a great collaboration between Analog Devices Inc. and MIT, where I was able to see best practices of engineering and was able to apply my take on it. Thank you also for making us feel always welcome at the Analog Devices and letting us try different things and offering us the opportunity to work with the best engineers in the world. Thank you Peter Cardillo and Pam Petzold for always helping us in the fab and guiding us through its and bits of the fab. Thank you Jose J. Pacheco for managing our program, Masters of Engineering. Lastly thank you to my friends and family who showed incredible support and love throughout the challenges I have faced.

## TABLE OF CONTENTS

Page

<b>Abstract</b> .....	<b>3</b>
<b>Acknowledgements</b> .....	<b>5</b>
<b>Table of Contents</b> .....	<b>6</b>
<b>List of Figures and Tables</b> .....	<b>9</b>
<b>1. Introduction</b> .....	<b>13</b>
<b>1.1. Background Information on Analog Devices Inc.</b> .....	<b>13</b>
<b>1.2. General Semiconductor Fabrication Process</b> .....	<b>14</b>
<b>1.3. Plasma Ashing Process</b> .....	<b>15</b>
1.3.1. Gasonics A3010 Tool .....	16
1.3.2. Partial and Forming Recipes .....	18
1.3.3. Data Collection and Logging .....	20
1.3.4. Calculation of Basic Statistics .....	22
<b>1.4. Problem Statement</b> .....	<b>24</b>
<b>1.5. Outline of the Thesis and Allocation</b> .....	<b>24</b>
<b>2. Theoretical Review of Key Concepts</b> .....	<b>25</b>
<b>2.1. Statistical Process Control (SPC).</b> .....	<b>26</b>
2.1.1. Origin of SPC.....	26
2.1.2. Shewhart Control Charts .....	26
<b>2.2. Analysis of Variance</b> .....	<b>28</b>
<b>2.3. Design of Experiments</b> .....	<b>29</b>
<b>2.4. Hypothesis Testing</b> .....	<b>31</b>
2.4.1. Z-Test for Detecting Mean Shift.....	31
2.4.2. F-test .....	33
2.4.3. Bartlett's Test.....	34
<b>3. Theoretical Review of Machine Matching</b> .....	<b>35</b>
<b>3.1. Chamber Matching Issues in Semiconductor Manufacturing</b> .....	<b>35</b>
<b>3.2. Chamber Matching Methodology</b> .....	<b>36</b>

<b>3.3. Performance Metric and Key Variables Selection .....</b>	<b>38</b>
<b>3.4. Golden and Inferior Chambers .....</b>	<b>38</b>
<b>3.5. In-Situ Sensor Monitoring .....</b>	<b>40</b>
<b>3.6. Run-to-Run Control .....</b>	<b>40</b>
<b>4. Chamber Matching Methodology and its application to Plasma Ashing .....</b>	<b>43</b>
<b>4.1. Issues in Plasma Ashing Chambers .....</b>	<b>43</b>
<b>4.2. Our Approach to Chamber Matching .....</b>	<b>44</b>
<b>4.3. Performance Metric and Key Variables Selection .....</b>	<b>46</b>
<b>4.4. Identification of Relevant Data .....</b>	<b>46</b>
<b>4.5. Identification of Golden and Inferior Chambers .....</b>	<b>47</b>
4.5.1. Control Charts.....	48
4.5.2. Normality Test.....	50
4.5.3. Box and Whisker Plots.....	51
4.5.4. Hypothesis Testing.....	52
<b>4.6. Design of Experiments and Multivariate Analysis .....</b>	<b>52</b>
<b>4.7. Global Process Modeling .....</b>	<b>55</b>
<b>4.8. Diagnosis of the Causes of the Mismatch .....</b>	<b>59</b>
4.8.1. Diagnosis with Instrumented Wafers .....	59
4.8.2. Diagnosis with Surface Mapping.....	62
4.8.3. Diagnosis with In-Situ Sensors.....	64
<b>4.9. Run-to-run Control .....</b>	<b>69</b>
<b>5. Results and Benefits .....</b>	<b>75</b>
<b>5.1. Results and Discussion .....</b>	<b>75</b>
<b>5.2. Benefits .....</b>	<b>75</b>
5.2.1. Decreased Yield Variability.....	76
5.2.2. Greater Routing Flexibility .....	76
5.2.3. Faster Tool Qualification .....	77
5.2.4. Reduced Time for Root Cause Analysis .....	77

<b>6. Conclusion, Future Work and Recommendations.....</b>	<b>78</b>
<b>References .....</b>	<b>79</b>
<b>Appendix A .....</b>	<b>82</b>
<b>Appendix B .....</b>	<b>83</b>
<b>Appendix C .....</b>	<b>84</b>

## List of Figures and Tables

<b>Figures</b>	<b>Page</b>
Figure 1: An overview of Analog Devices Inc.'s manufacturing operations .....	14
Figure 2: Wafer fabrication steps.....	15
Figure 3: Plasma Ashing Process Schematic .....	16
Figure 4: Gasonics Aura 3010 machine .....	17
Figure 5: Display screen of the Gasonics tool with a sample recipe .....	18
Figure 6: Spatial distribution and co-ordinate positions of the nine sites.....	21
Figure 7: Data logging from the Nanospec 9200 tool.....	22
Figure 8: Areal representation of each site on a wafer .....	23
Figure 9: The roadmap of the project that the team completed together .....	25
Figure 10: Example of a Shewhart control chart .....	27
Figure 11: Performance differences in etch rates between the machines .....	36
Figure 12: A machine matching process and decision tree.....	37
Figure 13: Chambers operating in different windows .....	39
Figure 14: Malfunction diagnosis steps .....	39
Figure 15: Machine Matching Software .....	40
Figure 16: A basic double EWMA controller schematic.....	41
Figure 17: Total wafer runs per machine .....	44
Figure 18: The chamber matching methodology followed in this project.....	45
Figure 19: The statistical tools to determine the golden and inferior machines .....	47
Figure 20: The control charts of the machines.....	49
Figure 21: Normal plots of G53 and G73 .....	50
Figure 22: The box and whisker Plots .....	51
Figure 23: G53 factors affecting the output in comparison to G63 .....	54
Figure 24: The individual model fits for G53 and G63 for its site-1 .....	57
Figure 25: The global model fits for G53 and G63 .....	58
Figure 26: G53 Temperature when $T=235\text{ C}^{\circ}$ .....	60
Figure 27: G63 Temperature when $T=235\text{ C}^{\circ}$ .....	60
Figure 28: Temperature profiles on x and y axis of the wafer.....	61

Figure 29: Surface profile map of the wafer straight and rotated 270 degrees in G53 .....	63
Figure 30: Surface profile map of the wafer processed with Partial Ash in G53 and G63 .....	64
Figure 31: Square surface temperature profile of the wafers processed in G53 and G63 .....	64
Figure 32: The Partial ash recipe temperature profile from CLTC for G23,G33, G43 .....	66
Figure 33: G33 CLTC Temperature Profiles on June 9 <sup>th</sup> and 14 <sup>th</sup> .....	66
Figure 34: G33 CLTC Temperature profile of four consecutive wafers .....	67
Figure 35: Complete Ash recipe on four consecutive wafers in G23 .....	68
Figure 36: G43 Temperature profile from CLTC running a THICK recipe.....	68
Figure 37: Output noise change as a function of the EWMA weights terms $\gamma_1, \gamma_2$ .....	70
Figure 38: Double EWMA Control is simulated with individual G53 model.....	71
Figure 39: Double EWMA Control is implemented on G53 with a drift of 3 Å /wafer and normal noise $\mu = 0, \sigma = 50$ .....	72
Figure 40: Double EWMA Control is implemented on G63 with a drift of 3 Å /wafer and normal noise .....	73
Figure 41: Resultant strip thickness when G53 experiences an upward drift and G63 downward drift with and without control being implemented .....	74



<b>Tables</b>	<b>Pages</b>
Table 1: Machine parameters for the Partial recipe .....	19
Table 2: Machine parameters for the Forming recipe .....	20
Table 3: $2^3$ full factorial experimental design.....	30
Table 4: $2^{3-1}$ factorial experimental design .....	31
Table 5: The mean and standard deviation in Angstroms of each machine .....	51
Table 6: The average strip thickness models .....	53
Table 7: 9 months' data control chart values.....	54
Table 8: Streamed data types with their respective units.....	65

*This page is intentionally left blank.*

## **1. Introduction**

The work in this thesis presents a methodology to implement chamber matching for semiconductor dry etch equipment. It lays out the steps to follow in characterizing and modeling the chambers using statistical tools, diagnosing the reasons behind the performance mismatch and demonstrates chamber matching with run-to-run control techniques. This is an industrial thesis and the work was done in collaboration with Analog Devices Inc. in their Wilmington, MA fabrication center. Analog Devices Inc. is a world leader in the design, manufacture, and marketing of high performance analog, mixed-signal, and digital signal processing integrated circuits used in a broad range of electronic applications. ADI is headquartered in Norwood, MA. Currently, there is a need in the company to rigorously analyze various processes and machine capabilities in an effort to improve yield, throughput, and reduce machine downtime through early detection of equipment damage. The purpose of this chapter is to provide background on Analog Devices Inc., introduce the semiconductor dry etch process that was studied in this work, and state the problems around it that this thesis attempts to address.

### **1.1. Background Information on Analog Devices Inc.**

Analog Devices Inc. is an American multinational company that specializes in the design, manufacture, and marketing of high performance analog, mixed-signal, and digital signal processing integrated circuits used in a broad range of electronic applications. The company's products play a fundamental role in converting, conditioning, and processing real-world phenomena such as temperature, pressure, sound, light, speed, and motion into electrical signals to be used in a wide array of electronic devices.

The company was founded in 1965 by Ray Stata and Matthew Lorber and is headquartered in Norwood, MA. Analog Devices Inc. has operations in 23 countries and serves over 100,000 customers from various industries like consumer electronics, automotive, and defense to name a few. The annual revenue of the company in the fiscal year 2015 was approximately \$3.44 billion [1].

The manufacturing and assembly of Analog Devices Inc.'s products is conducted in several locations worldwide. Figure 1 shows an overview of the locations and functions of the company's manufacturing and assembly facilities.

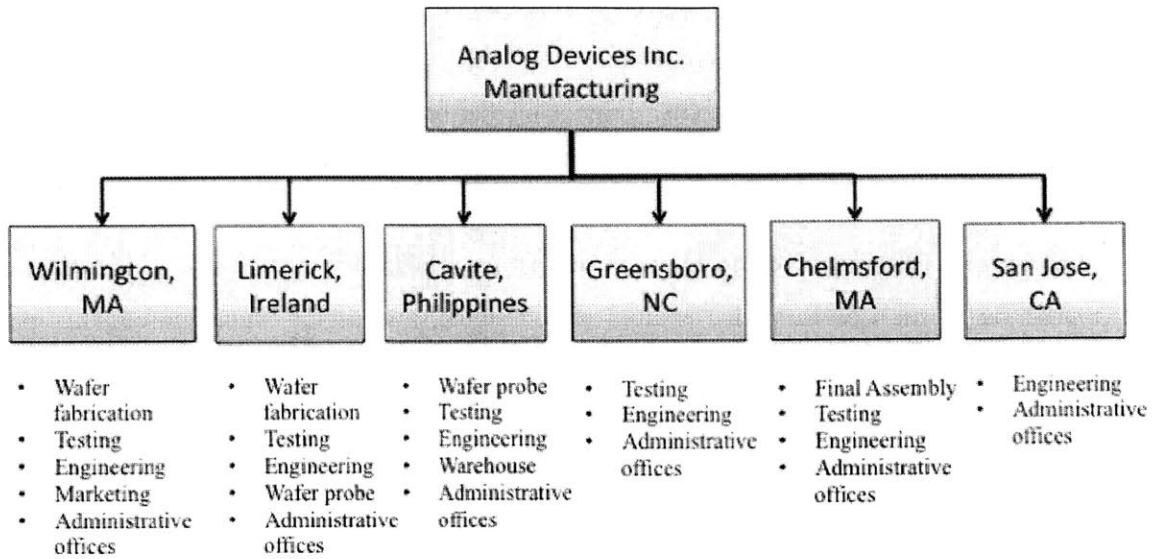


Figure 1: An overview of Analog Devices Inc.'s manufacturing operations.

The experiments in this thesis were carried out on a dry etch process in the Wilmington, MA fabrication center. This thesis is written in conjunction with the works of Tan Nilgianskul and Tanay Nerurkar, and several sections and descriptions in this thesis are written in common with their works [2,3].

### 1.2. General Semiconductor Fabrication Process

Figure 2 shows a number of semiconductor process steps repeatedly performed on a wafer in a fabrication facility. Photoresist acts effectively like a mold for other materials to be deposited in patterns, as shown in the diagram. After the metal has been deposited onto the desired parts of the wafer, the resist can then be stripped. Plasma ashing is a type of resist stripping technique, and is the focus of this thesis.

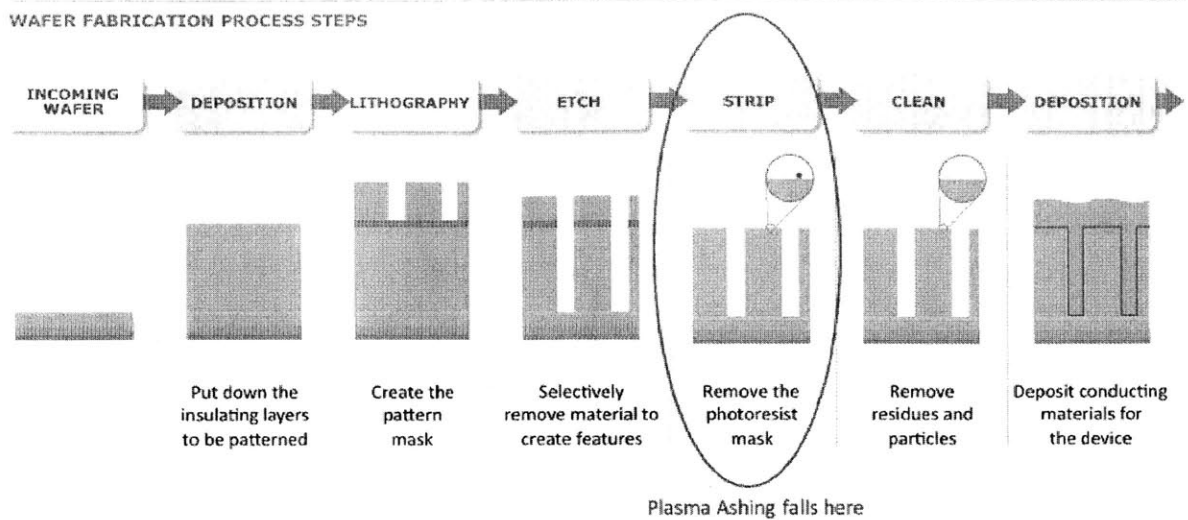


Figure 2: Plasma-ashing is a step among many other types of steps in wafer fabrication [4].

At Analog Devices, pre-doped wafers are supplied to the Wilmington, MA fabrication site as the starting material. The Wilmington, MA fabrication site is divided into five main sub-departments: thin-films, etch, photolithography, diffusion and CMP (chemical mechanical polymerization). A key procedure used at many points in the manufacturing of a device is photolithography where photoresist is deposited and patterned onto the desired parts of the wafer. This allows the diffusion team to selectively implant impurity ions, to etch or remove materials, the thin-films group to deposit metals onto the designated parts of the silicon wafer. Afterwards, the etch group then strips the resist off from these wafers. The function of the CMP group is to use chemical-mechanical reaction techniques to smoothen the surface of the deposited materials. Different types of devices will require a different set and configuration of material layers, with repeated sequences of photolithography, etch, implantation, deposition, and other process steps.

### 1.3. Plasma Ashing Process

For the purpose of this thesis, the plasma ashing process is investigated. This process is used to remove photoresist (light sensitive mask) from an etched wafer using a monoatomic reactive species that reacts with the photoresist to form ash, which is removed from the vicinity of the wafer using a vacuum pump. Exposing a gas such as oxygen or fluorine to high power radio or microwaves ionizes the gas, forming monoatomic and reactive species. Figure 3 shows a general schematic of the plasma ashing process with key components indicated.

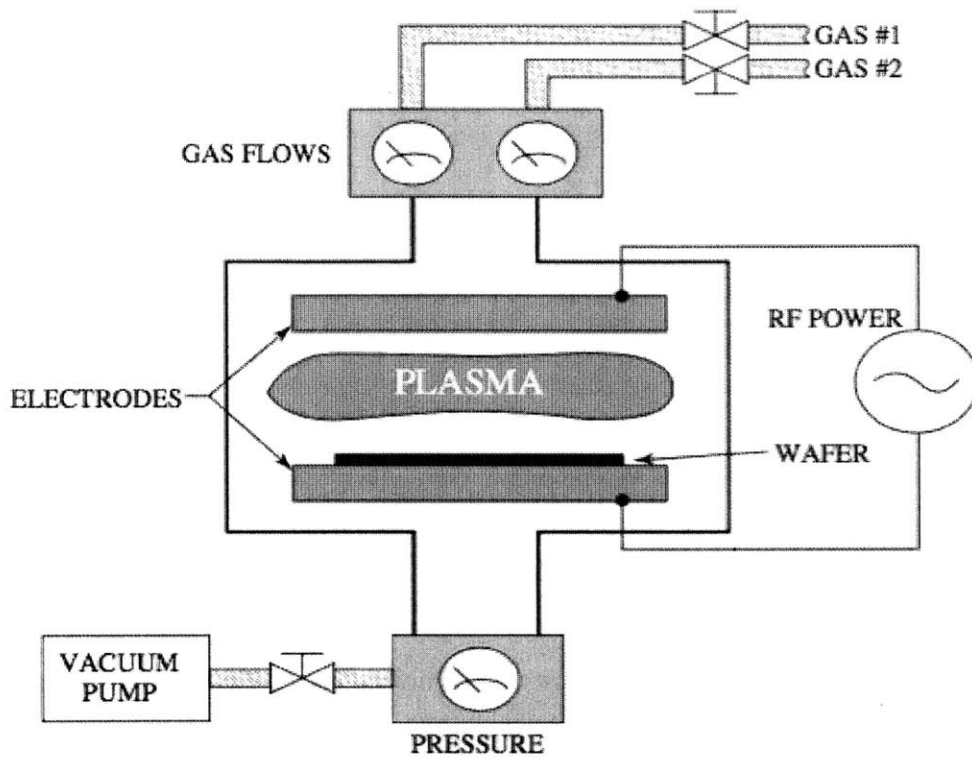


Figure 3: Plasma ashing process schematic[5].

Analog Devices Inc. uses the Gasonics A3010 tool to carry out the plasma ashing process. As the reactive gas, oxygen is used and microwaves are used to ionize the gas. The Gasonics A3010 tool allows for the change of several variables including temperature, chamber pressure, and power that make up a “recipe” to allow for different photoresist removal rates that may be needed for different products.

### 1.3.1. Gasonics A3010 Tool Components

The Gasonics Aura 3010 machine is used by Analog Devices Inc.’s Wilmington, MA fabrication center for photoresist ashing and cleaning of semiconductor wafers by creating a low-pressure and low-temperature glow discharge, which reacts chemically with the surface of the wafer. The Aura 3010 system is composed of three main components[6]:

- i. The reactor chamber that contains the system controller, the electro-luminescent display, wafer handling robot, the microwave generator, and the gas box.

ii. The power enclosure wall box.

iii. The vacuum pump.

Figure 4 shows a picture of the Gasonics Aura 3010 machine.

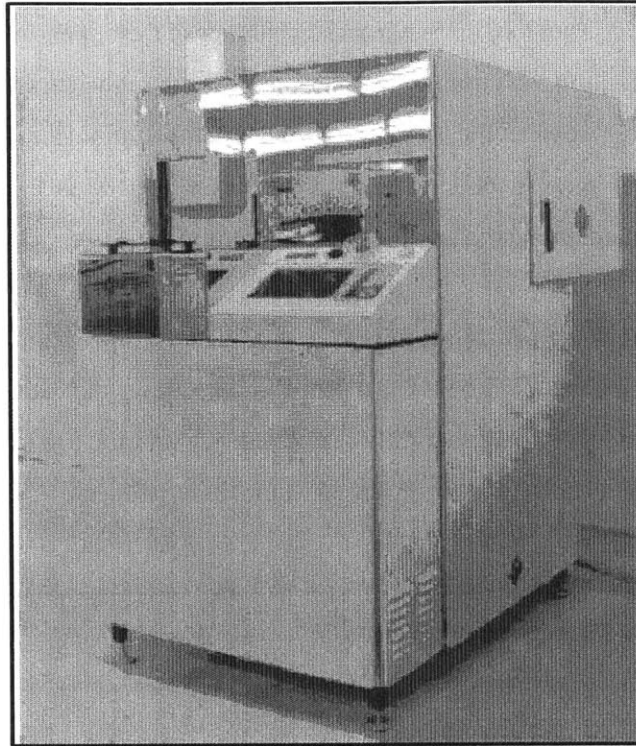


Figure 4: Gasonics Aura 3010 machine [7].

