# Development of an Information Discovery System

# to Achieve Process Optimization

TSANG Kai Fai

M. PHIL.

THE HONG KONG POLYTEHNIC UNIVERSITY

2006

The Hong Kong Polytechnic University

Department of Industrial and Systems Engineering

# Development of an Information Discovery System to Achieve Process Optimization

By

TSANG Kai Fai

A thesis submitted in partial fulfillment of the requirements

for the Degree of Master of Philosophy

December 2006

## CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

TSANG Kai Fai

## ABSTRACT

Quality is now acting in the essential role of helping manufacturers to survive in the competitive industrial market, thus ensuring the final products fulfill the required standard while keeping manufacturing costs and time to a minimum. The traditional inspection-oriented quality control methods can only provide limited information as to why defective products are produced. Furthermore, as this information is so limited, how to correct the fault cannot be identified. In an actual manufacturing environment, projects and productions are changed from time to time. Due to the enormous complexity of many processes and the large number of influencing parameters, conventional approaches to modeling and optimization are no longer sufficient. Because of this, a systemized and well-designed approach was formulated, that performs the desired task of problem identification based on a specific process.

Two techniques were employed. The first, a data warehousing technique called On-Line Analytical Processing (OLAP) converts complex data into useful analyzed information. The second is an Artificial Intelligence (AI) techniques including decision tree classification and Artificial Neural Networks (ANNs) that extrapolates probable outcomes based on available patterns of events. These were both deployed together to perform data drilling and analysis which are essential procedures for identifying factors of quality discrepancy in terms of dimensional accuracy and dynamic performance. The proposed integrated approach of Neural-OLAP was built as the Information Discovery System (IDS) which made use of the AI and experience of engineers in identifying foreseeable failure modes of a process or a series of distributed processes and planning for its elimination. This intelligent integrated system aims to process huge amount of production data from multi-processes in multi-manufacturing sites by using the object-oriented database design and multi-dimension data cube structure. Quality prediction can be achieved for performing manufacture feasibility evaluation, identifying potential quality problems and providing guidance for improvement through the decision tree-based ANNs, thereby achieving continual and real-time process enhancement. With this IDS, essential support for users who wish to identify the cause and source of problems can be provided so that immediate action can be taken for rectification.

The system prototype was implemented in a factory, which manufactures magnetic heads for Hard Disk Drives (HDDs), in order to validate the workability of the proposed methodology and the feasibility of the adoption of IDS. An application of the mass production environment of the slider fabrication process, Reactive Ion Etching (RIE), was studied. Substantial improvements in terms of the product quality defect enhancement and process optimization through the corresponding machine settings prediction can be achieved. There is a significant decrease in total quality costs. The frequency of rework and scrap are also reduced resulting in better customer satisfaction, which proves the effectiveness of the proposed methodology.

## PUBLICATIONS ARISING FROM THE THESIS

(Two conference papers were published, and one journal paper was accepted for publication)

1. Lau, H.C.W. and Tsang, K.F. Development of a Distributed Process Mining System for Continual Quality Improvement. *Proceedings of the 19th International Conference on Computer-Aided Production Engineering (CAPE 2005)*, 2005.

2. Tsang, K.F., Lau, H.C.W. and Kwok, S.K. Development of a Distributed Process Mining System for Reactive Ion Etching Enhancement. *Proceedings of the 4th International IEEE Conference on Industrial Informatics (INDIN'06)*, 2006.

3. Tsang, K.F., Lau, H.C.W. and Kwok, S.K. Development of a Data Mining System for Continual Process Quality Improvement. *Special Issue on Computer-Aided Production Engineering, Proceedings of the Institution of Mechanical Engineers, Part B, Journal of Engineering Manufacture*, 2007, Vol. 221, No. B2, 79-193.

## ACKNOWLEDGEMENTS

I would like to express my thanks to my chief supervisor, Dr. Henry C.W. Lau for his excellent guidance, innovative advice and valuable comments on my research work over the entire project. Sincere thanks also go to my co-supervisor, Dr. S.K. Kwok for providing the research direction and encouraging independent thinking in my professional growth. I would like to express my gratitude to my colleagues in the Department of Industrial and Systems Engineering for their thoughtful discussion and assistance.

I would like to give a special thanks to the supervisors of my company, Mr. Stephen C.C. Yu, Mr. Bobby K.W. Fung and Mr. Gerald S.W. Lau for their professional suggestions, great supervision and perpetual encouragement throughout the years. I would also like to acknowledge Mr. K.F. Ng. for supporting my research project as his coordination role in all matters; especially as some issues were actually very complicated. I express my deep appreciation to Mr. Patrick K.M. Lam and Mr. Martin Y.F. Chui for providing me with this golden opportunity and the case of process optimization in production respectively. Thanks should also be given to all my colleagues in SAE Magnetics (H. K.) Ltd. for their great help and kind support.

I am grateful to my family for their love and whole-hearted support and also Miss L.L. Yao for her continuous encouragement and precious suggestions. Without them, I would not have been able to finish my research study.

The research described in this thesis is supported by a grant from the Teaching Company Scheme under the University-Industry Collaboration Programme (UICP) of the Innovation and Technology Fund of the Government of the Hong Kong Special Administrative Region, and SAE Magnetics (H. K.) Ltd. (UIT/065).

## TABLE OF CONTENTS

## LIST OF FIGURES

## LIST OF TABLES

# CHAPTER 1. INTRODUCTION

## 1.1 Research Background

Companies are experiencing growing pressure to reduce production costs and to save time in today's increasingly competitive market. Each company wants to get the largest production from the given raw materials and the greatest profit from a fixed investment (Rao, 1996). Quality issues are therefore essential for manufacturers so that they can ensure that their products are up to the highest standard as the mean of lower production cost and greater profit. However, due to the various causes associated with the operational processes, it is virtually impossible to identity the sources of quality problems. The traditional inspection-oriented quality control methods can only provide limited information as to why defective products are produced; how to correct the faulty process still cannot be identified (Fan and Wu, 1992).

In an actual manufacturing environment, projects and productions are changed from time to time. Due to the enormous complexity of many processes and the large number of influencing parameters, conventional approaches to modeling and optimization are no longer sufficient (Westkämper and Schmidt 1998). When a change occurs in the system's physical structure, such as a machine failure, material variance, etc, the model becomes inaccurate and needs to be remade (Haouani *et al.*, 2000). Rietman *et al.* (1996) stated that "the objective of process modeling is to learn something about the process or to make real work prediction based on the behavior of the model. The traditional methods of statistics, expert systems and first principle are excellent for specific tasks, but they are limited." More research expertise also claimed that simple statistical models are not capable of dealing effectively with more than a few variables and it is hard to generate appropriate rules without prior experience or knowledge (Jemkins *et al.* 1986 and Fahn *et al.* 1999). Transferring the experts' knowledge into computer programming can be extremely difficult and expert system models built from a knowledge base often are too brittle because of over specialization of the rules generated. Dynamic models based on differential / difference equations would be very difficult for large nonlinear systems with many variables.

On the other hand, a number of global manufacturing networks have been established for the global and borderless manufacturing market, taking the advantages of fast-growing networking and Information Technology (IT) (Ho *et al.*, 2005). With the help of IT, Total Quality Management (TQM) can be achieved. Successful quality management practices by the TQM organization can prove the effectiveness of IT implementation (Besterfield, 2004). The utilization of IT is fast developing to meet the challenges of quality enhancement and process optimization. Data mining, Artificial Intelligence (AI) and distributed object technology have achieved significant attention as the key techniques for quality enhancement and process optimization in different industrial applications. They can help to provide critical information about the source of known defects in the process so as to identify the related operation needed to achieve proactive quality control and enhancement, providing the remedial actions and plan for the future within the production process flow and thereby greatly reducing the production costs. The research therefore aims to develop a "predictive" with quality problem-solving capability data mining system to achieve process optimization by using the combination of existing AI techniques.

## 1.2 Research Company Background

Slider fabrication, a common term in the HDD industry, represents a series of processes that manufacture thousands of Integrated Circuits (ICs) within a single wafer ranged a few inches. By using the most advanced technology, Giant Magneto-resistive (GMR) (Wang and Taratorin, 1999), Perpendicular Magnetic Recording (PMR) (Khizroev, 2004) and Tunneling Magneto-resistive (TMR) (Kagami *et al.*, 2006), the slider is manufactured and used as the magnetic head for the HDD to record and read data into the high density disk media. There is at least one HDD which contains several magnetic heads, in every single computer or even an electronic device such as a portable MP3 player.

SAE Magnetics (H.K.) Ltd., which is one of the leading contributor and manufacturers in the magnetic head industry, is the cooperative company for this research project. It is a wholly owned TDK subsidiary, and is one of the world's leading independent manufacturers of magnetic recording heads, Head Gimbals Assemblies (HGAs) and Head Stack Assemblies (HSAs) for computer disk drivers. SAE also supplies high performance recording heads for video tape recorders. SAE Magnetics (H.K.) Ltd. was established in October 1980 to meet the increasing demand for magnetic recording heads. In August 1986, SAE joined the TDK Corporation Group, the world's largest and most advanced ferrite manufacturer. Since then, the combination of SAE's technologies in head manufacturing and TDK's expertise in magnetic materials has brought SAE to a leading position in head manufacturing. SAE continues to provide the best support to customers by developing and producing the best magnetic recording heads, including the GMR and TMR technologies. More details of the research company background are included in the Appendix A.

### 1.2.1 Background of Magnetic Head Industry

The world’s first commercial HDD was designed by IBM back in June 1957 (Stevens, 1999). This product, IBM 350, was shipped inside the new machine called Random Access Method of Accounting and Control (RAMAC). The disk’s magnetic recording technology, based on magnetic drums, permitted an areal density of 2000 bits/in$^2$. This level of areal density was possible because the magnetic read / write head was supported above the disk surface by a hydrostatic pressurized air bearing.

Since then, the use of the magnetic storage devices has grown substantially within the last decade. The growth is based, to a large degree, on the invention and implementation of hard magnetic disc drivers, where a magnetic head is supported in close proximity to the rotating magnetic discs. In order to improve access to the stored information, it is essential to record at a higher areal density which is the product of linear transition density and track density in the direction of motion. According to Grochowski (2003), the HDD capacity and areal density will reach approximately 1 Tetra Bytes (TB) and 100 GB/in$^2$ by the end of 2006 respectively (Figure 1.1 and 1.2) that is 50 million times higher than 50 years ago in RAMAC.

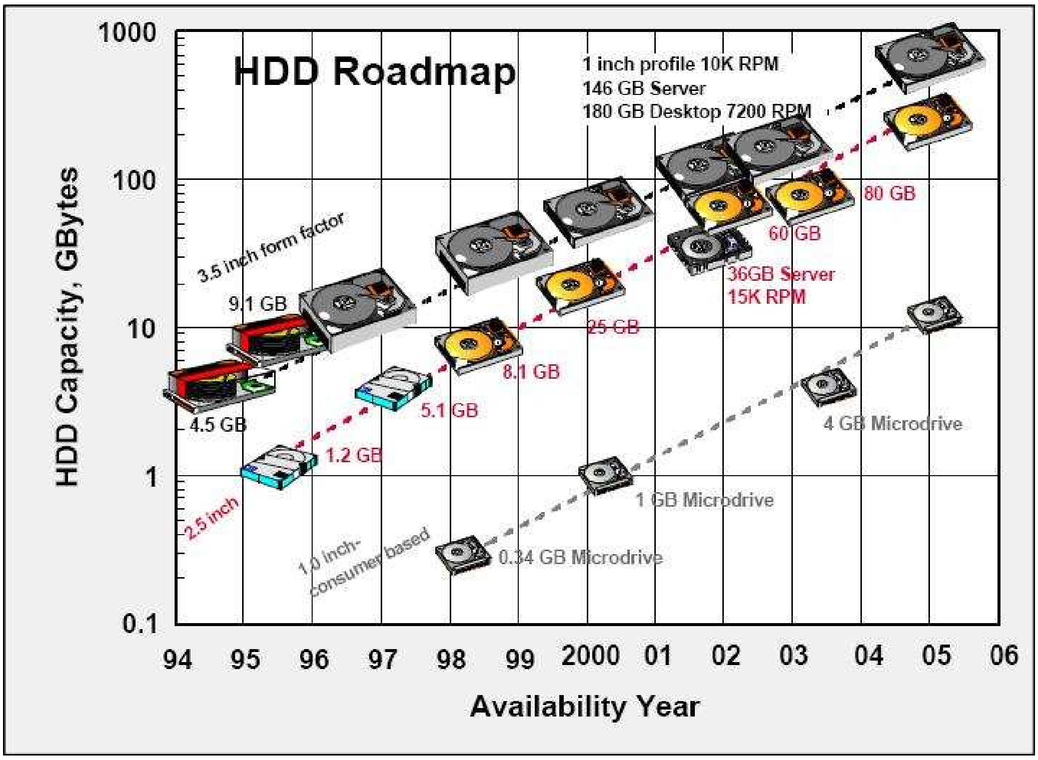



Figure 1.1 HDD Roadmap (Grochowski, 2003)

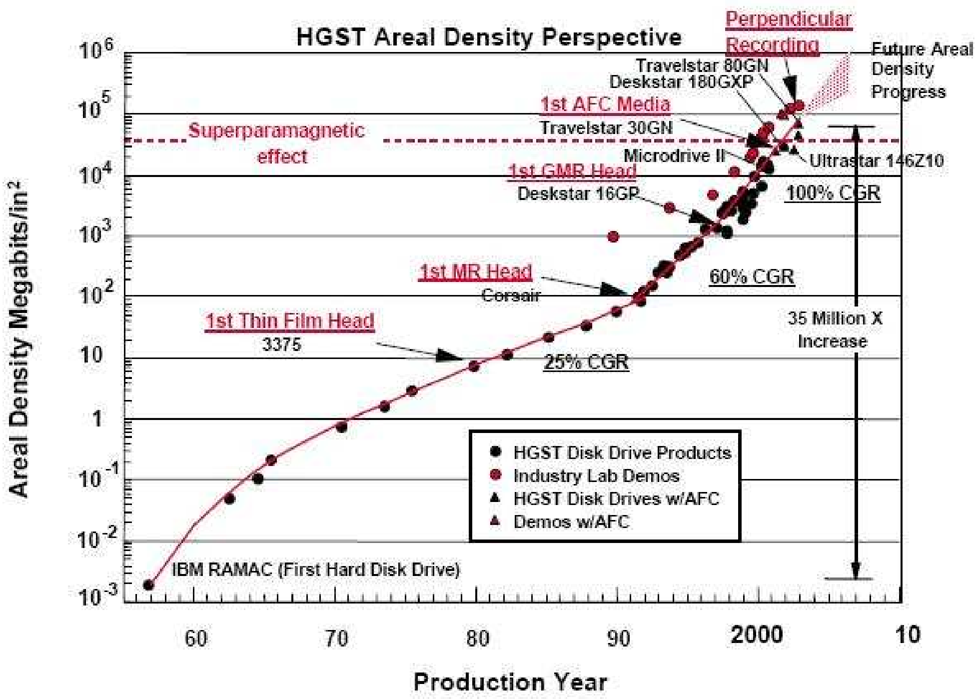



Figure 1.2 Areal Density Perspective (Grochowski, 2003)

The increasing recording density in areal and higher reliability endurance becomes the way to fulfill the needs of higher storage capacity and lower manufacturing costs of HDD. From the previous research by Pang (1998), the high frequency of magnetic flux transition causes a diminishing of the signal output during the high data transfer rate with fast rotational speed of disk medium. This leads to the need for continuing to scale down the head-to-media spacing.

Today's HDD all employ a slider that flies over the disk on a self-pressurized air bearing and have a magnetic spacing of 50 nanometers or lower. The head / medium interface is designed so that the magnetic head is separated from the disc by a thin air film to avoid wear. The air film must be suitable enough to prevent excessive material interactions and also give a sufficiently strong magnetic read-back signal. The designs of the air bearing and also the processes to manufacture it are essential to maximize the linear density of the recording system.

### 1.2.2 Introduction of HDD Components

To aid a better understanding of the research project's background, the basic design and its components in a standard HDD are discussed in following as Head Disk Assembly (HDA), Head Gimbol Assembly (HGA) and slider.


Figure 1.3 Basic HDD Components

**Head Disk Assembly**

The major core of the HDD is the HDA, which usually describes the HDD component without the Printed Circuit Board (PCB) assembly. It is combined with other elements, including spindle (motor), disk media, HSA (HGA, slider / head), VCM, and HDD base, etc in clean-room environment (Figure 1.3). After the PCB is attached to the HDA, it is then called an HDD and is put into the burn-in test to final check the performance and reliability of the HDD. The HDA contains from 1 to 4 magnetic disks mounted on a common spindle shaft which rotates an in anti-clockwise direction at several thousand of revolutions per minute (4,200 – 10,000 rpm). Through the disk rotation, a thin film of air is generated over the disk surface. The air flows from the leading edge of the magnetic head and induces an up-trust against the air bearing surface at the head / media interface. This allows the magnetic head to "fly" over the disk surface. The state-of-the-art fly height is on the order of 1 micro-inch or 25nm, while the relative speed between slider and disk is extremely high around 10 m/s or even higher. A good analogy is of the flying slider being like a Boeing 747 airplane flying constantly near the ground constantly (Wang and Taratorin, 1999).

