Control of a Semiconductor Dry Etch Process using Variation and Correlation Analyses

by

Tan Nilgianskul

B.S. Materials Science and Engineering
Cornell University (2015)

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

Master of Engineering in Manufacturing

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
September 2016

© 2016 Massachusetts Institute of Technology. All rights reserved.
This page is intentionally left blank.
Control of a Semiconductor Dry Etch Process using Variation and Correlation Analyses

by

Tan Nilgianskul
B.S. Materials Science and Engineering
Cornell University (2015)

Submitted to the Department of Mechanical Engineering on August 19, 2016 in partial fulfillment of the requirements for the degree of Master of Engineering in Manufacturing

Abstract

Statistical process control (SPC) is one of the traditional quality control methods that, if correctly applied, can be effective to improve and maintain quality and yield in any manufacturing facility. The purpose of this project is to demonstrate how to effectively apply SPC to a dry etch process (in this case plasma ashing), at Analog Devices, Inc., a company that runs large-scale fabrication sites in the Boston area. This thesis focuses on spatial and run-to-run variation across multiple measurement sites on a wafer and validates the assumptions of normality and correlation between sites within a wafer in order to justify and confirm the value of employing SPC theories to the plasma ashing process. By plotting control charts on past data, outlier data points are detected using Analog’s current monitoring system. Further, irregularities in the process that would not have been detected using traditional x-bar Shewhart charts are detected by monitoring non-uniformity. Finally, cost analysis suggests that implementing SPC would be a modest investment relative to the potential savings.

Thesis Supervisor: Duane Boning, Professor of Electrical Engineering and Computer Science
Acknowledgements

I would like to thank Prof. Duane Boning and Prof. David Hardt for advising us on process control theories throughout this project. I would like to thank Ken Flanders and Jack Dillon, heads of the process engineering and manufacturing operations teams, for providing close technical guidance and for their supervision throughout our time working at Analog Devices, Inc. Furthermore, all the experimental work done leading up to this thesis would not have been possible if not for other process engineers in the etch group including Pamela Petzold, Peter Cardillo, Rich DeJordy, Dale Shields along with many others who were of great support in performing experiments the cleanroom. I would also like to thank Dr. Brian Anthony for initiating this collaboration between MIT and Analog Devices and Jose Pacheco for coordinating with all parties involved throughout the M.Eng. program.

Finally, I would like to especially thank my project teammates, Tanay Nerurkar and Feyza for their contribution to the work presented in this thesis as well as for their incredible support throughout the entire course of this program.
# Table of Contents

Abstract ................................................................................................................................. 3

Acknowledgements .................................................................................................................. 4

List of Figures ............................................................................................................................ 7

List of Tables ............................................................................................................................. 9

1. Introduction ......................................................................................................................... 10

1.1 Background Information on Analog Devices, Inc............................................................. 10

1.2 General Semiconductor Fabrication Process ...................................................................... 11

1.3 Plasma Ashing Process ....................................................................................................... 12

1.3.1 Gasonics Aura 3010 Plasma Asher .............................................................................. 13

1.3.2 Partial and Forming Recipes ...................................................................................... 15

1.3.3 Data Collection and Logging .................................................................................... 17

1.3.4 Calculation of Basic Statistics .................................................................................... 19

2. Theoretical Review of Key Concepts .................................................................................. 23

2.1 Statistical Process Control (SPC) .................................................................................... 23

2.1.1 Origin of SPC .......................................................................................................... 23

2.1.2 Shewhart Control Charts ......................................................................................... 23

2.2 Analysis of Variance ......................................................................................................... 25

2.3 Design of Experiments ...................................................................................................... 26

2.4 Hypothesis-Testing .......................................................................................................... 29

2.4.1 Z-Test for Detecting Mean-shifts .............................................................................. 29

2.4.2 F-Test ...................................................................................................................... 31

2.4.3 Bartlett’s Test .......................................................................................................... 32

3. Statistical Process Control: Methodology .......................................................................... 34

3.1 Source of Data .................................................................................................................. 34

3.2 Test for Normality ............................................................................................................ 34

3.3 Fundamental Analysis Procedure .................................................................................... 36

4. Control Charts: Analysis .................................................................................................... 38
4.1 Raw Shewhart Charts .................................................................38

4.1.1 Analog’s Original Run Charts ...............................................38

4.1.2 Three Sigma Control Charts ...............................................39

4.1.3 Principle Component Analysis .............................................41

4.2 Separation of Variation Components .......................................44

4.3 Additional Modified Parameters to Monitor .............................45

4.3.1 Weighted Average Thickness .............................................45

4.3.2 Non-Uniformity .................................................................45

5. Benefits for Analog Devices, Inc. .............................................50

5.1 Choosing “Control Groups” .....................................................50

5.1.1 Monitoring Weighted Average Thickness .............................51

5.1.2 Monitoring Non-Uniformity ...............................................54

5.2 Recap of Possible Paths to Take ............................................58

5.3 Cost-Benefit and Tradeoffs ....................................................61

5.3.1 Cost-Benefit and Tradeoffs ...............................................61

5.3.2 Comparing the Costs of Three Options ...............................62

5.3.3 Recommendations .............................................................63

6. Conclusions and Future Work ................................................65

6.1 Conclusions ...........................................................................65

6.2 Suggestions for Further Studies ............................................66

References ..................................................................................67

Appendix .....................................................................................70
List of Figures

Figure 1-1: ADI’s manufacturing facilities. .................................................................11
Figure 1-2: Semiconductor processing steps [4]........................................................12
Figure 1-3: Ashing process schematic [5]. .................................................................13
Figure 1-4: Gasonics Aura 3010 machine [7]............................................................14
Figure 1-5: Display-screen of Gasonics tool [6].........................................................15
Figure 1-6: Spatial distribution and coordinate positions of the nine sites..............18
Figure 1-7: Data logging from Nanospec 9200. .........................................................19
Figure 2-1: Example of a Shewhart control chart.....................................................24
Figure 3-1: Histogram of the nine-site average thickness from G53000.....................35
Figure 3-2: Normal probability plot of nine-site average from G53000. .................36
Figure 4-1: Original control chart on G53000. .........................................................39
Figure 4-2: Shewhart x-bar control chart with control limits calculated using the average nine-site standard deviation as $s$.................................................................41
Figure 4-3: Data from site 2 plotted against data from site 3...............................42
Figure 4-4: Weighted average thickness run chart. Grand mean and ±3σ calculated based on all 127 runs.............................................................46
Figure 4-5: WIWNU run chart..............................................................................48
Figure 5-1: WAT control limits calculated from the lumped nine months of data ....52
Figure 5-2: WAT control limits calculated from samples 45 to 60.........................53
Figure 5-3: WAT control limits calculated from samples 64 to 89.........................54
Figure 5-4: WIWNU control limits calculated from the lumped nine months of data ....56
Figure 5-5: WIWNU control limits calculated from samples 40-70........................57
Figure 5-6: WIWNU control limits calculated from samples 98-111....................57
Figure 5-7: three methods in setting control limits.........................................................59

Figure A1: Control limits for WAT calculated from the lumped nine months of data under the forming recipe.................................................................70

Figure A2: Control limits for WIWNU calculated from the entire nine months of data under the forming recipe.................................................................71

Figure A3: Control limits for WIWNU calculated from the samples 50 through 80 under the forming recipe.................................................................72
List of Tables

Table 1-1: Machine parameters for the Partial recipe .................................................................16
Table 1-2: Machine parameters for the Forming recipe ...............................................................17
Table 2-1: 23 full factorial experimental design .............................................................................28
Table 2-2: 23-1 factorial experimental design .................................................................................28
Table 4-1: Variance distribution across nine principle components .............................................43
Table 4-2: Process variance broken down into the within-wafer or “error” component and the between-wafer or “group” component .................................................................44
Table 4-3: ANOVA table to test the significance of spatial variation ...........................................45
Table 5-1: Parameters and results for the $F$-test done on WAT of samples 45-60 vs. samples 64-89 ..................................................................................................................................52
Table 5-2: Parameters and results for the $t$-test done on WAT of samples 45-60 vs. samples 64-89 ..................................................................................................................................52
Table 5-3: Parameters and results for the $F$-test done on the WIWNU of samples 40-70 vs. samples 98-111 .......................................................................................................................55
Table 5-4: Parameters and results for the $Z$-test done on the WIWNU of samples 40-70 vs. samples 98-111 .......................................................................................................................56
Table 5-5: Summary of maintenance costs .......................................................................................62
Table 5-6: Costs and savings required to break-even under each option ......................................63
Chapter 1: Introduction

The work in this thesis presents a methodology to systematically perform statistical process control on a high volume semiconductor dry etch process. It lays out the necessary tools required to build statistical regression models and demonstrates potential implications and analyses that could be used to better understand, monitor and ultimately improve the performance of the machines involved. This is an industrial thesis, and the work was done in collaboration with Analog Devices Inc. (will also be referred to as ADI or Analog) at their fabrication facility in Wilmington, MA. 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. The company 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. The rest of this chapter will provide background information on Analog Devices Inc. as well as an introduction to the semiconductor dry etch process that was studied in this work. Finally, the chapter ends by explicitly stating the problem statement will be addressed later on throughout this thesis.