The machine is equipped with a three axis of motion wafer handling robot that picks up a single wafer from a 25-wafer cassette and places it in the process chamber to execute the photoresist stripping process. After a particular recipe is executed, the robot removes the wafer and places it on a cooling station if required before returning the wafer back to its slot in the cassette. Inside the process chamber, the wafer rests on three sapphire rods and a closed loop temperature control (CLTC) probe that includes a thermocouple to measure the temperature of the wafer during the ashing process. Twelve chamber cartridges embedded in the chamber wall heat the process chamber. During the plasma ashing process, eight halogen lamps heat the wafer to the required process temperature. The process gases (oxygen, nitrogen, or forming gas) are mixed and delivered to a quartz plasma tube in the waveguide assembly where microwave energy is generated by a magnetron that ionizes the gases into the monoatomic reactive species.

The machine is designed to only allow the lower-energy free radicals and neutrals to come in contact with the wafer surface as higher energy radicals can damage the wafer. After the wafer has been stripped, the halogen lamps, microwave power, and the process gas flows are turned off and the process chamber is then purged with nitrogen before being vented to the atmosphere for wafer removal. The door to the process chamber is then opened and the robot removes the wafer to either place it on the cooling station or put it back in the cassette slot.

Analog Devices Inc.'s Wilmington, MA fabrication center has seven Gasonics Aura 3010 machines that have a codename of GX3000 where X is a number between 1 and 7. The experiments and analysis that are presented in this work were conducted on G53000 and G63000 machines.

**1.3.2. Partial and Forming Recipes**

A recipe can be defined as a set of input settings that can be adjusted on a tool or machine to execute a desired manufacturing process. For example, Figure 5 shows a sample recipe on the display screen of the Gasonics Aura 3010 machine.

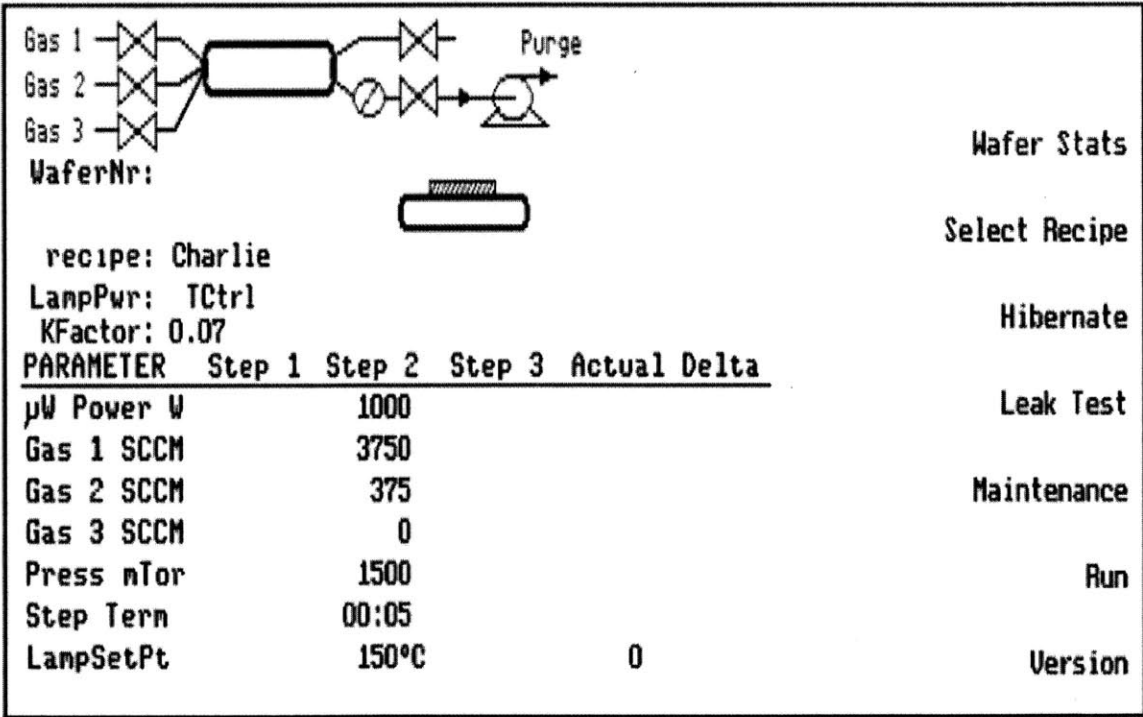


Figure 5: Display screen of the Gasonics tool with a sample recipe.



The machine allows the operator to vary the quantities under the column “PARAMETER”. The process engineers in the company are responsible for proposing and executing an optimal recipe taking into account product quality, throughput and cost constraints. In addition to designing recipes for production wafers, Analog Devices Inc. also designs recipes to run qualification tests. Qualification tests are used to periodically monitor product quality and verify machine calibrations. In this thesis, two qualification test recipes are studied which Analog Devices Inc. has named “Partial” and “Forming”. These two qualification tests are a representative of two different classes of production recipes that differ in conditions based upon the wafer chemistry. The details of these two recipes are as follows:

*i. Partial Recipe:* The Partial recipe is used for a qualification test to verify the rate of photoresist removal on a Gasonics Aura 3010 machine. The recipe simulates the machine conditions needed in the production recipe “Complete” which is used to completely strip photoresist from a production wafer that has not undergone prior harsh treatments like ion implantation. However, during the qualification test using the Partial recipe, photoresist is not completely stripped off from the dummy wafer. This is intentionally done so that the amount of photoresist removed and the time taken to do so can be recorded. An ideal Gasonics Aura 3010 machine would remove 6000 Angstroms of resist in eight seconds. The entire process with the partial recipe takes approximately 63 seconds with the first 20 seconds being allocated to heating the wafer to the necessary conditions and bringing the machine to steady state (Step-1), the next eight seconds being allocated to the stripping process (Step-2) and the last 35 seconds being allocated to cooling the wafer. Table 1 shows the necessary machine parameters needed for the partial recipe.

<b>Machine Parameter</b>	<b>Step-1</b>	<b>Step-2</b>
Wafer Temperature (Celsius)	215	235
Chamber Pressure (mTorr)	2000	2000
Microwave Power (Watts)	0	1400
Blower Vacuum Pumping Speed (kWh)	6	6
Main Vacuum Pumping Speed (kWh)	5.5	5.5
Oxygen Gas (SCCM)	3750	3750
Nitrogen Gas (SCCM)	375	375
<b>Step Term (seconds)</b>	<b>20</b>	<b>8</b>

Table 1: Machine parameters for the Partial recipe.

*ii. Forming Recipe:* The Forming recipe is also a qualification test used to verify the rate of photoresist removal on the Gasonics Aura 3010 machine, but this recipe simulates the machine conditions in a different production ashing recipe which is known as the “Implant” ash. The Implant ashing recipe is used to strip photoresist from a production wafer that has undergone harsh treatments like ion implantation. The necessity to use a different recipe for wafers that have undergone harsh treatments comes from the fact that the chemistry of the photoresist mask may have changed during those treatments, and not accounting for these changes can damage the wafer and product. As in the case of the Forming recipe, the ideal machine will remove 6000 Angstroms but the time taken to do so in this recipe is 60 seconds. The entire process with the forming recipe takes approximately 115 seconds with the first 20 seconds being allocated to heating the wafer to the necessary conditions and bringing the machine to steady state (Step-1), the next 60 seconds being allocated to the stripping process (Step-2) and the last 35 seconds being allocated to cooling the wafer. Table 2 shows the necessary machine parameters needed for the forming recipe.

<b>Machine Parameter</b>	<b>Step-1</b>	<b>Step-2</b>
Wafer Temperature (Celsius)	150	150
Chamber Pressure (mTorr)	2000	2000
Microwave Power (Watts)	0	1400
Blower Vacuum Pumping Speed (kWh)	6	6
Main Vacuum Pumping Speed (kWh)	5.5	5.5
Oxygen Gas (SCCM)	3750	3750
Forming Gas (Nitrogen+Hydrogen) (SCCM)	375	375
Step Term (seconds)	20	60

Table 2: Machine parameters for the Forming recipe.

### 1.3.3. Data Collection and Logging

The key parameter that needs to be measured in the plasma ashing process is the amount of photoresist removed from the wafer after the process has been completed. The amount of photoresist removed divided by the time for which the Gasonics A3010 tool was set to function gives the photoresist removal rate, which Analog Devices Inc. uses to infer machine health. The tool used to measure the amount of photoresist removed in Analog Devices Inc.’s Wilmington fabrication center is the Nanospec 9200. The Nanospec 9200 tool has the capability to accurately measure wafer thicknesses in the Angstrom range. The Nanospec 9200 tool is programmed to

measure nine sites on each wafer. Figure 6 shows the spatial distribution as well as the coordinate measurements of the nine sites on each wafer. In the spatial distribution diagram, the blue dots indicate the sites where the measurements are taken.

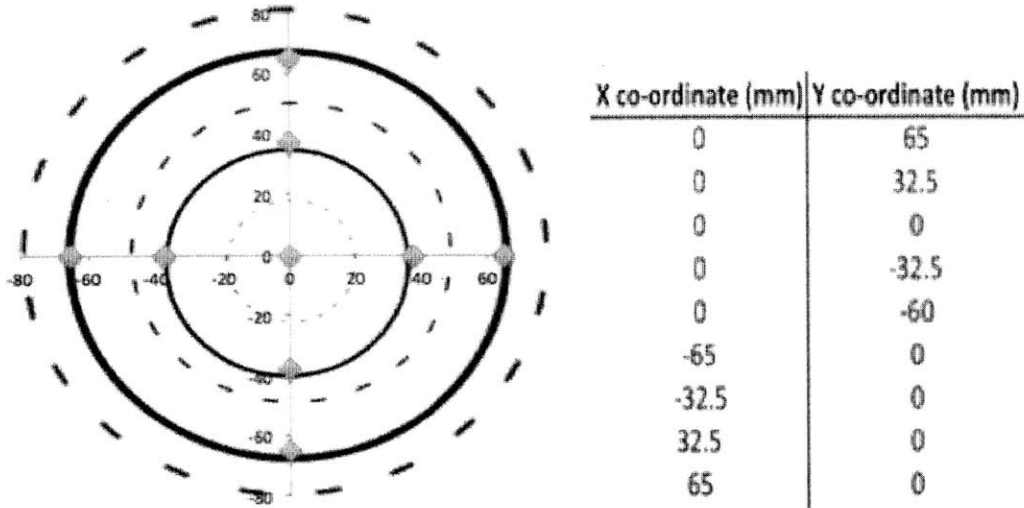


Figure 6: Spatial distribution and coordinate positions of the nine sites.

The measurement procedure of the thickness of the photoresist in each of the nine sites is as follows:

- i) The thickness of the photoresist is measured and recorded before the wafer undergoes the plasma ashing process. These are known as pre-measurements.
- ii) The thickness of the photoresist is measured and recorded after the wafer undergoes the plasma ashing process. These are known as post-measurements.
- iii) The difference between the pre-measurements and post-measurements gives the amount of photoresist removed during the process.
- iv) The amount of photoresist removed can be divided by the duration of the plasma ashing process, which is included as an input and monitored by the Gasonics A3010 tool.

The amount of photoresist removed for each of the nine sites on a single wafer is recorded in an excel spreadsheet on which further analysis can be done. An example of the excel spreadsheet can be seen in Figure 7. In Figure 7, the columns in the spreadsheet represent the measurements taken on the nine sites within a single wafer while the rows represent different wafers measured. The Nanospec 9200 tool also logs the date and time of the measurement, which is very useful in anomaly detection.

Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Site 7	Site 8	Site 9	ET_TIME
5549.35	5659.8	5763.96	5824.96	5877.34	5501.67	5699.48	5739.73	5540.88	6/20/15 8:07
5410.93	5524.85	5631.59	5688.9	5707.29	5356.96	5555.87	5612.81	5428.14	6/23/15 8:59
5493.44	5571.64	5671.82	5760.63	5794.19	5456.9	5618.04	5647	5462.18	6/27/15 7:57
5436.46	5552.35	5689.99	5784.46	5802.86	5460.92	5645.25	5648.27	5389.47	7/6/15 9:01
5595	5682	5796	5878	6004	5450	5724	5892	5699	7/10/15 9:39
5587.58	5693.32	5829.78	5954.44	6019.11	5571.12	5772.03	5817.37	5567.42	7/13/15 20:12

Figure 7: Data logging from the Nanospec 9200 tool

#### 1.3.4. Calculation of Basic Statistics

The raw data collected from the Nanospec 9200 tool as shown in Figure 7 needs to be manipulated further in order to make meaningful implications of the underlying trends and patterns. This section introduces the method that was used to calculate three statistical quantities:

- i. The mean thickness of the nine sites on a single wafer ( $\bar{x}^*$ )
- ii. The standard deviation of the nine sites on a single wafer ( $s$ )
- iii. The within wafer non-uniformity parameter ( $NU$ )

The nine sites that the Nanospec 9200 tool measures on a single wafer are distributed in a radial pattern from the center as can be seen in the spatial distribution diagram in Figure 8. Previous work by Davis *et al.* has shown that in a radial distribution pattern, the calculation of any statistics on the sites measured on a wafer have to take into account the wafer area represented by each site for accurate analysis [8]. Figure 8 shows the wafer areal representation of each site on a nine site radial distribution pattern. The wafers used for the purposes of this study have a diameter of 80 mm or 6 inches.

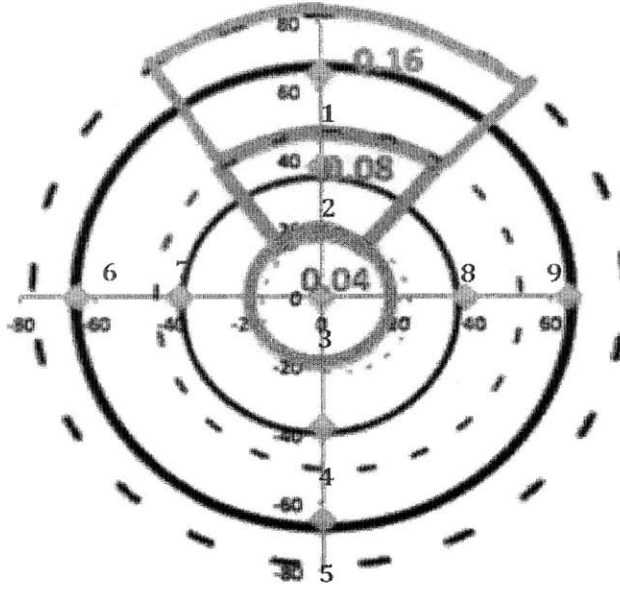


Figure 8: Areal representation of each site on a wafer.

In Figure 8, site 3 represents the area bounded by the green circle (4% of the total wafer area), sites 2, 4, 7, and 8 each represent the area bounded by the red segments (32% of the total wafer area), and sites 1, 6, 5, and 9 each represent the area bounded by the orange segments (64% of the total wafer area).

The mean ( $\bar{x}^*$ ) taking into account the areal representation of each site is calculated as follows:

$$\bar{x}^* = \frac{\sum_{i=1}^N w_i x_i}{\sum_{i=1}^N w_i} \quad (1)$$

where  $x_i$  is the wafer thickness measured at each site,  $w_i$  is the weighted area associated with that site and  $N$  is the number of sites.

The standard deviation ( $s$ ) taking into account the areal representation of each site is calculated as follows:

$$s = \sqrt{\frac{\sum_{i=1}^N w_i}{(\sum_{i=1}^N w_i)^2 - \sum_{i=1}^N w_i^2} \cdot \sum_{i=1}^N w_i (x_i - \bar{x}^*)^2} \quad (2)$$

where  $x_i$  is the wafer thickness measured at each site,  $w_i$  is the weighted area associated with that site,  $N$  is the number of sites, and  $\bar{x}^*$  is the mean. The derivation for Eq. (1.3) can be found in one of NASA's Giovanni documents [9].

The within-wafer non-uniformity parameter ( $NU$ ) taking into account the areal representation of each site is calculated as follows:

$$NU = \frac{s}{\bar{x}^*} \quad (3)$$

where  $s$  is the standard deviation and  $\bar{x}^*$  is the mean.

#### **1.4. Problem Statement**

This section presents the motivation behind the work done by our team in Analog Devices to improve the yield and increase throughput of the plasma ashing process. As the dimensions in which the semiconductor companies are operating are shrinking (<28 nm) more sophisticated control systems need to be used to achieve higher yields. Analog Devices is also adapting to these changes by implementing Internet of Things and Advanced Control Systems. The need that emerged from observing the pilot project for Internet of Things (IoT) is being implemented: the already existing control system does not facilitate steps to keep the process in control and the machines operate at different values of the critical dimensions measured for the process. Analog Devices stated that such differences might result in yield losses that appear later at the end-of-the-line after many process steps. Such losses can become more problematic at the smaller scale and for more expensive processes. Therefore the work we outline in this paper, an improvement plan for the use of statistical control, design of experiments for process modeling and lastly machine matching, would be a critical step to the implementation of Advanced Control Systems such as Predictive Maintenance (PM) that Analog Devices, similar to the rest of the semiconductor industry, is moving towards with their Internet of Things project.

#### **1.5. Outline and Task Allocation**

The problem statement presented to the team was regarding the photoresist strip (plasma ashing) process and chambers that run this process. However, both Analog Devices and our team saw this as a huge opportunity to create solutions and methodologies that are flexible enough to be customized for various processes. While defining our projects, the team aimed at bringing lasting solutions and recommendations that are easily applicable and ready to be rolled out. The team executed the entire project together; however, the discussion topics will differ. Figure 9 illustrates the roadmap of the team according to the order each module was executed and the team member who will be discussing each. The work in this thesis makes extensive use of the

work of T. Nilgianskul that focuses on setting the control limits to detect the quality issues early on, and the work of T. Nerurkar that focuses on the theory and the methodology of design of experiments and findings from the experiments that the team ran. Chamber matching was the last phase of the project where the focus is on rigorous machine matching of both statistical and dynamic behaviors of the machine.

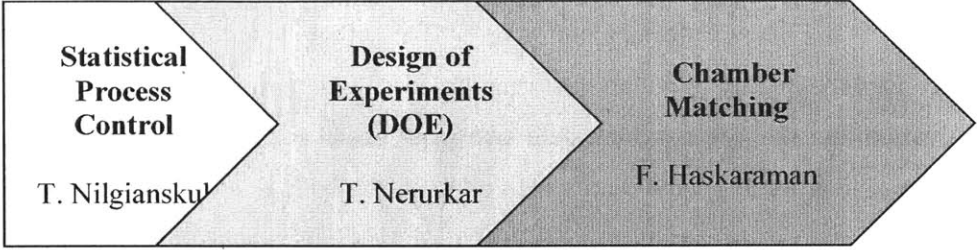


Figure 9: The roadmap and task allocation of the project that the team completed together aims to improve productivity in ADI facilities. The discussion topics for each member are indicated.

## **2. Theoretical Review of Key Concepts**

This chapter will introduce the mathematical concepts and models that are relevant to the construction of this thesis. This includes both theoretical SPC background from textbooks by Montgomery and May and Spanoss as well as prior research that has applied those concepts in both academic and industrial settings [5,10].

### **2.1 Statistical Process Control (SPC)**

Statistical Process Control or SPC is an applied statistics concept used to monitor and control the quality of a manufacturing process by minimizing process variability. With decreased variability, the rate at which defective parts occur also decreases, thereby reducing waste. Key topics that are applied towards the collaboration with Analog Devices include Shewhart control charts, analysis of variance (ANOVA), design of experiments (DOE) and hypothesis testing.

#### **2.1.1 Origin of SPC**

The SPC method was introduced by Walter A. Shewhart at Bell Laboratories in the early 1920s. Later in 1924, Shewhart developed the control chart and coined the concept of “a state of statistical control” which can actually be derived from the concept of exchangeability developed by logician William Ernest Johnson in the same year in one of his works called *Logic, Part III: The Logical Foundations of Science* [11]. The theory was first put in use in 1934 at the Picatinny Arsenal, an American military research and manufacturing facility located in New Jersey. After seeing that it was applied successfully, the US military further enforced statistical process control methods among its other divisions and contractors during the outbreak of the Second World War [12].

#### **2.1.2 Shewhart Control Charts**

A Shewhart control chart essentially plots an output parameter or an indicator of the process performance over a measure of time [13]. These plots are then bounded by control limits that are, as a rule-of-thumb, three standard deviations away from the mean on either side. An example of a control chart is shown in Figure 10 [14].



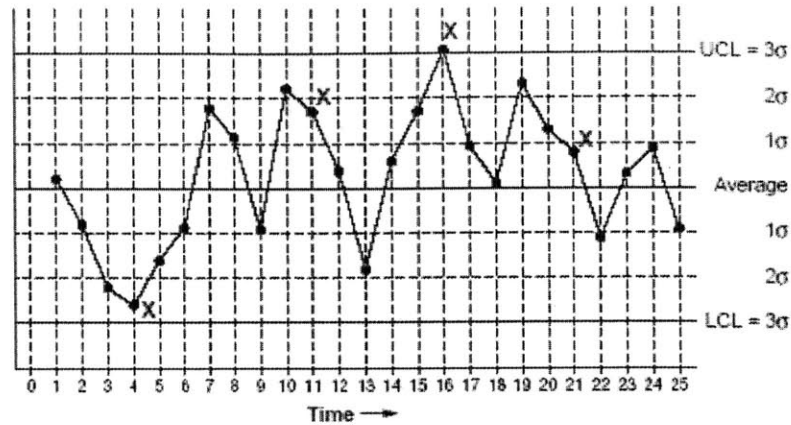


Figure 10: Example of a Shewhart control chart. Points marked with X's are points that would be rejected based on Western Electric Rules.