**Head Gimbol Assembly**

HGA is an assembly of slider, suspension and integrated circuit. They are all mounted on a single actuator that provides the accessing motion. Each magnetic head is spring loaded against the disk by a soft spring provided by the suspension. Rotational stiffness is provided by the flexure, which connects the slider to the suspension. During operation, the head glides above the disk, separated by a thin self acting film of air generated by the lubrication effect of the air bearing. During the starting and stopping phases, there is head-disk contact. The amount of wear of the magnetic head is thus of paramount importance (Pang, 1998). Several HGAs are then installed with the actuator arm and voice coil to form the HSA, one HSA contains up-to 8 HGAs depending on the design and required capacity.

**Slider**

The slider or magnetic head is a major element that carries the read / write transducer and is considered as the most precise and complicated part for recording and receiving a magnetic signal from the surface of the disk media. It is mounted on the end of a stainless gimbel-suspension. Sliders are designed to form an air bearing that gives the lift force to keep the slider-mounted head flying at the desired head-medium spacing. Properly designed, the slider follows any irregularities in the disk surface and maintains spacing as it moves outward in radius to fly at a higher linear velocity. Tribological issues, as well as contaminants in the environment, are of constant concern (Monson, 1999).

Sliders are actually manufactured from wafer through the fabrication processes. The wafer is first formed that is similar to other semiconductor wafers. It is then cut into row bars through the slider fabrication process which in itself involves several processes among which is Reactive Ion Etching (RIE) which is the major case study in the research. The row bar is finally cut as a slider as shown in Figure 1.4.

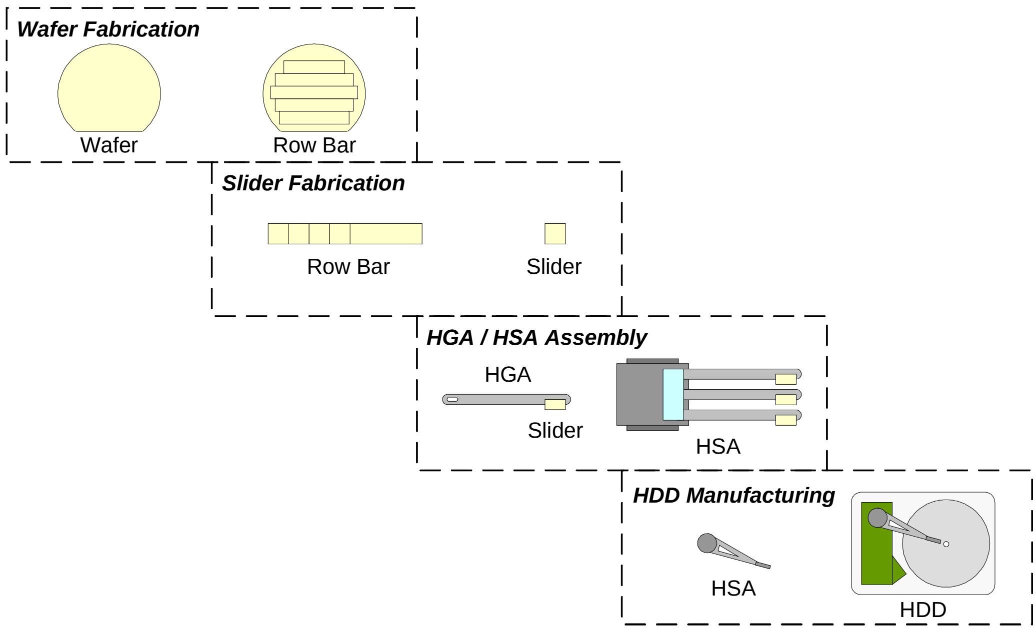



Figure 1.4 From Wafer to Hard Disk Drive

## 1.3 Background of Slider Fabrication

Since the proposed IDS is implemented under the cooperative company, the case study of the research is focused on only one of the processes in the slider fabrication, Reactive Ion Etching (RIE). This section covers the basic design of the slider and air bearing surface, the general slider fabrication process and the principles of RIE.

### 1.3.1 Designs of Slider and Air Bearing Surface

There are five major parts in a slider to provide the features to read / write from the magnetic disk medium by flying on the surface at a very high speed. They are the Air Bearing Surface (ABS), air groove, pole tip, bond pad and trailing surface (Figure 1.5).

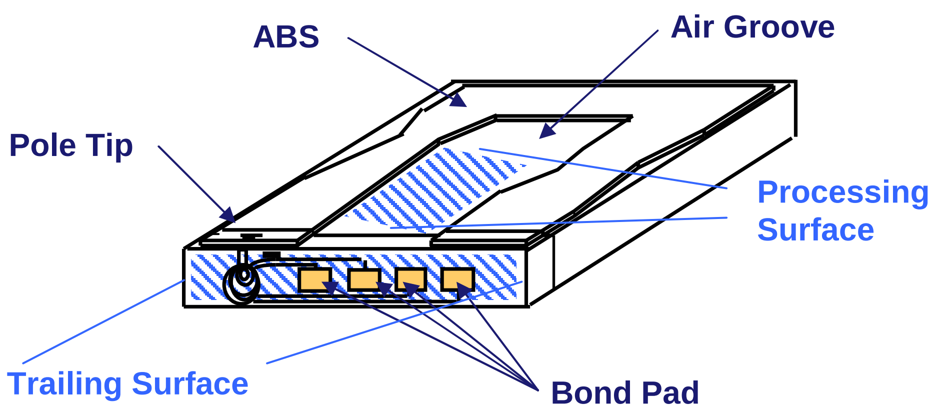



Figure 1.5 Slider Parts with ABS and Air Groove

There are several types of ABS design available on the market. They are mainly in two catalogues, positive and negative pressure air bearings (Pang, 1998). The most popular designs used nowadays, include:

- Positive pressure ABS

  - Machined Taper Flat (MTF)
  - Two-rails Positive Pressure Design
  - Three-rails Positive Pressure Design
  - Transverse Pressure Contour (TPC)
- Negative pressure ABS
  - Seagate’s Advanced Air Bearing (AAB)
  - Fujitsu’s Guppy
  - IBM’s NPAD
  - TDK’s XNP

Negative ABS have a better overall performance and so are more widely used nowadays. They have the advantage of low gram-load capability for improved head / disk interface reliability and result in reasonably constant fly height over the disk. For the mass production of magnetic heads, the RIE in the slider fabrication is necessary to etch the surface shallow of the slider to create the air groove for the ABS.

### 1.3.2 Slider Fabrication Processes

As discussed before, the slider is manufactured from (i) wafer, then (ii) block, (iii) row bar and finally forms (iv) a slider. Slider fabrication is s series of manufacturing processes (up-to 50 to 60 steps depending on the design and customer requirements). It takes approximately 3 to 5 working days for a wafer to finish all the processes in a clean room.

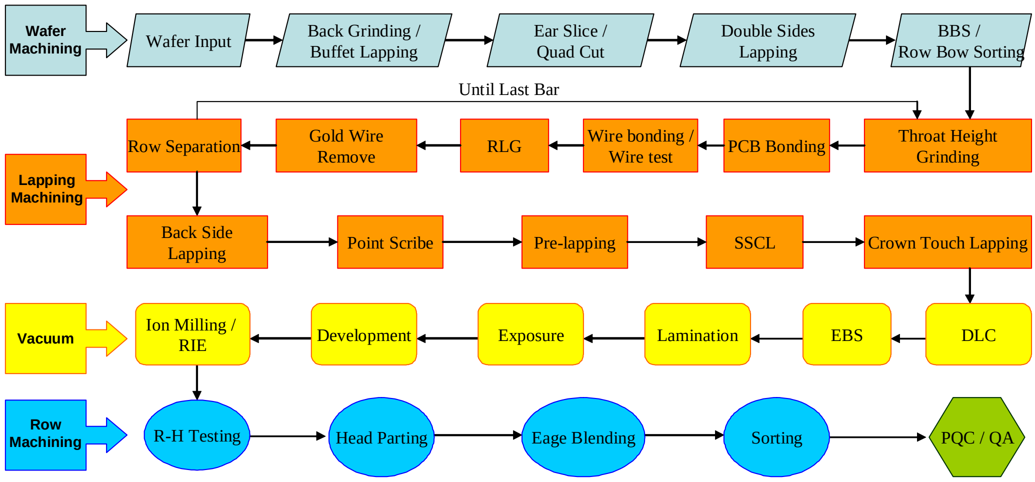



Figure 1.6 Slider Fabrication Processes

Figure 1.6 indicates the major process flow in slider fabrication. Five different colors identify the five major groups of processes. They are: (i) wafer machining – by cutting the input wafer into row bars for further processing, (ii) lapping machining – by lapping the row bar back and ABS into the required dimension, MR resistance, roughness, etc, (iii) vacuum – by coating the DLC to protect the ABS surface and etching the shallow and corresponding shape of the ABS, (iv) row machining – by cutting the row bar into a slider and (v) PQC / QA – by controlling the production quality and quality assurance of the slider in the final step.

### 1.3.3 Principles of Reactive Ion Etching

The manufacturing processes in the slider fabrication can be divided into several catalogues; the one used in the case study is the Reactive Ion Etching (RIE). As shown in previous Figure 1.5, Etching is the process of removing the top layer(s) from the slider surface through the openings in the resist pattern (Zant, 2000). RIE is one of the most accurate dry etch techniques in which gases are the primary etching medium. It combines both plasma and ion beam etching principles in resulting in a high selectivity ratio than in other methods. The etched surface in the slider is used as the air groove for the magnetic head to fly on the disk media in HDD. Both ion milling and RIE are used to produce the desired ABS. Ion milling is a technique of fine manufacture thin film by physics sputtering a substrate surface with inert ion. The etching part without any film cover forms the shallow of the air groove. RIE is used to get remain recess. This research study is focused mainly on this RIE process.

RIE is used to transfer a pattern from a photo-resist to the ABS of the substrate so as to produce an appropriate negative pressure cavity by removing a layer of opened areas. Traditionally, the fabrications of air-bearing surfaces on read / write heads have been achieved by using either RIE or ion milling. In SAE Magnetics (H.K.) Ltd., the RIE machine is the major tool to form the air bearing surface, which includes PTI (RIE system) and STS (ICP system) in Figure 1.7 and 1.8 respectively. In fact, the fabrication of air-bearing surfaces for Magneto-resistive (MR) and Giant Magneto-resistive (GMR) heads is widely adopting RIE technology all over the world.

Figure 1.7 PTI RIE System

Figure 1.8 STS ICP System

RIE is a powerful dry etching process which combines the advantage of ion milling and plasma etching, so both physical and chemical effects contribute to etching. The instrumentation required for this kind of dry etching includes a reaction chamber, an RF power supply and a vacuum source (Figure 1.9). The sample is introduced into the reaction chamber, which is evacuated by the vacuum pump. Reactive gases such as $CF_4$ / $C_2F_6$ and inert gases (e.g. He and Ar) are introduced into the chamber and converted to reactive plasma by the RF power supply. The plasma reacts with the surface of the Al·TiC substrate and resulting volatile by-products, like $TiF_4$, are extracted by the vacuum pump. The residue $AlF_3$ is either attached to the chamber everywhere or stays on the etching surface of the substrate (Figure 1.10). The six major steps of the RIE are:

a. Generation of reactive species (e.g. free radicals)
b. Diffusion to surface
c. Adsorption on surface
d. Chemical reaction
e. Description of by-products
f. Diffusion into bulk gas

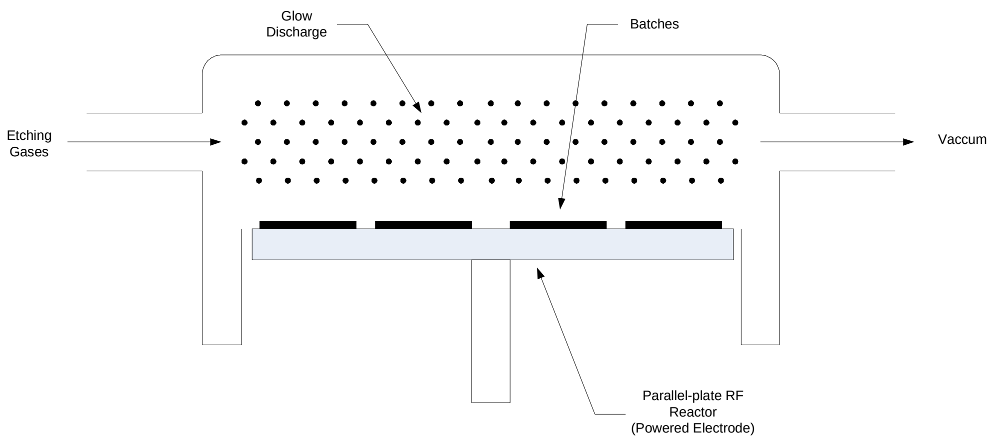



Figure 1.9 RIE Reaction Chamber

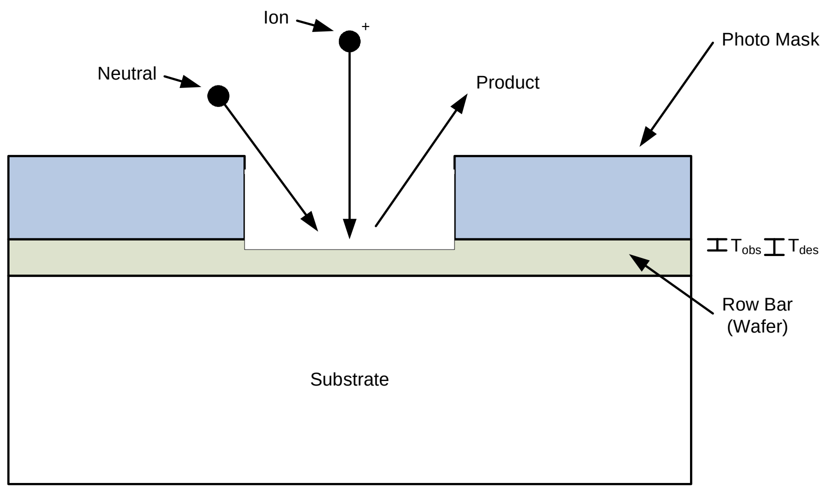



Figure 1.10 Process Diagram of RIE

The specification of the RIE reaction with the corresponding chemical equations is listed below:

- Condition:
  - Beam voltage: ~ 700 V
  - Beam current: ~ 1000 A
  - Incident RF power: ~ 400 W
  - RF frequency: 13.56 MHz
  - Pressure: 10 – 490m Torr
  - Gas Flow Rates:
    - $CHF_3$: 30 – 78 sccm
    - He / Ar: 1- 17 sccm
  - Fixture angle: - 45°
  - Chamber vacuum: ~ 5 x$10^{-5}$ Torr
- Inflate inert gas: He / Ar
- Forming plasma of He / Ar by using RF power

- Changing energy between ion and gas of $CF_4$ / $C_2F_6$ and ionizing F ion
- Chemical reactions between F ion and substrate are as below:

(i) $C_2F_6 + e^- \rightarrow 3CF_3 + e^-$

$CF_4 + e^- \rightarrow CF_3 + F + e^-$

(ii) $Al_2O_3 + F \rightarrow AlF_3 \downarrow$

$TiC + F \rightarrow TiF_4 \uparrow$

$Al_2O_3 \cdot TiC + F \rightarrow AlF_3 \downarrow + Ti_xF_y \uparrow + CO_2$

The following factors also affect the performance and the final quality of the products in RIE:

- Temperature (etching rate, spontaneous chemical reaction, etching directivity)
- Pressure (ion density, ion directivity)
- Power (ion density, ion kinetic energy)
- Others, including gas flow rate, reactor materials, reactor cleanliness, loading (micro-loading), mask materials.

Moreover, surplus ions collide with the substrate and form a physical etching process so the reaction products can be taken away. It can also be done through controlling etching time and depth. The etching time therefore becomes the most critical controlling parameter in RIE. The detailed process flow of RIE and the related steps after the process are listed in Table 1.1. Other than RIE, the next major steps are cleaning and inspection both visually under the microscope, and dimensionally through a testing machine.