1.1 Background Information on Analog Devices Inc.

Analog Devices Inc. is an American multinational company that specializes in the design, manufacturing, 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 that would be then used in a wide array of electronic devices.

The company was founded in 1965 by Ray Stata and Matthew Lorber. It has operations in 23 countries and serves over 100,000 customers ranging from those in the consumer electronics and automotive industry to the defense industry. Analog’s revenue in the fiscal year 2015 was approximately $3.4 billion [1].
The manufacturing and assembly of ADI’s products is conducted in several locations worldwide. Figure 1-1 shows an overview of the location and functions of the company’s manufacturing and assembly facilities.

Figure 1-1: ADI’s manufacturing facilities.

This 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 Feyza Haskaraman 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

Pre-doped wafers are supplied to the Wilmington fabrication site as the starting material. The Wilmington 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 desired parts of the wafer. This allows the diffusion team to selectively implant impurity ions, the etch group to remove materials, or the thin-films group to deposit metals onto the designated parts of the silicon wafer. Afterwards, the etch group strips the resist off from these wafers. The
function of the CMP group is then to use chemical-mechanical reaction techniques to smoothen the surface of the deposited materials.

![Wafer Fabrication Process Steps](image)

**Figure 1-2:** Semiconductor processing steps [4].

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. In fact, the flexibility of this process setup enables ADI to produce customizable electronic parts on a customer’s short-term order.

### 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. The reactive species is generated by exposing a gas such as oxygen or fluorine to high power radio or microwaves, which ionizes the gas to form monoatomic reactive species. Figure 1-3 shows a general schematic of the plasma ashing process with key components indicated.
ADI uses the Gasonics A3010 tool to carry out the plasma ashing process. The reactive gas used by the company is oxygen. Microwaves are used to ionize the gas. The Gasonics A3010 tool allows for changes to be made to 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 chemically reacts with the surface of the wafer. The Aura 3010 system is composed of three main components [6]:

i. The reactor chamber which contains the system controller, the electro-luminescent display, the wafer handling robot, the microwave generator, and the gas box.
ii. The power enclosure wall box.

iii. The vacuum pump.

Figure 1-4 shows a picture of the Gasonics Aura 3010 machine.

![Image](image.png)

**Figure 1-4: Gasonics Aura 3010 machine [7].**

The machine is equipped with a wafer-handling robot that picks up a single wafer from a 25-wafer cassette and places it into the photoresist stripping process chamber. 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 [6]. Inside the process chamber, the wafer rests on three sapphire rods and a closed loop temperature control (CLTC) probe. CLCTC is a thermocouple that measures the temperature of the wafer during the ashing process. Twelve chamber cartridges embedded in the chamber wall then 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 generated by a magnetron ionize the gases into the monoatomic reactive species. The machine is designed in a way 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 [6]. 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 which 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 the G53000 machine.

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 1-5 shows a sample recipe on the display screen of the Gasonics Aura 3010 machine.

![Figure 1-5: Display-screen of Gasonics tool [6].](image)
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 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. They are named “Partial” and “Forming”. The details of these two recipes are as follows:

i. **Partial**: The Partial recipe is used for a qualification test to calibrate the rate of photoresist removal on a Gasonics Aura 3010 machine. The recipe is designed such that the photoresist mask is not completely removed from the wafer after the process. 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, the next eight seconds being allocated to the stripping process and the last 35 seconds being allocated to cooling the wafer. Table 1-1 shows the necessary machine parameters needed for the Partial recipe.

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

*Table 1-1: Machine parameters for the Partial recipe.*
ii. *Forming*: 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 recipe which is known as the “Implant” ash. The Implant 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 to the stripping process (Step-2) and the last 35 seconds to cooling the wafer. Table 1-2 shows the necessary machine parameters needed for the forming recipe.

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

Table 1-2: 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 the Gasonics Aura 3010 tool was set to function gives the photoresist removal rate. Analog uses this parameter to monitor
machine health. The tool used to measure the amount of photoresist 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 1-6 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.

<table>
<thead>
<tr>
<th>x-coordinate (mm)</th>
<th>y-coordinate (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>65</td>
</tr>
<tr>
<td>0</td>
<td>32.5</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>-32.5</td>
</tr>
<tr>
<td>0</td>
<td>-60</td>
</tr>
<tr>
<td>-65</td>
<td>0</td>
</tr>
<tr>
<td>-32.5</td>
<td>0</td>
</tr>
<tr>
<td>32.5</td>
<td>0</td>
</tr>
<tr>
<td>65</td>
<td>0</td>
</tr>
</tbody>
</table>

**Figure 1-6:** 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 to give the resist-removal rate, 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 conducted. An example of the spreadsheet can be seen in Figure 1-7. In Figure 1-7, the columns in the spreadsheet represent the measurements taken on the nine sites within a single wafer while the rows represent the different wafers measured. The Nanospec 9200 tool also logs the date and time of the measurement, which is very useful in detecting output anomalies.

<table>
<thead>
<tr>
<th>Site 1</th>
<th>Site 2</th>
<th>Site 3</th>
<th>Site 4</th>
<th>Site 5</th>
<th>Site 6</th>
<th>Site 7</th>
<th>Site 8</th>
<th>Site 9</th>
<th>ET_TIME</th>
</tr>
</thead>
<tbody>
<tr>
<td>5549.35</td>
<td>5599.8</td>
<td>5763.96</td>
<td>5824.96</td>
<td>5877.34</td>
<td>5501.67</td>
<td>5699.48</td>
<td>5799.73</td>
<td>5540.88</td>
<td>6/20/15 8:07</td>
</tr>
<tr>
<td>5410.93</td>
<td>5524.85</td>
<td>5631.59</td>
<td>5688.9</td>
<td>5707.29</td>
<td>5356.96</td>
<td>5555.87</td>
<td>5612.81</td>
<td>5428.14</td>
<td>6/23/15 8:59</td>
</tr>
<tr>
<td>5493.44</td>
<td>5571.64</td>
<td>5671.82</td>
<td>5760.63</td>
<td>5794.19</td>
<td>5456.9</td>
<td>5618.04</td>
<td>5647</td>
<td>5462.18</td>
<td>6/27/15 7:57</td>
</tr>
<tr>
<td>5436.46</td>
<td>5552.35</td>
<td>5689.99</td>
<td>5784.46</td>
<td>5802.86</td>
<td>5460.92</td>
<td>5645.25</td>
<td>5648.27</td>
<td>5389.47</td>
<td>7/6/15 9:01</td>
</tr>
<tr>
<td>5595</td>
<td>5682</td>
<td>5796</td>
<td>5878</td>
<td>6004</td>
<td>5450</td>
<td>5724</td>
<td>5892</td>
<td>5699</td>
<td>7/10/15 9:39</td>
</tr>
<tr>
<td>5587.58</td>
<td>5693.32</td>
<td>5829.78</td>
<td>5954.44</td>
<td>6019.11</td>
<td>5571.12</td>
<td>5772.03</td>
<td>5817.37</td>
<td>5567.42</td>
<td>7/13/15 20:12</td>
</tr>
</tbody>
</table>

**Figure 1-7:** Data logging from Nanospec 9200.

### 1.3.4 Calculation of Basic Statistics

The raw data collected from the Nanospec 9200 tool, as shown in Figure 1-7, needs to be processed 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 weighted average thickness of the nine sites on a single wafer ($\bar{x}^*$)

ii. The area-weighted standard deviation of the nine sites on a wafer ($s$)

iii. The within wafer non-uniformity parameter (WIWNU)

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 1-6. Davis *et al.* has shown that in a radial distribution pattern, the calculation of any statistics on the sites measured on a wafer should take into account the
wafer area represented by each site for accurate analysis [8]. Figure 1-8 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 1-8: Areal representation of each site on a wafer.