Control charts can either be plotted as a run chart or an x-bar chart. The run chart plots each measurement separately on the chart while the x-bar control chart plots the average of several measurements. Because the thickness measurements associated with the plasma ashing process do not come in batches and are sampled individually over a period of time, only the run chart will be relevant in subsequent analyses here.

The goal of plotting control charts is to monitor the manufacturing process and detect when it is out of control. Assuming that the data plotted is normally distributed, which is usually the case for most processes, the chance that any single point would lie above the upper control limit  $UCL$  or below the lower control limit  $LCL$  (in the case of the typically used three standard deviations above or below the mean) would be less than 0.3%. Assuming that a set of data is normally distributed with mean  $\mu$  and variance  $\sigma^2$ ,  $UCL$  and  $LCL$  can be expressed as:

$$UCL = \mu + 3\sigma \tag{4}$$

$$LCL = \mu - 3\sigma \tag{5}$$

With that, the probability of a point lying beyond outside the limits for any normally distributed data set can be solved for:

$$P(X > UCL) = P\left(Z > \frac{UCL - \mu}{\sigma}\right) = P(Z > 3) \cong 0.0013 \tag{6}$$

$$P(X < LCL) = P\left(Z < \frac{\mu - LCL}{\sigma}\right) = P(Z < -3) \cong 0.0013 \tag{7}$$

$$P(X > UCL | X < LCL) = P(X > UCL) + P(X < LCL) \cong 0.0027 \tag{8}$$

$$P(\text{point lies outside control limit}) \sim 3\%$$

Design of experiments (DOE) is a systematic method to determine how factors affecting or the inputs to a process quantitatively relate to that process's output. It is a powerful tool for identifying cause-effects within a certain manufacturing process. This information could then be used to tune the process inputs in order to optimize the outputs of that process to achieve production goals. The focus of DOE is not on figuring out how to perform individual experiments but rather on planning the series of experiments in order to obtain the most information in the most efficient manner. This leads to the concept of designing fractional factorial experiments.

Besides the upper and lower control limit rule, there are other Western Electric rules that could be used as guidelines to suspect when the process is out of control. These include 1) if two out of three consecutive points lie either two standard deviations above or below the mean, 2) four out of five consecutive points lie either a standard deviation above or below the mean, 3) nine consecutive points fall on the same side of the centerline/mean [14].

## 2.2 Analysis of Variance

Analysis of variance or ANOVA is a collection of statistical models used to analyze the differences among group means and variance between and within sets of data. This would thus indicate the difference in the process associated with those data. ANOVA only came into substantial use in the 20th century, although mathematicians have been passively implementing parts of it in prior academic work, the earliest of which dates back to when Laplace conducted hypothesis testing in the 1770's [15].

In semiconductor processing, extra attention will be paid to nested analysis of variance. This is the analysis that is done when data can be broken down into groups, subgroups, etc. Nested variance analysis will determine the significance of the variance between and within groups and subgroups of data [16]. For instance, say there are  $W$  groups of data with  $M$  data in each of those groups; the mean squared sum between groups ( $MS_W$ ) and within groups ( $MS_E$ ) can be calculated as follows [5].

$$MS_W = \frac{SS_W}{W-1} \quad (9)$$

$$MS_E = \frac{SS_E}{W(M-1)} \quad (10)$$

where:

$SS_W =$  squared sum of deviations of group means from grand mean

$SS_E =$  squared sum of deviations of each data point from its group mean

Note that  $SS_W$  sums up the grand-group mean deviation for every individual point. Therefore in this case, each squared difference between grand to group mean difference is multiplied by M before summing them together. The significance of the between-group variation is then determined, given that the ratio  $MS_W/MS_E$  approximately follows the F-distribution.

It is important to take into account that the observed variance of the group averages does not reflect the actual wafer-to-wafer variance because of the existence of sub-variation (group variance). The observed variation between the group averages  $\sigma_w^2$  can be written as a linear combination of the true variance  $\sigma_w^2$  and the group variance  $\sigma_M^2$  [5].

$$\sigma_w^2 = \sigma_w^2 + \frac{\sigma_M^2}{M} \quad (11)$$

Hence the true group-to-group variance can be expressed as:

$$\sigma_w^2 = \sigma_w^2 - \frac{\sigma_M^2}{M} \quad (12)$$

From this, both the group-to-group component and the within-group component can be expressed as a percentage of the total variance. This variance decomposition enables us to differentiate between measurements among wafers and within silicon wafers.

### 2.3 Design of Experiments

Factorial experiments allow for both individual factor and multiple-order interactions (effect of varying multiple factors simultaneously) to be evaluated from one experiment. Single factor relationships are also termed “main effects.” Experimental design is built upon the foundation of analysis of variance and orthogonality. Analysis of variance is used to break down the observed variance into different components while orthogonality is, in other words, the relative independence of multiple variables which is vital to deciding which parameters can be simultaneously varied to get the same information [10].

Our experiments were done based on pre-designed half-factorial experiments, as reducing from full-factorial to half-factorial experimental designs requires fewer experimental combinations at the expense of aliasing or confounding main effects with multiple-order interactions that can be assumed negligible. These interactions usually include some second degree or higher than third degree order interactions that are typically less significant than lower-degree interactions. For example, Table 3 shows the full factorial ( $2^3$ ) experimental design for a two-level test with three variable input parameters (A, B and C) as the main effects. “-1” indicates a low setting while “+1” represents the high setting of the input parameter. The two levels mean that each main effect will only be varied between two values, the high value and the low value [10].

Table 3:  $2^3$  full factorial experimental design. -1 indicates a low setting while +1 represents the high setting of the input parameters.

	A	B	AB	C	AC	BC	ABC
(1)	-1	-1	+1	-1	+1	+1	-1
a	+1	-1	-1	-1	-1	+1	+1
b	-1	+1	-1	-1	+1	-1	+1
ab	+1	+1	+1	-1	-1	-1	-1
c	-1	-1	+1	+1	-1	-1	+1
ac	+1	-1	-1	+1	+1	-1	-1
bc	-1	+1	-1	+1	-1	+1	-1
abc	+1	+1	+1	+1	+1	+1	+1

By defining the following identity relation and aliases:

$$I = ABC$$

$$A + BC$$

$$B + AC$$

$$C + AB$$

a half factorial experimental design can be designed. Table 4 shows the half factorial design. This is extremely powerful when there are several factors to consider as it can immensely reduce the number of experiments needed.

Table 4:  $2^{3-1}$  Factorial experimental design

Run	Factors		
	A	B	C
1	-1	-1	+1
2	+1	-1	-1
3	-1	+1	-1
4	+1	+1	+1

## 2.4 Hypothesis-Testing

A statistical hypothesis test compares at least two sets of data that can be modeled by known distributions. Then assuming that those data follow the proposed distributions, the probability that a particular statistic calculated from the data occurs in a given range can be calculated. This probability is also referred to as the P-value and is ultimately the basis to either accept or reject the current state or the null hypothesis. The acceptance/rejection cutoff is marked by a rather arbitrary “significance level.” Generally, the decision as to what significance level to use would depend on the consequences of either rejecting a true null hypothesis (type I error) versus accepting a false null hypothesis (type II error). The three upcoming sections will outline the three tests around which this project revolves. Each of these tests centers on a different distribution [13].

### 2.4.1 Z-Test for Detecting Mean Shift

The Z-test technically refers to any hypothesis test whereby the distribution of the test statistic under the null hypothesis is modeled by the normal distribution. This becomes useful in many cases (including this project) because of the central limit theorem. With the central limit theorem, means of a large number of samples of independent random variables approximately follow a normal distribution. Mathematically, the sample mean of any distribution of mean  $\mu$  of

sample size  $n$  and standard deviation  $\sigma$  would be normally distributed with the same mean and standard deviation  $\frac{\sigma}{\sqrt{n}}$ , or  $\sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$  [13].

For instance, when testing for whether the mean of a given process (with default mean  $\mu$  and standard deviation  $\sigma$ ) has shifted, the following hypotheses can be formed [10].

$$H_0: \mu = \mu_0 \quad (13)$$

$$H_1: \mu \neq \mu_0 \quad (14)$$

The null hypothesis  $H_0$  is assumed to hold with the true mean  $\mu$  being equal to the assumed mean  $\mu_0$  to begin with. Now given a set of data or observations with sample mean  $\bar{x} > \mu_0$ , the test statistic  $Z_0$  could be calculated.

$$Z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \quad (15)$$

The  $P$ -value can then be deduced as follows.

$$P_{value} = P(\bar{x} > \mu_0) = P(z > Z_0) \quad (16)$$

Given a significance level  $\alpha$ , the null hypothesis would be rejected if  $P_{value} < \alpha/2$  or, equivalently, if  $Z_0 > Z_{\alpha/2}$  then the alternative hypothesis  $H_1$  would be accepted, that the mean has shifted.

The probability of encountering a type I error would be the significance level  $\alpha$  itself, i.e.,  $P(\text{Type I Error}) = \alpha$ . Given an alternative mean  $\mu_1$ , the distribution of the alternative distribution could be written as  $\bar{x} \sim N(\mu_1, \sigma/\sqrt{n})$ . Hence the probability of making a type II error could be calculated as

$$P(\text{Type II Error}) = P(\bar{x} < \bar{x}_{critical}) \quad (17)$$

where  $\bar{x}_{critical}$  is the  $\bar{x}$  that corresponds to  $Z_{1-\alpha/2}$  under the old mean  $\mu_0$ .

$$\bar{x}_{critical} = \mu_0 + Z_{1-\frac{\alpha}{2}} \cdot \sigma \quad (18)$$

Therefore, continuing from Equation (16)

$$P(\text{Type II Error}) = P\left(Z < \frac{\mu_0 - \mu_1}{\sigma} + Z_{1-\frac{\alpha}{2}}\right) \quad (19)$$

As previously mentioned, the significance level would depend on the tolerance for these two errors. For instance, if the detection of a mean shift would trigger an alarm and it is very costly to encounter a false alarm, then a lower  $\alpha$  would be desired in order to minimize  $P$  (Type I Error). However, if it is very crucial to detect the mean shift even at the cost of incurring several false alarms, then a higher  $\alpha$  would be desirable to minimize  $P$  (Type II Error).

Note that the example presented is a two-sided test because the  $P$ -value is tested against the probability of the sample mean being too far from the mean on either side. If it was a one sided test, with the alternative hypothesis would be  $H_1: \mu > \mu_0$  or  $H_1: \mu < \mu_0$ , the  $P$ -value would be compared to  $\alpha$  and the null hypothesis would be rejected if  $Z_0 > Z_\alpha$  (no  $\frac{1}{2}$  factor on  $\alpha$ ). The format of other tests will more or less follow the same structure as the example above but with different formulas for calculating the test statistics and their probabilities.

#### 2.4.2 F-test

Rather than detecting a mean shift, the F-test tests whether the ratio of the variance of two sets of data are statistically significant. Following the same method as in the previous Z-test example, the F-test begins with formulating hypotheses around the variance ( $s_1^2$  and  $s_2^2$ ) of two sets of data [13].

$$\begin{aligned} H_0: s_1^2 &= s_2^2 \\ H_1: s_1^2 &\neq s_2^2 \end{aligned} \tag{20}$$

The test statistic  $F_0$  in this case is simply the ratio of the variance where the numerator is the greater of the two variances,  $s_1^2 > s_2^2$ .  $F_0$  can approximately be modeled by the F-distribution.

$$F_0 = \frac{s_1^2}{s_2^2} \tag{21}$$

With that, the null hypothesis  $H_0$  would be rejected under a certain significance level  $\alpha$  if  $F_0 > F_{n_1-1, n_2-1, \alpha}$  where  $n_1$  and  $n_2$  represent the sample sizes of the first and second data sets respectively. Alternatively, the  $P$ -value could be calculated and tested directly against the significance level. The calculation of the  $P_{value}$  is shown in Equation (22) below.

$$P_{value} = P(F > F_0) \tag{22}$$

This is a one-sided test as can be seen intuitively. To modify this into a two-tailed test,  $F_0$  would simply be compared with  $F_{n_1-1, n_2-1, \alpha/2}$  where  $n_1$  and  $s_i$  represent the first data set (i.e.

$s_1^2$  is not necessarily larger than  $s_2^2$ ). Typically for testing whether or not two variances ( $s_n$ ) are different, a two-tailed test would not be used.

### 2.4.3 Bartlett's Test

Bartlett's test is used to determine whether  $k$  samples are sampled from distributions with equal variances. The null and alternative hypotheses can be formulated as follows.

$$\begin{aligned} H_0: s_1^2 &= s_2^2 = s_3^2 \dots = s_k^2 \\ H_1: s_i^2 &\neq s_j^2 \quad \text{for at least one pair } (i, j) \end{aligned} \quad (23)$$

Given the  $k$  samples with sample sizes  $n_i$ , and sample variances  $s_i^2$ , the test statistic  $T$  can be written as follows [17].

$$T = \frac{(N-k) \ln(s_p^2) - \sum_{i=1}^k (n_i - 1) \ln(s_i^2)}{1 + \frac{1}{3(k-1)} \left( \sum_{i=1}^k \left( \frac{1}{n_i - 1} \right) - \frac{1}{N-k} \right)} \quad (24)$$

where  $N$  is the total number of data points combined and  $s_p^2$  is the pooled estimated variance.

$$\begin{aligned} N &= \sum_{i=1}^k n_i \\ s_p^2 &= \frac{1}{N-k} \sum_{i=1}^k (n_i - 1) s_i^2 \end{aligned} \quad (25)$$

$T$  can be approximated by the chi-squared distribution.  $H_0$  would therefore be rejected under a significance level  $\alpha$  if  $T^2 > \chi_{k-1, \alpha}^2$



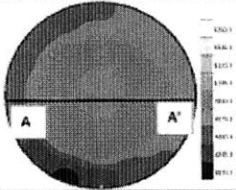
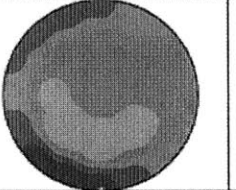
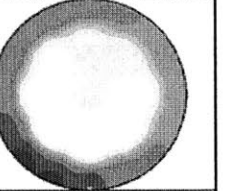
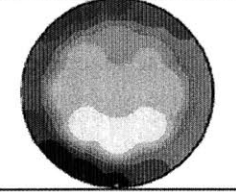
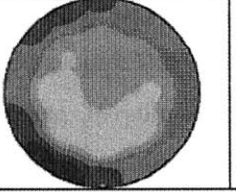
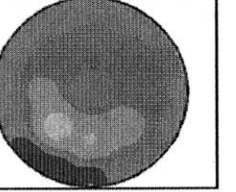
### **3. Theoretical Review of Chamber Matching**

In this section, the motivation behind chamber matching in the semiconductor industry will be introduced from the perspective of the manufacturers and scientists; common issues in chamber matching will be underlined and various implementation approaches will be discussed. Furthermore, current advancements in control methods that are instrumental in chamber matching will be discussed.

#### **3.1. Chamber Matching Issues in Semiconductor Manufacturing**

Today, building semiconductor manufacturing facilities can cost several billion dollars along with the additional costs of machine maintenance and upgrades. Some fabrication sites employ more than 30 fabrication tools and chambers running the same processes. The chamber-to-chamber performance differences between these machines can be substantial. These chambers consist of several thousands of parts, many of which might be replaced during Preventive Maintenance (PM). From run to run, behavior of the machines can dynamically move in different directions [18]. Unequal distribution of the tasks per chamber can be another reason of such differences. As the chambers run different number of tasks, the level of wear changes and causes drifts.

Chamber performance deviations have been studied widely in the context of semiconductor manufacturing. For plasma etching processes, Baek *et al.* observed that these differences cause etch rate variations and result in yield losses [18]. The topography and etch rate across the wafers change from chamber to chamber. Figure 10 shows the comparison of two chambers to the golden chamber in their etch rate and the non-uniformity. The findings show that all of the chambers have different operating etch rate and non-uniformity values.

Chamber	Golden	1	2
Etch rate ( $\text{\AA}/\text{min}$ ) : Average (Range)	4646.7 (8.4%)	4740.2 (9.0%)	5069.3 (13.5%)
Wafer map			
Chamber	3	4	5
Etch rate ( $\text{\AA}/\text{min}$ ) : Average (Range)	4750.4 (15.0%)	4823.3 (10.7%)	4740.2 (9.0%)
Wafer map			

(a)

Figure 11: Performance differences in etch rate and non-uniformity between the golden (best performing) and the inferior machines are shown using each rate surface map [8].

### 3.2. Chamber Matching Methodology

Today, many fabrication sites implement partial matching which concentrates on statistical matching and control. A more rigorous matching involves both statistical and dynamic matching [19]. The matching extends from process setup and configuration, hardware and software audit, to matching tool sensors and data collection. Matching chambers during production with advanced control techniques can move the resultant thickness to the target values, yet if they are not combined with matching of the operating states (e.g, optimized recipe), it can introduce further process yield variability. OEMs like Applied Materials therefore concentrate on achieving a machine matching criteria as “chambers are matched if their states of operation are matched” [20].

The chamber matching process defined by Baek *et al.* is an example of such matching using recipe optimization and feedback controller implementation. It starts with diagnosing the differences of the chambers and determining whether or not the machines are in different states. If the states are not different, it might mean that some of the machines are suffering from a malfunction and therefore, a root cause analysis can be run to understand the reasons behind the

mismatch. A possible procedure, a decision and implementation flow, is outlined in Figure 11 to attack machine-matching issues [18]. If the hypothesis that the machines are performing at different states is accepted, advanced equipment control can be implemented. In order to implement feedback control loops, firstly, the controlled variables (CVs), which characterize the performance of the chamber and the inputs or Manipulated Variables (MVs) are selected. For instance, etch rate or non-uniformity can be the CVs of this process and temperature and pressure would be the manipulated variables (MVs). Using these controlled or performance variables, a series of statistical analyses is used to determine the golden and inferior machines. A data driven (empirical) model is then created. Once the models are created, recipe optimization for matching the chambers can be implemented along with the Run-to-Run algorithm for better control [18].

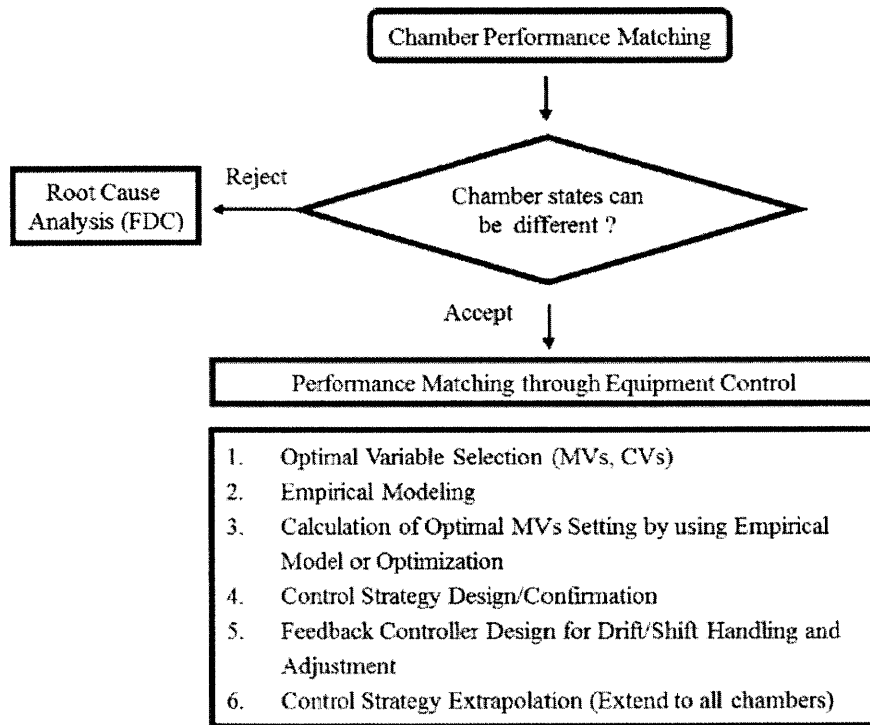


Figure 12: A machine matching process and decision tree [8]

### **3.3. Performance Metric and Key Variables Selection**

There are various approaches to select performance metrics. An approach involves selecting multiple inputs and possible metrics using expert knowledge. Here, besides etch rate, other controlled variables or inputs are also taken into account such as uniformity. A linear regression model can be created with a step change test to analyze the linearity of the output metrics with the inputs. The inputs that do not change linearly with the response are discarded at this step. Using a step change test data and employing a Raw Gain Matrix, the inputs and the output metrics can be paired to proceed with the experimental modeling [18].

Performance metric and key variable selection can also differ according to the resources of the team. If the end-of-line yield data is available, the performance metric can be chosen to be the yield. If the sensor data is also accessible for the inputs, the signals can be traced and data profile mismatch can be identified. After these variables are identified, the end-of-line yield data can be related to these variables to understand which key variables have the most effect on the end-of-line quality. This approach requires yield data available at the end of the line [21].