Table 1.1 Process Flow of RIE with Cleaning and Inspection

| **List** | **Process** | | **Description** |
|---|---|---|---|
| 1 | RIE | Incoming inspection | Check uniformity of package / wafer no., up / down |
| 2 | | Putting on tray | Putting jigs on fixture plate |
| 3 | | Etching | Etching main cavity recess |
| 4 | Film Removal | DI water cleaning | Clean row bars surface |
| 5 | | $N_2$ dryer | ~ 30 seconds |
| 6 | | U/S ethanol cleaning | ~ 45 minutes, in cleaning tank |
| 7 | | Ethanol brushing | ~ 1 minute, in cleaning tank |
| 8 | | DI water cleaning | ~ 30 seconds |
| 9 | | $N_2$ dryer | ~ 30 seconds |
| 10 | ABS Cleaning | Auto brushing cleaner | NaOH 0.02% |
| 11 | | $N_2$ dryer | ~ 30 seconds |
| 12 | Visual Inspection | 40x visual | For E8 defect |
| 13 | Dimension Inspection | ABS pattern | TiC dimension X & Y |
| 14 | | Depth dimension | For recess height with a-step machine |
| 15 | Debonding | U/S cleaning in ACE | ~ 35 minutes for row bar divorced from jig |
| 16 | | Transfer tray | In ACE |
| 17 | Row Cleaning | U/S Cleaning in ACE | ~ 45 minutes, cleaning row bars contamination |
| 18 | | Manual brushing on ABS | ~ 1 minute |
| 19 | | DI water cleanliness | ~ 30 seconds |
| 20 | | $N_2$ Dryer | ~ 30 seconds |
| 21 | Visual Inspection | 40x visual and Q-tip | With ACE for ABS, housing surface, bond pad and wafer no. |
| 22 | | Outgoing | For final QA |

### 1.3.4 Needs of Process Optimization

Because of the complexity of a customer's requirements, the process procedures and related process settings are strongly related to the process performance and the finished product quality. In fact, the whole jig of sliders will be wasted or required to be reworked if there is just a single failure, such as an under-etch, where the opening depth is not enough, or an over-etch which the causes damage of the row bar. However, the causes of failure take time to be identified and the result cannot be seen immediately when the RIE is still work-in-process. Therefore, the IDS is proposed for helping the engineers to find out the optimum factors in a shorter time period. In the case of magnetic head manufacture, the system aims to (i) identify the sources of critical process information to the related quality problem and (ii) forecast the corresponding machine settings to optimize the process and improve product quality, including the prediction of the ideal etching time / rate and classification of the corresponding model for new conditions, through the OLAP and classification-based ANNs.

## 1.4 Research Objectives

The aim of this project is to develop a data mining system which is able to provide useful information for analyzing quality problems associated with the components produced by a magnetic head manufacturing company. This data mining system needs to be adaptive enough to predict the future trend of events and the ability to recognize data patterns, both of which are important for improving quality performance and achieving process optimization. Local Small and Medium Enterprises (SMEs) can also achieve the benefits through applying the proposed methodology.

The objectives of this research project are:

(i) To develop an adaptive process mining system which provides critical information about the process source of known defects so as to identify the related operation;

(ii) To introduce a “hybrid” approach for enhancing machine self-learning by using the decision tree-based ANNs’ principle augmented with OLAP technology, culminating in the formation of Neural-OLAP which can emulate human workers with problem-solving capabilities;

(iii) To provide an infrastructural framework, which because of its adaptive ability, can be used in identifying the quality data pattern which in turn can be used to support identification of the causes of the quality problem normally related to distributed operational processes.

## 1.5 Assumptions and Limitations

The following assumptions and limitations are made for IDS implementation in SAE Magnetics (H.K.) Ltd.:

1. This research is company-based, most of the sensitive data is considered as the confidential commercial information. Only the research approach and methodology with non-sensitive data can be published.

2. This research focuses on the environment of slider fabrication business in Hard Disk Drive (HDD) industry. The literature reviews, approach proposed and modeling are based on the process and production of related product and its requirement. It is believe that the same methodology can benefit other similar businesses for process optimization.
3. The case study of Reactive Ion Etching (RIE) is based on the current environment and condition in SAE Magnetics (H.K.) Ltd. Those related input / output parameters are limited to the business needs and machine / equipment availability.
4. The foundation software packages and techniques of IDS are fixed by business decision, the data warehouse and business intelligence providers are Oracle and SAS Institute respectively. Those AI approaches are limited to the current available in SAS Enterprise Miner.
5. System performance and benefit measurement are done by company-side. Mainly the ANN models verification and the related approaches used by company-side are discussed.

## 1.6 Thesis Outline

This thesis consists of seven chapters. Chapter One is an introduction in which the current problems of production quality defect due to process failures in general and the research objectives are described. In Chapter Two, research literature related to Statistical Process Control (SPC), Process Modeling, Artificial Intelligence (AI) techniques and the successful industrial applications is reviewed. The system design and development of the proposed intelligent hybrid system are defined in Chapter Three. The roadmap for system implementation is stated in Chapter Four. Chapter Five is a case study of Reactive Ion Etching (RIE) in a collaborative company for validating the proposed system. Discussions of the validated result compared with the traditional approach, implications, limitations and challenges are given in Chapter Six. Finally, conclusions and recommendations for future research are presented in the last chapter.

# CHAPTER 2. LITERATURE REVIEW

## 2.1 Introduction

This chapter aims to review the research in the area of quality enhancement and process optimization starting from the introduction of popular Statistical Process Control (SPC), followed by with the definition and functionality of the process modeling. After that Artificial Intelligence (AI) techniques are discussed as well as the successful intelligent applications in the Hard Disk Drive (HDD) industry.

## 2.2 Statistical Process Control

Two areas of Statistical Process Control (SPC), the foundation and the needs of Information Technology, are discussed in the following paragraphs.

### 2.2.1 Foundation of Statistical Process Control

Since the popular Statistical Process Control (SPC) was first introduced by Vilfredo Pareto (1848-1923) in the early 20th century (Thompson and Koronacki, 2002), it has been widely used in industry as a standard tool for process control and quality improvement. According to Pareto, “the many failures in a system are the result of a small number of causes”. In order to improve a system, skilled investigators are required to find and correct the causes of the ‘Pareto glitches’”. Deming (1982) then suggested the famous fourteen points for the modern paradigm of quality control after World War II for manufacturers in Japan and noted that “the failures in systems can be viewed, mathematically, as a problem in contaminated distributions. This fact provides us with a tool that can lead to the replacement of the quasi-feudal managerial systems Pareto predicted by the nurturing system of continual improvement". Another common term, Statistical Quality Control (SQC), also known as control charting, has evolved into SPC to reflect the move away from product control to a system focus which started with Walter Shewhart's work at Western Electric in the 1920s (Gruska and Kymal 2006). They are all seen as the foundation of process control nowadays.

### 2.2.2 Needs of Information Technology in Statistical Process Control

These traditional approaches, however, may not be able to ensure the mass production of finished goods of a high enough quality and low enough cost in this dynamic market. They are more or less seen as simply quality management methodology and basic statistical procedures. Manufacturing processes are now monitored by SPC and any non-random departure from the desired target is detected using a Shewhart control chart. Some pioneer researches enhanced SPC further by giving the modern industry an insight. The Next Generation Quality Control (NGQC) (Wu *et al.*, 1989) proposed to increase the inspection sampling size to 100%, thus ensuring the product quality with zero sampling error. Two tasks are essential in NGQC: real-time data collection and modeling the manufacturing process. Spanos *et al.* (1992) also proposed a real time SPC using time series filters and multivariate statistics techniques. This scheme is capable of generating alerts based on the machine parameters that have no specified target value but are related to how the equipment is functioning. Moreover, it is able to correlate quality data with process operating data in order to provide a mean of fast fault source identification and guidance in correcting the faulty process. It therefore has had a big impact on industry – replacing the slow performance for analyzing a large of amount of quality data by humans and traditional approaches.

As currently practiced, processes are re-tuned only when SPC indicates an out-of-control situation occurring in the routine operation. The re-tuning is generally based on the operator's experience (Wang, 2003) which was found not to be so accurate or efficient under certain conditions. The assignable causes are usually not easy to identify and sometimes are also costly to remove. Guh (2005) indicated that a control chart is the key part of SPC implementation. The power of a control chart lies in it ability to separate out out-of-control special disturbances from in-control inherent variability in the process. The traditional control chart however could identify only when to seek a disturbance, it could not show where and what to look for directly. Implementing the effective Information Technology (IT) system was suggested to recognize the unnatural patterns in the control chart. However, a more powerful statistical capability is required. Computerized SPC and AI were then introduced as the solutions.

## 2.3 Process Modeling

The definition, purposes and the Artificial Intelligence (AI) in process modeling are discussed below.

### 2.3.1 Definition and Purposes of Process Modeling

According to Bequette (1998), a process model is defined as "a set of equations (including the necessary input data to solve the equations) that allows us to predict the behavior of a process system". There are many reasons for developing process models, the major objective is to improve or understand the operation of the process by developing the dynamic process model. It can be defined mathematically as the search for the settings of the n variables of a function, such as $f(x_1, x_2, \ldots, x_n)$ that optimize $f(x)$ (Shaffer, 1996). Papalambros and Wilde (2000) defined a model that is an abstract description of the real world giving an approximate representation of highly complex functions of physical operations. Modeling has long been an integral component in organizing, synthesizing and rationalizing observations of and measurements from real processes and in understanding their causes and effects (Nirmalakhanda, 2002). Baquette (1998) pointed out that those models are often used for (i) operator training, (ii) process design, (iii) safety system analysis or design, and (iv) control system design.

By discovering and simulating the process, experiment is one of the best ways to know better how the process operates. However, direct experimentation is time-consuming, expensive, even dangerous and simply impossible. The price of mistakes and wrong decisions is unacceptably high (Samarskii and Mikhailov, 2002). The aim in developing the model is to help and answer the needs for that process as a mean to answer the questions. Generally speaking, the goals and objectives of modeling are twofold: research-oriented or management-oriented. Nirmalakhandan (2002) suggested that the specific goals of modeling efforts can be one or more of the points such as interpreting the process operation, analyzing its behavior, managing, operating or controlling it to achieve desired outcomes, designing methods to improve or modify it, testing hypotheses or forecasting its response under varying conditions. Therefore, models are now important for manufacturers to have a better understanding of the dynamic behavior of their processes from both process design and process control perspectives.

### 2.3.2 Process Modeling with Artificial Intelligence

As modeling is used to describe the real world, nearly all physical systems in the existence are nonlinear. Unless physical insight and the laws of physics can be applied, establishing an accurate nonlinear model using measurement data and the system identification method is difficult in practice (Leung, 2004). The relationships between the process variables and *f(x)* may not be well understood and indirect methods need to be employed. The setting of process variables is modified in some logical way and then presented to an objective function to determine whether this particular combination of process variables is an improvement. Such an objective function would include numerically the performance of a particular combination of process variables. In many applications, the calculation of the derivate is not possible analytically or it is very time consuming or numerically inaccurate (Ning, 2003). Therefore, an alternative to this statistical approach, the Artificial Intelligence (AI) approach, is suggested. According to Westkämper and Schmidt (1998), the bases for computer-based optimization and control of process chains are numerically assessable interdependencies for the partial models. These models have to represent all characteristics of the manufacturing process necessary for optimization and control with the accuracy needed to perform the task. Researchers in different engineering disciplines concentrated on identifying the generic nature of problem-solving tasks and on the application of the AI techniques to solve the tasks in a generic manner (Takeda *et al.*, 1990). Such an approach gave rise to a number of generic problem-solving models depending on the nature of the knowledge required and the nature of the information being processed.

## 2.4 Computational Approaches in a Manufacturing Environment

Many remarkable research outcomes were gained by using the powerful computational tools. These computer processing techniques, which include different data mining and AI algorithms, are the key techniques for quality enhancement and process optimization in different industrial applications. A brief review of the specific technical literature reveals that there are many cases of successful applications of AI techniques and information technologies to the quality evaluation area.

### 2.4.1 Data Mining and Artificial Intelligence

It has already been seen that different tasks in engineering problem solving require different computational tools. Inference or deduction from a set of facts, which simulate intelligent decision making, plays a major role in many problem solving tasks (Krishnamoorthy and Rajeev, 1996). Those computational tools that assist designers should make the designers more creative to the intelligence and experience that the designer has as well as be able to use their expert knowledge of the problem domain for decision making. Depending on the type of task, the knowledge and processing required may involve use of numerical models, database systems, visualization tools or decision-making models to provide solutions that need human expertise. Thus to address a wide spectrum of tasks, AI and expert system technologies provide the much-needed software tools to integrate the various processes to build knowledge-based systems for computer aided engineering (Green, 1993).

Other than AI, data mining is another common term to describe the computational process involved in finding information from a huge amount of data. Data mining is the process of automatically discovering useful information in large data repositories. Data mining techniques are deployed to scour large databases in order to find novel and useful patterns that might otherwise remain unknown. They also provide one with the capability to predict the outcome of future observations (Tan *et al.*, 2006). According to Two Crows (1999), data mining is a process that uses a variety of data analysis tools to discover patterns and relationships in data that may be used to make valid predications.

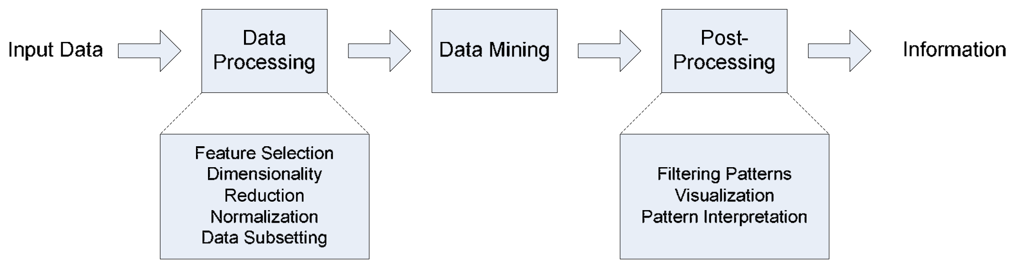



Figure 2.1 Knowledge Discovery in Database (KDD)

Tan *et al.* (2006) also indicated that data mining is an integral part of Knowledge Discovery in Databases (KDD), which is the overall process of converting raw data into useful information. This process consists of a series of transformation steps, from data preprocessing to post-processing of data mining results (Figure 2.1). On the other hand, Two Crows (1996) suggested the first and simplest analytical step in data mining is to describe the data – summarize its statistical attributes, visually review it using charts and graphs, and look for potentially meaningful links among variables. A predictive model based on patterns determined from known results is then built and tested on results outside the original sample which can be a useful guide to understanding the business. The final step is to empirically verify the model. The AI techniques, including Artificial Neural Networks (ANNs), are the ones in which the predictive model algorithms lay.

### 2.4.2 On-Line Analytical Processing

Turning data into knowledge was a difficult task in the past as it relied on manual analysis and interpretation. Such manual data analysis, in fact is becoming impractical in many domains as data volumes grow exponentially (Fayyad *et al.*, 1996). Due to the rapid development in IT and e-commerce, people want to get more from their database systems. Good examples are Wal-Mart and NASA. Wal-Mart generates around 20 million transactions a day while NASA's earth-observing system launched in 2000 produced 50 GB of image data per hour (Cios *et al.*, 1998). Veloso and Borrajo (1994) explained that modern industrial processes require advanced computer tools that should adapt to the user requirements and to the tasks being solved. Popular application software, such as Excel, can perform statistical analysis based on the quality raw data. However, it is time-consuming and not real-time processing. According to Tan *et al.* (2006), "a number of database systems support such a viewpoint, most notably, On-Line Analytical Processing (OLAP) systems. Indeed, some of the terminology and capability of the OLAP system has made their way into spreadsheet programs that are used by millions of people. OLAP systems also have a strong focus on the interactive analysis of data and typically provide extensive capabilities for visualizing the data and generating summary statistics". Therefore the data warehouse technique OLAP was introduced.

On-Line Analytical Processing (OLAP) is a technique which provides a service for accessing, viewing and analyzing large volumes of data with high flexibility and in a timely manner through the multi-dimensional view (Chaudhuri and Dayal, 1997). It is mainly used for describing knowledge discovery in databases, knowledge extraction, data exploration, data pattern processing and information harvesting, thereby providing sophisticated analysis for business application to support decision making (Lee and Siau, 2001; Forcht and Cochran, 1999). The essential characteristic of OLAP is that it performs a numerical and statistical analysis of data and the data is organized multi-dimensionally. OLAP's functionality includes: (i) calculation and modeling applied across dimensions via hierarchies and / or across values, (ii) trend and seasonal analysis over time periods, (iii) slice and dice data in almost any manner, (iv) drill down into data to get deeper level of detail and (v) reach-through of underlying detail data (Dahr and Stein, 1997). With the combination of extensively indexed and combined data in the data warehouse, OLAP tools allow quality engineers to dig through megabytes or gigabytes of quality data without waiting hours for results.