In Figure 1-8, 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 [9]:

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

(1.1)

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 weighted standard deviation taking into account the areal representation of each site can be derived as follows starting from the fundamental standard deviation formula:

\[ s_{\text{non-weighted}} = \sqrt{\frac{\sum_{i=1}^{N} (x_i - \bar{x})^2}{N - 1}} \]  

(1.2)

The unbiased estimate for the area-weighted standard deviation, replacing \( \bar{x} \) with the weighted mean \( \bar{x}^* \), can be written as:

\[ s = \sqrt{\frac{\sum_{i=1}^{N} w_i \left( x_i - \bar{x}^* \right)^2}{\left( \sum_{i=1}^{N} w_i \right)^2 - \sum_{i=1}^{N} w_i^2 \left( \sum_{i=1}^{N} w_i \bar{x}^* \right)^2}} \]  

(1.3)

In both equations, \( x_i \) is the wafer thickness measured at each site, \( w_i \) is the weight associated with that individual site, and \( \bar{x}^* \) is the weighted average mean. The derivation for Eq. (1.3) can be found in one of NASA’s Giovanni documents [9].

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

\[ WIWNU = \frac{s}{\bar{x}^*} \]  

(1.4)

where \( s \) is the weighted standard deviation and \( \bar{x}^* \) is the weighted average mean [8].

1.4 Problem Statement

This section outlines the motivation behind the work done alongside this thesis with Analog Devices, Inc. to improve yields and increase throughput of the plasma ashing process. As manufactured semiconductors are shrinking in various dimensions, more sophisticated control systems need to be implemented to achieve higher yields. Analog is also adapting to these changes by implementing Internet of Things (IoT) and
Advanced Control Systems. From observing Analog’s IoT pilot project, it is clear that 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. ADI stated that such differences that might result in yield losses appear later at the end-of-the-line after many process steps. Such losses can become more problematic in high-resolution manufacturing and for more expensive large-scale processes. Therefore, the work outlined in this thesis including an improvement plan for the use of statistical control, process modeling and eventually machine matching, would be a critical step towards the implementation of Advanced Control Systems such as Predictive Maintenance (PM).

1.5 Outline of Thesis

This chapter has just discussed the background information of Analog Devices as well as the dry etch process and data collecting conventions around which this thesis revolves. Chapter 2 outlines the fundamental SPC procedures and details, with equations and examples, the mathematical tools that are relevant to making recommendations to the company. Chapter 3 discusses the SPC methodology and how it will be applied to the nine-month resist-thickness data from the fabrication facility. Chapter 4 will then delve into these analyses in a quantitative manner. Control charts and correlation graphs will be presented accordingly. Chapter 5 will consolidate the results, presenting a solution on how to proceed and evaluate the potential benefits of adopting the solution. Finally, the last chapter (Chapter 6) will draw a conclusion from the statistical analyses done in prior chapters. Suggestions for further studies will also be discussed in the final section.
Chapter 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 May and Spanos and Montgomery 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 some indicator of the process performance over a measure of time [13]. These plots are then bounded by control limits which 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 2-1 [14].
Points marked with X’s are points that would be rejected based on Western Electric Rules (set of rules that indicate when process is out of control). 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 periods 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% [14]. Assuming that a set of data is normally distributed with mean $\mu$ and variance $\sigma^2$, the $UCL$ and $LCL$ can be expressed as:

$$UCL = \mu + 3\sigma$$
$$LCL = \mu - 3\sigma$$

(2.1)

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

$$P(X > UCL) = P\left(Z > \frac{UCL - \mu}{\sigma}\right) = P(Z > 3) \approx 0.0013$$
\[ P(X < UCL) = P \left( Z < \frac{\mu - LCL}{\sigma} \right) = P(Z < 3) \approx 0.0013 \]
\[ P(X > UCL | X < LCL) = P(X > UCL) + P(X < LCL) = 0.0013 + 0.0013 \]
\[ P(\text{point lies outside control limit}) \sim 3\% \]

(2.2)

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 variances 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 the nested analysis of variance. This is the analysis that is done when data can be broken down into groups, subgroups and 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].

\[ MS_w = \frac{SS_w}{W - 1} \]  
\[ MS_E = \frac{SS_E}{W(M - 1)} \]  
(2.3)  
(2.4)
where:

\[ SS_w = \text{squared sum of deviations of group means from grand mean} \]
\[ SS_E = \text{squared sum of deviations of each datapoint from its group mean} \]

Note that \( SS_w \) sums up the grand-group mean deviation for every individual point. So in this case, each squared difference between grand and group mean is multiplied by \( M \) before summing them together. The significance of the between-group variation could then be 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].

\[
\sigma_w^2 = \sigma_w^2 + \frac{\sigma_M^2}{M}
\]

(2.5)

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

\[
\sigma_w^2 = \sigma_w^2 - \frac{\sigma_M^2}{M}
\]

(2.6)

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 one to differentiate between measurements among and within silicon wafers.

2.3 Design of Experiments

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.

Factorial experiments allow for both individual factor and multiple-order interactions (effect of varying multiple factors simultaneously) to be evaluated from one set of experiments. 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]. For this work, experiments were done based on pre-designed half-factorial experiments and it would extort from the main objectives of this project to stress all the theoretical background that led up to the designs.

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. These interactions, effects of which are assumed negligible, usually include some second degree interactions and third degree (or higher) order interactions that are typically less significant than lower-degree interactions. For example, Table 2-1 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]
By defining the following identity relation and aliases:

\[ I = ABC \]
\[ A + BC \]
\[ B + AC \]
\[ C + AB, \]

a half factorial experimental design can be made. Table 2-2 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.
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 of a test 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-shifts

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 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].

\[
H_0: \mu = \mu_0 \\
H_1: \mu \neq \mu_0
\]

(2.7)

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

\[ p\text{-value} = P(\bar{x} > \mu_0) = P(z > Z_0) \]  

(2.9)

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 hypothesis 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:

\[ P(\text{Type II Error}) = P(\bar{x} < \bar{x}_{\text{critical}}) \]  

(2.10)

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

\[ \bar{x}_{\text{critical}} = \mu_0 + Z_{1-\alpha/2} \cdot \sigma \]  

(2.11)

Therefore, continuing from Eq. (2.11)

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

(2.12)

As previously mentioned, the significance level would depend on the tolerance for these two errors. For instance, if the detection of a mean shift triggers a very costly alarm, then a lower $\alpha$ would be desired in order to minimize $P(\text{Type I Error})$ or practically the probability of a false alarm. 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 indicates whether the ratio of the variances of two sets of data is statistically significant. Following the same method as in the previous $Z$-test example, the $F$-test begins with formulating hypotheses around the variances $(s_1^2$ and $s_2^2$) of two sets of data [13].

$$H_0: s_1^2 = s_2^2$$

$$H_1: s_1^2 \neq s_2^2$$

(2.13)

The test statistic $F_0$ in this case is simply the ratio of the variances 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.

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

(2.14)

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 Eq. (2.15).
\[ p\text{-value} = P(F > F_0) \]  

(2.15)

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 \( s_1 \) 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 are different, a two-tailed test would not be used.

2.4.3 Bartlett’s Test

Bartlett’s test is used to determine whether \( k \) sets of numbers of were sampled from distributions with equal variances. The null and alternative hypotheses can be formulated as follows:

\[
H_0: s_1^2 = s_2^2 = s_3^2 \ldots = s_k^2 \\
H_1: s_i^2 \neq s_j^2 \text{ for at least one pair } (i,j)
\]

(2.16)

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].

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

(2.17)

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

\[
N = \sum_{i=1}^{k} n_i
\]

(2.18)

\[
s_p^2 = \frac{1}{N - k} \sum_{i=1}^{k} (n_i - 1)s_i^2
\]

(2.19)
$T$ can be approximated by the chi-squared distribution. $H_0$ would therefore be rejected under a significance level $\alpha$ if $T > \chi^2_{k-1, \alpha}$. [17]
Chapter 3: Statistical Process Control Methodology

This chapter discusses the data analysis procedure used to support the statistical process control methodology. Based on the background SPC theories and literature reviews presented in Chapters 1 and 2, a plan was developed to first make sense of the data that comes out of the Gasonics tool and then to use that information to make valuable conclusions that will lead to improving the performance of the machine.