### **3.4. Golden and Inferior Chambers**

For every process, there is a range in which the machine can perform in order to have an acceptable yield. The variations in between these values can be tolerated. However, some machines might be in different states and do not abide with these limits although they have the same recipe. The machines, then, can be grouped according to their performances as can be seen in Figure 12. A golden machine is a machine that has a broad range of input conditions and yet results in an acceptable output value. Anything that is beyond the golden machine will be called an inferior chamber. Inferior chambers can have different behaviors too. For example, unstable machines have very high variation and therefore, limited tolerance for any variation in the input. Difficult to adjust machines have a tighter range of inputs that they can tolerate and might even have different response characteristics. Lastly, incapable machines perform mostly out of control and outside of the limits.

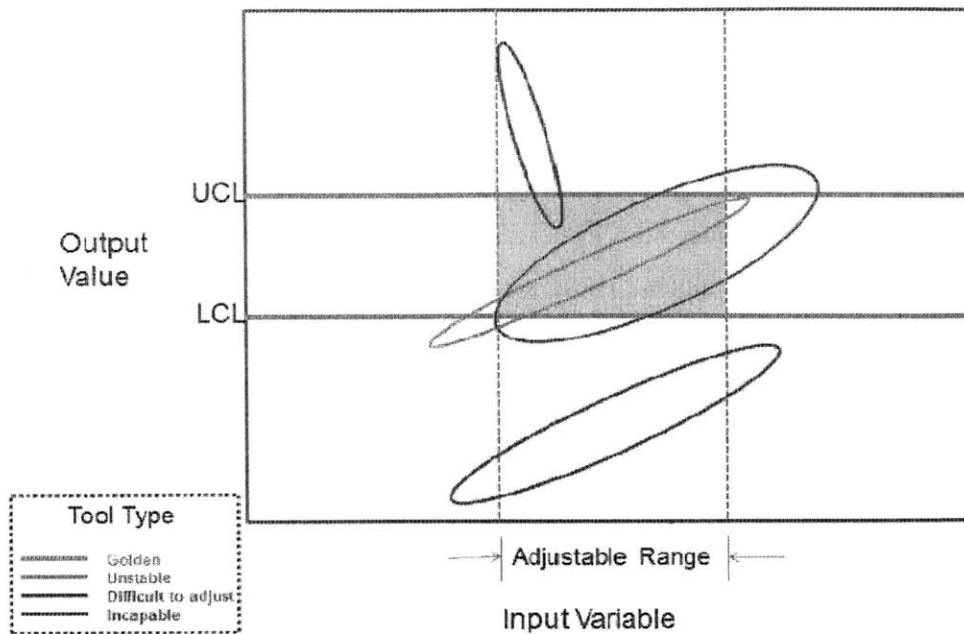


Figure 13: Chambers operating in different windows [19].

This diagnostic can be thought similar to a malfunction diagnostics as can be seen in Figure 13. Maintenance, in-line and on-line data can be analyzed to understand the performance differences between the machines. For example, an approach taken by Cosway involves illustrating visually the historic on-line data and grouping the machines according to their range of outputs, or using in-line pressure, temperature data and power or gas flow signals to tell whether the machines are matched or not [19].

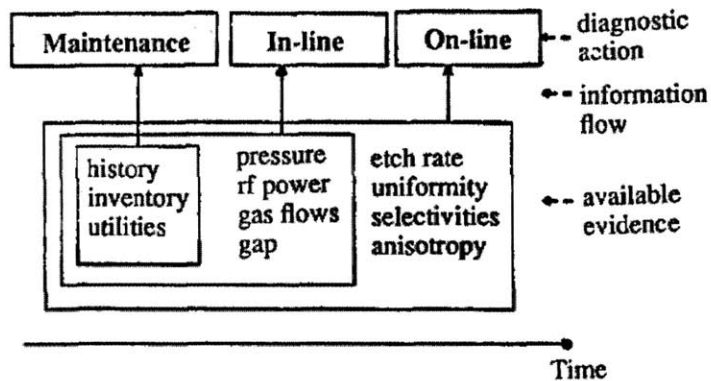


Figure 14: Malfunction diagnosis process steps can be used to group machines according to their performances [22].

### 3.5. In-Situ Sensor Monitoring

In-situ sensors are usually used for implementation of real-time feedback control, to detect process yield issues and for equipment or process diagnosis [6]. In-situ sensors can be grouped in three categories: equipment state, process state and wafer state sensors. The type of sensors that we are interested in here is the process state sensors, which monitor the process environment and the inputs. These sensors provide closed loop control and keep the process variables stable. With the development of the advanced control techniques like run-to-run control, sensors also extend to monitoring process shifts and help make comparisons in real time in the differences between different productions. The latest software like Applied Materials' E3 software for chamber matching enable real time monitoring of the inputs to be visible to the manufacturers, enabling them to compare how the important variables of the process behave as they also monitor the end-line statistics and yields [20].

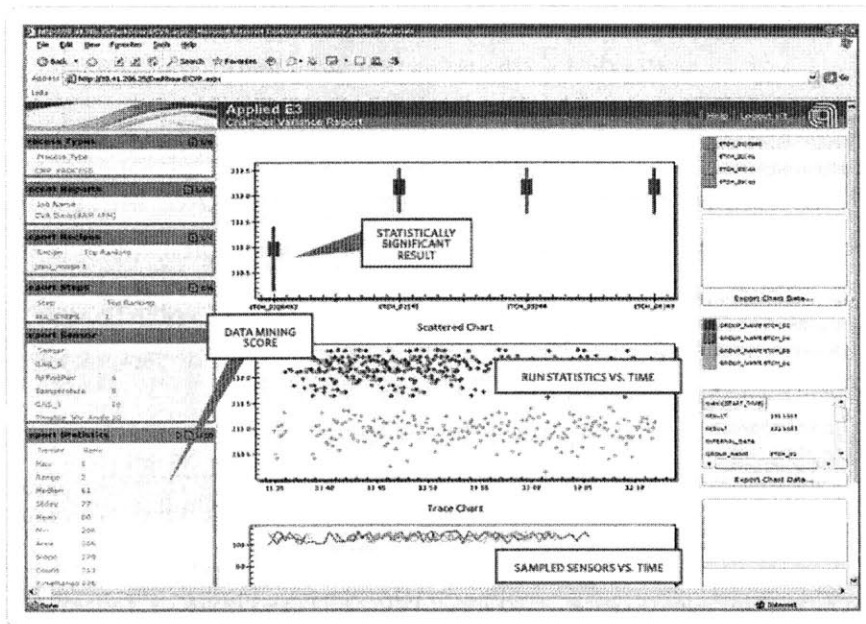


Figure 15: Machine matching software. The data statistics and sensor raw data are analyzed to find the root cause of the mismatch [20].

### 3.6. Run-to-Run Control

Run-to-run control has become increasingly popular in the semiconductor industry due to shrinking dimensions of integrated circuits. As the dimensions narrowed down, the need for precision increased. Previously, the processes were run with “fixed” recipes and were not being

retuned. However, to reduce the variability of the processes, a promising advance in the industry has been made which uses feedback control to adjust the recipe settings to correct for the mean shifts in the process as depicted in Figure 16. This type of approach requires an accurate model of the process that is usually a result of the regression of the experimental data that is updated at each run with the on-line measurements [23].

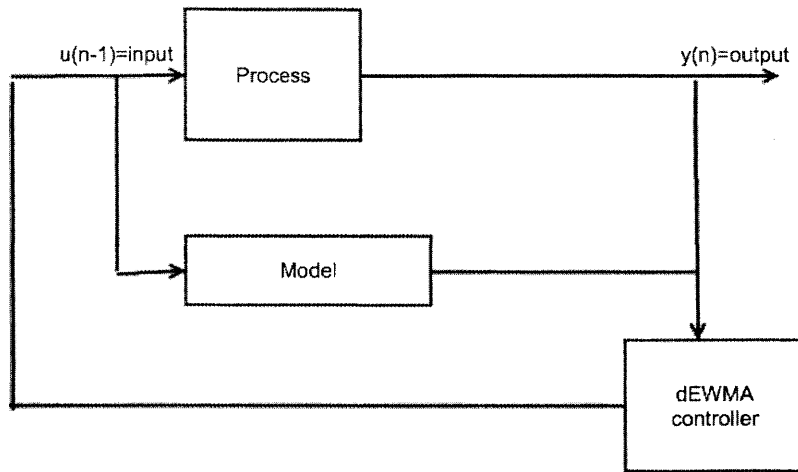


Figure 16: A basic double EWMA controller schematic [24].

One of the most widely used controllers employs an exponentially weighted moving average (EWMA) method to adjust process inputs. In controlling drifting processes, a predictor corrector control algorithm, a small modification of double EWMA (dEWMA) was shown to perform well. By adjusting two equations, the “predictor-corrector,” the drifts in the mean are compensated with feedback control [24]. Because most processes require multiple inputs and have multiple outputs, a multiple input and multiple output (MIMO) version has received a lot of attention. In this thesis, we will use only the single input single output (SISO) case to illustrate how run-to-run control can be implemented to match chambers by only modifying one input. Butler and Stefani use Equation (26) for SISO application [25].

$$y_t = \alpha + \beta u_{t-1} + \delta t + \epsilon_t \quad (26)$$

where  $y_t$  is the measured quality characteristic of run  $t$ ,  $u_{t-1}$  is the level of the controllable factor set at the end of run  $(t - 1)$ ,  $\delta$  is the average drift per run, and  $\{\epsilon_t\}_{t=1}^{\infty}$  is a white noise sequence. The model assumes a deterministic trend of disturbance, and a pooled value for the drift and noise terms ( $\delta t + \epsilon_t$ ). The parameter  $\alpha$  models any offset or bias from target,  $\beta$  is the input-output gain parameter and  $b$  is the estimator for  $\beta$ , acquired from the empirical model. Then the dEWMA controller is given by Equation (28) that modifies the intercept of the model and Equation (29) which predicts the offset that will be introduced in the next step.

$$u_t = \frac{(T - a_t - D_t)}{b} \quad (27)$$

$$a_t = \lambda_1(y_t - bu_t - 1) + (1 - \lambda_1)a_{t-1}; \quad 0 < \lambda_1 \leq 1, \quad (28)$$

$$D_t = \lambda_2(y_t - bu_{t-1} - a_{t-1}) + (1 - \lambda_2)D_{t-1}; \quad 0 < \lambda_2 \leq 1. \quad (29)$$

Since the offset term is coupled with the noise as can be seen in Equation (29), the second filter for offset might add some variability to the control if the noise inherent in the process is significant [26].



#### **4. Chamber Matching Methodology and its application to Plasma Ashing**

The chamber matching project took place concurrently with Analog Devices Internet of Things project. Analog Devices Inc. is concentrating its efforts to make its fabrication sites smarter by implementing Internet of Things. As a progressive company, ADI sees smart factories as providers of the flexibility and autonomy for manufacturing to adapt to the changing market conditions and customer needs. As a pilot project, real-time sensors and online data streaming capability were implemented in Gasonic plasma-ashing chambers. These new capabilities motivated us to illustrate how to evaluate the tools and match them using a feedback controller. Our efforts concentrated less on the implementation of these sensors and data streaming and more on how to diagnose the major differences in the machines, how to model the chambers, and how to use these models to match the machines and implement run-to-run controls.

##### **4.1. Issues in Plasma Ashing Chambers**

Gasonic machines G23, G33, G43, G53, G63 and G73 run the plasma ashing process in ADI's semiconductor device production facilities. Although the models of the machines are the same, there are differences between the physical and non-physical properties of these chambers. For example, the length of the gas exhaust of one machine can be different than the other, which might change the behavior of the gas and chamber.

Secondly, the machines go through major updates occasionally. For instance, the hardware and software have been updated to a newer model. Some of the machines look the same from outside but use a heated platen to heat the wafer as opposed to infrared lamps.

Despite the model standardization, each machine has parts replaced during preventive maintenance (PM) that are from a different vendor or different in version depending on the item's availability in the inventory. For example, the closed loop thermocouple (CLTC) probe, which is used to measure the temperature of the wafer during each step, is replaced frequently and calibrated by the maintenance team by measuring the amount removed, a repetitive process that requires maintenance team judgment.

Even if every chamber is at the same period in their lifetime, the number of wafers they process differs significantly. Figure 16 shows how many wafers each machine has produced so far. G73 has the most tasks and it is expected to perform worse than the G43 due to machine wear or aging.

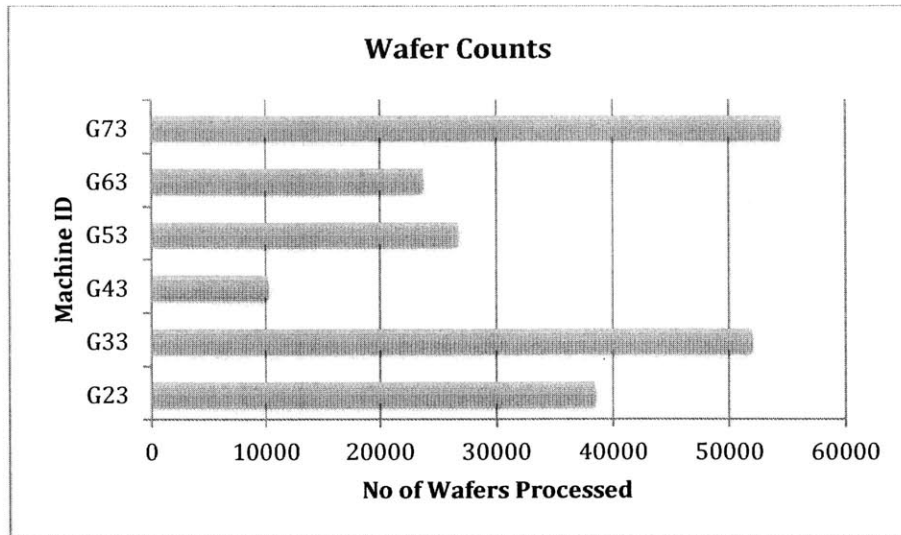


Figure 17: Every machine has processed a different number of wafers since they have been in operation. G73 has processed the most number of wafers and G43 the least.

Due to these differences, the performances of these machines are not exactly the same. In order to understand the differences in these machines, the first approach is to look at the historic data and performance of these machines, then to choose the best performing machine as a golden machine and compare the other machines to match it to the golden machine.

#### 4.2 Our Approach to Chamber Matching

Our approach to machine matching will concentrate on understanding the behaviors of the Gasonics machines by using statistical analysis and modeling utilizing the experimental data. Only one recipe, the partial recipe as described in Section 1.3.2, will be used for the experimental runs. The output of the stripping process is an amount of photoresist removed in the units of Angstroms ( $\text{\AA}$ ). We will refer to the process output the strip thickness or amount removed. This is a difference between the prior thickness and remaining thickness of the photoresist, before and after the partial strip process.

The process we designed is very similar to Baek's machine matching process flow, except that the recipe optimization will not be performed due to the vast variety of recipes run per machine and the limited time [8]. However, a run-to-run controller will be simulated to show how machine matching can be achieved in the future. Figure 17 shows the step-by-step methodology that was used in this project.

In addition to the empirical modeling of individual machines, we also answer if the machines are similar enough that we can represent the behavior of various machines in one combined model and decrease the number of estimators. In order to understand the root causes of these differences better, methods like instrumented wafers and surface plotting will be used to help to understand the reasons behind these differences.

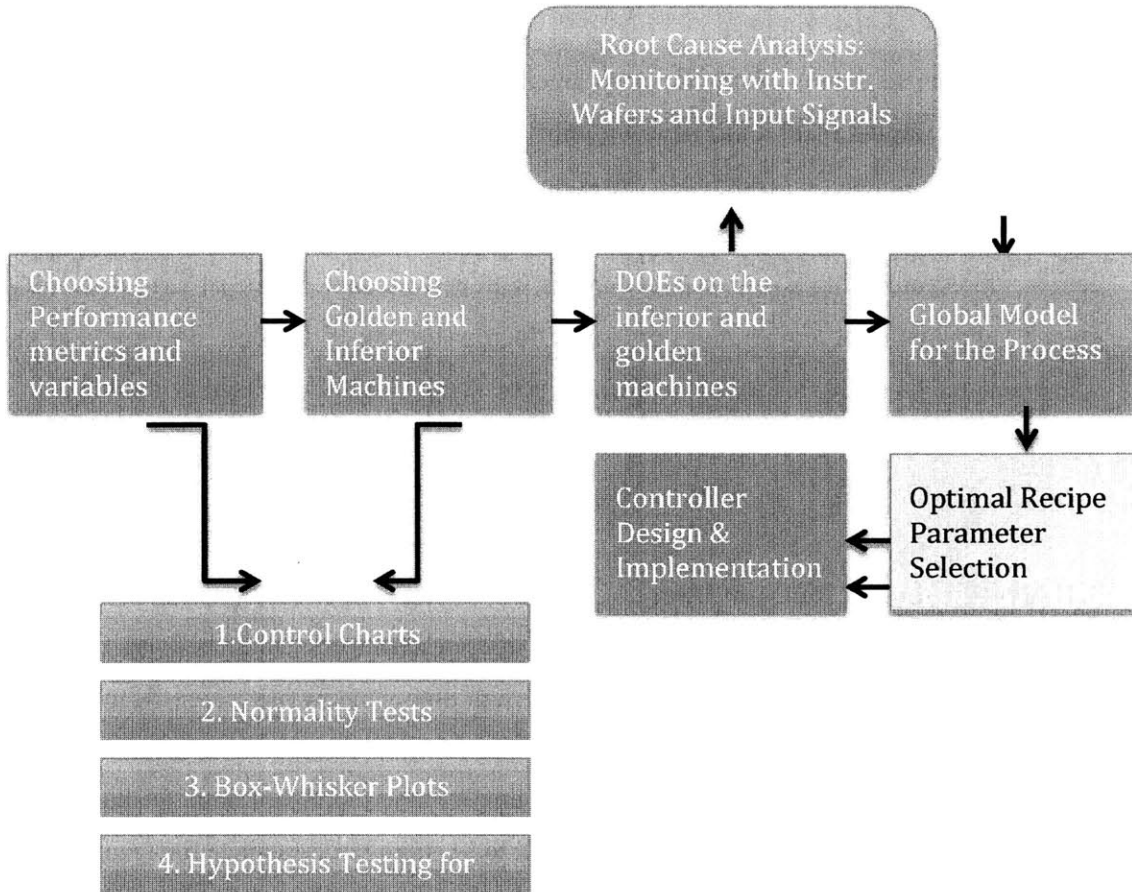


Figure 18: The chamber matching methodology followed in this project.

na

### 4.3. Performance Metric and Key Variables Selection

We use the target strip thickness as our target or controlled variable. Although uniformity is an important measure, the uniformity even in the extreme cases does not significantly affect the yields and therefore will not be considered for the machine matching comparison and modeling purposes. Due to the lack of end-of-line yield data, the strip thickness mean and variability will be used as a performance metric unlike what is proposed by Pan *et al.* as discussed in Section 3.3 [21].

In order to determine the input parameters, we consulted process experts who helped us define the six inputs that would have the most effect in the process. Because there is only one performance parameter considered, a raw gain matrix is not needed. A detailed discussion of the variables selected and the experimentation can be found in Nerurkar's thesis [2]. Each machine can be represented with a regression equation with performance metric of strip thickness of the layer  $y_1$  and inputs of temperature, pressure, microwave power, O<sub>2</sub> flow, H<sub>2</sub> flow and pumping speed denoted as  $u_1, u_2, u_3, u_4, u_5, u_6$  respectively. A noise factor of  $e$  is also included.

$$y_1 = f(u_1, u_2, u_3, u_4, u_5, u_6) + e \quad (30)$$

### 4.4. Identification of Relevant Data

In order to designate the golden and inferior chambers using historic data, the time window of the data needs to be identified. A certain time window has a significant effect on the choice of the machines that the team would like to match. For instance, three months of data can be a snapshot containing a mean shift in the course of nine months of data. Discussing with the ADI team, three months data window was found to be more relevant since the team believed that the latest maintenance issues should have more influence on the decision. Although three months of data was used to evaluate the statistical differences in the next couple of phases, nine months of data was kept in perspective to see whether it suggests otherwise. Among these machines, G13 was not in production and G33 was down most of the time, and therefore G13 and G33 are excluded from this analysis.

#### 4.5 Identification of Golden and Inferior Chambers

The comparison of the machines will use different tools to bring out the differences between the machines. Control charts have the most visible information as to the out-of-control points that each machine produced for the last three months. The visibility of control charts to the engineers makes them an appealing tool to develop an intuition about the stability for each machine. Secondly, the machines should be performing without any individual malfunction, so that they should have a normal distribution of outputs. The normality assumption was tested using a normality test to eliminate the machines that are producing dimensions not following a normal distribution.

Following the normality test, a non-parametric approach will be taken by producing box-and-whisker plots to visually see where the machines operate at. Barlett's test will then be used to test the hypothesis of whether or not all of the machines were performing at the same variation or if any of them is different. If the hypothesis of this test is negative, a machine-by-machine F-test approach will be taken to verify whether or not the variances are significant across different machines.