A multi-dimensional database is a type of database that is optimized for data warehouse and OLAP applications. They are frequently created using input from existing relational databases. Whereas there are typically accessed using a

Structured Query Language (SQL) query, a multi-dimensional database allows a user to ask questions related to summarizing business operations and trends. An OLAP application that accesses data from a multidimensional database is known as a Multi-dimensional OLAP (MOLAP) application A multi-dimensional database or a Multi-dimensional Database Management System (MDDBMS) – implies the ability to rapidly process the data in the database so that answers can be generated quickly (Bob and Coronel 2002). A number of vendors provide products that use multi-dimensional databases. Approaches to how data is stored and the user interface vary. Multi-dimensional database technology is a key factor in the interactive analysis of large amounts of data for decision making purposes (Paul, 2004). In contrast to previous technologies, these databases view data as multi-dimensional cubes that are particularly well suited for data analysis. Multi-dimensional models categorize data either as facts with associated numerical measures or as textual dimensions that characterize the facts. Queries aggregate measure values over a range of dimension values to provide results. This database technology is being applied to distributed data and to new types of data that current technology often cannot adequately analyze.

OLAP and data mining are two very different tools but which can complement each other (Two Crows, 1999). Analysis from OLAP generates a series of hypothetical patterns and relationships and uses queries against the database to verify them or disprove them. Data mining is different from OLAP because rather than verifying hypothetical patterns, it uses the data itself to uncover such patterns. It is essentially an inductive process. As for the term ‘data warehouse’, this “is a subject-oriented, integrated, time-variant, and nonvolatile collection of data in support of management's decision making process" according to Inmon (1996). The construction of a data warehouse, which involves data cleaning and data integration, can be viewed as an important preprocessing step for data mining. Moreover, data warehouses provide OLAP tools for the interactive analysis of multi-dimensional data of varied granularities, which facilitate effective data mining. Furthermore, many other data mining functions can be integrated with OLAP operations to enhance the interactive mining of knowledge at multiple levels of abstraction (Han and Kamber, 2001). Hence, the data warehouse has become an increasingly important platform for data analysis and OLAP and will provide an effective platform for data mining.

Leung *et al.* (2003) indicated the effectiveness of the OLAP-based fuzzy-cum-GA approach for use in supporting the optimization of the ion plating process. The multi-dimensional database structure could support the data mining process through the capturing of relevant knowledge for future decision-making. Kaya and Alhajj (2005) used the OLAP for building the fuzzy data cube OLAP architecture which facilitates the effective storage and processing of the state information reported by the system agents. This integrated OLAP mining enhances the power and flexibility of data mining and makes mining an interesting exploratory process. Ho *et al.* (2004) suggested using OLAP as the back-end service for analyzing large volumes of data with high flexibility and performance, detecting opportunities and suggesting business strategies. They also pointed out that OLAP is only able to provide numerical and statistical analysis of data in an efficient and timely way. It lacks a predictive capability for solving complex problems. A certain "ingredient" of intelligence element needs to be added to OLAP to enable the self-learning capability. OLAP therefore was used as the core architecture for handling the data processing in a flexible multi-dimensional way in this research. More AI techniques, however, are required to enhance the prediction capability for solving the process quality problems in the proposed system.

### 2.4.3 Decision Tree Classification

Classification problems aim to identify the characteristics that indicate to which group each case belongs. This pattern can be used both to understand the existing data and to predict how new instances will behave. A decision tree creates classification models by examining already-classified cases and inductively finding a predictive pattern. According to Tan *et al.* (2006), a series of questions, and their possible answers, can be organized in the form of a decision tree, which is a hierarchical structure consisting of nodes and directed:

- A root node that has no incoming edges and zero or more outgoing edges.
- Internal nodes, each of which has exactly one incoming edge and two or more outgoing edges.
- Leaf or terminal nodes, each of which has exactly one incoming edge and no outgoing edges.

In a decision tree, each leaf node is assigned a class label. The non-terminal nodes, which include the root and other internal nodes, contain attribute test conditions to separate records that have different characteristics. Classifying a test record is straightforward once a decision tree has been recorded and follows the appropriate branch based on the outcome of the test. This will lead either to another internal node, for which a new test condition is applied, or a leaf node. The class label associated with the leaf node is then assigned to the record. Decision trees can easily be converted to classification rules by this flow-chart-like structure (Han and Kamber, 2001).

Decision trees have been used in many application areas ranging from medicine to game theory and business (Mitchell, 1997). They are the basis of several commercial ruled induction systems. ID3 (Quinlan 1986), a well-known decision tree induction algorithm is the basic algorithm for decision tree induction. It is a greedy algorithm that constructs decision trees in a top-down recursive divide-and-conquer manner. The algorithm uses an entropy-based measure known as information gain as a heuristic for selecting the attributes that will best separate the samples into individual classes. This attribute becomes the "test" or "decision" attribute at the node. In order to classify an unknown sample, the attributes values of the same are tested against the decision tree. A path is traced from the root to a leaf node that holds the class prediction for that sample.

Guh (2005) and Wang *et al.* (2005) proposed using decision tress for the real-time pattern recognition in SPC and for selecting dispatching rules of a semiconductor final testing factory respectively. The results proved that the decision tree approach is efficient in helping a model to classify the case in the relative situation. First, compared to a neural network or Bayesian classification based approach, a decision tree is easily interpreted and comprehended by humans (Breiman *et al.*, 1984). Second, while training neural networks can take large amounts of time and thousands of iterations, inducing decision trees is efficient and is thus suitable for large training sets. (Rastogi and Shim, 2000) Also, decision tree generation algorithms do not require additional information besides that already contained in the training data (Fayyad, 1991). Finally, as shown by Mitchie *et al.* (1994), decision trees display good classification accuracy compared to other techniques. For inductive learning, decision tree learning is attractive for three reasons (Utgoff and Brodley, 1990):

- A decision tree is a good generalization for unobserved instance, only if the instances are described in terms of features that are correlated with the target concept.
- The methods are efficient in computation that is proportional to the number of observed training instances.
- The resulting decision tree provides a representation of the concept that appeals to humans because it renders the classification process self-evident.

Therefore the decision tree classifier was selected as the major classification approach for the pre-processing of the manufacturing process prediction module. It will now be compared with the Naïve Bayesian classification in the following section.

### 2.4.4 Naïve Bayesian Classification

Bayesian classifiers are statistical classifiers. They can predict class membership probabilities, such as the probability that a given sample belongs to a particular class (Han and Kamber 2001). Bayesian classification is based on

Bayes theorem (Bayes, 1763). The Bayesian theorem is applicable in the case of events with known outcomes whose occurrence depends on factors preceding them. The calculated probability provides a given factor of the cause of the particular event (the effect). Studies comparing classification algorithms have found a simple Bayesian classifier known as the Naïve Bayesian classifier to be comparable in performance with both the decision tree and neural network classifiers. The use of the Naïve Bayesian classifier also requires the probability calculations on the basis of an appropriately prepared training data set, which consists of examples described by means of attributes. These classifiers have also exhibited high accuracy and speed when applied to large databases.

Naïve Bayesian classifiers assume that the effect of an attribute value on a given class is independent of the values of the other attributes. This assumption is called class conditional independence. It is made to simplify the computations involved and, in this sense, is considered “naive” (Tan *et al.*, 2006). In theory Bayesian classifiers have the lowest error rate in comparison with all other classifiers. However, in practice this is not always the case owing to inaccuracies in the assumptions made for its use, such as class conditional independence, and the lack of available probability data (Mitchell, 1997). Marcin *et al.* (2005) compared the Naïve Bayesian classifier and Artificial Neural Networks (ANNs) based on their prediction errors and relative importance factors of input signals and founds that the Naïve Bayesian is more effective in some applications than ANNs. This approach is relatively less popular than the decision tree or neural network based classifiers, however. To find a more effective and appropriate method for the pre-processing, Naïve classification was compared with the performance of decision classification.

### 2.4.5 Artificial Neural Networks

The study of Artificial Neural Networks (ANN) was inspired by attempts to simulate biological neural systems. Analogous to human brain structure, an ANN is composed of an interconnected assembly of nodes and directed links (Tan *et al.*, 2006). According to the DARPA Neural Network Study (1988), the definition of a neural network can be regarded as a system composed of many simple processing elements operating in parallel whose function is determined by network structure, connection strengths, and the processing performed at computing elements or nodes. Not only that, ANN is also a massively parallel distributed processor made up of simple processing units, which has a natural propensity for storing experiential knowledge and making it available for users (Haykin 1999). It resembles the brain in two respects:

- Knowledge is acquired by the network from its environment through a learning process.
- Interneuron connection strengths, known as synaptic weights, are used to store the acquired knowledge.

ANNs are able to learn the relationships between data sets by simply having sample data represented to their input and output layers with the ability to generate relevant data for certain phenomenon based on what the system has been trained on (Haykin, 1999). Currently a wide variety of ANN architectures are being studied in different application areas (Negnevitsky, 2005). Each network has its own structure, training and learning approach. Examples of the architectures include Back-propagation Networks (BPNs), recurrent networks, Hopfield networks, Radial Basis Function (RBF) networks, adaptive resonance theory networks, restrict coulomb energy networks, probabilistic networks, and modular networks, etc. Their most important advantage is in solving problems that are too complex for conventional technologies – problems that do not have an algorithmic solution or for which an algorithmic solution is too complex.

Perceptron is the major element in ANNs. As demonstrated by Minsky and Papert (1969), the single layer artificial neural networks, often called perceptron, were incapable of solving many simple problems. Tan *et al.* (2006) defined that perceptron consists of two types of nodes: input nodes, which are used to represent the input attributes, and an output node, which is used to represent the model output. The nodes in ANN architecture are commonly known as neurons or units. In a perceptron, each input node is connected via a weighted link to the output node. The weighted link is used to emulate the strength of synaptic connection between neurons. By adding one or more hidden layers, the network is enabled to extract higher-order statistics. In a rather loose sense the network acquires a global perspective despite its local connectivity due to the extra set of synaptic connections and the extra dimension of neural interactions (Churchland and Sejnowski, 1992). The ability of hidden neurons to extract higher-order statistics is particularly valuable when the size of the input layer is large.

Referring to the studies from Haykin (1999), error back-propagation learning basically consists of two passes through the different layers of the network: a forward pass and a backward pass. In the forward pass, an activity pattern (input vector) is applied to the sensory nodes of the network, and its effect propagates through the network layer by layer. Finally, a set of outputs is produced as the actual response of the network. During the forward pass the synaptic weights of the networks are all fixed. During the backward pass on the other hand, the synaptic weights are all adjusted in accordance with an error-correction rule. Specifically, the actual response of the network is subtracted from a desired (target) response to produce an error signal. This error signal is then propagated backward through the network, against the direction of synaptic connections – hence the name “error back-propagation”. The synaptic weights are adjusted to make the actual response of the network move closer to the desired response in a statistical sense. It is often said that the design of a neural network using the back-propagation algorithm is more of an art than a science in the sense that many of the numerous factors involved in the design are the results of one’s own personal experience. Nevertheless, there are methods that will significantly improve the back-propagation algorithm’s performance, and these are described below:

- Sequential versus batch update
- Maximizing information content
- Activation function
- Target values
- Normalizing the inputs
- Initialization

- Learning from hints
- Learning rates

As for the training of ANNs, a set of input-output pairs, with each pair consisting of an input signal and the corresponding desired response, is referred to a set of training data or training sample. The property that is of primary significance for a neural network is the ability of the network to learn from its environment and to improve its performance through learning. A neural network learns about its environment through an interactive process of adjustments applied to its synaptic weights and bias levels. Ideally, the network becomes more knowledgeable about its environment after each iteration of the learning process (Haykin 1999). According to Mendel and McLaren (1970), learning is a process by which the free parameters of a neutral network are adapted through a process of simulation by the environment in which the network is embedded. The type of learning is determined by the manner in which the parameter changes take place. The learning process may also be viewed as a “curve-fitting” problem. The network itself may be considered simply as a nonlinear input-output mapping (Negnevitsky, 2004). The network performs useful interpolation primarily because multilayer perceptrons with continuous activation functions lead to output functions that are also continuous.

The use of ANNs is strongly recommended in the industrial areas for forecasting and prediction, which recognize and evaluate the potential failure of a product before production. For process modeling, neural networks have been used to model complicated manufacturing processes, such as the autoclave curing of composites, Low-Pressure Chemical Vapor Deposition (LPCVD), plasma etching, selective laser sintering and others (Joseph *et al.*, 1992; Himmel *et al.* 1993; Han *et al.* 1994; Rietman *et al.*, 1996, Boillat *et al.*, 2004). These researchers show that the ANNs exhibit superior accuracy over the statistical models and also required fewer training experiments. In recent years ANN's have emerged as a promising tool in the modeling, optimization and control of manufacturing processes (Narendra and Parthasarathy, 1990; Nugyen and Widrow, 1990; Fan and Wu, 1992; Hong and May, 2004). Research by Kim *et al.* (2003) constructed the predictive model for plasma etching processes by ANNs. The high prediction ability shown the possibility to model and control the complex plasma processes which gave a remarkable perception for the major case study in this research – Reactive Ion Etching (RIE). Several active researchers found out that BPNs with multi-hidden nodes were capable of modeling the control design and prediction of manufacturing processes (Haouant *et al.*, 2000 and Liu *et al.*, 2006).

### 2.4.6 Hybrid Intelligent Approaches

The above studies have demonstrated the ability of the ANN to model a variety of processes. However, the need for process optimization, an important consideration in manufacturing, has not been fully investigated. An intelligent system with one single approach is found to be neither so accurate nor flexible enough for solving some complex problems under the dynamic situations. More and more researches into hybrid intelligent systems has been introduced in the recent years.

Kim and Kim (2005) used the hybrid approach of GA-optimized back propagation neural network to construct the plasma etch process by optimizing multi-parameterized neuron gradients in genetic algorithm. “Although first principle models are beneficial to understanding detailed plasma dynamics, they are subject to many simplifying assumptions due to the lack of understanding of the physical and chemical processes along with insufficient diagnostic data.” The results therefore frequently show a large different between the predictions and actual measurements. It is common that all neurons in either the hidden or the output layer are of the same value. It was proved that better predictive models could be achieved once the gradients (or temperature) were assigned differently to each neuron. The gradient effect in the model was optimized by means of a genetic algorithm for a significant prediction performance improvement. Guh (2005) proposed the hybrid systems integrating the neural network and the decision tree-based approach for the real-time pattern recognition in statistical process control. The neutral network and decision tree are for coarse and fine classification of recognition respectively. The paper proved this integrated approach can solve the problem of false recognition and increase the accuracy of control chart pattern classification in different situations. It was proposed that the integrated approach of the decision tree-based classification and the BPNs in ANNs be used in this research.

## 2.5 Successful Applications in Industry

Several case studies were conducted by different solutions providers, including Microsoft Corporation and SAS Institute Inc. This is the best and only possible way for academic researchers can get to know the actual applications in industry nowadays and how the industry applies the knowledge and methodology in a systematic way through the commercial tools in large scale implementation. Three studies from the popular Hard Disk Drive (HDD) and wafer fabrication / semiconductor manufacturers were conducted.

The production system or Factory Information System (FIS) is one of the most important elements in electronic component manufacturering. According to the case study conducted at Western Digital (Microsoft, 2003), it will cause production error and some failures result in significant production loss if there is a problem in the enterprise database, for example, slow response times. The powerful central data warehouse was required to handle the peak transaction rate for database connection up-to 12,000 per minute in Western Digital. This study showed a 100% uptime and return of investment after implementing the new central data warehouse for six months. The application was just the beginning and data analysis for process optimization and quality improvement could be started after collecting a huge amount of valuable production data.

The intelligent system developed by IBM Storage (now called Hitachi Global Storage Technology after August 2002) brought advanced technology to actual application in the HDD industry (Rutledge, 2000). The data warehouse combines operational data from the disk, wafer and head stack. Analyses, such as yield analysis and yield sensitivity study can be performed by OLAP and data mining algorithms including decision trees and neural networks. The system allows engineers to quickly identify the factors causing yield loss, evaluate the cost / benefit of proposed changes, and operate their plants with optimal processes and specifications. Although the concept of Just Enough Database (JED) can gain benefits by fast and effective yield analysis, some hidden sources of failure may not be able to be discovered and the database structure is not flexible enough for a changing environment involving new processes and testing parameters.

Another successful implementation from Motorola in semiconductor production, Engineering Data Analysis System (EDAS), also showed the possibility of analyzing a large amount of engineering data from all the global manufacturing sites (Whitney and Fowler, 2000). Any engineering and final test database anywhere in the world can be accessed remotely through the web-based system. Several popular analyses, such as Analysis of Variance (ANOVA), correlation analysis, descriptive statistics, probability tests, regression analysis and T-test, can be performed. These three cases provide helpful reference and usable knowledge for the Information Discovery System (IDS) development in real industry application.

## 2.6 Discussions

Recent studies on the relevant topics indicates that whilst a number of approaches and expert systems were designed and implemented, this is still an area that requires more in-depth study and investigation for the fabrication / HDD component industry. This research focuses on the combination of two computational intelligence techniques, OLAP and ANNs, to form an integrated model for solving quality problems, not only in the experimental or laboratory environment, but also modeling for the mass production environment in an actual manufacturing company. That is our research target. Both decision tree and Naïve Bayesian classification were included to formulate the hybrid intelligent systems and evaluate the accuracy and efficiency by comparing the prediction performance with the traditional methodology.