3.1 Source of Data

Data is obtained from the Gasonics A3010 machines at Analog’s facility in Wilmington. For simplicity and to isolate our investigation to as few external variables as possible, SPC analysis is limited only to the data from the Partial recipe from the G53000 machine. This was the etching process that was responsible for monitoring the removal rate by only partly removing coated resist from silicon wafers. The data extracted consists of the thickness of the photoresist layer before and after the etching process. The difference between the two numbers is the amount (in thickness) of photoresist removed. Dividing this thickness by the time spent etching is a measure of resist-removal rate (typically in Å/min). Each wafer undergoes this measurement at nine sites. The average of resist-removal rate on the nine sites is then used to plot the control charts.

The data analyzed in this thesis is extracted from the end of August 2015 all the way through to the middle of March 2016. The measurements were taken approximately once every two days. The general procedure begins by applying some of the simplest methods such as the traditional run-chart, which then inspires bringing in more complicated analytical techniques to make sense of the data which will be discussed in subsequent sections in this chapter.

3.2 Tests for Normality

Before any of the conventional SPC theories can be applied, the first step is always to test the data for normality. To evaluate data distribution for normality, either a histogram or a normal probability plot could be used. In a histogram, one observes how close the distribution of the data is to the “bell-curve” shape that would be seen if the data was perfectly normally distributed. In a normal probability plot, one looks for how
linearly the sorted data fits to the selected values. The concept of a normal probability is derived from what is called the quantile-quantile plot (or Q-Q plot) that graphs the quantiles of two distributions against each other [10].

A histogram of the data (for G53000) is shown in Figure 3-1. This is the average of the nine sites in each wafer. Histograms for the nine individual sites also exhibit similar bell-shaped trends. This supports the assumption that the output thickness could approximately be modeled by a normal distribution.

Figure 3-1: Histogram of the nine-site average thickness from G53000.

Figure 3-2 is a normal probability plot of the same data shown in the previous histogram in Figure 3-1. The bulk of the data on the plot lies very closely to the linear fit, hence also supporting the assumption of a normal approximation on the distribution of the output thickness. Note that it is expected for the data on both ends of the distribution to deviate farther from the fit than those nearer to the mean. This is why there is more deviation from the linear fit for the thicknesses less than 5700 Å and greater than 7000 Å as seen in the plot shown in Figure 3-2.
3.3 Fundamental Analysis Procedure

Comparing the control limits on Analog’s current run charts to the three sigma (±3σ) limits shows that the three sigma limits are slightly tighter than those on ADI’s current SPC run charts. Shewhart x-bar and s-charts were also plotted, assuming that the each of the nine-site measurements were independent replicates (i.e., plotted x-bar and s-charts with n = 9). From the control charts, a few outliers that were outside of the control limits could be seen. These were later found to be erroneous data and were removed from the analysis. These charts, however, appeared to produce overly tight control limits. From this observation, the team chose to respectively explore 1) the difference between the variation within wafer (spatial variation) and between wafers (temporal variation), and 2) the correlation between the nine-site measurements.

Figure 3-2: Normal probability plot of nine-site average from G53000.
The analysis of nested variance was initially used to separate the components of variation from one measurement to the next and, more specifically, to determine whether the variation between measurements or wafers was even significant at all. Results from this provided motivation to further look into the correlation between the nine sites. The concept of principal component analysis was used to relate the measurements on these nine sites. Throughout this process, it was noted that the correlation and variation between the 9 sites can largely be explained by a simplified factor, namely the non-uniformity element. Finally, the run charts were re-plotted using the parameters and factors deduced from those analyses. This would allow for those charts to be juxtaposed against the traditional SPC method. Each of these steps is presented in further detail in Chapter 4.
Chapter 4: Control Charts: Analysis

This chapter incorporates all of the data analyses done on the nine months of data in a coherent order. The purpose is to take readers step-by-step through the logic of how and why certain tests and analyses were done subsequent to the results summarized in Chapter 3.

4.1 Raw Shewhart Charts

One of Analog’s major issues is they are not able to detect as many of the problematic data points as they would like to and as early as they would hope to. And so maintenance would not occur at optimal times. The first milestone is therefore to refine these control charts in such a way that would allow ADI to effectively detect problems with the Gasonics tool in advance of when it would inflict further harm on its operations.

The Shewhart charts that were plotted assuming each of the nine sites to be independent replicates generated no out-of-control points as will be seen in Section 4.1.2 on the x-bar chart in Figure 4-2.

4.1.1 Analog’s Original Run Charts

Analog’s original control limits for the plasma ashing process were calculated collectively for all of the machines with the means of the resist removed on all machines pooled together. The population standard deviation of the data was estimated to be half of the range of those means. And it follows that the upper and lower control limits were three times that (estimated) standard deviation above and below the mean. With this method, the control limits on all of the machines would be identical. The original run chart with control limits for machine G53000 is shown in Figure 4-1. As seen, the current control limits allow for practically every out-of-control state to go undetected. Note also that the ADI’s original control charts plot the non-weighted average of the thicknesses removed on each wafer.
Analog Devices considers data points to be out of control points and would require machine shut down and component inspection in either of the following two cases:

i. Any data point that crosses the 3-sigma control limit

ii. Two or more consecutive data points crossing the 2-sigma control limit

4.1.2 Three Sigma Shewhart Control Charts

On the other hand, contrary to ADI’s original run charts, the control limits calculated using the traditional Shewhart x-bar chart seems to be too tight for the process and it would not make sense financially for ADI to shut down and maintain the machine as frequently as indicated by the x-bar chart. Note the x-bar chart also does not take into account the areal weighting of the within-wafer measurements. The control limits on the traditional x-bar chart were calculated as follows [13]:

i. The mean of the nine sites on each wafer and the within-wafer standard deviation of the nine sites were calculated:
\[
\mu_i = \frac{\sum_{j=1}^{9} x_j}{9}, \quad s_i = \sqrt{\frac{1}{8} \sum (x_j - \mu_i)^2}
\]

(4.1)

ii. The mean of the within-wafer standard deviations or s-bar was then calculated as follows:

\[
\bar{s} = \frac{\sum s_i}{m}
\]

(4.2)

where \( m \) is the total number of wafer runs.

iii. The 3-sigma control limits for the \( \bar{x} \)-chart can then be written out as:

\[
UCL = \bar{x} + \left( \frac{3\bar{s}}{c_4} \right)
\]

\[
LCL = \bar{x} - \left( \frac{3\bar{s}}{c_4} \right)
\]

(4.3)

where \( c_4 \) is a constant used to estimate the standard deviation of the process and \( \bar{x} \) is the grand mean (\( c_4 = 0.9693 \) for \( n = 9 \)). The Shewhart x-bar control chart is shown in Figure 4-2. As seen, almost a third of the points lie beyond the \( UCL \) and \( LCL \). On further investigation on the comments recorded by the operators of the machine, it was noted that many of these out-of-control points, including those that are circled in green in the graph shown in Figure 4-2, would have merely been false alarms as the machine seemed to be operating as expected and the wafers were well within the specification limits. After discussing with the manufacturing team, it was determined that the control limits generated using this method are now overly tight and that this method would not be feasible for ADI to implement.
Figure 4-2: Shewhart x-bar control chart with control limits calculated using the average nine-site standard deviation as $\bar{S}$.

To construct the x-bar chart, the average of the within-wafer standard deviation was used as $\bar{S}$ for calculating the control limits. This would usually be done when the measurements on the nine sites are uncorrelated and with the assumption that the measurements of all the sites on all wafers come from the same overarching distribution. One hypothesis as to why the control limits on the x-bar chart were not an accurate representation of the state of the process was that the resist removal rate on the nine sites within a single wafer were not independent; some (if not all) sites may have been strongly correlated with each other.

4.1.3 Principle Component Analysis

A preliminary test for this correlation hypothesis was to make a scatter plot of the data from any two sites and look for trends. The scatter plot of site 2 vs. site 3, as shown in Figure 4-3, clearly shows a strong correlation pattern and preliminarily validates the hypothesis that the two sites are indeed correlated. The linear fit has a slope of very close to one, and the data on the plot lies very closely to the linear fit. The R-squared value is also very close to one, indicating a high quality of fit. One might even be able to infer this
as an indication that sites 2 and 3 are not that much statistically different from each other, though that is not as relevant at this stage.

A more formal approach to test for redundancies in the nine-site data is to analyze the variation of the data set along its principle components. This method is known as principle component analysis. The principle components are a set of orthonormal axes that capture the maximum variation in each dimension of the data set. If the variation of the data set is equally distributed across all the principle components, then the entire data set is uncorrelated. If most of the variation is captured by only some of the principle components, then those principle components that capture the least variation are redundant and the number of dimensions or variables can be reduced.