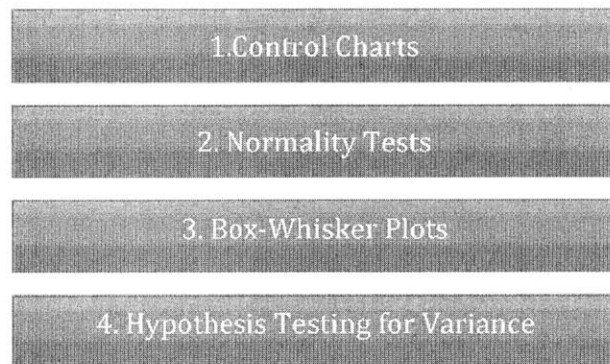


Figure 19: The statistical tools used to determine the golden and inferior machines.

#### **4.5.1. Control Charts**

Control charts are the most visible tools for manufacturing teams with which they monitor the process and make decisions to take action on any abnormal behavior. For the purposes of detecting machine differences, we evaluated the three months data with UCL (upper control limit) and LCL (lower control limits) set using the statistical analysis proposed by Nilgianskul [3].

It can be seen in Figure 20 that G53 and G63 have gone through a mean shift and did not quite stabilize since then. G23 is inherently operating at a lower strip thickness or amount removed than the rest of the machines. Looking at the log information for the past one month, it was seen that the engineers and maintenance team uncovered the root cause of this problem, and it was that the heat lamps were reversed and caused discrepancy in the temperature readings. Thus G23 can be categorized as an inferior machine due to its lower operating levels than the target around 6000 Å.

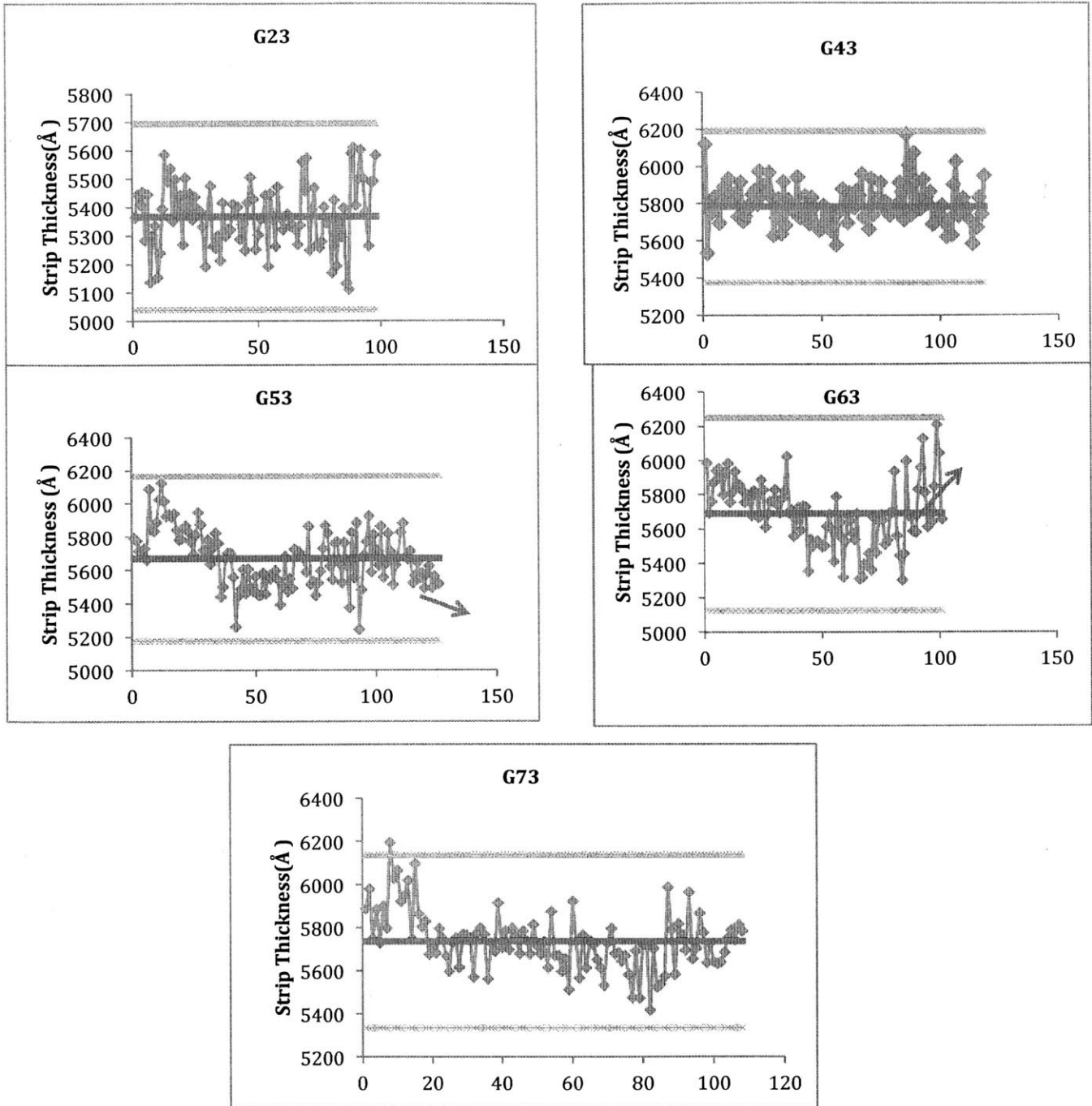


Figure 20: The control charts of the machines. G23 has been performing at the lowest mean and away from the target. G53 is showing downward trend and G63 upward trend for the last couple of days.

#### 4.5.2. Normality Test

A Lilliefors test has a null hypothesis that says that a population belongs to a normal distribution [27]. Using MATLAB software, the normality test was run all the machines. The normality test was run on all of the three month data points of G73 was 0.008 and the null hypothesis of normality needs to be rejected, whereas for G23, G43, G53 and G63 the normality is accepted with 0.05 significance level. Although it might be useful to investigate why the deviation has occurred, for the purposes of machine matching, G73 will not be eligible because the machines need to be at least following a normal distribution or else individual machine maintenance issues need to be addressed.

Normal probability plots display a normal probability plot overlaid on to the data to be tested to visually suggest whether the machines are following normal distribution. Figure 21 shows the comparison of the normal behavior of G53 in contrast with the G73 behavior. The deviation from the normal line shows that the behavior of the G73 is not normal.

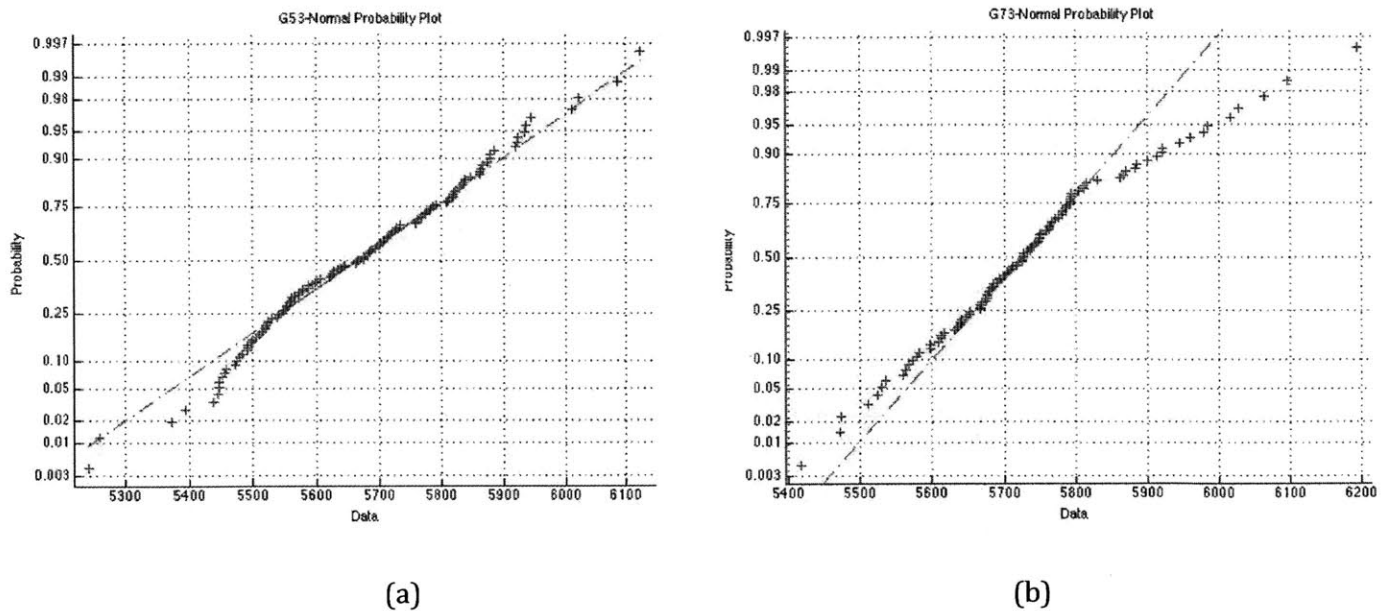


Figure 21: Normality plots of G63 and G53. (a) G53 exhibits normal behavior, (b) G73 deviates from the normal behavior line.



### 4.5.3. Box and Whisker Plots

Box and Whisker Plots depict data in quartiles and is used in this analysis to display at what point each machine operates. G23 has a mean lower than the target and a small standard deviation. G43 has a mean closest to the target value and has 50% of the data points very close to its operating point. G43 can be described as a golden machine looking at the plot in Figure 21. The variations and the means of the machines are summarized in Table 5 below. The calculations confirm that G73 has a standard deviation abnormally higher than the rest of the machines. G43 operates in the tightest output window.

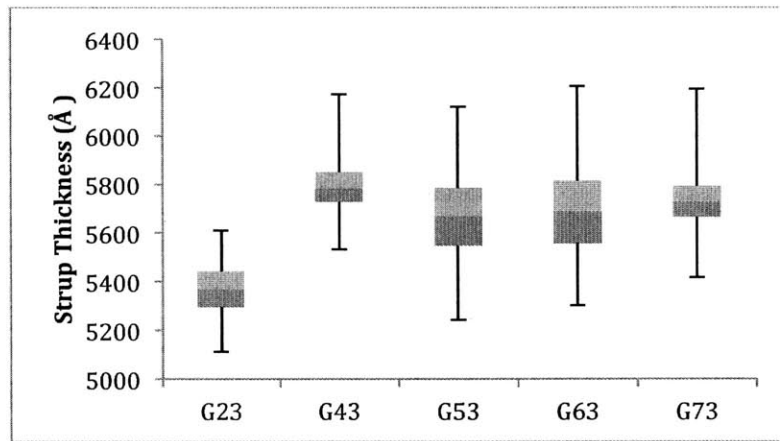


Figure 22: The box and whisker graphs. The quartile representation of machine data is used to decide which machine can be assigned as golden machine.

Table 5: The mean and standard deviation in Angstroms of each machine in the last three months.

Equip	Mean	STDEV
G13	5716.63	188.83
G23	5369.45	110.89
G43	5792.99	107.76
G53	5693.95	202.00
G63	5690.67	187.47
G73	5931.39	327.95

#### 4.5.4. Hypothesis Testing

In order to see if the standard deviations of the machines are significantly different, first Bartlett's Test is used and it is found with 0.05-confidence level that the null hypothesis, which assumes that these machines have the same variation, needs to be rejected. This result leads to using a piecewise F-Test to see which machines have similar variations. The result of the F-Tests reveals that the G23-G43 and G53-G63 can be grouped as having same variations with a 0.05 significance level.

Using the box plots and the operating levels in Figure 21, we choose G43 as the golden machine due to its low variation and high accuracy of thickness strip. However, due to the maintenance and production restrictions on machine G43, the team decided to focus on G53 and G63 instead. Although three months variations seem to be the same and the operating means are in the 50-70 Å range, the last month shows a visible upwards mean shift in G63 and downwards mean shift in G53. Due to the availability of the machines and the higher variation seen in the control charts for G63 in the previous month, G53 was chosen as the golden machine and G63 as the inferior machine.

#### 4.6. Design of Experiments and Multivariate Analysis

Once the key variables are selected, screening experiments using a fractional factorial  $2^{6-2}$  design with center points were run. An extensive explanation of the design of these experiments can be found in the work of Nerurkar [2]. Using the data from these designs of experiments, a process model was constructed to analyze and predict the manufacturing process behavior. The experiments which were run to understand the inputs and their effects on the process outputs were run solely on the machines G53 and G63. A regression equation with dependent variable of strip thickness in Angstroms (Å) in terms of uncorrelated independent variables including temperature, pressure, power, pumping speed, oxygen flow, nitrogen flow was attained for every site on the wafer with 0.01 significance level. Some second order interactions as well as main effects appear in the regression models as can be seen in Table 6. The model of the average amount removed was not modeled directly since it is not measured but calculated by appropriate weight of the wafer site strip thicknesses once these site models are attained. This method has proven more reliable than finding the average thicknesses and fitting a surface to these averages, because the weighted average is not a dimension we measure [6].

Table 6: The average strip thickness models are obtained by the weight of the site according to the wafer surface area each site represents.

Effect	Amount Removed Average	
	G53	G63
<b>Intercept</b>	5925.31	5505.57
<b>Temperature(220,250)</b>	1077.84	950.43
<b>Pressure(1600,2400)</b>	-5.82	-4.42
<b>Power(1100,1400)</b>	246.10	244.86
<b>Pumping Speed(4.8,7.2)</b>	-16.01	0
<b>O2(3000,4500)</b>	-4.94	63.88
<b>N2(300,450)</b>	-34.00	0
<b>Temperature*Pressure</b>	-48.55	-59.16
<b>Temperature*Power</b>	98.63	89.47
<b>Temperature*Pumping Speed</b>	0	0
<b>Temperature*O2</b>	0	7.02
<b>Temperature*N2</b>	0	0
<b>Pressure*Power</b>	0	0
<b>Pressure*Pumping Speed</b>	0.00	0.00

These findings suggest that temperature is the most significant factor for the process and for both of the machines, followed by power as the second most important factor. The effects that are significant for G53 can become insignificant for other sites as shown in the Figure 22. The magnitude of the main effects is much lower for G63 than G53. This is a result of the amount removed being higher for G53 than for G63 under the same settings. It is also important to note that although G53 and G63 were historically performing around the same mean, the DOE results revealed a difference in the means of around 500 Å.

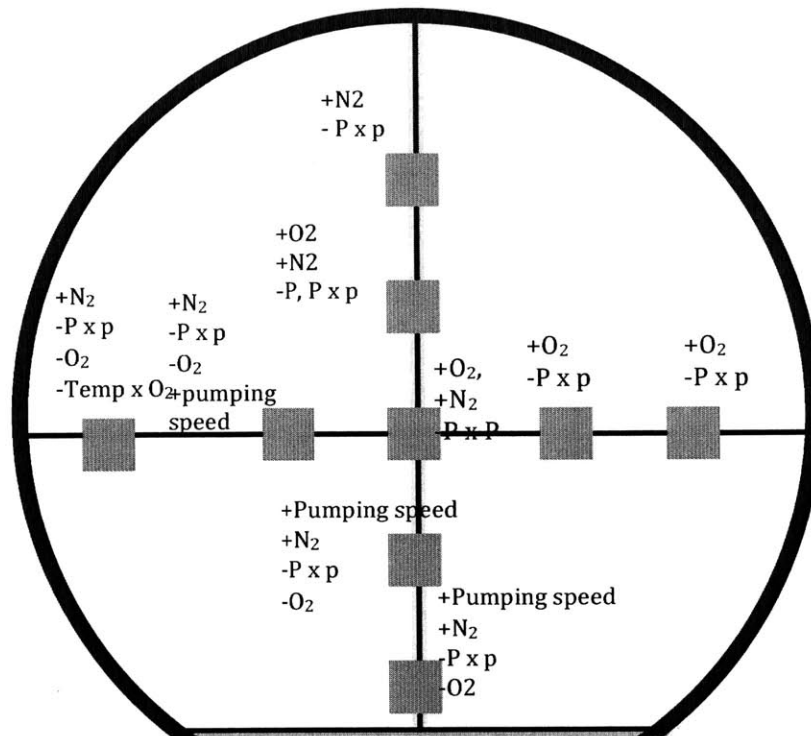


Figure 23: G53 factors affecting the output in comparison to G63. For every site, (+) means that the factor becomes effective for G53 but not in G63 and (-) means that factor becomes ineffective in G53 vice versa.

Table 7: Nine months of data control chart values.

Machine	Average X-BAR	UCL X-BAR	LCL X-BAR	AVERAGE NU
G23000	5367.18	5695	5039	13.88
G43000	5783.8	6193	5375	4.53
G53000	5669.75	6166	5175	7.9
G63000	5688.2	6251	5126	5.66
G73000	5735.13	6139	5332	6.05

#### 4.7. Global Process Modeling

The individual models for each machine are helpful to compare two machines in terms of their sensitivity to input changes. However, we hypothesize that the two machines can be combined into one model that can eliminate multiple regression coefficients that may be more than what is necessary to create a robust model to optimize the recipe parameters and to be used in a controller. A global model, which has a good enough fit, also suggests that the machines are performing around the same state.

Three possible cases might be valid in a combined model. The groups can differ (1) only in its intercept, (2) in the slopes or model coefficients for some of the factors or (3) in all of the factors in the final combined model [29].

The first approach is that only the intercepts may differ across the groups. To test this assumption, we compare the confidence intervals for the slopes (model factor coefficients) for each machine and see that they do not intersect, which tells us that the slopes are statistically different. However, our hypothesis is that we can introduce a categorical variable that can offset the intercept and force these machines into one global model with a good enough fit, where the only difference between the models is in the estimator for this categorical intercept variable. The slopes will be the same across these machines but y-intercept will be different and will be accounting for the offset between the machines. In order to test this assumption, two categorical variables are introduced as a factor. These categorical variables differ from the numerical variables in that they are either true or false for the categories they represent [29]. The categorical variables G53 and G63 are 0 or 1 depending on which machine is being run. It is crucial to notice that there is not a global intercept; rather, the y- intercept changes according to the machine that is being run. The resulting model shown in Equation (31) represents both G53 and G63 and has a regression fit ( $R^2$ ) ranging from site to site between 0.990 and 0.994 with significance level 0.01.

$$\begin{aligned} \text{strip thickness} = & 5834.93 * G53 + 5606.96 * G63 + 1066.80 * Temp + 245 * Power + \\ & 92.96 * Temp * Power - 50.60 * Temp * Pressure - 36.28 * N_2 + 33.61 * O_2 \end{aligned} \quad (31)$$

For each machine, the actual values of the experiments are compared with the global model prediction values. Figure 23 shows the site 1 model fits for G53 and G63 and Figure 24

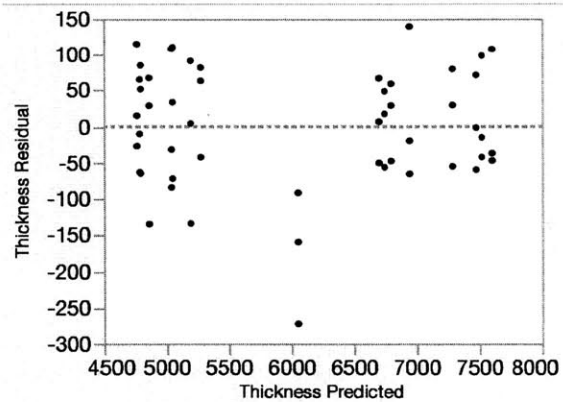
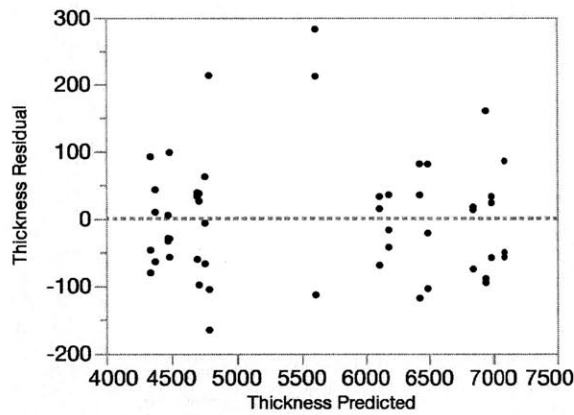
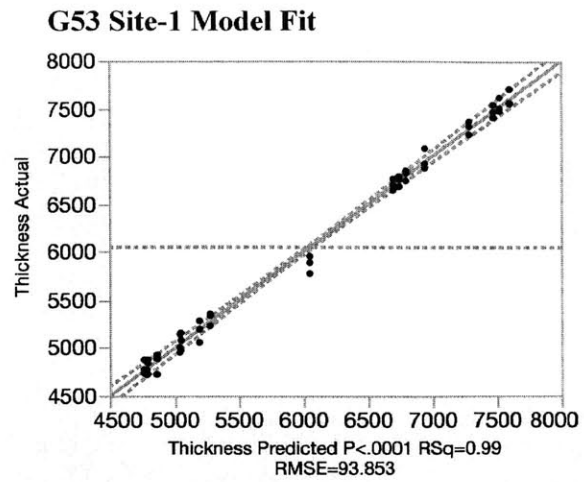
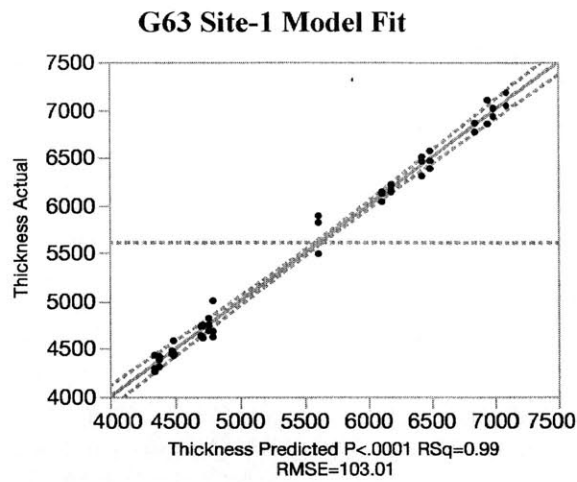
show the global model attained by introducing only unique offsets in y-intercepts. Both of these figures are presented with the residual versus predicted output values. Any trend in residual versus predicted graph such as an increasing value with the predicted value or vice versa would suggest that the model is not right, and a function of the output or input might need to be introduced. The random behavior of the residuals versus predicted values along with the  $R^2$  greater than 0.99 indicate that these models are good fits for the data.