# CHAPTER 3. SYSTEM DESIGN

## 3.1 Introduction

The system design is divided into eight sections. It begins with the architecture framework of the proposed Information Discovery System (IDS) which describes the overall research approach and methodology, then follows with the common problems in database system design. After that, the four modules inside the IDS design are discussed; they are the Object-technology Database Module, the Multi-dimensional Database Module, the Process Intelligence Module and the Web Reporting Module. There is also a summary section at the end of this chapter.

The overall research approach is using the IDS with the case study of slider fabrication and RIE in SAE Magnetics (H.K.) Ltd. to validate the proposed methodology. All the modules of IDS are integrated and cooperated together to form a platform for aiding the users to solve the quality problem and achieve process optimization. Examples of the source of quality defect identification and etching rate prediction in RIE are discussed in Chapter 4. Appropriate production data is used to validate the IDS, the optimized result is then compared with the result from traditional approach and actual production performance.

The original idea of developing the data warehouse in SAE Magnetics (H.K.) Ltd. was proposed by the user department and top management of SAE in 2000. The first prototype was failed; the company then invited the academic side to cooperate for this project. The research student with the supervisors from the university side proposed the new approach of IDS to the company in 2003. There are more than 10 members from IT team to work together for the whole development. The research student acted as the roles of a system architect who proposed the IDS, a business analyst between IT team and user departments who collected and analyzed the business logic & user requirement, and a developer of the major core of IDS in Neural-OLAP engine and ANN models.

## 3.2 Architecture Framework of IDS

A long-term strategy for quality improvement is not well planned in most of the industries nowadays. Several large standalone databases are kept for individual purposes including incoming quality control, product performance, quality assurance, finished good and material tracking, process monitoring, enterprise resource planning, etc, under different database platforms in different locations. The sources of any quality problems are found difficult to identify because of a lack of supporting information.

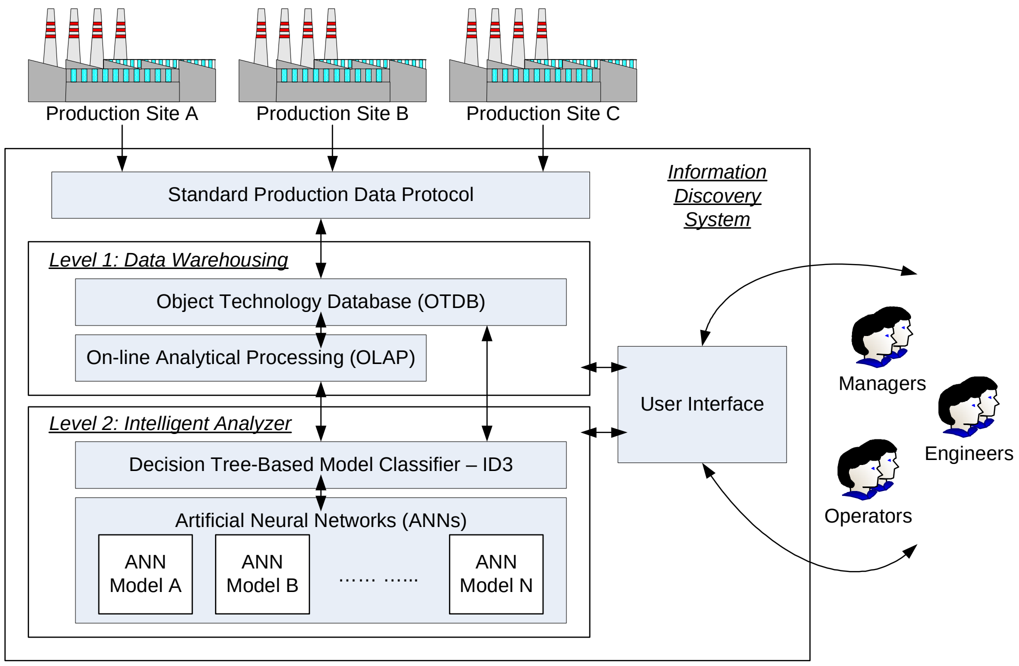



Figure 3.1 Infrastructure of IDS

The infrastructure of IDS was based on the concept of a central data warehouse (Figure 3.1) which collects all the relevant data from the local and remote factories. There are two levels within the IDS: the data warehousing and the intelligent analyzer, which are integrated as the core of Neural-OLAP for the OLAP cube construction and process configuration prediction respectively. The standard production data protocol was proposed to be used to collect all the data from different production sites. Production data is first collected in this standardized format and is then inserted as the object in the array table of the Object Technology Database (OTDB).

OLAP helps the data mining engine to handle the multi-dimensional cubes and process a huge amount of data up to several gigabytes in timely manner. Analysis from OLAP generates a series of hypothetical patterns and relationships and uses queries against the database to verify or disprove them. The data mining module then uses the data itself to uncover those patterns. OLAP cubes are then built based on the pre-defined structure with the multi-dimensional data from the array table. Summary and index can be processed at the same time speeding up the query of the data warehouse. Users can access the user interface to explore the OLAP cubes. Selected critical production raw data and summarized data are used in the intelligent analyzer for specific process classification and prediction. This OLAP system also has a strong focus on the interactive analysis of data and typically provides extensive capabilities for visualizing the data and generating summary statistics for ad-hoc quality failure analysis. Those important summary data in slider fabrication is transferred from OTDB to OLAP for aggregation.

Since one single neural network may not be able to fulfill the needs and conditions of process prediction, a certain number of neural networks are therefore built based on the specific situation and environment for that prediction. By finding the most accurate and reliable network model, the decision tree-based classification is involved in classifying the case to the most suitable predictive network. The hybrid Intelligent system by combining the decision tree and ANNs, which was proposed in the data mining engine in IDS, is then able to help and provide a more accurate and flexible solution for some complex problems under the dynamic situations especially the etching rate prediction in RIE.

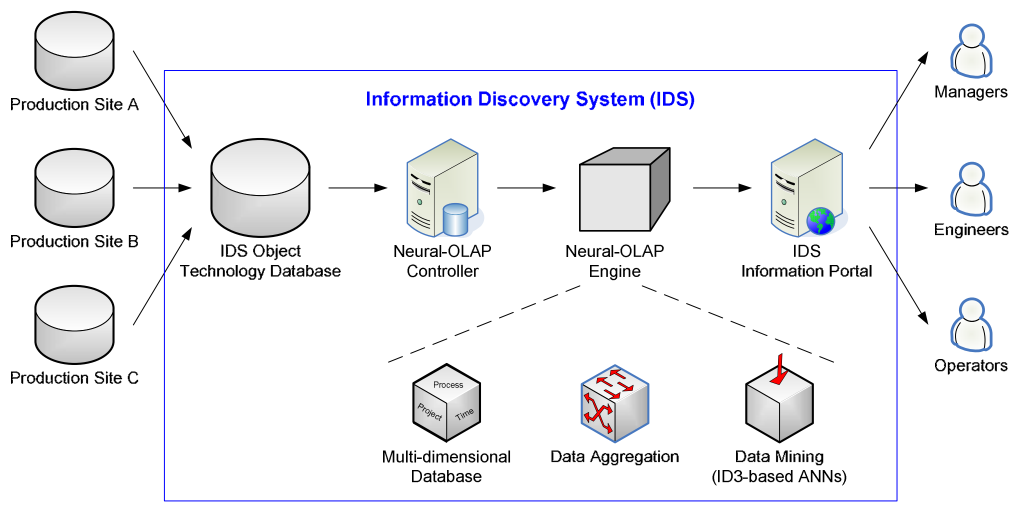



Figure 3.2 IDS Modules

From the system design perspective, the system was developed into several modules including the IDS object technology database, Neural-OLAP controller, Neural-OLAP engine and IDS information portal (Figure 3.2). All production related databases were connected with the central database, and the data is regularly replicated to this database in specific standardized format. A system management module, called the Neural-OLAP controller, was introduced that aimed to schedule and activate the engine for data cleaning and cube processing. The innovative use of OLAP technology could be applied in the real time quality monitoring. Whilst the quality information is stored in the data warehouse, a further step is to conduct data analysis by using the prediction module. The Neural-OLAP engine is the solution by creating the OLAP cubes as well as summarizing the huge amount of production data for intelligent forecasting in the data mining module. Reporting in order to alert or assist quality or process engineers in solving quality problems can be accessed through the web-based information portal. Not only operators, but also engineers and managers are able to get the latest information and suggested corrective actions in terms of optimized control parameters from the system anytime anywhere.

There are no limitations for the database system platform; the system design was based on the common platforms for general purposes available in the market. The database can be replaced by Oracle, DB2, SQL Sever, MySQL, etc. However, some minor modifications may be required during the actual implementation.

## 3.3 Common Problems and Needs in Database System Design

A Factory Information System (FIS) plays an essential role as collector and storer of all the production information of an enterprise. Due to the increasing amount of data, together with the need for data integrity and data traceability, the traditional relation database system may not be able to fulfill these roles in the near future, so a more flexible and reliable Database Management System (DBMS) is necessary.

By using the example of the production system database in the research cooperated company, SAE Magnetics (H.K.) Ltd., the traditional one had been in use as the data storage for many years and focused mainly on the storage of raw production data. Compared with today's more technical point of view, only low-level features, similar to those provided from a spreadsheet, were being used. A lot of potential functionalities had not been discovered yet. And the database was neither well-designed nor well-structured, which further limited the possibility of development. Therefore, the first step was to reconsider the database structure and data collection mechanism. By reviewing and evaluating the existing database system, some critical problems were discovered that might be limiting the data usage and affecting the system development. Figure 3.3 is a list of the existing tables and the related fields in the database system. The following examples of table structure and table relationship are used to highlight the problems:

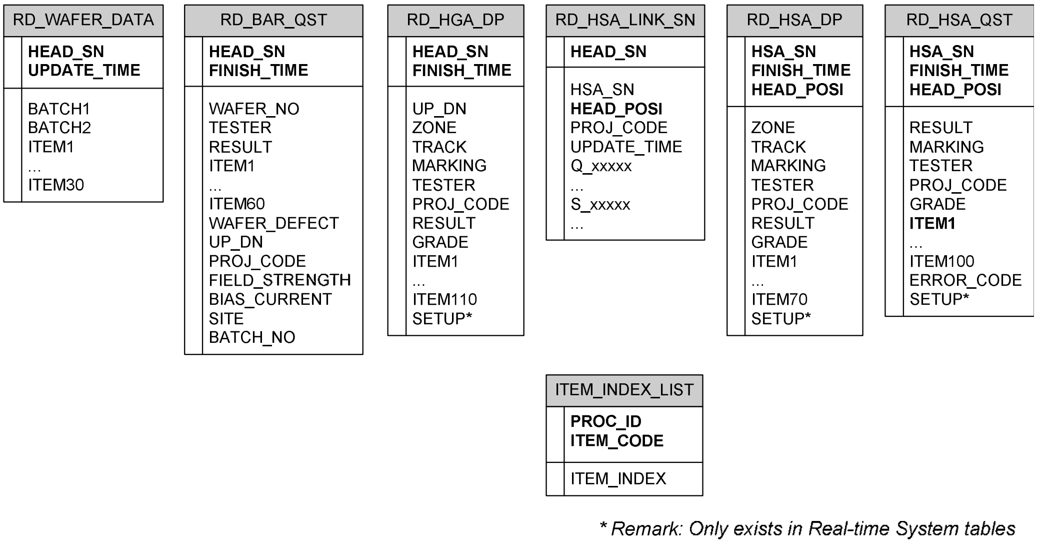



Figure 3.3 Existing OCR Database Structure in SAE Magnetics

## Table Structure

(A) No Primary Keys (No Traceability)

Although there were HEAD_SN or HSA_SN + HEAD_POSI as the keys of each record imported to the tables, they were both non-unique, which meant that there were no primary keys at most tables. An example was that retest and rework data were stored in the same table, this being a serious problem for data traceability.

(B) Fixed Table Structure (No Extendibility)

Since there were lots of different projects working on the production shop floor at the same time, the number of inspection data (fields of data) required was different depending on the projects or its status (e.g. Prototype / Mass Production / Rework). The existing solution was to use ITEMs for storing the extra data required. However, table definitions need to be changed if additional ITEMs are required.

(C) Empty Fields (Over Estimate)

The number of ITEMs varied from table to table; the minimum and maximum number of fields were 30 to 110 respectively (There were more than 110 in some special cases), which were over-estimated. Nearly two-thirds of ITEMs were empty thereby reducing the performance and wasting the computer resources.

**Table Relationship**

(D) No Relationship between Tables (No Traceability)

A table relationship was in fact built by connecting the keys from different tables. No relationship could be defined if there were neither primary keys nor composite keys. The functionality of the database cannot be used effectively, or it is used as a spreadsheet instead of a database.

(E) Data Redundant

A data redundant problem did exist in the current database, which led to a decrease in the speed and an increase in the time spent on data query. Fields starting with “Q_” and “S_” in table RD_HSA_LINK_SN were copied from table RD_HSA_QST and RD_HSA_DP respectively. Another example was WAFER_NO which is the first five characters of HEAD_SN.

**Others**

(F) Data Integrity

Data integrity relied on the data collected from different sources. In fact, errors did exist in all individual components. The analysis result was not reliable if the error data was used. The best solution was to eliminate or ignore those data which are not traceable. Only data, which could be traced, was used as the reference for data analysis.

Other than the above problems, the major reason for building a new database is the needs of data analysis. In fact, data analysis, especially the data mining technique is the key for extracting the hidden predictive information from a large database, which requires a well-prepared foundation, including the well-structured relational and OLAP / Multi-dimensional Database (MDDB). A new DBMS was designed based on the below three rules below, which were aimed at solving the existing problems in the database and providing a well-equipped foundation for the data modeling:

1. Traceable
2. Flexible
3. Extendable

By using the concept of Relational Database (RDB) and the techniques from Object Orientation (OO), a new generic database model was designed with extendable modules, which means that more functionalities and further applications can be added to this core system. This new design allows users to add, remove and modify the data fields easily without changing the database structure. The build-in data relationship can also increase the traceability of data. Furthermore, in the future developments – data analysis can be constructed on this database foundation.

### 3.4 Object-technology Database Module

The new database design was based on the concept of RDB combining with the techniques used in Object Orientation (OO). There are two parts, core and sub-module, in the proposed DBMS. The core contains the most important component, Data collection / data director, which is used to collect and manage all engineering data. The data director is the tool to put the right data into the right table in the right field. In considering the future extension, the design includes modules for adding new features and applications based on the system foundation. There are four packages suggested for the engineering data warehouse, including process flow / process director, business rule, table generator and adaptor (Figure 3.4). More modules can be added and developed based on this architecture.