**Figure 4-3:** Data from site 2 plotted against data from site 3.

To perform a principle component analysis on the nine-site data set, an $m \times n$ matrix $X$ was formulated where $m$ is the total number of wafers and $n$ is the number of measurements on each wafer. The covariance matrix ($C$) of $X$ was then calculated as follows [18]:

\[
y = 1.09x - 445.61
\]

\[
R^2 = 0.95
\]
\[ C = \frac{1}{n} XX' \]  

(4.4)

The principle components of the nine-site data set are the eigenvectors associated with the covariance matrix \( C \) and the variance across each principle component is the eigenvalue associated with that principle component/eigenvector. This data set will have nine principle components, as there are nine measurement sites on each wafer. These values were computed in MATLAB. The percentage of variation captured by each principle component is shown in Table 4-1 and can be calculated using the following formula:

\[
\text{%Variance} = \frac{\lambda_i}{\sum_{i=1}^{9} \lambda_i} \times 100
\]

(4.5)

where \( \lambda_i \) is the eigenvalue associated with the principle component \( i \) [18].

<table>
<thead>
<tr>
<th>%Variance Captured by Each Principle Component</th>
</tr>
</thead>
<tbody>
<tr>
<td>91.242</td>
</tr>
<tr>
<td>4.260</td>
</tr>
<tr>
<td>1.967</td>
</tr>
<tr>
<td>1.314</td>
</tr>
<tr>
<td>0.995</td>
</tr>
<tr>
<td>0.096</td>
</tr>
<tr>
<td>0.055</td>
</tr>
<tr>
<td>0.052</td>
</tr>
<tr>
<td>0.020</td>
</tr>
</tbody>
</table>

Table 4-1: Variance distribution across nine principle components.
4.2 Separation of Variation Components

It is apparent from the nature of the etching process data and the control charts that there has to be a differentiation between the spatial (within-wafer) variance and the temporal (between-wafer) variance. Therefore, ANOVA was performed, results of which are shown in Tables 4-2 and 4-3. The MS ratio (between-wafers to within-wafer or MS_{GROUP}/MS_{ERROR}) or $F$-value was found to be 6.54. This translates into a $p$-value on the order of $10^{-68}$ which is way below any reasonable significance level. Hence it can be concluded with virtually full confidence that the variation between each wafer run is significant to the process. Note that the reverse conclusion could not be drawn with this method without affirming that there is no spatial correlation within each wafer.

<table>
<thead>
<tr>
<th>Variation Source</th>
<th>MS</th>
<th># data in SS</th>
<th>Observed Variance</th>
<th>Estimated Variance</th>
<th>% Est Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERROR (site to site)</td>
<td>38253</td>
<td>1</td>
<td>38253</td>
<td>38253</td>
<td>61.89</td>
</tr>
<tr>
<td>GROUP (wafer to wafer)</td>
<td>250256</td>
<td>9</td>
<td>27806</td>
<td>23556</td>
<td>38.11</td>
</tr>
<tr>
<td>TOTAL</td>
<td>288509</td>
<td>1</td>
<td>288509</td>
<td>61809</td>
<td></td>
</tr>
</tbody>
</table>

Table 4-2: Process variance broken down into the within-wafer or “error” component and the between-wafer or “group” component.

The total variation of the etching process was further separated into their components. This was done by analyzing the nested variance in the data. The results of the calculation are shown in Table 4-2. SS and MS values are calculated using formulas presented in Eq. (2.3) through Eq. (2.6) in Chapter 2. The spatial variation accounted for 62% of the total variation while the temporal variation accounted for only 38% of the total variation. In other words, more than half of the variation associated with this process can be explained by the within-wafer variation. However, because of the observable trends in the temporal run charts and the fact that the ANOVA $F$-test indicated an exceedingly high significance level, there is reason to suspect that a large part of the spatial variation could be owed to strong correlations between the nine sites.
4.3 Additional Modified Parameters to Monitor

The monitoring strategy proposed by the classical Shewhart control strategy using the nine-site sample standard deviation would not be feasible for Analog’s plasma ashing operation. Control limits would be too narrow and would constantly trigger unwanted alarms. On the other hand, with the original control chart method, the system would still not have detected problematic data points during the past nine months when there definitely was a significant amount of improperly etched wafers. Analyzing the variation components and within-wafer correlation suggests that a wafer-to-wafer run chart might be most suited for Analog’s purpose.

The significance of the spatial variation indicates that it might also be beneficial to look into more parameters that would account for the within- versus between-wafer distribution of the output thickness. The next two sections will re-visit the concepts of weighted average thickness and within-wafer non-uniformity mentioned in the first two chapters, plotting them on control charts to see what other benefits could be realized.

4.3.1 Weighted Average Thickness

This section is partly a continuation from Section 1.3.4 of this thesis. The weighted average thickness (WAT) is calculated from Eq. (1.1) where the measurement of each site is roughly weighted according to the size of the area it represents. The control chart plotting the weighted average thickness (instead of non-weighted) is shown in Figure 4-4. The average nine-site standard deviation $\bar{s}$ (weighted or not) would not be relevant in this case so a plain run chart of the individual weighted average values were plotted. Note that from this point forward, all data plotted on control charts will account for areal weighting. The central line or mean $\mu_{WAT}$ and sigma $\sigma_{WAT}$ are the average and

<table>
<thead>
<tr>
<th></th>
<th>SS</th>
<th>DOF</th>
<th>MS</th>
<th>F-value</th>
<th>P-Value (Pr&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSG (Temporal)</td>
<td>31282053</td>
<td>125</td>
<td>250256</td>
<td>6.54</td>
<td>6.09×10^{-69}</td>
</tr>
<tr>
<td>SSE (Spatial)</td>
<td>38558931</td>
<td>1008</td>
<td>38253</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSD (Total)</td>
<td>69840984</td>
<td>1133</td>
<td>61643</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4-3: ANOVA table to test the significance of spatial variation.
standard deviation of all 127 runs of the WAT values. This is represented formulaically in Eq. (4.6) and (4.7) as follows.

\[
\mu_{WAT} = \frac{\sum_{i=1}^{127} WAT_i}{127}
\]

(4.6)

\[
\sigma_{WAT} = \sqrt{\frac{\sum_{i=1}^{127} (WAT_i - \mu)^2}{127 - 1}}
\]

(4.7)

where in both cases, \(WAT_i\) represents each of the individual run values the weighted average resist-thickness removed.

In this chart, the \(\pm 3\sigma\) control limits become wider than that in the Shewhart control chart of the non-weighted average thickness (Figure 4-2). The trends, however, are almost identical.

![Figure 4-4: Weighted average thickness run chart. Grand mean and \(\pm 3\sigma\) calculated based on all 127 runs.](image)
Even though there are no points that lie outside the upper or lower control limit in this chart, it is clear that the process is not in control. Especially if the Western Electric rules were to be applied, it could be concluded that many or most of the runs plotted are not in control. Starting at sample 12, there is a clear drift downwards. Performing a formal pairwise Z-test to test for a mean shift between samples 1-27 and 37-63 (assuming the population variance is equal to the variance of the entire nine months of data), it can be concluded that the mean has shifted with a 1% significance level, with a Z-value of 7.18 which translates into a p-value of essentially zero. Referring to the logged observation, the drift was due to an adjustment in the magnetron voltage that creates the microwave power. Shortly after, there was a leak found in the O-ring which may have worsened the tool’s condition.

Defining specific criteria to detect out-of-control output requires a more elaborated discussion with ADI’s manufacturing team. Slight tweaks can be made on the control chart to make many out-of-control data points more apparent, though these techniques will involve cost-benefit tradeoffs to consider. These different options will be discussed in more depth in Chapter 5. In the meantime, there is another parameter that cannot be neglected when dealing with semiconductor manufacturing processes. This is the within-wafer non-uniformity.