The  $R^2$  values for individual model fits is different from the global model with individual machine y-intercepts only by 0.2% for G63 and 0.4% for G53. By sacrificing only 0.4% of goodness in fit, the global model with fewer coefficients can be effectively used.

Although this first approach where we assume there are no machine specific differences in the input coefficients gives a good enough model, we can improve the fit with the second and third cases where the estimators are also different for each model. We introduced some interaction terms with categorical variables to the equation. The least square fit results in a regression model that not only has categorical terms for G53 and G63 but also G63 interacting with the temperature, the biggest effector in the individual models. The following regression in Equation (32) with the significant effectors is obtained for Site 1:

$$\begin{aligned} \text{strip thickness} = & 5834.93 * G53 + 5606.96 * G63 + 1066.80 * T + 245 * Power + 92.96 * Temp * \\ & Power - 50.60 * Temp * Pressure + \mathbf{80.90 * Temp * (G53 - 0.5)} - 36.28 * N_2 + 33.61 * O_2 \end{aligned} \quad (32)$$

Although global models are inferior to the individual models as expected, they are good enough models and indicate that these two machines are at similar states. As the historic data suggests, these machines are not substantially different than each other. These particular machines were chosen due to the task availability and fine-tuning purposes. The global model can also be used in recipe optimization and use of global models in different cases and machines can surface the differences between machines that might be in similar or totally different states.



#### Summary of Fit

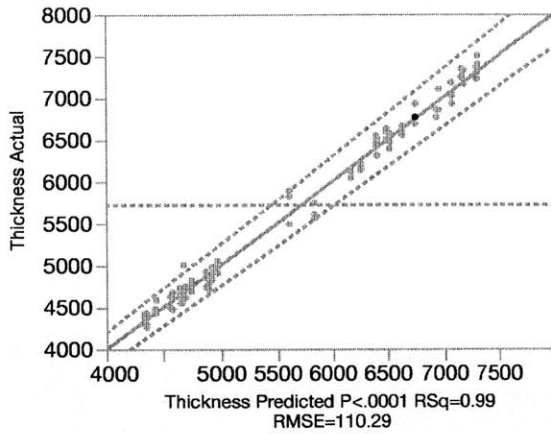
RSquare	0.992629
RSquare Adj	0.990301
Root Mean Square Error	103.0128
Mean of Response	5606.956
Observations (or Sum Wgts)	51

#### Summary of Fit

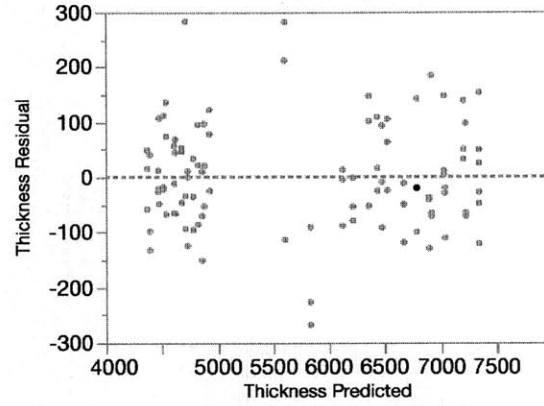
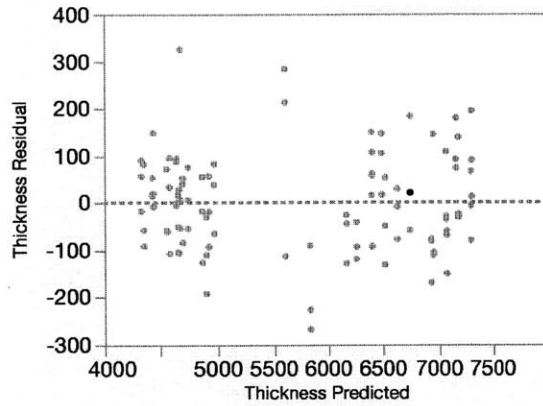
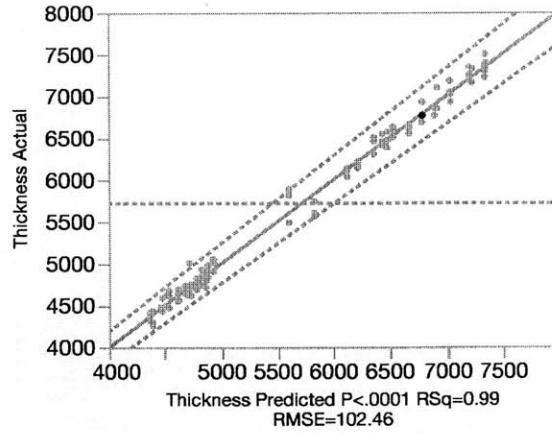
RSquare	0.99445
RSquare Adj	0.992697
Root Mean Square Error	93.85303
Mean of Response	6048.979
Observations (or Sum Wgts)	51

Figure 24: (Left) The individual model fit for G63 for its site-1. Actual vs. predicted has a slope close to 1. Residual vs. predicted is distributed randomly and no trend can be observed. The summary of fit shows that the model fit is very good. (Right) The individual model fit for G53 for its site-1. Actual vs. predicted has a slope close to 1 except the center points are well below the model. Residual vs. predicted is distributed randomly and no trend can be observed. Summary of fit shows  $R^2 > 0.99$ .

**GM: only intercepts different**



**GM: slopes & y-intercept different**



**Summary of Fit**

RSquare	0.990994
RSquare Adj	0.989664
Root Mean Square Error	110.2899
Mean of Response	5720.944
Observations (or Sum Wgts)	102

**Summary of Fit**

RSquare	0.992316
RSquare Adj	0.991079
Root Mean Square Error	102.4596
Mean of Response	5720.944
Observations (or Sum Wgts)	102

Figure 25: (Left) The global model fit when G53 and G63 experiment outputs are pooled and a categorical variable is introduced for the intercept. (Right) The global model fit with not only the intercept but also an interaction of a main effector and the categorical variable. The model predicts the variability around 0.2% better than the global model with only y-intercept change, yet 0.2% worse than the individual model.



#### **4.8. Diagnosis of the Causes of the Mismatch**

The models we acquired above suggest that there is a significant mismatch between G53 and G63. It was shown that on a site-by-site basis, different factors become effective and their effect on the process output differs. Moreover, the magnitude of the effects of the factors differs from machine to machine. Before attempting to implement a controller, a root cause analysis approach was taken to understand how the process improvement should be shaped. Because temperature was found to be the most significant factor, the diagnosis focuses on analyzing the “real” temperature on the wafer measured by an instrumented wafer and temperature signals from the temperature sensor that was implemented as part of the pilot Internet of Things (IoT) project.

##### **4.8.1 Diagnosis with Instrumented Wafers**

Instrumented wafers can be effective in determining the causes of the chamber mismatch [30]. KLA-Tencor’s SensArray SensorWafers can be used to make measurements like temperature and RF voltage. In the analysis made with these wafers, the etch shows a reasonable correlation of around 0.7 between the temperature and etch rate. Higher temperatures result in higher etches rates.

Because the temperature effect for G53 is different than G63 by around  $120 \text{ \AA/ C}^\circ$ , we hypothesize that real temperatures sensed on the wafers are different for G53 and G63. Using KLA-Tencor SensArrays, which has nine built-in sites that measures the temperature directly on the wafer, the temperature maps in Figure 25 and 26 were generated for the partial recipe for which the temperature setting is  $235 \text{ C}^\circ$ . This measurement reveals that although the temperature dials are set to  $235 \text{ C}^\circ$ , the temperature on the wafer that is being processed is  $13.3 \text{ C}^\circ$  lower than in G53.

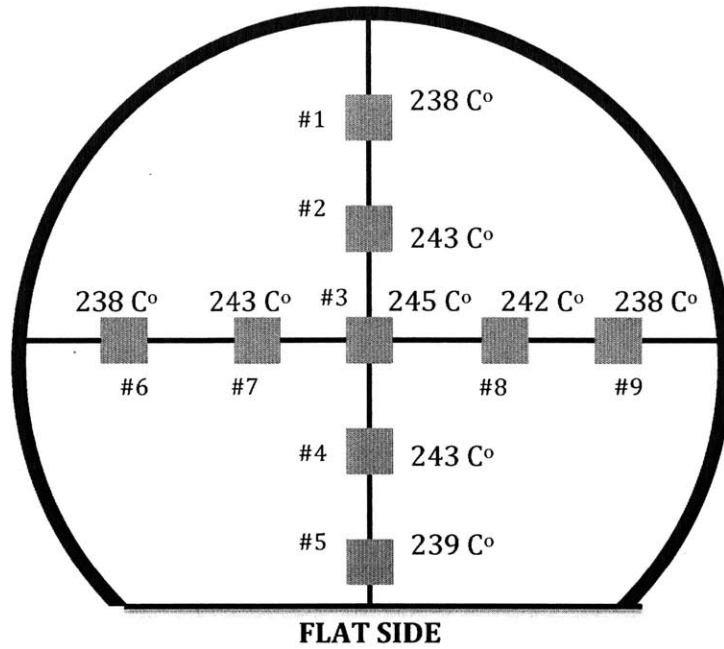


Figure 26: G53 temperature readings when T=235 C° with instrumented wafer.

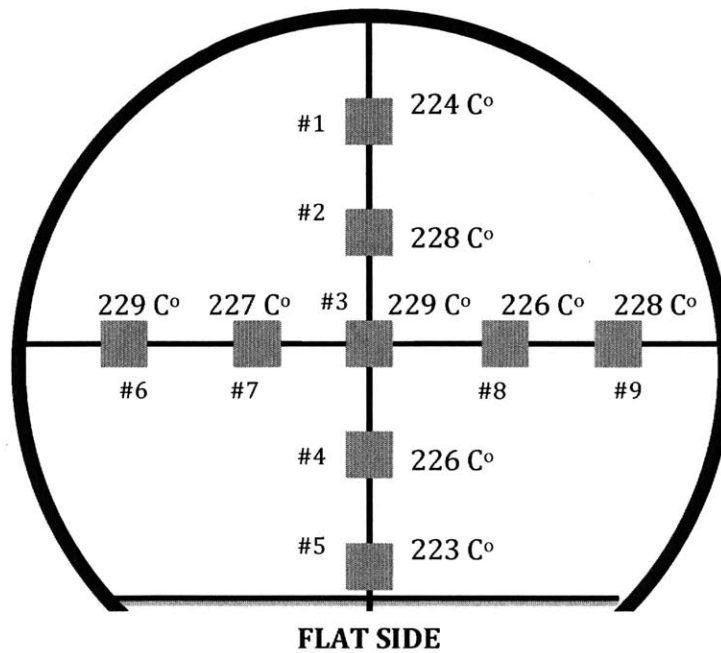


Figure 27: G63 Temperature readings when T=235 C° with instrumented wafer.

Another problem that was surfaced is that the temperature profile is not the same across the y-axis. Sites 7 and 8 seem to have lower temperatures than the edges for G63, in contrast with the dome look of the G53 data that has decreasing etch rates towards the corners for both x and y-axes as seen in Figure 28.

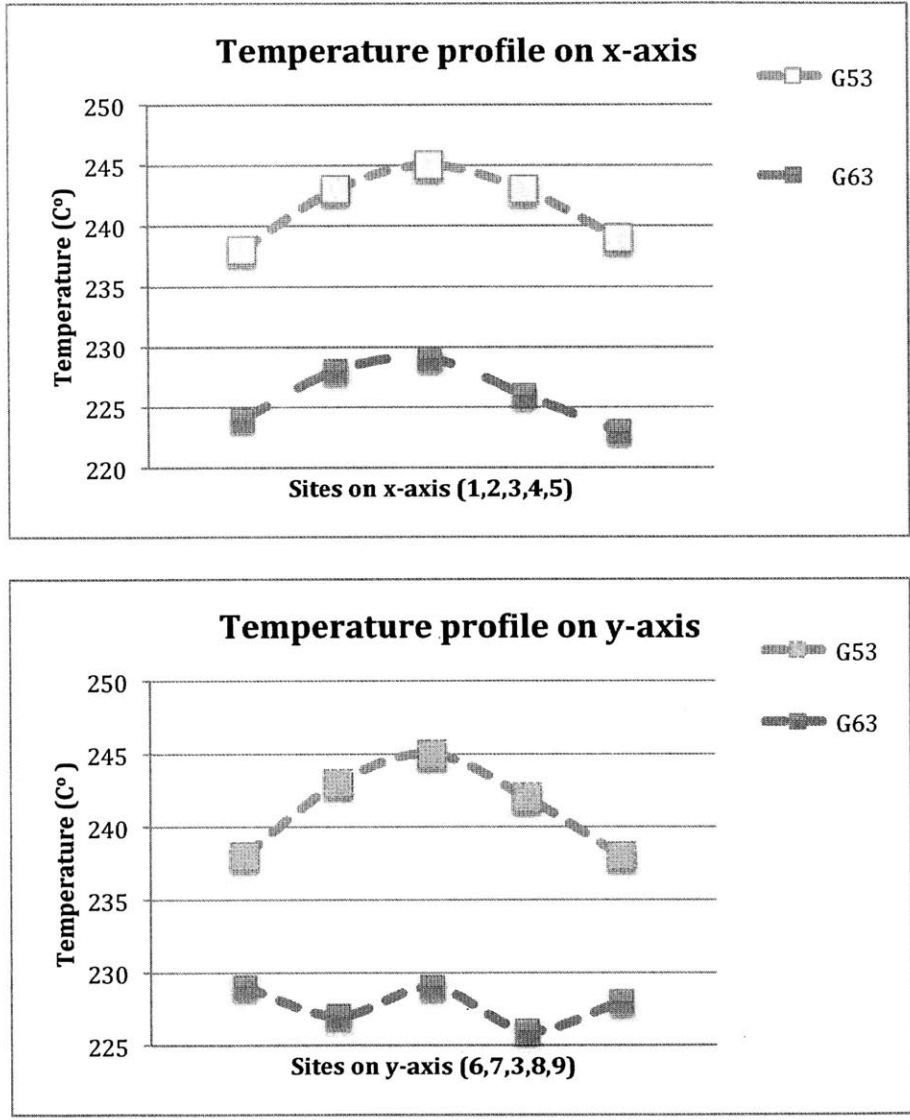


Figure 28: Temperature profiles on x and y axis of the wafer, respectively. G63 performs at a lower temperature on both axes. Y-axis temperature profile has also a different shape than x-axis.

The Gasonics machine contains a closed loop temperature control (CLTC ) probe which senses the temperature and controls the temperature on the wafer. CLTC is one of the more frequently changed machine parts and comes in different models from various manufacturers. Every CLTC probe is calibrated by running a series of tests on wafers to see how close the output is to the results. For instance, while one machine has a constant of  $K = 0.1$  another can have a CLTC tuning /controller constant of  $K = 0.28$ . The hardware and setting differences between the controllers might be the underlying cause of the temperature mismatch. The calibration of these CLTC probes should be standardized.

#### **4.8.2 Diagnosis with Surface Mapping**

The temperature and the non-uniformity relation across the wafer is more challenging to understand and correlate than the average temperature effects since many other effects might be contributing to the non-uniformity [30]. In order to help with understanding these effects, we created a spatial surface model running a partial recipe on one of the wafers.

For accuracy of the spatial modeling, rather than nine point measurements, 49 points are picked radially away from the center as seen in Figure 29 below. Using the pre and post measurements on these 49 points, a TriScatterInterpolant with corresponding coordinates on the wafer X, Y and strip thickness as Z were used (please refer to the coding scheme in Appendix B). The TriscatterInterpolant function in MATLAB uses Barycentric Interpolation to assign thicknesses for each coordinate in the X-Y space using the values of the thickness measured at the 49 locations.

Figure 29 shows the surface of a wafer after the partial recipe is run. The red color suggests that more photoresist was removed in the middle of the wafer. This surface is expected since the temperature is not even across the wafer. Instead, it decreases as towards the edges of the wafer. However, on the negative y -axis, the strip thickness is still high unlike on the other top edge. To investigate this issue, we worked with the maintenance team to generate some hypotheses. The most plausible explanation is that the lower edge is the wafer flat and due to the short distance of the gases to travel before being removed and the sharp edges of the wafer, turbulent flow and vortices might be causing the gas to flow differently which might increase the rate of reaction in this region.

Rather than delving into the theory behind this hypothesis, we designed an experiment that will keep all of the other effects like the vacuum pump and sapphire pin locations constant

but change the location of the flat in the chamber. The wafer was placed at  $270^\circ$  rotation from its original position. The same 49-point measurement procedure was followed and a surface map is generated. The resultant surface map that was shown in Equation (29) suggests that the behavior is still seen at the flat edge, proving that the “wafer flat effect” is the root cause of the unexpected behavior, rather than nonuniform temperature distributions.

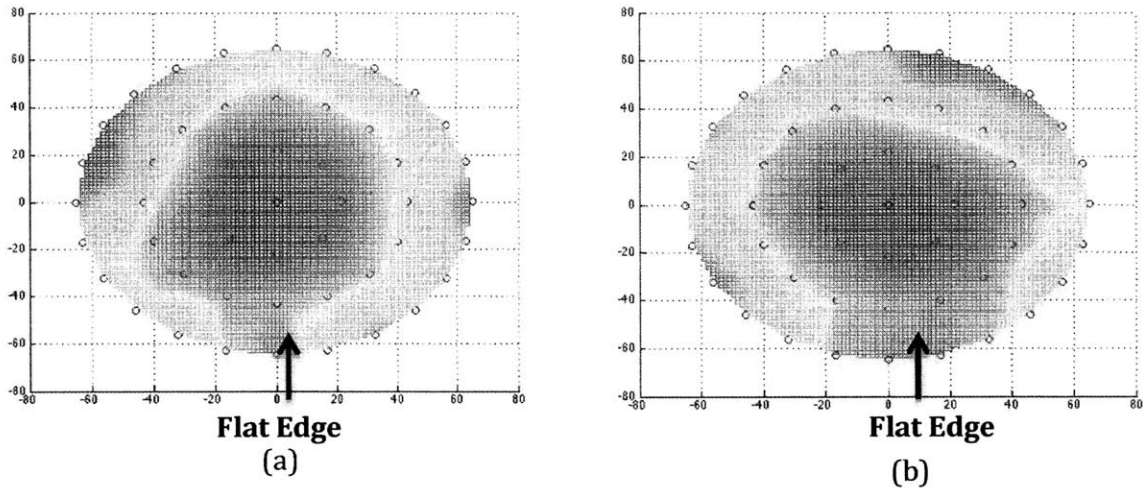


Figure 29: (a) Surface profile map of the wafer after processed by the partial ash recipe in G53. (b) Surface profile map after the wafer is rotated 270 degrees and processed with the same partial recipe in G53.

To understand if the wafer surface would be the same for machine G63, another wafer was processed with the same recipe using G63 and the surface map was generated. As seen in Figure 30, this time right side of the wafer had higher strip thickness that is not on the flat side. The flat side effect could still be arguably seen on the bottom edge. Other physical differences between machines such as the position of the RF lamps or the vacuum pump outlets can be used to explain the cause of the behavior. One of the common reasons is a gas leak. Figure 31 shows 49-point square measurements made from -46 to 46 mm on both x and y axes and reveals the possibility of a leak in G63 that might be causing higher etch rates on the right side of the wafer. The nonuniformity for the wafer processed in G53 and G63 is 2.3% and 2.5% respectively, which is not a big difference.

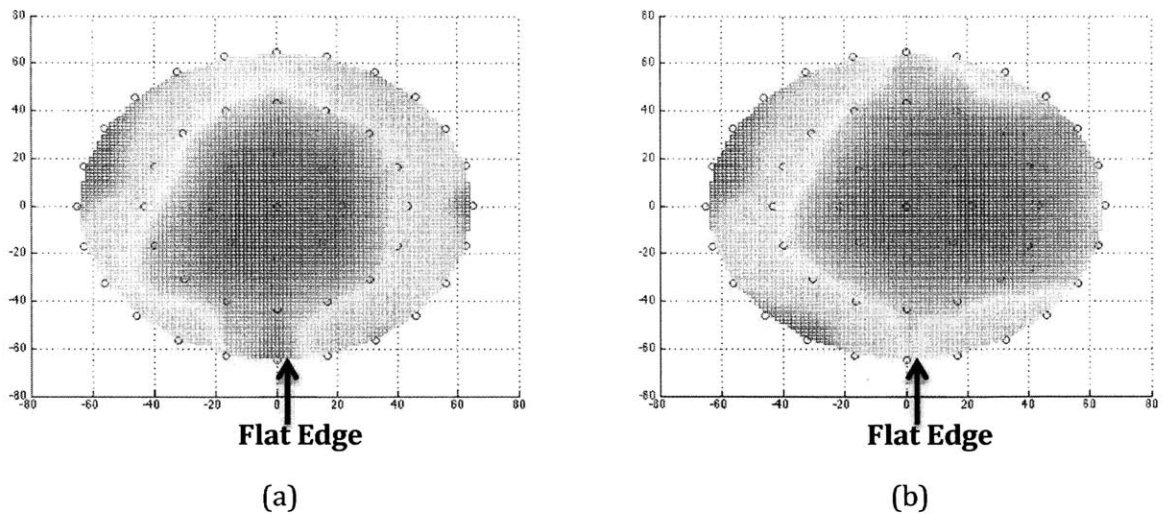


Figure 30: (a) Radial surface profile map of the wafer processed by the partial ash recipe in G53. (b) Radial surface profile map of the wafer as processed by the partial recipe in G63.