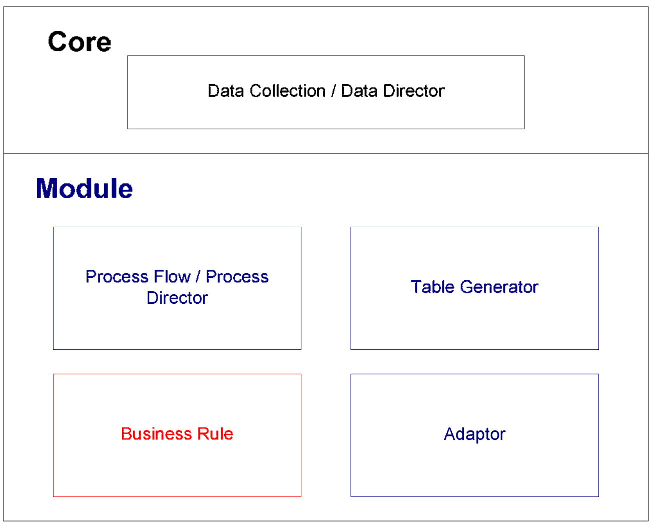
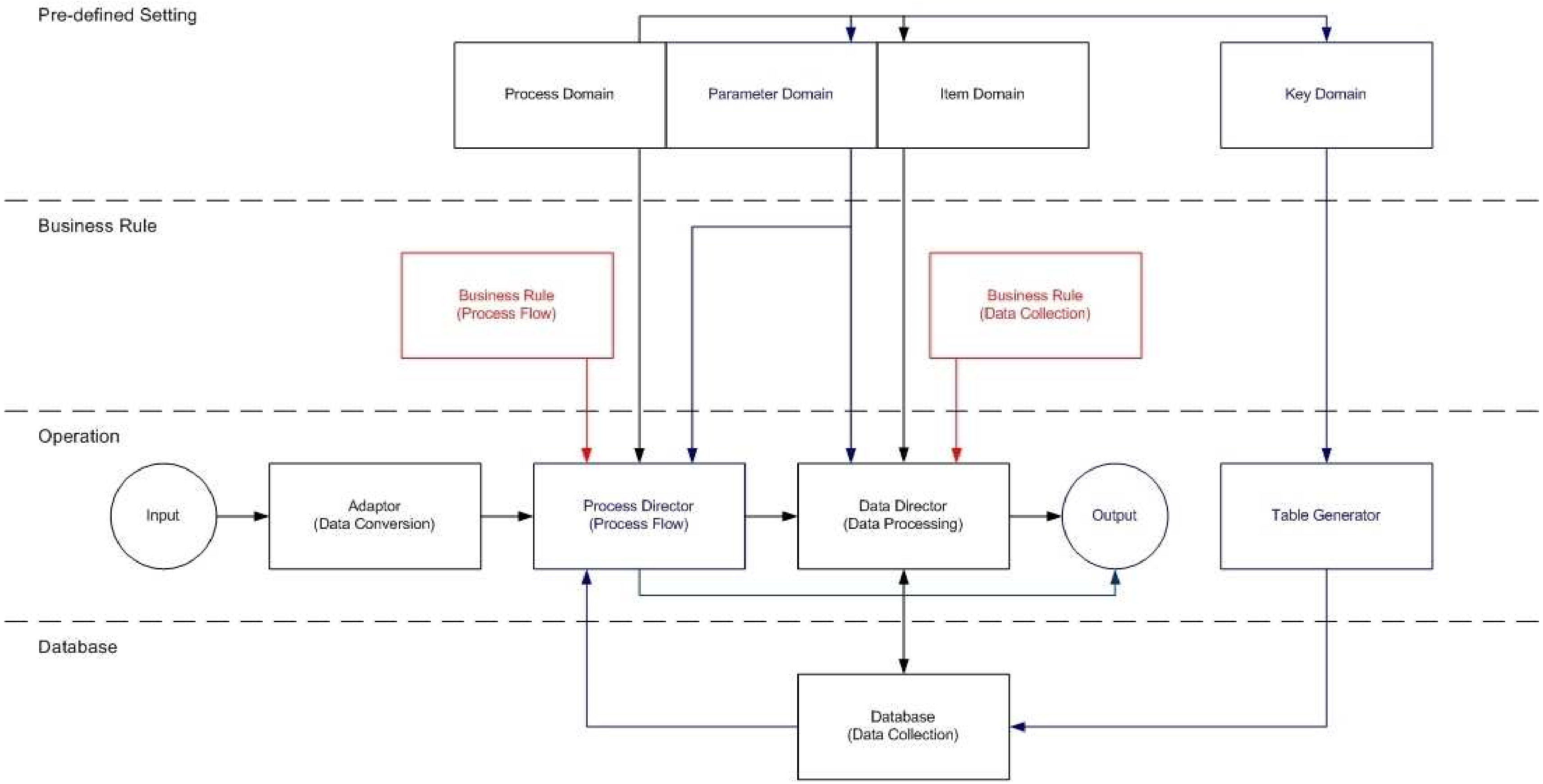



Figure 3.4 Database Management System (DBMS) Components

Detail design of the DBMS architecture is presented in the following two diagrams, the operation flow diagram and the Entity Relationship Diagram (ERD). Figure 3.5 shows the four layers operation flow of the new DBMS, including pre-defined setting layer, business rule layer, operation layer and database layer. Development is also divided into three different phases, Phase 1 – OCR Database System (OCR is the internal database in the cooperated company), Phase 2 – Process Flow System and Phase 3 – Business Rule System. Figure 3.6 is the Entity Relationship Diagram (ERD) of the new database system, which is a generic design for the FIS used in data collection. By reducing the complexity, the attributes will not be listed in all ERDs in this proposal. As for the details of attributes, please refer to the Database Model Diagram (DMD).




**Phase 1 - OCR Database System**

**Phase 2 - Process Flow System**

**Phase 3 - Business Rule System**

Figure 3.5 Four layer Operation Flow of New DBMS

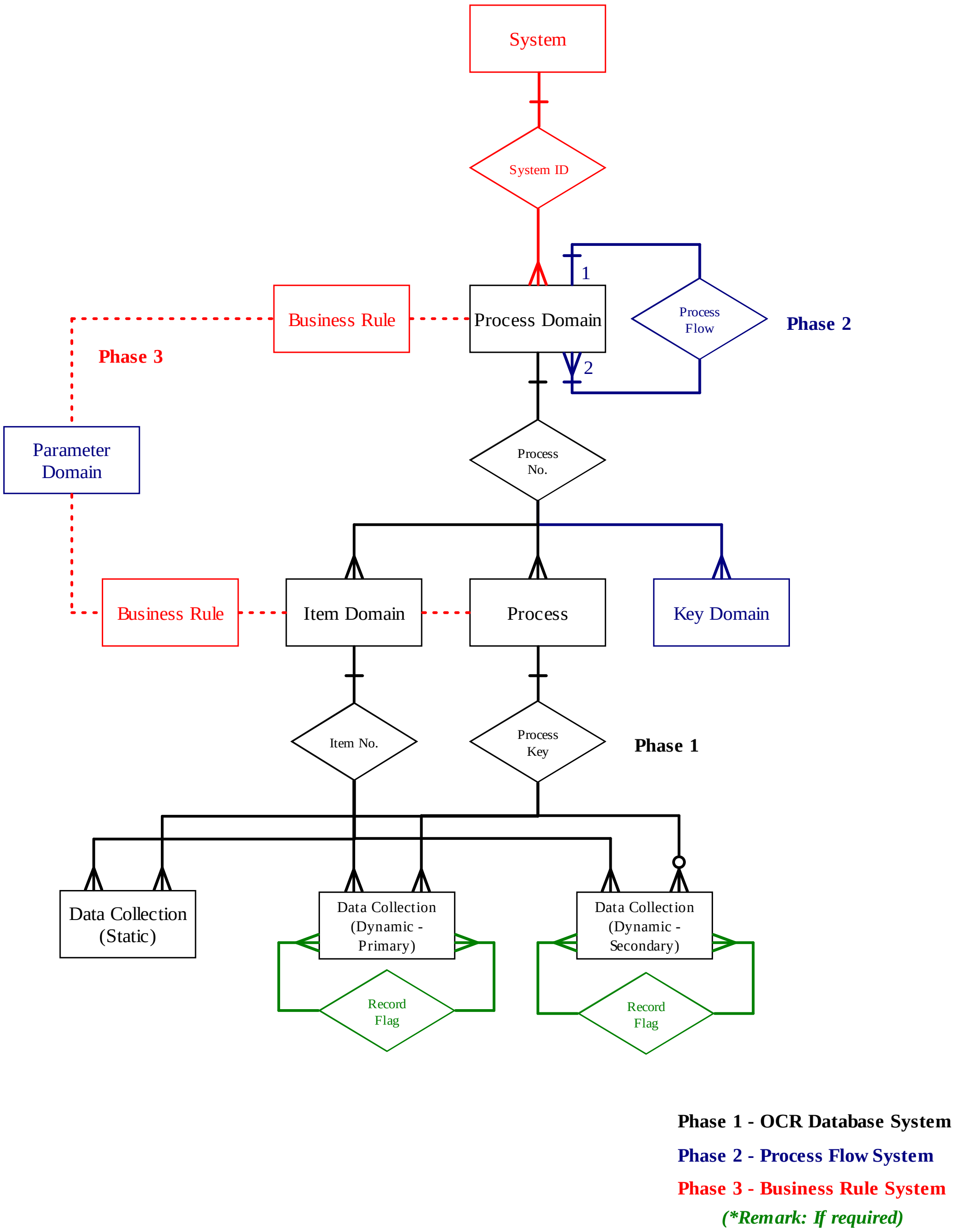



Figure 3.6 ERD of New Database System (Data Collection Only)

The "Business Rule" of the proposed system architecture will not be implemented in this research project. Parts of the business rules, such as the pre-defined parameters / settings, were "hard-coded". Detailed system process mechanisms and settings are discussed in the case study chapter.

### 3.4.1 Real-time Three-Tier Data Collection Architecture

Since the number of steps and the amount of data are both relatively very large, inspection data located in different places vary from the text files in the testing machines or standalone databases (e.g. Microsoft Access or Common Separated Variable (CSV)). It was suggested, therefore, that a generic standard production data protocol be developed. By using the Distributed Component Object Model (DCOM) with the 3-tier architecture design, the data was sent to the central database in real time environment (Figure 3.7). The DCOM is responsible for distributing the data conversion tasks to the idle data adaptor for processing and uploading data to the object technology enabled central database guided by the domain tables. Production data could be found and forecasted in the IDS within a few seconds after the process had just been completed.

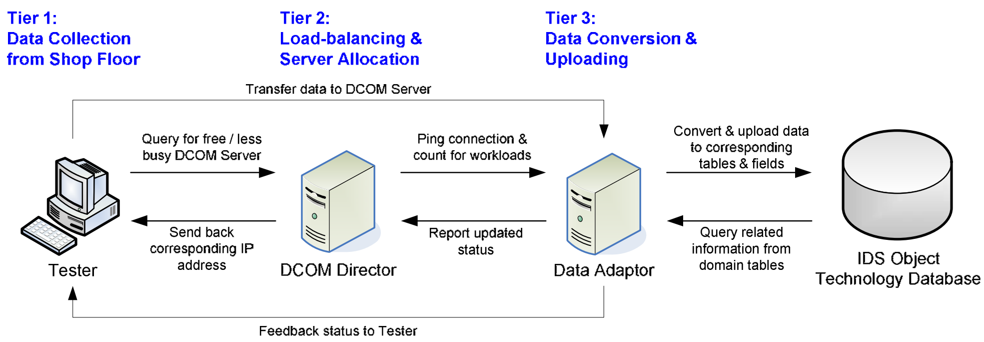



Figure 3.7 Three-tier DCOM Data Collection Architecture

The major role of the first tier on the production shop floor is data collection directly from the testing machines to the next tier. It queries the free or less busy DCOM servers on the second tier, and the DCOM director on this tier then acts as the load-balance checker that looking at the workload of each data adaptors on the third tier and reporting his findings to the tester with the corresponding IP address of the available data adaptor. After that the direct connection between the tester and the data adaptor is created and the data is transferred. There is a feedback signal from the data adaptor back to the tester to notify him that the process has been completed. The data adaptor has another function and that is to convert the data into specific standardized format and upload it to the corresponding databases and tables.

### 3.4.2 Relational Database by Using Object Technology

Product testing is the key to measuring the quality of finished goods in a manufacturing environment. However, these inspections are based on the groups of products, the incoming materials, the types of production, the process operation flows, the testing machines, the testing approaches and customer requirements, all of which are difficult to be fixed or standardized, especially in the electronic industry. What's more the traditional relational database is not able to handle this dynamic changing environment which requires that database definitions should be changed when adding, modifying and deleting fields in the tables. All of which create a serious problem for data analysis.

| Key(s) | | | Fixed Fields | | | | | Variable |
|---|---|---|---|---|---|---|---|---|
| *WAFER* | *BK_N* | *ROW* | *TESTER* | *PROJ_CO* | *LOT_NO* | *MKG* | *ZONE* | … … |
| 504H6 | C2 | KA | 1D#2E | TSK740 | D2027-434 | MP | ID | |

| Array Table | |
|---|---|
| *ITEM* | *VALUE* |
| MRR_ | 40.53 |
| TAAA | 20.30 |
| PW50 | 1.056 |
| ASYM | 0.4576 |
| HLA_ | -0.967 |
| … | … |

Figure 3.8 Table Structure of Object Technology Database

Object technology was implemented in the relational database to eliminate this barrier. This new mechanism put inspection items as an object in the table array, thereby allowing users to freely change the number of items (Figure 3.8 shows the example in slider fabrication). In addition, the OLAP integration was also applied so that cubes could be constructed by using the dynamic testing parameters automatically. Several domain tables were stored in the database to keep the field definitions and their relationships with the array tables. The systems therefore can identify the table structure and construct the OLAP cubes on-the-fly.

## 3.5 Multi-dimensional Database Module

The multi-dimensional database module acts as the essential storage for the production raw data in a well-structured multi-dimensional database hierarchy and turns those data into meaningful and helpful information. The central data warehouse and OLAP are the sub-modules that handle the data storage and data processing respectively.

### 3.5.1 Central Data Warehouse

According to Connoly and Begg (2005), the data warehouse data flows have these below five catalogues:

- Inflow: The processes associated with the extraction, cleansing, and loading of the data from the source systems into the data warehouse.
- Upflow: The processes associated with adding value to the data in the warehouse through summarizing, packaging, and distribution of the data.
- Downflow: The processes associated with archiving and the backup of data in the warehouse.

- Outflow: The processes associated with making the data available to the end-users.
- Metaflow: The processes associated with the management of the metadata.

The inflow is the one related to the data capture which, in the case of IDS, is handled by the previous module, OTDB. The central data warehouse collects the data directly from the OTDB through the DBLink, the direct connection between two databases in Oracle.

The upflow is the central warehouse and OLAP together which process and summarize the data and keep it in the database in a multi-dimensional structure. The down flow of archive and the backup of data is also handled by the database through the application software, Oracle Backup, and the physical archive tape device. The IDS information portal, which is discussed in the following chapter, is the user interface that makes the data available to the end-users - including operators, engineers and managers - as outflow. The metaflow is also handled by the database management software, Oracle Enterprise Manager, for the management of data and its processes.

The IDS central data warehouse design also fulfills the requirements for data warehouse DBMS (Connoly and Begg, 2005) so that are all the essential elements that a good data warehouse should include are found in it. These ten major elements are listed below:

1. Load performance

2. Load processing
3. Data quality management
4. Query performance
5. Terabyte scalability
6. Mass user scalability
7. Networked data warehouse
8. Warehouse administration
9. Integrated dimensional analysis
10. Advanced query functionality

### 3.5.2 On-Line Analytical Processing

The essential characteristic of the OLAP technique facilitates the timely access and manipulation of the quality data with the functionality drilling down into further information. It performs a numerical and statistical analysis of data in multi-dimension. A control chart for keeping a continuous record of a particular quality characteristic can benefit by using OLAP. Engineers can easily view the quality information based on the time-to-time, piece-to-piece and within-piece variation.

An example of the wafer thickness measurement at slider fabrication in Figure 3.9 shows the cube structure, the hierarchy like how the relationship between the wafer and row bar is created in the dimension and fact tables with the data queried from the OTDB. The required inspection items from the array table are inserted into the fact table in new columns based on the pre-defined settings by engineers. The aggregation in case of the summary of wafer thickness can be processed, which is the important information for the etching depth in RIE prediction.

| Wafer_Dim | | |
|---|---|---|
| JIG | WAFER | ... |
| J0768D | 504H6 | ... |
| J0768D | 51073 | ... |
| ... | ... | ... |

| Row_Bar_Fact | | | | | |
|---|---|---|---|---|---|
| WAFER | BK_N | ROW | MEASURE_A | THICK_A | ... |
| 504H6 | B | 013 | 1D#2E | 0.1273 | ... |
| 504H6 | B | 013 | 1D#2E | 0.1275 | ... |
| 504H6 | B | 013 | 1D#2E | 0.1269 | ... |
| 504H6 | B | 013 | 1D#2E | 0.1270 | ... |
| ... | ... | ... | ... | ... | ... |

| Wafer_Thick_Summary | | |
|---|---|---|
| JIG | WAFER | M_THICK_A |
| J0768D | 504H6 | 0.127175 |
| ... | ... | ... |

Figure 3.9 OLAP Cube Structure

The OLAP is used to build all other tables and cubes of the database in a multi-dimensional hierarchy with the corresponding pre-defined aggregations, such as sum, mean, standard deviation / sigma, process capability (Cpk), percentage related to certain group of data, etc.

## 3.6 Process Intelligence Module

The process intelligence module includes the Neural-OLAP prediction engine which includes a controller to manage and process the system operation in a systemic way instead of ad-hoc analysis thereby increasing the efficiency for large amounts of data prediction. Two prediction methodologies are discussed in this section: decision tree-based and Naïve Bayesian-based Artificial Neural Networks.

### 3.6.1 Neural-OLAP Prediction Engine and Controller

A hybrid intelligence system, the Neural-OLAP engine, was designed as the heart of IDS. Figure 3.10 represents the functionality of this engine within the whole system. The system flow design is constructed on the traditional SPC approach (Oakland, 1996) by first identifying the quality problem and possible sources of related data. The Neural-OLAP controller was then responsible for controlling the engine processing.

The system first checks whether the new problem is in a known problem area or not. If there is an existing one, it checks whether the data and information exist in the tables of the central database. Otherwise the data collection process is started. If the problem is new and there is not a suitable reference, it needs experts to help define and evaluate the process before putting it into IDS. Once all the related data is available, the engine starts the trigger for the process including the cube processing activation, engine monitoring, load-balancing and task distribution. However, if either the cube or the ANN model does not exist, the rule building is started instead by defining the cube structure, the ANNs model, the cleaning rules, and the task schedules, etc.