4.3.2 Non-uniformity

Similar to the last section on weighted average thicknesses, this chapter also continues from Section 1.3.4. The within-wafer non-uniformity (WIWNU) is defined as the weighted coefficient of variation of the sites within a wafer. Monitoring the spatial uniformity is crucial to maintaining lean and high quality semiconductor manufacturing and will come in especially useful when trying to optimize for the most effective process parameters [19]. Process optimization of this same process can be read more about in the corresponding thesis written by Nerurkar [2]. Note that because of the areal adjustment concept, the mean of the nine sites will be the weighted average as per what was explained in the previous section. The control chart of the WIWNU is plotted and shown in Figure 4-5. Again, the average nine-site standard deviation \( \sigma \) would not be relevant in this case so a plain run chart of the non-uniformities were plotted, now using the mean
\( \mu_{\text{WIWNU}} \) and standard deviation \( \sigma_{\text{WIWNU}} \) calculated using the 125 individual wafer-level run values to define the center line and \( \pm 3\sigma \) limits.

\[
\mu_{\text{WIWNU}} = \frac{\sum_{i=1}^{127} WAT_i}{127}
\]

(4.8)

\[
\sigma_{\text{WIWNU}} = \sqrt{\frac{\sum_{i=1}^{127} (WAT_i - \mu)^2}{127 - 1}}
\]

(4.9)

![Figure 4-5: WIWNU run chart.](image)

There is a much different trend in the WIWNU run chart than in the WAT run chart. Just by inspection, it can be seen that WIWNU is more in-control than the thickness removed over the past nine months. However, there are some out-of-control points that would be detected in the WIWNU chart that otherwise would not have been detected with the WAT run chart alone, one of which is sample 89 (circled in Figure 4-5) which singly lies out of the upper and lower control limits. This run signifies the leak across the door of the Gasonics machine which was a result of the O-ring breakdown (detected earlier by the thickness chart). Another example is the group of samples from sample 97 to sample 109, all of which consecutively lie on the lower side of the mean.
According to the respective logged comments, this was due to a power glitch, causing at least a temporary mean shift (there was no further data readily available to confirm whether or for how long this mean shift persisted). This problem is one that would not have been detected by the WAT run chart, even if all the Western Electric Rules were to be enforced. This underlines the importance of monitoring the WIWNU in addition to monitoring the weighted average removed resist thickness.
Chapter 5: Benefits for Analog Devices, Inc.

To evaluate the benefits of the statistical process control analyses done throughout this thesis, all the methods and options must be compared. Chapter 5 will focus on consolidating the results obtained from Chapter 4. In addition, a few different cases of the previously presented analyses will be re-introduced to exemplify the associated costs and benefits. Specific recommendations will be proposed together with the possible tradeoffs that should be considered between each of the choices. Finally, this will lead to the conclusion of how Analog should consider its options on how to proceed from this point forward, as well as what further investigations could be conducted to make more informed decisions.

5.1 Choosing “Control Groups”

To systematically implement the SPC methods that have been discussed in this thesis, the focus question is how one chooses the control group off of which to calculate the control limits. In other words, Analog needs to decide, based on its needs for maintenance, how “in control” it wants this etching process to be. For instance, looking back at the run chart in Figure 4-4, Analog can conclude that the entire process is sufficiently in control, assuming that all the observed fluctuation is just part of the natural variation of the entire system. This will be referred to as the “lumped data/trend” method. On the other hand, Analog can also be as strict as to take only a small group of points with the least variation, for instance samples 45 to 60, to be the control group, and determine that any points or group of points that lie out of the control limits calculated from that control group is out of control. This will be referred to as the “small-window control limits” method. In an ashing process, it is beneficial to detect control violations as early as possible to prevent compounded downstream effects [20]. Yet devoting too much attention to insignificant fluctuations can also result in unnecessary maintenance costs.

Cassidy showed that simulation models can be used to optimize control limits to achieve the desired optimal cost-saving decisions [21]. However, because precise information on yield and maintenance benefits is not readily available, a more simplified approach is taken here. A few extreme cases and methods were examined on the nine
months of data and the investment prospects of each were evaluated against each other to come up with a final most reasonable recommendation for ADI.

### 5.1.1 Monitoring Weighted Average Thickness

Basing the control limits off of the entire nine months of data, the control chart would look like what was shown in Figure 4-4 in the previous chapter. Three different scenarios of the WAT run chart are presented in Figures 5-1 to 5-3. Figure 5-1 shows the control chart when the control limits are calculated from the lumped nine months of data. Figure 5-2 shows the control chart when the control limits are calculated from a smaller window, in this case samples 45 through 60. Finally, Figure 5-3 shows the control chart when the control limits are calculated from samples 64 through 89, a slightly larger and more lenient control group. These last two sample groups were selected to display the extreme and moderate cases (control group of small and moderate variance).

A one-tailed $F$-test and a student $t$-test were done to test for unequal variances and mean respectively. The $F$-test results are shown in Table 5-1. The test statistic or $F$-value was calculated to be 3.65, and the $p$-value for the test was 0.0059. Hence it can be concluded with greater than 99% confidence level that the two samples come from distributions with different variances.

<table>
<thead>
<tr>
<th>$s_{45-60}$</th>
<th>67.89</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_{64-89}$</td>
<td>129.62</td>
</tr>
<tr>
<td>$n_{45-60}$</td>
<td>16</td>
</tr>
<tr>
<td>$n_{64-89}$</td>
<td>26</td>
</tr>
<tr>
<td>$F$ (test statistic)</td>
<td>3.65</td>
</tr>
<tr>
<td>$p$-value</td>
<td>0.0059</td>
</tr>
</tbody>
</table>

*Table 5-1:* Parameters and results for the $F$-test done on WAT of samples 45-60 vs. samples 64-89.

With the result from the preceding $F$-test, the $t$-test would have to be done assuming unequal variances. Results from the $t$-test are shown in Table 5-2. The test statistic $t$ was calculated to be 3.84, and the $p$-value was 0.00022. Based on the two tests, it can be concluded with better than 1% significance level (greater than 99% confidence) that both samples come from distributions of different variances and different means. Hence choosing which control group could be an important choice.
Table 5-2: Parameters and results for the $t$-test done on WAT of samples 45-60 vs. samples 64-89.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_{45-60}$</td>
<td>5528.76</td>
</tr>
<tr>
<td>$X_{64-89}$</td>
<td>5646.16</td>
</tr>
<tr>
<td>$n_{45-60}$</td>
<td>16</td>
</tr>
<tr>
<td>$n_{60-89}$</td>
<td>26</td>
</tr>
<tr>
<td>$\nu$</td>
<td>39.26</td>
</tr>
<tr>
<td>$t$ (test statistic)</td>
<td>3.84</td>
</tr>
<tr>
<td>$p$-value</td>
<td>0.00022</td>
</tr>
</tbody>
</table>

With the lumped data method (Figure 5-1), none of the points would lie outside of the upper or lower $\pm 3\sigma$ limit. If all of the Western Electric Rules were to be enforced, however, a large portion of the data would be considered out-of-control, as circled in red on the graph in Figure 5-1. It would be very difficult and costly for Analog Devices to conduct maintenance on the Gasonics tool every time there is a violation of one of the Western Electric Rules, not only because of the maintenance cost but also because it would unnecessarily slow down production. In the lumped control chart, more than 40 out of 127 points or roughly a third of the points were out of control. Note that in all the control charts from this point forward, starting with Figure 5-1, $\pm 1\sigma$ and $\pm 2\sigma$ bounds will be represented dotted lines, aside from the solid lines which are used to represent the centerline and the $\pm 3\sigma$ control limits.

![Figure 5-1: WAT control limits calculated from the lumped nine months of data.](image)
It was therefore agreed with the manager in manufacturing operations that not all of these rules would be enforced. Specifically, the rules that would be of interest were reduced to include only the first two rules:

- **Rule 1**: Any data points outside the ±3σ UCL and LCL
- **Rule 2**: Two out of three consecutive points outside ±2σ

With this, the flagged points would only include samples 11, 12 and 13. If samples 45 to 60 were selected as the control group as presented in Figure 5-2, 51 points alone would lie outside the ±3σ control limit. With Rule 2 enforced, an additional 16 points would have to be flagged, totaling to 67 out of 127 flagged points. This means that more than half of the samples tested will call for maintenance. Due to the scale of the plot, points that broke the second rule were not circled in Figure 5-2.

![Figure 5-2: WAT control limits calculated from samples 45 to 60.](image)

Alternatively, a more lenient control group with a greater sample variance could be chosen, like what is presented in Figure 5-3. With the control limits calculated based off of samples 64 through 89, nine out of 127 data points over the past nine months would have been flagged, five for breaking Rule 1 and four for breaking Rule 2.
Ultimately, there is an obvious tradeoff between enforcing tight control limits to rigorously control the process, and incurring maintenance costs and delays in production associated with getting the process back in control. This will be discussed in more detail in Section 5.2 of this chapter. The next section (Section 5.1.2) will be similar to this past section but will discuss non-uniformity rather than the weighted average resist thickness removed during plasma ashing.