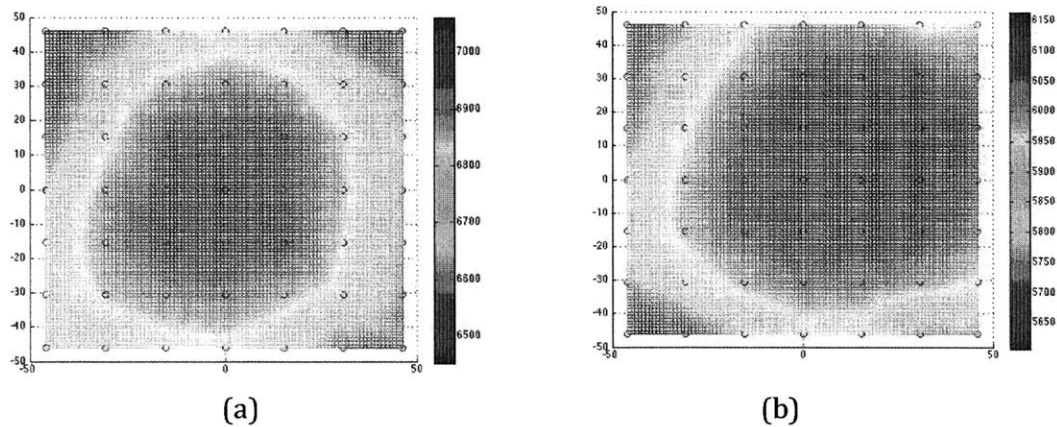


Figure 31: (a) Square surface profile map of the wafer processed by the partial ash recipe in G53. (b) Square surface profile map of the wafer as processed by the partial recipe in G63.

#### 4.8.3. Diagnosis with In-Situ Sensors

With advances in sensors and connectivity, it is easier to track the tool health and yield issues by not only looking at the end-of-line measurements but also at the in-situ process. advanced control systems like predictive maintenance use in-situ sensor monitoring to predict and avert any maintenance issues that are likely to occur in the chamber, looking at the trends in

these signals. In this section, we will illustrate how signal monitoring can enhance chamber matching by diagnosing the root causes of differences before matching; and the deviations from matched state if any after chamber matching.

Using Core Tegral software, sample data as shown in Table 8 from the sensors were streamed from the Gasonics machines. In order to stream data, the machines should not be running in local mode. Because G53 and G63 were taken out of production to run design of experiments for the analysis described above and the Core Tegral Software has only been running for a month, there was not sufficient data available from these machines to support or refute the arguments made above to explain the differences between the machines.

Table 8: Streamed data types with their respective units.

RF Forward Power	watts
MFC1, MFC2 and MFC3	sccm
Wafer Temperature	°C
Lamp Power	%
Pressure	Torr

For example, Analog Devices use closed loop temperature control probe (CLTC) as an in-situ sensor to control the temperature of a wafer that is known to be the most significant factor in process output and yields. Therefore, in this section, we will use the temperature signals from machines G23, G33 and G43 to illustrate monitoring at input signal level.

Figure 31 illustrates the temperature signal during a “complete” process recipe that lasts around 2 minutes. MATLAB software was used to align the data and signals. Sampling frequency of G23 was slower than for the other two machines. The signals were matched to the signaling frequency of the G23 (please refer to Appendix A). G23 signal does not seem to stabilize throughout the process and instead oscillates with amplitude of 5 C<sup>o</sup> while G33 and G43 seem to have stabilized quickly. This behavior difference might be stemming from the fact that there is no standardized way to calibrate the probes. The calibration constants of the probes can be different and causing the problem seen in Figure 32.

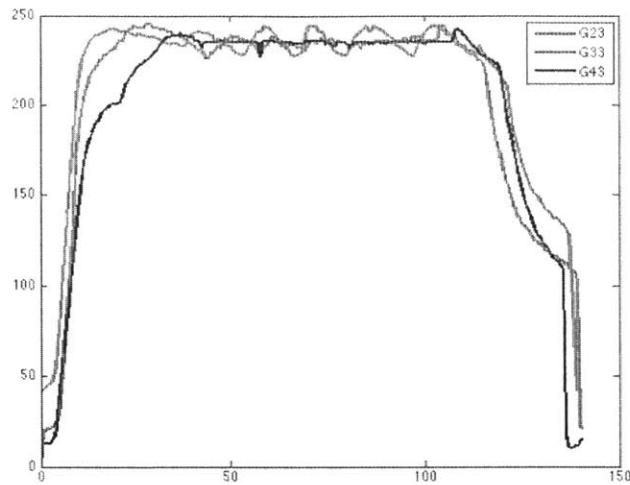


Figure 32: The partial ash recipe temperature profile from CLTC raw data for G23, G33 and G43. G23 overshoots and oscillates around the target, whereas G43 and G33 settle quickly.

However, a deeper look into the behavior of G33 at a different time in Figure 33 reveals that between June 9<sup>th</sup> and June 14<sup>th</sup>, there is a significant change in signal behavior. On June 14<sup>th</sup>, the signal overshoot the target and kept oscillating. Since the probe has not been recalibrated or changed between these two dates, another hypothesis around the behavior was proposed. The variability might be due to the type of batch and how many layers of processing every wafer went through. This means that for a new wafer batch coming in, the signal might behave differently.

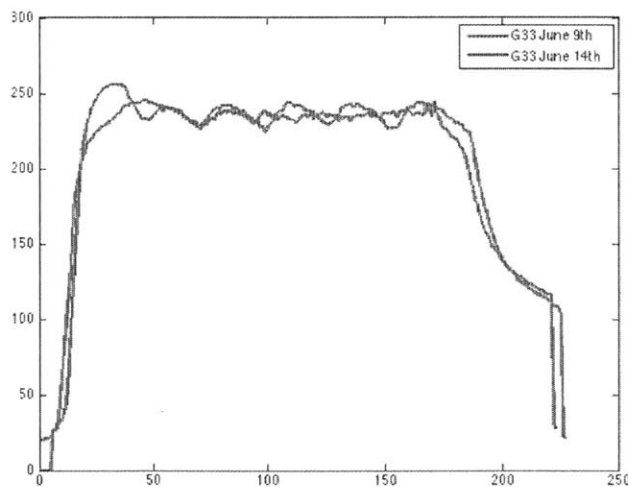


Figure 33: G33 CLTC temperature profile on June 9<sup>th</sup> and 14<sup>th</sup> is shown. A different overshoot and oscillatory behavior is observed. This behavior can be explained by either a time or batch dependent phenomena.



Taking a look at the G33 data stream on June 14<sup>th</sup>, we see that the oscillatory behavior continued for four consecutive wafers as shown in Figure 34. We are unable to say whether there was a change after the four consecutive wafers or in another batch due to the unavailability of data. However, the batch dependency can be caused by refractory surfaces. For instance, wafers, which have abundant metal atoms in the strip material, can create refractory surface and cause the CLTC probe with a certain calibration constant to behave differently.

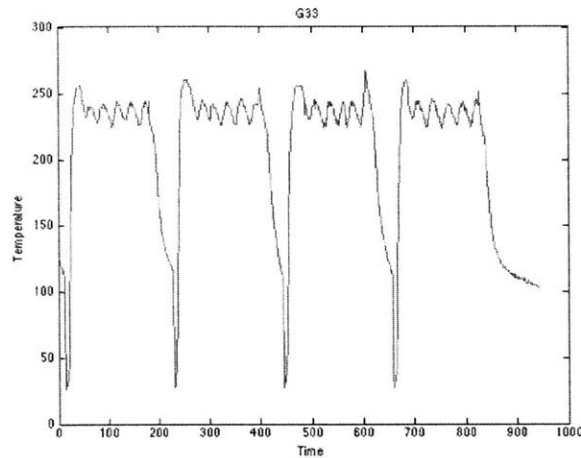


Figure 34: G33 CLTC temperature profile of four consecutive wafers showing the same oscillation characteristic.

The same recipe above was also run in G23 generating Figure 35, with the process time becoming smaller and smaller. This unnatural behavior could be a sample rate variation problem. While monitoring in real-time, sampling rates need to be uniform or trends that are not inherent in the process can appear and contaminate the features that are used in recognizing trends and predicting the need for maintenance.

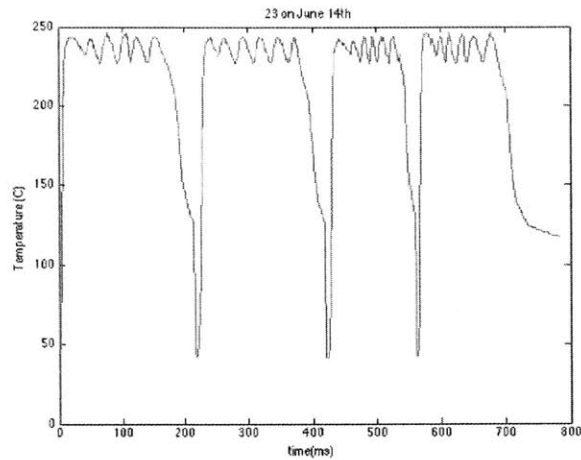


Figure 35: Complete ash recipe on four wafers in G23. Sample rate variation on the third and last wafer run is revealed.

However, more obvious malfunction alarms are detected using in situ sensor data when the malfunction was partially due to or affected the temperature measurement. For example, digging into the PROMIS log data for Out-of-Spec (OOS), the temperature readings were retrieved. Figure 36 shows that clearly the temperature was not able to ramp up for the second wafer processed; this can be due to many reasons, but this in-situ sensor data point at problems that would involve temperature.

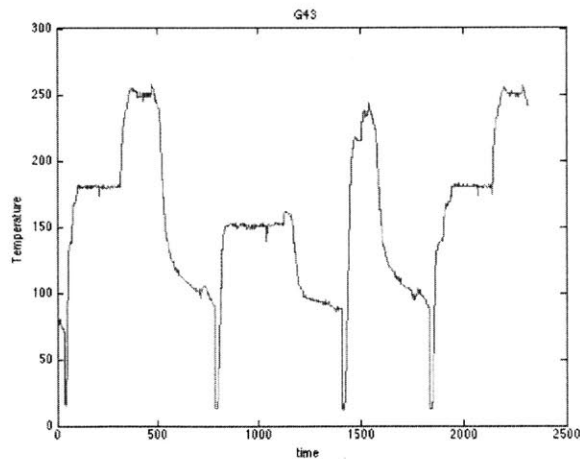


Figure 36: G43 temperature profile from CLTC running a “thick” ash recipe. The PROMIS log showed that the machine had out-of-spec error that corresponds to this input signal.

After technical problems like data sampling variations are addressed, the in-situ sensor monitoring can become a robust tool in both predicting malfunctions, diagnosing the causes of the malfunction and finally uncovering the machine mismatch. More dynamic decision-making can also be possible. For instance, the chamber itself can update the calibration factor in the recipe type according to the surface material and layer type of the wafers and potentially eliminate the oscillatory behavior of the CLTC.

#### 4.9. Run-to-run Control Application

The individual models created off-line were used to implement SISO double EWMA with both predictor and run-to-run control. A single input in this case was picked because single most effective input of the process is temperature and because it was found that the root cause of the machine performance difference is often the temperature control discrepancies between the machines. Moreover, some effects such as pumping speed are not easily controlled in the Gasonics machines. For practical reasons, temperature will be the single input for this process and interaction terms will be ignored.

Monte Carlo simulation with 100 wafer runs was performed using the individual models of the machines. We introduced a drift to the output up to 5% of the target thickness. To mimic noise and drift cases that are very similar to the real phenomena in these machines, a normal noise with standard deviation ( $\sigma$ ) of 30 and 50 Å was respectively introduced on top of the deterministic change of  $\delta = 3$  Å per wafer run. The deterministic drift went until 5% of the target thickness (T) of 6000 Å is reached. The weights for the intercept ( $\gamma_1$ ) and the offset ( $\gamma_2$ ) EWMA updates were picked as 0.1. The model factors except for temperature are lumped in a term we will call effects, and the second order interactions will not be taken into account for the time being. Equation 34 and 35 illustrates how the model is used for EWMA control.

$$effects = \beta_{power} * Power + \beta_{press} Pres + \beta_{P.speed} P.speed + \beta_{N_2} N_2 + \beta_{O_2} O_2 \quad (33)$$

$$strip\ thickness = intercept + \beta_{temp} * Temp + Effects + \delta t + \epsilon_t \quad (34)$$

The EWMA weight is critical to the process because higher values result in faster response but also increase the output noise [31]. Standard deviation of the output gives us an estimate of the output noise. In order to find the optimal operating point, different weights were

used and the standard deviation of the output was tracked for G53. Figure 37 shows that the optimal weight to minimize the errors is 0.1. Although the response time will increase, the noise reduction will be taken in to account when the weights are assigned as 0.1.

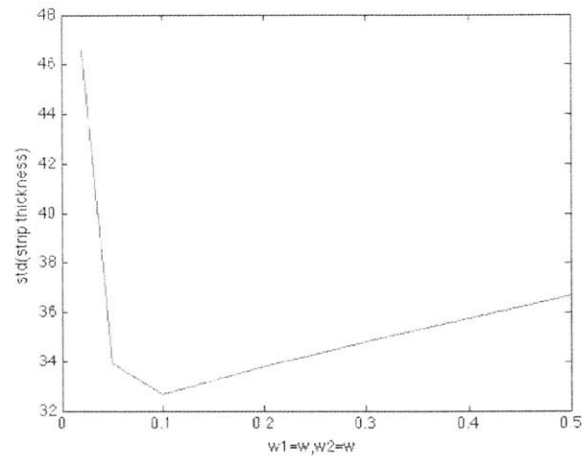


Figure 37: Output noise change as a function of the EWMA weight terms  $\gamma_1, \gamma_2$ .

MATLAB software was used to run the simulation (please refer to Appendix C). The controller simulation shows that the controller successfully adjusts the G63 output that is historically higher than in G53 down to behave at the target levels by decreasing its temperature linearly. Implementing this control, the temperature control stays within the acceptable temperature range window of 210-250 C° as shown in Figure 38.

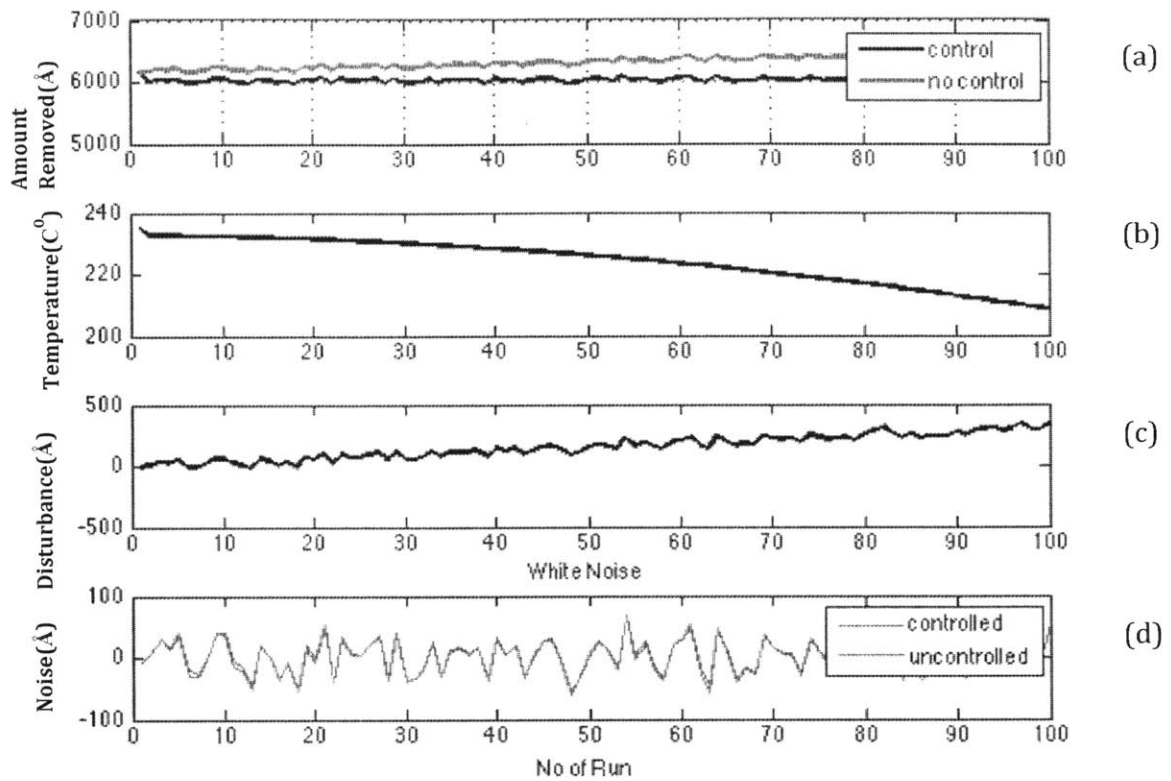


Figure 38: Double EWMA Control is simulated for 100 wafer runs with individual G53 model. The drift of  $3 \text{ \AA} / \text{wafer}$  and normal noise with  $\mu = 0, \sigma = 30$ . 1.a. The process comes back to the target value due to the control. 1.b. The temperature decreases abruptly at first and then exponentially to account for the linear upwards drift. 1.c. The linear upward drift introduced to the process up to  $300 \text{ \AA}$  or 5% of the target value. 1.d. The output noise does not change significantly, showing that the controller is reliable at low levels of noise.

The control scheme was found to have coupled noise and offset which can cause the input to react to the noise if it is significant. To see the effect of noise magnitude, noise standard deviation was amplified from 30 to 50 Angstroms. Figure 39 shows an input noise (no control case) and the output noise (control) when the noise is amplified from 30 to 50 Å to illustrate the output noise that is not that aligned.

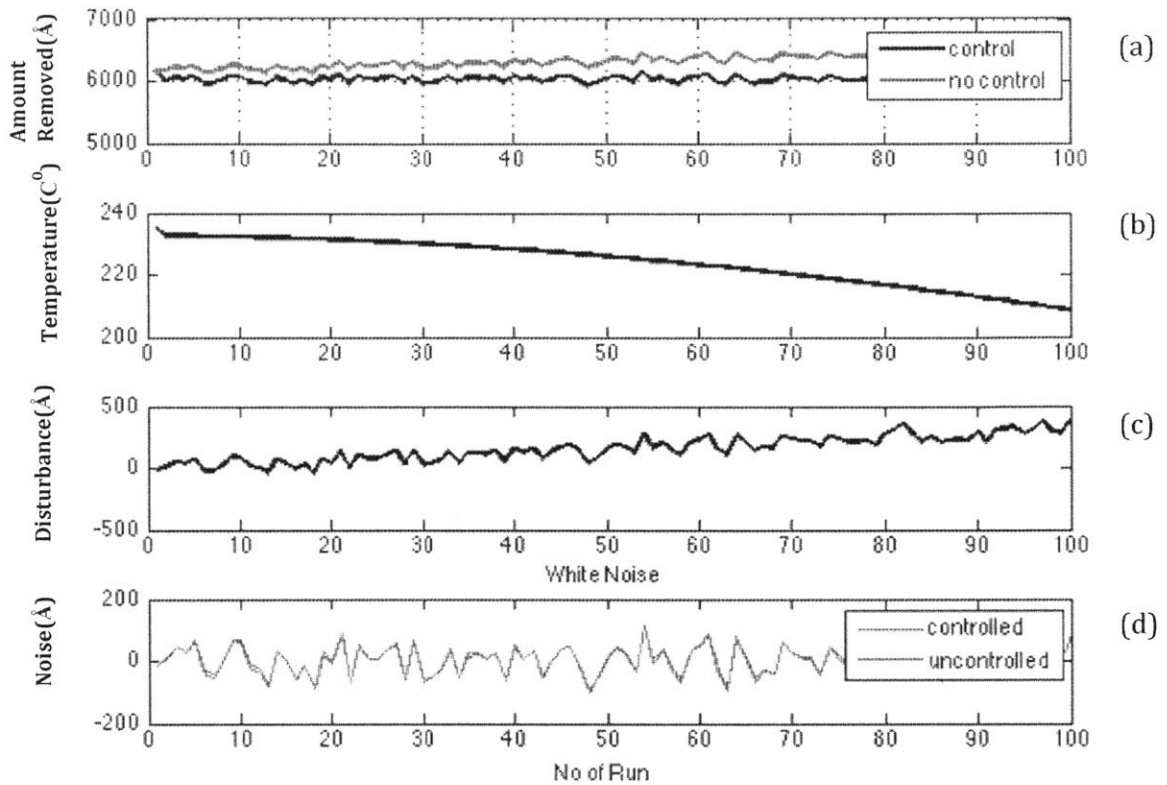


Figure 39: Double EWMA control is implemented on G53 with a drift of 3 Å /wafer and normal noise  $\mu = 0, \sigma = 50$ . 1(a) The process comes back to the target value thanks to the control. 1 (b) The temperature decreases to account for the linear upwards drift. 1(c) The linear upward drift introduced to the process of up to 300Å or 5% of the target value. 1(d) The difference in output noise between the controlled and uncontrolled process increases only slightly.