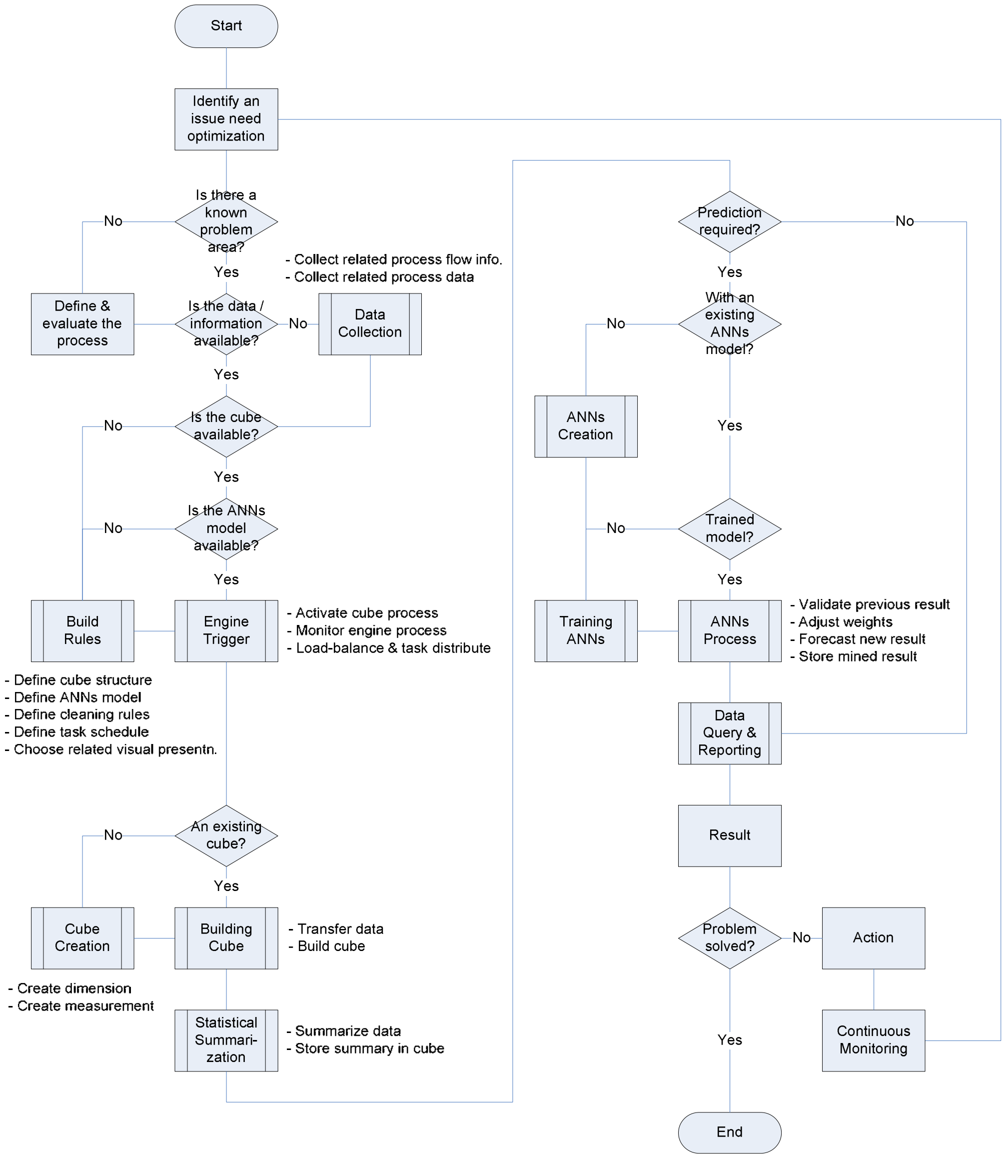



Figure 3.10 System Flow Chart of IDS

The OLAP cubes are built by using the production data collected in the central database based on the predefined rules from the Neural-OLAP controller. Statistical summarization is then processed for calculating those quality related factors. The data is transferred through scheduled tasks from OTDB to OLAP into multi-dimensional structure or by creating its dimensions and measurements if a new cube is applied. Aggregation such as summarization of the wafer thickness is then started and the summarized data is stored in the cube. Some of the raw and summarized data are used directly in the presentation to users, but other important information may need further analysis by the prediction module before being provided to users.

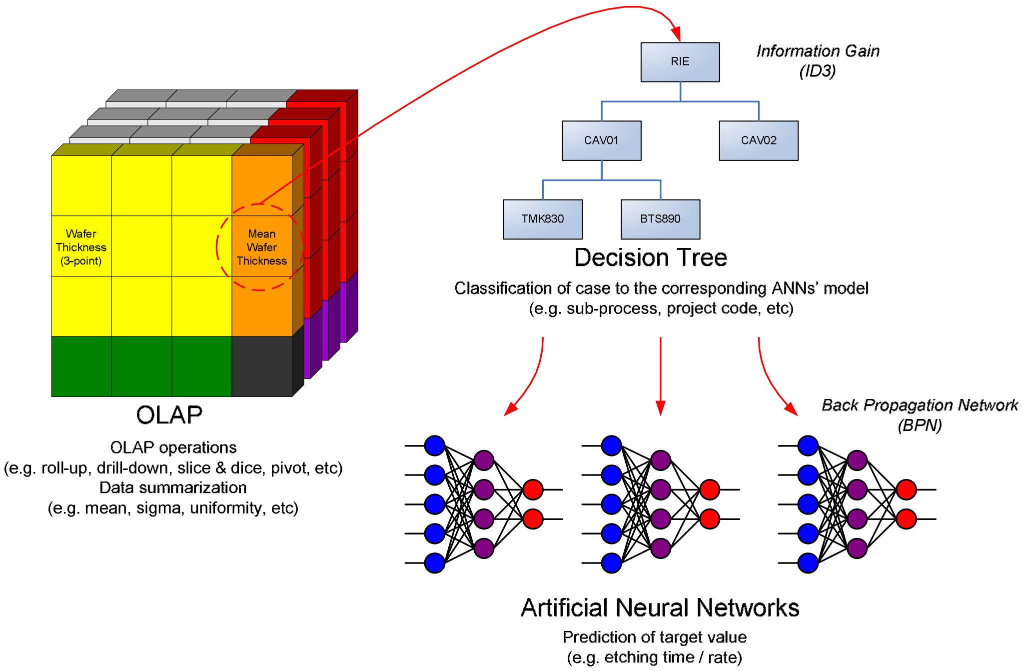



Figure 3.11 Methodology of Neural-OLAP Engine

Similar to OLAP, new ANNs model may be required to be created. Model training by the suitable amount of corresponding historical data may also be needed. The Back Propagation Network (BPN) was used for the ANNs in the IDS thus gaining benefits in addressing any problems requiring the recognition of complex patterns and performing nontrivial mapping functions. It was designed to operate as a multilayer, feed forward network, using the supervised mode of learning. The proactive quality prediction can be conducted by using the ANNs to learn complex quality data log obtained in each process adaptively without any formulation of the casual relationship between the input and output patterns. The decision tree-based (or Naïve Bayesian-based) classification methodology helps the users to investigate the best suitable ANN model for the corresponding case according to the background information and the historical data provided. A more appropriate and accurate predictive model can be suggested. The decision tree-based ANNs take part in the prediction of the possible outcomes according to the current inputs and previous reference data as shown in Figure 3.11. The process performance and trend can be forecasted combining with the summarized data. The model and its weights are tuned and adjusted respectively every time to find out a more optimized and reliable result for the decision support of process improvement. Details of the ANNs process mechanism are discussed in the following sections of the decision tree-based and Naïve Bayesian-based ANNs approaches. The ANNs prediction then is processed including these below four steps:

1. Validation of previous result
2. Weight adjustment
3. New result forecasting
4. Storing of mined information / data

Users can now query the system and get the required data and information in the specific presentation and reporting formats. The Neural-OLAP engine is a closed-loop system that validates the model and continuously monitors each individual cube and model to see whether they are still suitable and accurate enough for the related problems to achieve process optimization.

### 3.6.2 Decision Tree-based ANNs Approach

Each ANNs' model in the IDS was designed and trained for a specific case only with the related conditions to increase the accuracy of prediction.

Therefore, more than one model was required to reflect the actual situation in a single process. On the basis of the discovered error, the most appropriate model would be suggested by using the information gained from the decision tree. The basic algorithm for the decision tree is a greedy algorithm that constructs decision trees in a top-down recursive divide-and-conquer manner. The advantages of the decision tree are the ease with which it can be understood through its flow-chart-like tree structure, and its ability to handle missing class labels.

The ID3 version is used in this research. The basic strategy of the decision tree according to Han and Kamber (2001) is as follows:

- The tree starts as a single node representing the training samples (step 1).
- If the samples are all of the same class, then the node becomes a leaf and is labeled with that class (steps 2 and 3).
- Otherwise, the algorithm uses an entropy-based measure known as information gain as a heuristic for selecting the attribute that will best separate the samples into individual classes (step 6). This attribute becomes the “test” of “decision” attribute at the node (step 7). In this version of the algorithm, all attributes are categorical, that is, discrete-valued. Continuous-valued attributes must be discretized.
- A branch is created for each known value of the test attribute, and the samples are partitioned accordingly (steps 8-10).
- The algorithm uses the same process recursively to form a decision tree for the samples at each partition. Once an attribute has occurred

at a node, it need not be considered on any of the node's descendents (step 13).

- The recursive partitioning stops only when any one of the following conditions is true:
    - o All samples for a given node belong to the same class (step 2 and 3), or
    - o There are no remaining attributes on which the samples may be further partitioned (step 4). In this case, majority voting is employed (step 5). This involves converting the given node into a leaf and labeling it with the class which is in the majority among the samples. Alternatively, the class distribution of the node samples may be stored.
    - o There are no samples for the branch test-attribute = $a_i$ (step 11). In this case, a leaf is created of the majority class among the samples (step 12).

Assuming there are two classes, *P* and *N*, and let the set of examples *S* contain *p* elements of class *P* and *n* elements of class *N*, the equation of information gain is defined as (1):

$$I(p,n) = -\frac{p}{p+n}\log_2\frac{p}{p+n} - \frac{n}{p+n}\log_2\frac{n}{p+n} \quad (1)$$

The equation for the entropy *E* (2) with the attribute *A* (3):

$$E(A) = \sum_{i=1}^{v}\frac{p_i+n_i}{p+n}I(p_i+n_i) \quad (2)$$

$$A \in \{S_1, S_2, \ldots, S_v\} \quad (3)$$

Where $S_1$ contains $p_i$ examples of $P$ and $n_i$ example of $N$.

The encoding information that would be gained by branching of $A$ (4) is:

$$Gain(A) = I(p,n) - E(A) \tag{4}$$

The information gain measure is used to select the test attribute / class label at each node in the tree. Such a measure is referred to as an attribute selection measure or a measure of the goodness of split. The attributes with the highest information gain or greatest entropy reduction are chosen as the test attribute for the current mode. This attribute minimizes the information needed to classify the samples in the resulting partitions and reflects the least randomness or “impurity” in these partitions. Such an information-theoretic approach minimizes the expected number of tests needed to classify an object and guarantees that a simple tree is found. The algorithm computes the information gain of each attribute. The attribute with the highest information gain is chosen as the test attribute for the given set. A node is then created and labeled with the attribute, branches are created for each value of the attribute, and the samples are partitioned accordingly.

### 3.6.3 Naïve Bayesian-based ANNs Approach

Other than the decision tree-based approach, the Naïve Bayesian classification is also used to enhance the flexibility of the ANNs model. The Naïve Bayesian is based on the Bayes theorem; and the posterior probability is listed below (5):

$$P(H \mid X) = \frac{P(X \mid H)P(H)}{P(X)} \tag{5}$$

Where $X$ is a data sample whose class label is unknown and $H$ is some hypothesis, such as that the data sample $X$ belongs to a specified class $C$.

The Naïve Bayesian Classifier works as follows (Han and Kamber, 2001):

- Each data sample is represented by an $n$-dimensional feature vector, $X = (x_1, x_2, \ldots, x_n)$, depicting $n$ measurements made on the sample from $n$ attributes, representing, $A_1, A_2, \ldots, A_n$.
- Suppose that there are $m$ classes, $C_1, C_2, \ldots, C_m$. Given an unknown data sample, $X$ (i.e., having no class label), the classifier will predict that $X$ belongs to the class having the highest posterior probability, conditioned on $X$. That is, the Naïve Bayesian classifier assigns an unknown sample $X$ to the class $C_i$ if and only if:

  $P(C_i \mid X) > P(C_j \mid X)$ for $1 \le j \le m, j \ne i$.

  Thus $P(C_i \mid X)$ is maximized. The class $C_i$ for which $P(C_i \mid X)$ is maximized is classed the maximum posterior hypothesis (6):

$$P(C_i \mid X) = \frac{P(X \mid C_i)P(C_i)}{P(X)} \tag{6}$$

- In order to reduce computation in evaluating $P(X \mid C_i)$, the Naïve assumption of class conditional independence is made. This presumes that the values of the attributes are conditionally independent of one another, given the class label of the sample, that is, there are no dependent relationships among the attributes. Thus (7):

$$P(X \mid C_i) = \prod_{k=1}^{n} P(x_k \mid C_i) \tag{7}$$

- In order to classify an unknown sample $X$, $P(X \mid C_i)\,P(C_i)$ is evaluated for each class $C_i$. Sample $X$ is then assigned to the class $C_i$ if and only if

  $P(X \mid C_i)P(C_i) > P(X \mid C_j)P(C_j)$ for $1 \le j \le m, j \ne i$.

  In other words, it is assigned to the class $C_i$ for which $P(X \mid C_i)\,P(C_i)$ is the maximum.

By finding the most appropriated approach in classification-based ANNs, the Naïve Bayesian approach is also integrated with the ANN models as the classification methodology which compares with the decision tree approach to identify the best suitable one. Details are presented in the case study section.

## 3.7 Web Reporting Module

The IDS information portal is a customized web-based user interface which aims to allow the accessing of quality data and predicted outcomes. Decision support is suggested through different kinds of visual presentations, such as the trend plot, wafer map, correlation chart, etc (Figure 3.12). Engineers can then find out the sources of any faults and take actions to improve the processes. Continuous monitoring can also take place to ensure the processes have been improved efficiently and effectively.

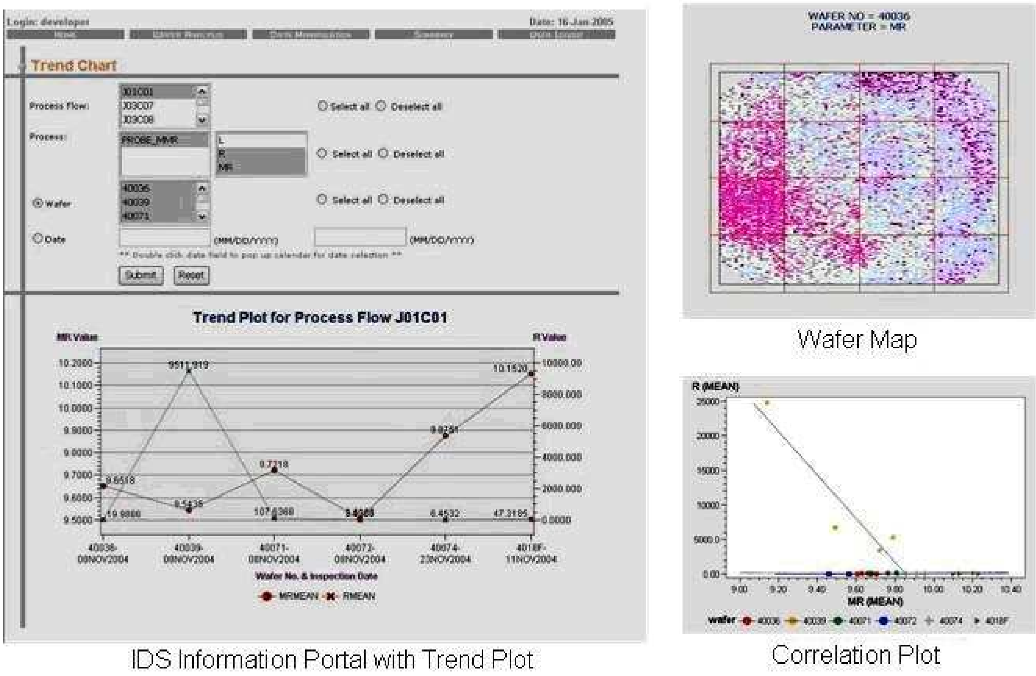



Figure 3.12 IDS Information Portal

## 3.8 Summary

The proposed IDS makes use of the AI and the experience of the engineers in identifying foreseeable failure modes of a process or a series of distributed processes and planning for its elimination. This “before-the-event” action can achieve proactive quality control and enhancement, providing remedial action and planning for the future within the production process flow and greatly reduces the production costs. Details of the validation of proposed IDS are discussed in the case study with the examples of slider fabrication and RIE in Chapter 4.