5.1.2 Monitoring Non-Uniformity

The non-uniformity is overall more in-control than the WAT. This can be observed by inspection from the run charts in Chapter 4. Three control charts with different control groups of the same data are presented in Figures 5-4, 5-6 and 5-7. Figure 5-7 shows the lumped trend run chart of the non-uniformity with the control limits calculated based on all nine months of data. Control charts shown in Figures 5-6 and 5-7 are plotted with control limits calculated from samples 40-70 and 95-110, respectively. Again, the $F$-test and $t$-test were used to test the difference in variances and means of the two control groups. The results of the two tests are shown in Tables 5-3 and 5-4.
According to the results, the variances can be concluded as unequal at a 5% significance level (95% confidence) but not at a 1% significance level (99% confidence). The same goes with the difference in the mean of the distributions from which the two samples were drawn (assuming population variance is equal to that of the nine months of data). Note, however, that if the variances are assumed to be different, it would not be possible to conclude that there is a mean difference in the two samples under even a 10% significance level. Because the difference in the means and the variances are not hugely significant, it might be most worthwhile just to use the lumped trend control chart (Figure 5-4). The control charts plotted with the two other control groups, however, are presented in Figures 5-5 and 5-6 for comparison.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Mean</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>40-70</td>
<td>0.00279</td>
<td>0.00192</td>
</tr>
<tr>
<td>98-111</td>
<td>0.0043</td>
<td>0.00312</td>
</tr>
</tbody>
</table>

**Table 5-3:** Parameters and results for the F-test done on the WIWNU of samples 40-70 vs. samples 98-111.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Mean</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>40-70</td>
<td>0.0341</td>
<td>0.0312</td>
</tr>
<tr>
<td>98-111</td>
<td>0.0043</td>
<td>0.00312</td>
</tr>
</tbody>
</table>

**Table 5-4:** Parameters and results for the Z-test done on the WIWNU of samples 40-70 vs. samples 98-111.
Figure 5-4: WIWNU control limits calculated from the lumped nine months of data.

With the lumped data control chart, one point (sample 90) was found to be outside the ±3σ control limit, violating Rule 1. No other points would be detected even if Rule 2 were to be enforced.

Figure 5-5 graphs the control chart with the control limits calculated from samples 40 through samples 60. As seen, six out of the 127 points over the past nine months would have been flagged for violating Rule 1. Two other points would have been flagged for violating Rule 2. Alternatively, an even “stricter” control group or a control group with a smaller variance can be used. A control chart illustrating this option is displayed in Figure 5-6. Control limits on this run chart were calculated based off of samples 98 through 111. With this option, 11 out of the 127 data points in the past nine months would have been detected for being outside the upper and lower ±3σ control limits or violating Rule 1. Another additional nine points would be detected for violating Rule 2.

It should be noted that control charts for the Forming recipe can be found in the appendix. These will be relevant when evaluating the implementation costs in Section 5.3.
Figure 5-5: WIWNU control limits calculated from samples 40-70.

Figure 5-6: WIWNU control limits calculated from samples 98-111.
5.2 recap of possible paths to take

Up until this point of this chapter, everything has been more or less presented in the order of how the mathematical analysis was conducted. However, from this point on, the perspective of the thesis will shift more to ADI’s perspective as a business. The purpose of this section is to provide a synopsis and itemize the possible directions in which the manufacturing operations department could take with the plasma ashing process. From the statistical analyses performed thus far, it is highly recommended for Analog to at least take the following steps towards improving the manufacturing process.

I. **Remove Erroneous Points from System:**
   Just by plotting the WAT data on a simple run chart, it was found that there are many erroneous or outlier data points. When the data of each individual site was plotted, more erroneous data points were found. These invalid outliers can substantially distort the SPC monitoring process, hence they should be removed to avoid inaccurate analyses.

II. **Adjust for Tighter Control Limits:**
   It is clear that Analog’s current control limits which are pooled across all the Gasonics machines are too wide to detect when the process is out of control. The method for calculating control limits therefore should be changed.

III. **Enforce Monitoring Rules:**
   This refers to the Western Electric Rules, the most significant of which are listed in section 5.1.1 as Rule 1 and Rule 2. These two rules should be used to set minimum requirements for the process to be considered out of control. It might be beneficial to enforce more of the listed Western Electric Rules when there are changes in the maintenance costs and/or yield requirements.

IV. **Add Parameters To Monitor**
   New parameters that would be valuable to monitor are the weighted average thickness (WAT) and the weighted average within-wafer non-uniformity (WIWNU). Control charts could be plotted for these parameters, and control limits and Western Electric Rules could then be applied.
V. Schedule for Maintenance:

When data indicates that the process might be out of control, there has to be a systematic procedure to first decide whether there is enough reason to conduct maintenance and second call for maintenance.

Steps II and III mentioned above will involve a number of decisions. These decisions are detailed in the flow chart in Figure 5-7.

![Flow Chart]

**What is “out of control”?**

- Lump trends into bigger noise model
- Small window control limits
- Use all data
- Moving control limits every next 15-25 runs
- Reset when tolerable shifts occur

*Option 1  Option 2  Option 3*

**Figure 5-7:** three methods in setting control limits.

When deciding whether a data point is out of control, one would refer to the Western Electric Rules as mentioned previously throughout this thesis. After discussing with experts in Analog’s manufacturing department, it was concluded that it would not be beneficial to enforce all of the Western Electric Rules at this moment (this can later change). With the current performance, there would be more alarms than can be handled if Analog were to strictly enforce the all rules on the Gasonics machine. This is why, as mentioned earlier, the Western Electric Rules were reduced to the two key rules noted in Section 5.1.1.

The next decision to make is, based on Rule 1 (any single point outside the $±3\sigma$) and Rule 2 (two out of three consecutive points outside $±2\sigma$), how to calculate the control
limits (or which “control groups” to use when calculating control limits). The three main methods for going about this are outlined in the flowchart in Figure 5-7. Some of the advantages and disadvantages of each of the three methods are detailed as follows:

Option 1 – Lumped Trend/data
This data lumps all of the past data together. In other words, all of the variation that has occurred in the past is presumed as simply being the natural variation of the system. With this method, control limits would be wide (compared to the other two options) and the process would rarely be flagged, hence the process would most likely not, strictly speaking, be in a state of full statistical control. However, if a point is flagged under this option, it is a strong indication that maintenance should be called for.

Option 2 – Smaller-window Control Limits
With control limits calculated based on a time-window of as small as 25 “in control” samples, control limits will be very narrow. Data will be frequently flagged as an indication that the etch process is out of control. The advantage here is that this option would allow for the photoresist to be etched to very close to the desired thickness with a low probability of the output wafers having been over- or under-etched. However, Analog would have to suffer from increasing maintenance costs and potentially lose revenue because of slowed production.

Option 3 – Reset Control Limits After Tolerable Shifts Occur
This option is similar to the previous option, except for that the control limits might not have to be updated as frequently. Maintenance will not be called for immediately after a data point breaches the control limit. Instead, when a point is flagged, the process will still be allowed to run for another week or so, after whence Z-tests and F-tests are conducted to determine whether there is a drift in the process, a mean shift, or abnormal variation. If there is, the next step would be to determine whether the process output is tolerable. If the output is still tolerable even after there has been a statistical change in the process, then the control limits would be reset to reflect the updated process condition. If not, then maintenance would have to be called for. This method will closely monitor the
process and make sure no unnecessary maintenance is performed. However, ADI will need to develop a procedure to consistently determine whether maintenance is needed.

5.3 Cost-benefit and Tradeoffs

In deciding how to proceed, Analog must consider the tradeoffs between enforcing tighter control requirements and incurring maintenance costs. In order to make an informed analysis, a few estimates and assumptions have to be made.

5.3.1 Cost Background and Assumptions

To illustrate the decision-making process, relevant hypothetical estimates and assumptions related to costs and volumes are listed as follows:

- Analog outputs roughly 125,000 wafers a year
  - Out of these wafers, 100 are lost during the ashing process
  - Each scrapped wafer is valued at 1,000 USD
- Seven Gasonics tools are currently up and running
- Presently, Analog incurs about 180,000 USD on maintaining the G53000 machine in one year

With this information, the current loss owed to scrapped wafers can be deduced to be 100,000 USD per year. Since there are five Gasonics tools that are responsible for this loss, the loss per machine can be valued at about 20,000 USD per year. The goal of implementing SPC monitoring methods ultimately will be to save on these two areas: cost due to scrapped wafers and maintenance cost. To analyze the cost associated with the proposed methods, further illustrative assumptions and estimates will need to be made.