When the same controller and simulation was run with the G63 Machine with noise  $\mu = 0, \sigma = 50$  Å, the controller accounts for the difference again by pushing the temperature up first and then down as the drift increases. It is important to realize, especially when the noise is large with respect to the accumulated drift, that the controller amplifies the noise (which is not desired) as can be seen in Figure 40.

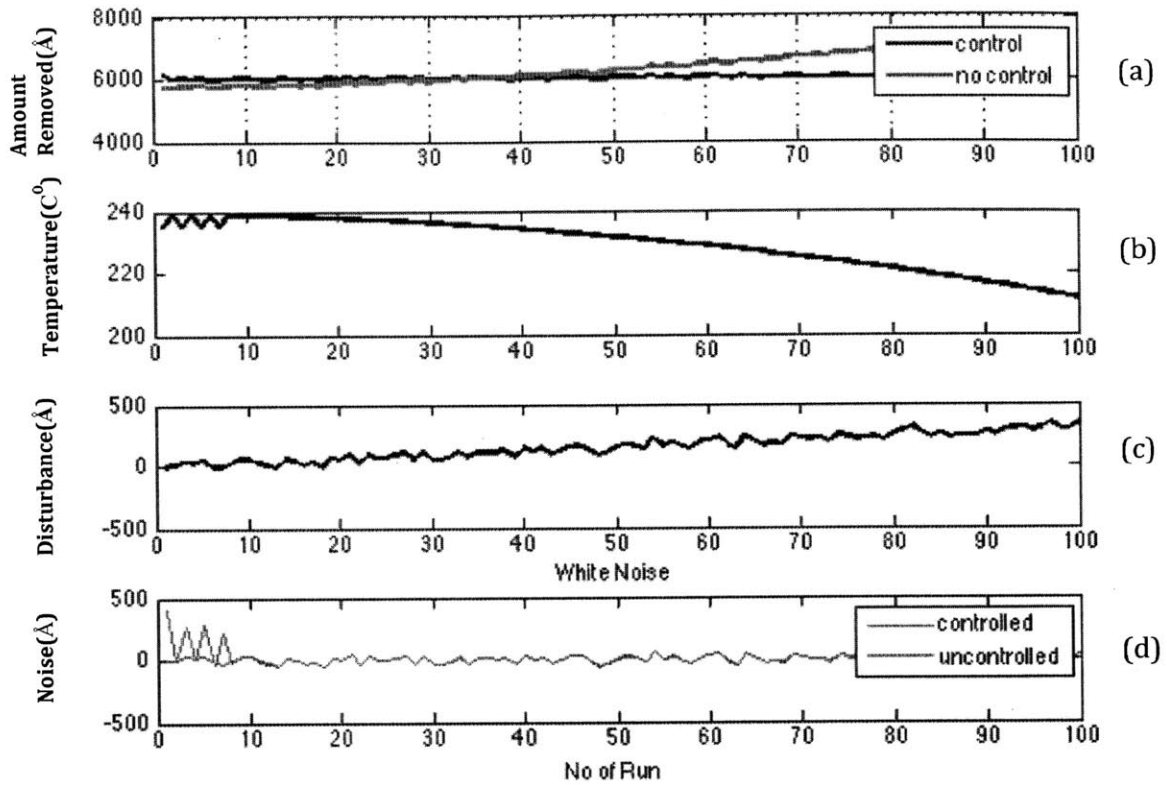


Figure 40: Double EWMA Control is implemented on G63 with a drift of  $3 \text{ \AA} / \text{wafer}$  and normal noise with  $\mu = 0, \sigma = 30$ . 1(a) The process comes back to the target value due to the control. 1(b) The temperature decreases to account for the linear upwards drift. 1(c) Linear upward drift introduced to the process of up to  $300 \text{ \AA}$  or 5% of the target value. 1(d) The difference in output noise between the controlled and uncontrolled process increases. Controlling scheme introduces more noise as the drift accumulates.

In order to simulate a case where the machines are drifting in different directions, G63 was subjected to a drift of  $-3 \text{ \AA} / \text{wafer run}$  instead of  $+3 \text{ \AA} / \text{wafer run}$ . The rest of the equations were kept the same. Figure 41 shows how machine matching can be achieved even with drift in different directions. Rather than having  $600 \text{ \AA}$  resultant strip thickness different as a result of the drift, the machines have  $\sim 50 \text{ \AA}$  mismatch in thickness.

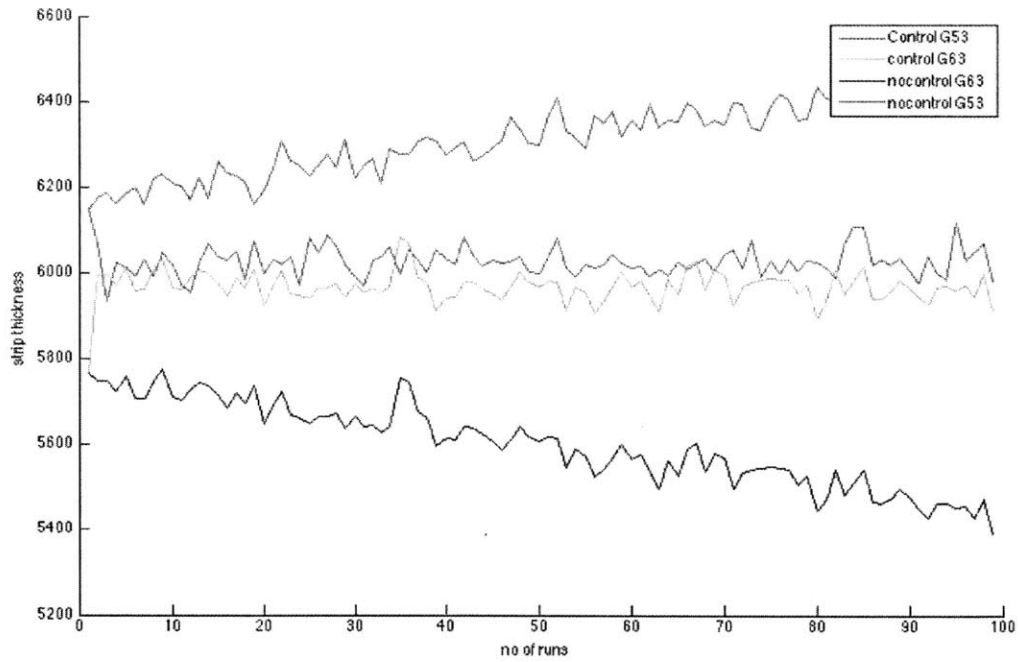


Figure 41: Resultant strip thickness when G53 experiences an upward drift and G63 downward drift with and without control being implemented.



## **5. Results and Benefits**

This chapter focuses on the results and benefits of the chamber matching methodology using statistical analysis and run-to-run control. Even though chamber matching was not implemented physically on the tools, this work can be used as a roadmap to chamber matching in the future.

### **5.1 Results and Discussion**

The methodology was developed taking into account ADI's capabilities and its progressive outlook for advanced techniques to improve their world-class manufacturing. Four statistical tools were identified to detect different states of machines and categorize them in terms of golden and inferior machines. These statistical tools are not limited to the Gasonics machines; they can be used in any process. A Design of Experiments was performed on two machines and individual regression models were obtained. These models showed that temperature is the most important factor affecting the target strip thickness for these machines. The models also revealed that the temperature factor is more effective in G53 than in G63. This finding led us to use two diagnosis techniques to identify the causes of the difference: instrumented wafers and spatial surface mapping. ADI was already familiar with the instrumented wafers but 49-point measurements were introduced along with a MATLAB utility for surface mapping as shown in Appendix 2.

Working alongside with the IoT team, we were able to analyze some of the first data streaming from the fab. The data was compared for different machines and conclusions of malfunction alarms and possible batch related problems were identified. Types of problems to watch out for were pointed out working with the team. Lastly, run-to-run control simulations were applied to the machine models and shown to decrease the variation and handle drifts as seen in Gasonics machines. It was shown to the ADI team that the R2R control will be instrumental as the dimensions shrink and wafer sizes increase. The process engineering team recognized the importance of the run-to-run control and has decided to start conversations with the manufacturers about the implementation of the controller.

### **5.2. Benefits**

Chamber matching is an important topic in the industry today. The chamber matching analyses and discussions helped ADI to have a quantitative understanding of each machines'

characteristics and have solidified the steps that need to be taken in the future for improvement. Our simulation results show that if implemented, chamber matching will decrease variability.

### **5.2.1. Decreased Yield Variability**

A batch that was processed in one chamber might exhibit a higher yield loss than a batch processed in another chamber. There are hundreds of steps involved in fabrication of ADI products. Even though plasma-ashing is not the most sensitive process, it is a model for improvements in more sensitive processes and machines. The methodology and the application we proposed in this thesis can be rolled out to other processes that can be vital in process performance. Even though end-line yield data was not available and it is hard to quantify the yield improvement directly, it can be inferred that decreasing the variations in the controlled variables will decrease the parametric yield loss that depends on the process variation.

### **5.2.2. Greater Routing Flexibility**

One of the findings of our research is that different machines have different levels of wafer outputs historically even though they are running the same process, as seen in Figure 16. The operators know about the behavior of the machines because they work closely with them, and operators often use the machines that give better outputs and require them to spend less time with out of spec measurements. The engineers were also aware of the situation. Although every machine might be available at any time, they have a predilection to use one of the machines rather than the other, which at times creates queues, and wait time. Our work underlined the differences between the machines and their operating points; these were somewhat known to the teams but had never been fully investigated. We turned the discussions of “that machine is a problem” to the more specific discussion of frequent mean shifts, high variability or historically lower means. This initial step helped us discuss with the ADI team the differences between the machines. The proposed chamber matching methodology and simulation showed that both G53 and G63 can be made to run close to the target values; in the case of drift, rather than having a significant difference between the output dimensions, the difference between the tools will be kept in a much narrower window. This promises to bring the machine variability and performance closer to the target values and will build confidence by the operators and the team on having seven healthy machines rather than a “problematic machine” or a “golden machine.”

### **5.2.3. Faster Tool Qualification**

A new recipe on Gasonic machines can take several months to develop. Electric tests will have to be run on split lots for each of the major process flows which will make use of a new recipe. If the recipe is for a limited set of products, it might take one to two months. For a recipe that will be used by many products, it might take four to six months. Every machine needs to go through the tests. When the machines are matched across hardware and software platforms, the processes and tool qualifications can be rolled out much more quickly on all of the machines. This can save cost and time of running several wafers to qualify each process and tool.

### **5.2.4. Reduced Time for Root Cause Analysis**

By running design of experiments on the tools, the inherent differences in the machines were surfaced along with the reasons behind these differences. Using instrumented wafers, it was shown that the inferior machine has temperature control issues. Talking to the maintenance team, it was found that the CLTC probes were from different companies and different models. Closer hardware matching was suggested to ADI for the future. Standardization across the machines is important for the maintenance efforts to be pooled and for building intuition about the probe settings.

## **6. Conclusion, Future Work and Recommendations**

An application of chamber matching methodology for semiconductor manufacturing using statistical analysis and run-to-run control was demonstrated in this work. This methodology starts with hardware and software matching but needs to be supported with in-situ sensors and feedback control, creating chambers that are capable of making decisions about the adjustments on the recipes to reach the target thickness. Chamber matching is more relevant to Analog Devices than ever with its growth and new fabrication sites. Ensuring the same quality and value on all of its fabrication sites is essential.

In this work, we concentrated on the target strip thickness as a performance metric; however, other metrics like non-uniformity can be significant in other processes. As future work, to include nonuniformity, the controller needs to have another regression equation that it needs to solve for, and a multi input multi output (MIMO) approach should be taken in that case.

This work did not include a recipe optimization step for chamber matching, which is one of the major steps for future work. We used the status quo recipe because the process engineering teams fine-tune the recipes and implementing change is difficult on the recipes. Once the chamber matching efforts are put into place, recipe optimization should be concurrently undertaken.

The work shown here on only two machines should be extended to all seven machines. DOEs will need to be run on each machine. The machine malfunctions that cause non-normal data should be resolved before the machines are being tested.

Finally, the full implementation of a chamber-matching project will involve the OEMs, IoT team and the process engineers. Analog Devices is recommended to create cross-functional teams to attack differences across fabrication sites.

## References

- [1] A. Husain, "Analog Devices Reports Fourth Quarter and Fiscal Year 2015 Results," Analog Devices, Inc., Nov. 24, 2015. [Online]. Available: <http://investor.analog.com/releasedetail.cfm?releaseid=944231>. Accessed: Aug. 12, 2016
- [2] T. Nerurkar, 2016, "Design of Experiments on a Semiconductor Plasma Ashing Process: Methodology and Analysis," M.Eng. Thesis, Massachusetts Institute of Technology, Cambridge, MA.
- [3] T. Nilgianskul, 2016, "Control of a Semiconductor Dry Etch Process using Variation and Correlation Analyses," M.Eng. Thesis, Massachusetts Institute of Technology, Cambridge, MA.
- [4] "Products | Lam Research", *Lamresearch.com*, 2016. [Online]. Available: <http://www.lamresearch.com/products/products-overview>. [Accessed: 04- Aug- 2016].
- [5] G. S. May and C. J. Spanos, *Fundamentals of Semiconductor Manufacturing and Process Control*. United Kingdom: Wiley-Blackwell (an imprint of John Wiley & Sons Ltd), 2006.
- [6] *Aura 3010 Operations Guide*, 1st ed. USA: GaSonic International Corporation, 1995, pp. 14-24, 31-58.
- [7] Justin BD&L, Inc., "Gasonics A3010 Asher", SEMI, Seoul, 2004.
- [8] W. F. Davis *et al.*, "Statistical IC simulation based on independent wafer extracted process and experimental designs," *Proc. Bipolar Circuits and Technol. Meeting, 1989*, pp. 262-265, 1989.
- [9] "Giovanni-3 Operation Technical Summary: Time Series Statistics," in *NASA Goddard Earth Sciences Data and Information Services Center*, 2009. [Online]. Available: [http://disc.sci.gsfc.nasa.gov/giovanni/documents/g3stats\\_with\\_proof](http://disc.sci.gsfc.nasa.gov/giovanni/documents/g3stats_with_proof). Accessed: Aug. 16, 2016.
- [10] D. C. Montgomery, *Introduction to Statistical Quality Control*, 6<sup>th</sup> ed. United Kingdom: Wiley, John & Sons, 2008.

- [11] W. E. Johnson, *Logic, Part III: The Logical Foundations of Science*. New York, NY: Cambridge University Press, 1924.
- [12] R. E. Barlow and T. Z. Irony, "Foundations of statistical quality control," *Current Issues in Statistical Inference: Essays in Honor of D. Basu*, pp. 99-112, 1992.
- [13] D. E. Hardt, 2015, "Control of Manufacturing Processes –SPC Methods: Control Charts," 2.830 class notes, Massachusetts Institute of Technology, Cambridge, MA.
- [14] Western Electric Company, Inc., 1958, *Statistical Quality Control Handbook*, 2<sup>nd</sup> ed., Easton: Mack Printing Co., 328p.
- [15] S. M. Stigler, *The history of statistics: The measurement of uncertainty before 1900*, 8th ed. Cambridge, MA: Belknap Press of Harvard University Press, 1986, p. 134.
- [16] D. Drain, *Statistical Methods for Industrial Process Control*. Boca Raton, FL: Chapman & Hall, 1997.
- [17] G. W. Snedecor and W. G. Cochran, *Statistical Methods*, 8<sup>th</sup> ed., Iowa State University Press, 1989.
- [18] K. H. Baek, et al., "Multiple input Multiple output controller design to match chamber performance in plasma etching for semiconductor manufacturing," in *Journal of Vacuum Science&Technology B*, 2013, p. 31.
- [19] R. G. Cosway *et al.*, "A Novel Technique for Epitaxy Tool-to-Tool and Chamber Matching and Optimization," in *27th Annual SEMI Advanced Semiconductor Manufacturing Conf. (ASMC)*, AZ, 2016, pp. 429-434.
- [20] M. Yedatore et al., "Improving Yield with Fleet Chamber Matching," [Online], Available:<http://www.appliedmaterials.com/files/nanochip-journals/nanochip-fab-solutions-december-2013-revised.pdf>. Accessed: June 13, 2016.

- [21] T. Pan, S. Jang and D.S.H. Wong, "Chamber Matching of Semiconductor Manufacturing Process Using Statistical Analysis," in *IEEE Transactions on Systems Man and Cybernetics C (Applications and Reviews)*, July 2012, vol. 42 no 5, pp. 571-6.
- [22] G. S. May, C. J. Spanos, "Automated Malfunction Diagnosis of a Plasma Etcher," in *IEEE/Semi Int'l Semiconductor Manufacturing Science Symp.*, CA, 1991, pp. 62-8.
- [23] J. Musacchio et al., "On the utility of run to run control in semiconductor manufacturing." in *IEEE International Symp. on Semiconductor Conf. Proc.*, San Francisco, CA., 1997, pp. D9-12.
- [24] E. Del Castillo and R. Ragojapar, "A multivariate double EWMA process adjustment scheme for drifting processes," in *IIE Transactions*, 2002, no.34, pp. 1055-1068
- [25] S.W. Butler and Stefani, J.A., "Supervisory run-to-run control of a polysilicon gate etch using in situ ellipsometry," in *IEEE Transactions on Semiconductor Manufacturing*, May 1994, vol. 7, no. 2, pp. 193-201.
- [26] C.A. Bode,. (2001). Run-to-Run Control of Overlay and Linewidth in Semiconductor Manufacturing. PhD thesis. The University of Texas at Austin.
- [27] W.J. Conover, *Practical Nonparametric Statistics*, 3<sup>rd</sup> ed., Wiley, NY,1999.
- [28] R. Williams, "Interaction Effects and Group Comparisons," 2015. [Online]. Available: <https://www3.nd.edu/~rwilliam/stats2/l51.pdf> Accessed: June 11, 2016.
- [29] M. P. Allen, *Understanding Regression Analysis*, New York: Plenum Press, 1997.
- [30] M. Tesauro, G. Roche, "Instrumented wafers enable etch chamber matching." in *Solid State Technology*, June 2008, vol. 51 issue 6, p. 25.
- [31] D. Boning et al., "Practical Issues in Run by Run Process Control," in *1995 IEEE/SEMI Advanced Semiconductor Manufacturing Conference*, 1995, pp. 201 - 208, DOI: 10.1109/ASMC.1995.484371.

## Appendix A

### Aligning Signals and Matching Sample Rate

```
[G33_A,G43_A] = alignsignals(G33_part,G43_part);  
plot((0:140),G23_part,'r','Linewidth',2);  
hold on  
plot((0:(size(G33_A)-1))/(228/141),G33_A,'b','Linewidth',2);  
hold on  
plot((0:(size(G43_A)-1))/(227/141),G43_A,'k','Linewidth',2);  
grid on
```



## APPENDIX B

### Surface Map Plotting

```
x2=A(:,1);
y2=A(:,2);
z2=A(:,3);
[qx2,qy2] = meshgrid(linspace(min(x2),max(x2)),linspace(min(y2),max(y2)));

F = TriScatteredInterp(x2,y2,z2);
qz2 = F(qx2,qy2);
mesh(qx2,qy2,qz2);
hold on
plot3(x2,y2,z2,'o');
view(2);
colorbar
```

## APPENDIX C

### Run-to-Run Control Algorithm

```
target=6000;
w=0.01;
intercept= repmat(5925.31,100,1);
b=1077.84;
bpressure=-5.82;
bpower=246.1;
bpump=-16.01;
bO2=-4.94;
bN2=-34;
temp= repmat(0,100,1);
pressure= repmat(0,100,1);
power= repmat(1,100,1);
pump= repmat(0,100,1);
O2= repmat(0,100,1);
N2= repmat(0,100,1);
offset=zeros(100,1);
w2=0.01;
for i=1:(length(disturbance)-1)
    effects=bpower*power(i)+bpump * pump(i)+bO2*O2(i)+bN2*N2(i)+bpressure* pressure(i);
    thickness_real(i,1)=intercept(i)+b*temp(i)+effects+disturbance(i);
    thickness_npcontrol(i)=5925.31+b*0+effects+disturbance(i);
    if abs(target-thickness_real(i))>0.001
        intercept(i+1)= w *(thickness_real(i)-b*temp(i)-effects)+(1-w)*intercept(i);
        offset(i+1)= w2*(thickness_real(i)-b*temp(i)-effects-intercept(i))+(1-w2)* offset(i);
        temp(i+1)=(target-intercept(i+1)-offset(i+1)-effects)/b;
    end
end
end
```