A day of maintenance could be considered to result in 500 USD in labor costs. Lost opportunity from not running production is a little harder to quantify, but from talking to the manufacturing team, it could be estimated that the company would incur an opportunity cost of 7,000 USD for every day that the ashing process is down. Note that because there are currently seven Gasonics machines up and running, it could be estimated that the company would incur 7,000/7 = 1,000 USD for every day that a single ashing tool (or specifically G53000 in this case) is down. A summary of the cost per
maintenance period is listed in Table 5-5. To estimate the total cost, it is assumed that maintenance takes one day (on average, maintenance actually takes less than once day but in order to remain conservative, an overestimate of the costs will be used in the further analyses.

<table>
<thead>
<tr>
<th>Maintenance Aspect</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labor</td>
<td>500 USD/day</td>
</tr>
<tr>
<td>Lost Production</td>
<td>1,000 USD/day/machine</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1,500 USD to bring back in control</strong></td>
</tr>
</tbody>
</table>

Table 5-5: Summary of maintenance costs.

The next subsection will now evaluate the different options by calculating how much additional cost (to the current cost) is associated with the implementation of each option.

5.3.2 Comparing the Costs of the Three Options

The cost associated with implementing the first option can be calculated using the approximations in the previous subsection. Recalling from Sections 5.1.1 and 5.1.2, the lumped data method flagged a total of 10 data points (five from the Partial recipe and another four from the Forming recipe) over the past nine months. Scaling this to a one-year period (by 12:9), it can be estimated that, on average, 12 data points will be flagged in a year. Now given that the average total cost to bring the process back in control is 1,500 USD (Table 5-5), it could be stated that Analog would incur a total maintenance cost of 18,000 USD in a year under option 1 or the lumped data method. This is less than 10% of the 200,000 USD that is lost every year on scrapped wafers and overall maintenance on this machine. In other words, the SPC implementation would have to save collectively only 10% of the current maintenance and scrapped wafers costs to make the implementation a worthwhile investment for Analog. Table 5-6 compares this break-even savings between the different options.
<table>
<thead>
<tr>
<th></th>
<th>Implementation cost</th>
<th>Current Cost</th>
<th>Percent of Cost that needs to be Saved</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Option 1</strong></td>
<td>$18,000</td>
<td>$200,000</td>
<td>9%</td>
</tr>
<tr>
<td><strong>Option 2 (high)</strong></td>
<td>$184,000</td>
<td>$200,000</td>
<td>92%</td>
</tr>
<tr>
<td><strong>Option 2 (low)</strong></td>
<td>$80,000</td>
<td>$200,000</td>
<td>40%</td>
</tr>
<tr>
<td><strong>Option 3</strong></td>
<td>N/A</td>
<td>$200,000</td>
<td>N/A</td>
</tr>
</tbody>
</table>

**Table 5-6:** Costs and savings required to break-even under each option. The low and high cases of **Option 2** refer to the extreme cases of how strictly the “control groups” are picked (refer back to Sections 5.1.1 and 5.1.2). The cost associated with **Option 3** is a little harder to quantify and will discussed later on.

Exploring option 2 more closely in Table 5-6, implementation costs can range from 80,000 to almost 184,000 USD, depending on how the control group is picked. As seen, choosing very strict control groups and enforcing very tight control limits would not be beneficial. It is highly unlikely for Analog to be able to cut down on 92% of its current maintenance and scrapped wafer costs (any savings over 100% is physically impossible to achieve). The low-end cost for this option still requires 40% savings. This means that going forward, the G53000 would have to reduce the scraps by almost half to make the implementation worthwhile.

The cost associated with option 3 cannot be accurately calculated at this point because the implementation is a little more sophisticated than the other two options. The bulk of the cost is expected to lie in the procedure to determine whether maintenance is needed every time a data point breaches the control limits. Whenever it is determined that mean shifts are tolerable and no maintenance is required, additional cost might be necessary to update control limits to what is appropriate following that mean shift.

**5.3.3 Recommendations**

Since it costs only an additional estimated amount of a little less than 20,000 USD to implement option 1, it is recommended for Analog to start with this lumped data option. Especially because the actual spec limits range from 4000 – 8000 Å, there does not seem to be any urgency to vastly tighten the control limits. Option 2 also seems to be an appealing alternative to maintaining a state of true statistical process control if the appropriate “control group” is chosen. It is important to note that there is no hard evidence to support whether scrapped wafers are caused by or strongly correlated with
out-of-control states. Hence with the current level of production, it would be safer to initially proceed with option 1.

However, it is also important to account for future growth and near-term expansions. As production grows, small scrap rates will rapidly translate into substantial losses. If this happens, it will be crucial for Analog to maintain rigorous control on even the more minor processes. When such growth could be predicted, Analog should seriously consider SPC monitoring methods like options 2 and 3. Option 3 is one that could be beneficial for Analog to develop in parallel with implementing the other options. If the manufacturing team can efficiently evaluate every statistical anomaly as it comes up in real time and update the control limits accordingly, this could be an attractive alternative. Yet this could require further studies as well as more tasks for the current team.
Chapter 6: Conclusions and Future Work

This chapter provides a conclusion of everything that has been discussed throughout the investigations and of this thesis, and recapturing the essential ideas of the statistical analyses conducted. The final section then provides guidelines on further research that could be done to build on the work of this thesis and improve Analog’s manufacturing operations.

6.1 Conclusions

In summary, it was demonstrated that implementing statistical process control charts on the weighted average removed thickness and within-wafer non-uniformity is a modest investment that would enable Analog to more quickly detect irregularities in its plasma ashing process, and therefore also more efficiently maintain its Gasonics tools. Since it was established that the process output can be approximated by the normal distribution, it was valid to use traditional SPC methods to analyze the nine months of data. Performing ANOVA on this data, it was found that 62% of the process variance was due to the within-wafer variation and 38% to run-to-run variation. PCA confirmed that a large part of the within-wafer variation is not entirely random but is owed to the correlation between each of the sites. This motivated the use of both the weighted average thickness and non-uniformity control charts. Closely examining the control charts and performing $F$-tests and mean-shift tests on selected regions, it was concluded that the plasma ashing process was not, strictly speaking, in a state of statistical control. However, from discussing with manufacturing experts in Analog’s engineering team, it would not be worthwhile to invest in bringing the process back into complete control.

It was ultimately concluded that the most feasible method is to use control limits calculated from the entire past nine months of data (option 1) and enforcing only the first two Western Electric Rules (any point out of $±3\sigma$ and two out of three consecutive points outside $±2\sigma$ on the same side of the mean). This would essentially be an approximately 18,000 USD yearly investment which is less than 10% of the combined cost of the current maintenance and losses on scrapped wafers. The goal of implementing this monitoring policy is not only to improve Analog’s wafer qualities but also to reduce the
maintenance period per year by detecting and dealing with errors in the tools before they generate considerably harmful effects.

6.2 Suggestions for Further Studies

One of the most substantial areas that would complement this thesis is the study of how states of in statistical control and out of statistical control actually correlate to the instantaneous yield or scrapped rate at that moment. The lack of information on this front prevents one from being able to quantitatively optimize control limits for maximum yield and minimum costs. This would not only open the door to evaluating more SPC methods but also enable the fine-tuning of specific control limits to achieve explicit economic goals.

Going forward, Analog can potentially make powerful use of live data gathered from their currently developing IoT system. With live or online-SPC data, the detection methods can be used in parallel with DOE and regression models of the tool’s input parameters (which are discussed more in depth in Nerurkar’s thesis [2]) to automatically and actively optimize the manufacturing process in real time. One could delve more deeply into machine learning and feedback control systems for more implementation options.
References


Appendix

A1. Control chart of weighted average thickness under forming recipe

Under the forming recipe, the thickness seems to be in a state of statistical control. No points are therefore flagged under either the lumped data method (option 1) or the smaller-window control limits method (option 2).

**Figure A1:** Control limits for WAT calculated from the lumped nine months of data under the forming recipe.
A2. Control chart of WIWNU under forming recipe using lumped trend

Under the forming recipe, the area-weighted WIWNU is not in a state of statistical control. Using the nine months of data to form the upper and lower control limits, a total of five points were flagged.

Figure A2: Control limits for WIWNU calculated from the entire nine months of data under the forming recipe.
A3. Control chart of WIWNU under forming recipe using smaller window control limits

Unlike the previous chart, this time only samples 50-80 were used to calculate the control limits. Using this method, a total of 12 points were flagged over the past nine months.

Figure A3: Control limits for WIWNU calculated from the samples 50 through 80 under the forming recipe.