# CHAPTER 4. CASE STUDY

## 4.1 Introduction

In this chapter, a case study into the background knowledge of the magnetic head / HDD industry is established. It describes the details of how this industry is facing production difficulties at the moment and what it needs to do to optimize the production process in order to achieve success in the competitive market. The implementation and validation are also included so as to be able to investigate the possibility of using the AI model and the systematic approach to help the industry. If the implementation is successful in the company, it can indicate that the hybrid Neural-OLAP with decision tree-based / Naïve Bayesian-based classification could be applied to other manufacturing processes and resulting in the same benefits.

IDS was developed in-house as the joint research project of SAE Magnetics (H.K.) Ltd. and the Hong Kong Polytechnic University, and it is considered being an intellectual property of the company. Part of the information in IDS is identified as the confidential commercial information that is protected from disclosure. We regret that it is unable to share the sensitive part of the application (i.e. the details of the input / output parameters) or any of its source code. The thesis includes the overall research approach, the step-by-step methodology and the case study of system validation.

## 4.2 Production Data Collection and Processing

Data collection is an essential starting step in all analysis; both the quantity and quality of data affect the final analyzed result. It therefore is the most important step and serves as the foundation of nearly all research nowadays. The research project implementation is first describes the flows of production data collection and processing.

### 4.2.1 Production Data Audit

Before starting production data collection, a large scale of production data audit was performed in SAE Magnetics (H.K.) Ltd. to study and ensure that the data collected are relevant and accurate. This data quality cannot be done by one department, it is necessary to have all the department leaders collaborating together with top management support. Otherwise, the data collected are not so accurate, the final result becomes meaningless and nothing can be benefited after a lot of work has been done. The company had experienced such a failure previor to this research project. The objective of this research is therefore focused on collaboration between departments and also the insight from the university side.

Nearly all the manufacturing units (including project engineering and quality assurance teams in both slider production and HGA / HSA production), the testing equipment unit, the machine engineering unit, the information technology unit, equipment suppliers, raw materials suppliers and even customers all work together to participate in the project. Several investigations were done on the production shop floor and six scopes for data collection were defined:

- Reduce all possible data failure, especially human error (Solution: Automation)

- Standardize all attributes / inspection items under certain testing conditions
- Set range and specification for fixed attributes
- Eliminate fields that may cause redundancy
- Discover and record the potential pattern, relationship and dependence between data
- Study and suggest other suitable data collection for future analysis

Lots of tasks were done at the same time within different departments and locations, the target being to make sure all the data collected is accurate and usable within the new standard data collection approach.

### 4.2.2 Re-design of Table Structure

To fulfill the needs of dynamic data collection in a changing environment, the OTDB module was implemented once several new enhancements were applied to it. These are horizontal to vertical, parent-child relationship and domain tables.

**Horizontal to Vertical**

By using RD_BAR_QST from OCR database as an example, the traditional database stores information horizontally. All the required fields should be pre-defined in the table. Therefore, ITEM1 to ITEM60 are kept in the RD_BAR_QST table and the corresponding ITEM_CODE and ITEM_INDEX are then filled in the ITEM_INDEX_LIST (Figure 4.1). The number of ITEMs listed in the table is more than the fields required; therefore more than one-third of the fields are empty. On the other hand, if an extra field is needed, a table's structure should be modified, which causes a very serious problem in the system operation.

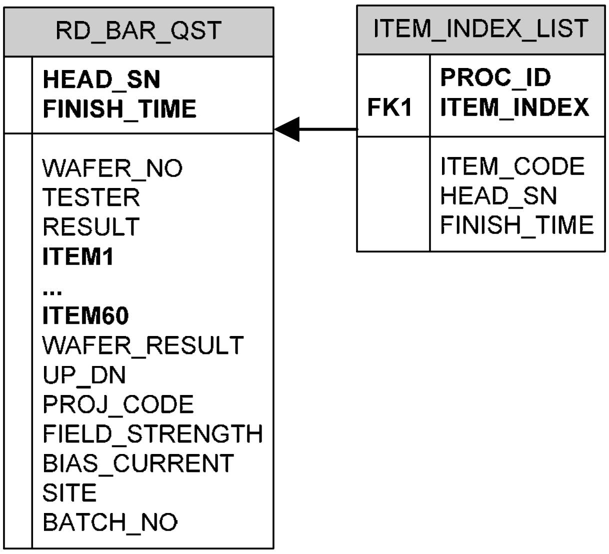



Figure 4.1 Existing Table Relationship of RD_BAR_QST

According to the rules of database design (Elmasri, 2000), a database should be designed as follows:

1. Database design is concerned with base relations, not user view;
2. Mixing attributes of multiple entities may cause a problem, so separating the multiple entities into different tables is strongly recommended;
3. Relationship should be designed such that their tuples (analogous to record) will have as few "NULL" values as possible.

Normalization is the standard for database design. It is a process of decomposing unsatisfactory "bad" relations by breaking up their attributes into smaller relations. And normal form is the condition using keys and functional dependencies of a relation to certify whether or not a relation schema is in a particular normal form.

The new table was then designed by changing the data field from horizontal to vertical as shown in Figure 4.2. The ITEM_NO and the corresponding information are pre-defined in the OCR_item table, and then the ITEM_VALUE with the ITEM_NO is stored in the BAR_QST_data. A well-defined relationship is built and the system is comparatively more flexible and extensible.

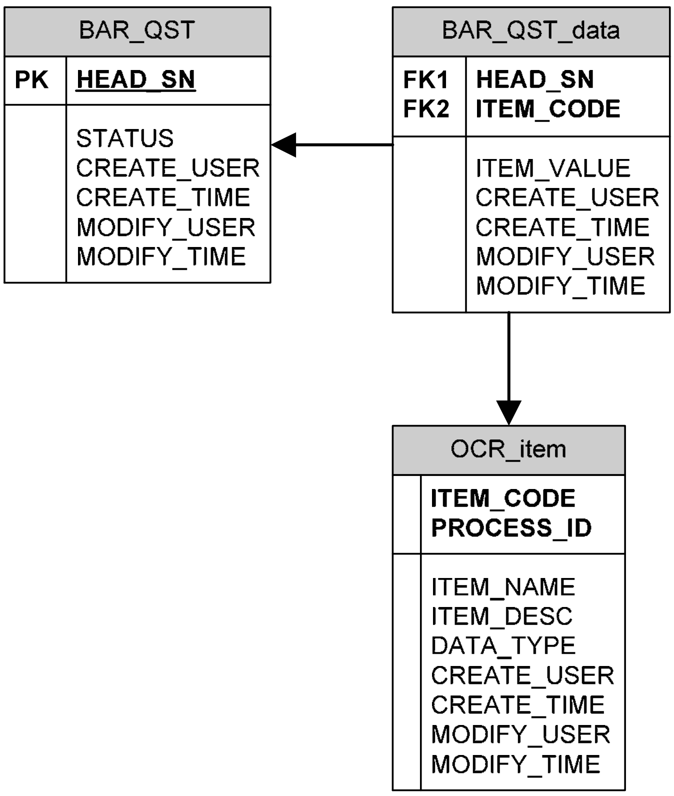



Figure 4.2 New Table Relationship RD_BAR_QST

VARCHAR2 (Oracle) / VARCHAR or CHARACTER (Others) is strongly recommended as the field type for ITEM_VAULE. It is because there is the possibility of having characters in items. A new field DATA_TYPE is added to define the exact type of data at that particular field. There are no primary keys in both BAR_QST_data and OCR_item, because retesting of the same slider (the same HEAD_SN) does occur, and the same ITEM_NO of different processes does appear in the same item table. The solution will be discussed in the following paragraph.

**Parent-child Relationship**

Retesting is one of the major problems for the system. The same slider, which uses the unique HEAD_SN, may contain several retest data entries to database. It therefore cannot be used as the primary key. Although a composite key can be used, it takes some time to trace back the data. On the other hand, it does create errors for the vertical database, because it is difficult to find the exact number of retests done for that particular ITEM.

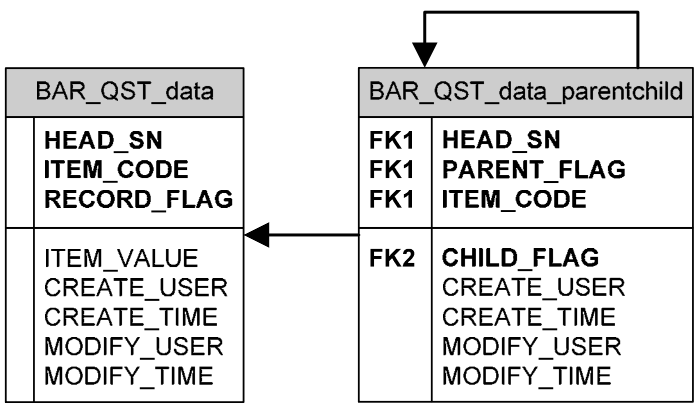



Figure 4.3 Parent-child Relationship at RD_BAR_QST

A new field RECORD_FLAG and table BAR_QST_data_parentchild are added (Figure 4.3). The RECORD_FLAG stores the corresponding retest sequence in that field, and the BAR_QST_data_parentchild collects the retest path. In the case of SAE Magnetics (H.K.) Ltd., FINISH_TIME is suggested. Otherwise, a numerical sequence or the Oracle sequence can be used. Two assumptions should be made, the date time of the server and all clients should be synchronized and all the FINISH_TIMEs of the inspection and the ITEM_VALUEs in a single test are the same.

Table 4.1 Parent-child Relationship Table of BAR_QST

| **BAR_QST_data** | | | | |
|---|---|---|---|---|
| **HEAD_SN** | **ITEM _NO** | **ITEM _VALUE** | **…** | **RECORD _FLAG** |
| 0190433 | 012 | 45,343.50 | | 12/23/03 11:40:20 PM |
| 0190433 | 013 | -5.75935 | | 12/23/03 11:40:20 PM |
| 0190433 | 014 | 34.894 | | 12/23/03 11:40:20 PM |
| : | : | : | | : |
| 0190433 | 012 | 45,392.01 | < Exist in all tests | 12/23/03 11:42:53 PM |
| 0190433 | 013 | -5.75982 | < Only in 1 & 2 | 12/23/03 11:42:53 PM |
| : | : | : | | : |
| 0190433 | 012 | 45,377.14 | < Exist in all tests | 12/23/03 11:46:07 PM |
| 0190433 | 014 | 34.729 | < Only in 1 & 3 | 12/23/03 11:46:07 PM |
| 0190433 | 107 | -12.06290 | < New ITEM | 12/23/03 11:46:07 PM |

| **BAR_QST_data_parentchild** | | | |
|---|---|---|---|
| **HEAD_SN** | **PARENT_FLAG** | **CHILD_FLAG** | **…** |
| 0190433 | 12/23/03 11:40:20 PM | 12/23/03 11:42:53 PM | |
| 0190433 | 12/23/03 11:42:53 PM | 12/23/03 11:46:07 PM | |
| 0190433 | 12/23/03 11:46:07 PM | NULL | < Last test / retest |
| *Remarks: If required (By using non-ordering sequence) | | | |

Referring to Table 4.1, the system will check the BAR_QST_data_parentchild first and see if the current HEAD_SN exists or not. If it is a new data input, a new record, which PARENT_FLAG and CHILD_FLAG are FINISH_TIME and “NULL” respectively, will be added into the BAR_QST_data_parentchild table and all ITEMs will mark the

FINISH_TIME in the RECORD_FLAG under the BAR_QST_data table. For the retest case, the system will find that the HEAD_SN exists in the BAR_QST_data_parentchild table. It will then search the latest record and update the “NULL” to the new retest FINISH_TIME. After that, a new record, which PARENT_FLAG and CHILD_FLAG are the latest FINISH_TIME and “NULL” respectively, will be added into the BAR_QST_data_parentchild table and all the ITEMs will mark the latest FINISH_TIME in the RECORD_FLAG under the BAR_QST_data table.

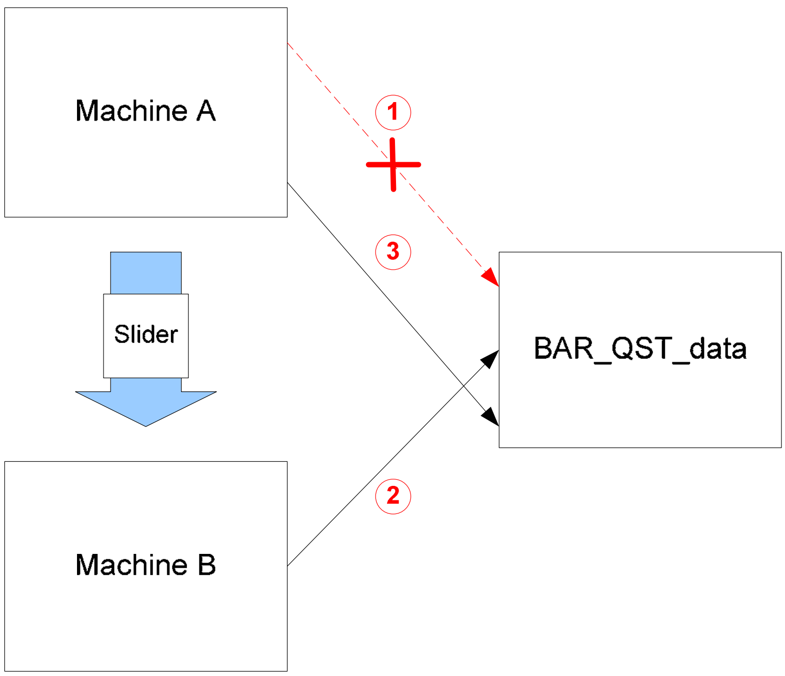



Figure 4.4 Network Failure

Imagining there two tests running in the production line with a single slider. The first test is done at machine A, the same as, in Figure 4.4. The connection between machine A and the BAR_QST_data server is suddenly broken and the testing data is then saved in the temporary table inside the memory of machine A. After that, the slider is moved to machine B and a retest is done for some purpose. The data is uploaded to the database server successfully as the first record in the BAR_QST_data table. However, the previous record in machine A is transferred to the central database again after the network problem has been repaired. Therefore, the previous record becomes the retest record in this situation.

By using the FINISH_TIME as the record flag, the right order of test data can be found by creating an index with FINISH_TIME. The later date / time, the larger number of the test is (Table 4.2).

Table 4.2 Parent-child Relationship during the Case of Network Failure

| **BAR_QST_data** | | | | |
|---|---|---|---|---|
| **HEAD_SN** | **ITEM _NO** | **ITEM _VALUE** | **…** | **RECORD _FLAG** |
| 0190433 | 012 | 45,343.50 | Second Test > | 12/23/03 13:04:20 PM |
| 0190433 | 013 | -5.75935 | | 12/23/03 13:04:20 PM |
| 0190433 | 014 | 34.894 | | 12/23/03 13:04:20 PM |
| : | : | : | | : |
| 0190433 | 012 | 45,392.01 | First Test > | 12/23/03 12:42:53 PM |
| 0190433 | 013 | -5.75982 | | 12/23/03 12:42:53 PM |
| 0190433 | 014 | 34.729 | | 12/23/03 12:42:53 PM |

| **BAR_QST_data_parentchild** | | | |
|---|---|---|---|
| **HEAD_SN** | **PARENT_FLAG** | **CHILD_FLAG** | **…** |
| 0190433 | 12/23/03 13:04:20 PM | NULL | |
| *Remarks: If required | | | |

| **BAR_QST_data_parentchild** | | | |
|---|---|---|---|
| **HEAD_SN** | **PARENT_FLAG** | **CHILD_FLAG** | **…** |
| 0190433 | 12/23/03 12:42:53 PM | 12/23/03 13:04:20 PM | < New record! |
| 0190433 | 12/23/03 13:04:20 PM | NULL | |
| *Remarks: If required | | | |

In fact, the parent-child table is not necessary in the situation of using a sequence like FINISH_TIME as the RECORD_FLAG. It is only required if a non-ordering sequence is used. This parent-child design is suitable for multi-item data entry. The above example is modified for ease of understanding when describing the design idea, the final implementation using the same design structure of parent-child relationship, but the array table is used instead which is described in Chapter 3. For other database solutions that do not have the same methodology of array table in Oracle 9i, the parent-child design can be used directly.

**Data Collection Table**

In the actual manufacturing, data collection is also affected by different type of production. These can be divided into prototype, mass production, rework, etc. The difference between these types of production is the amount of data / items required and the frequency of the data collected / recollected. They can be categorized into two areas, static and dynamic as shown in Table 4.3.

Table 4.3 Static and Dynamic of Data Collection

| **Catalog** | **Type of Production Data** | **Frequency of Collection** | **Number of Items** | **Record Flag / Parent-child Table Required** |
|---|---|---|---|---|
| Static | Information from Vendors / Suppliers | Once | Fixed | No |
| Dynamic | Prototype | Many | Changes from Time to Time | Yes |
| | Mass Production | Once or more | Fixed | Yes |
| | Rework | Once or more | Changes from Time to Time | Yes |

By using this concept and the normalization rule of database design, it is recommended to separate the tables based on the types of production. An example is listed in Figure 4.5. Corresponding PARENTCHILD tables are created for keeping the data entity path (if required).

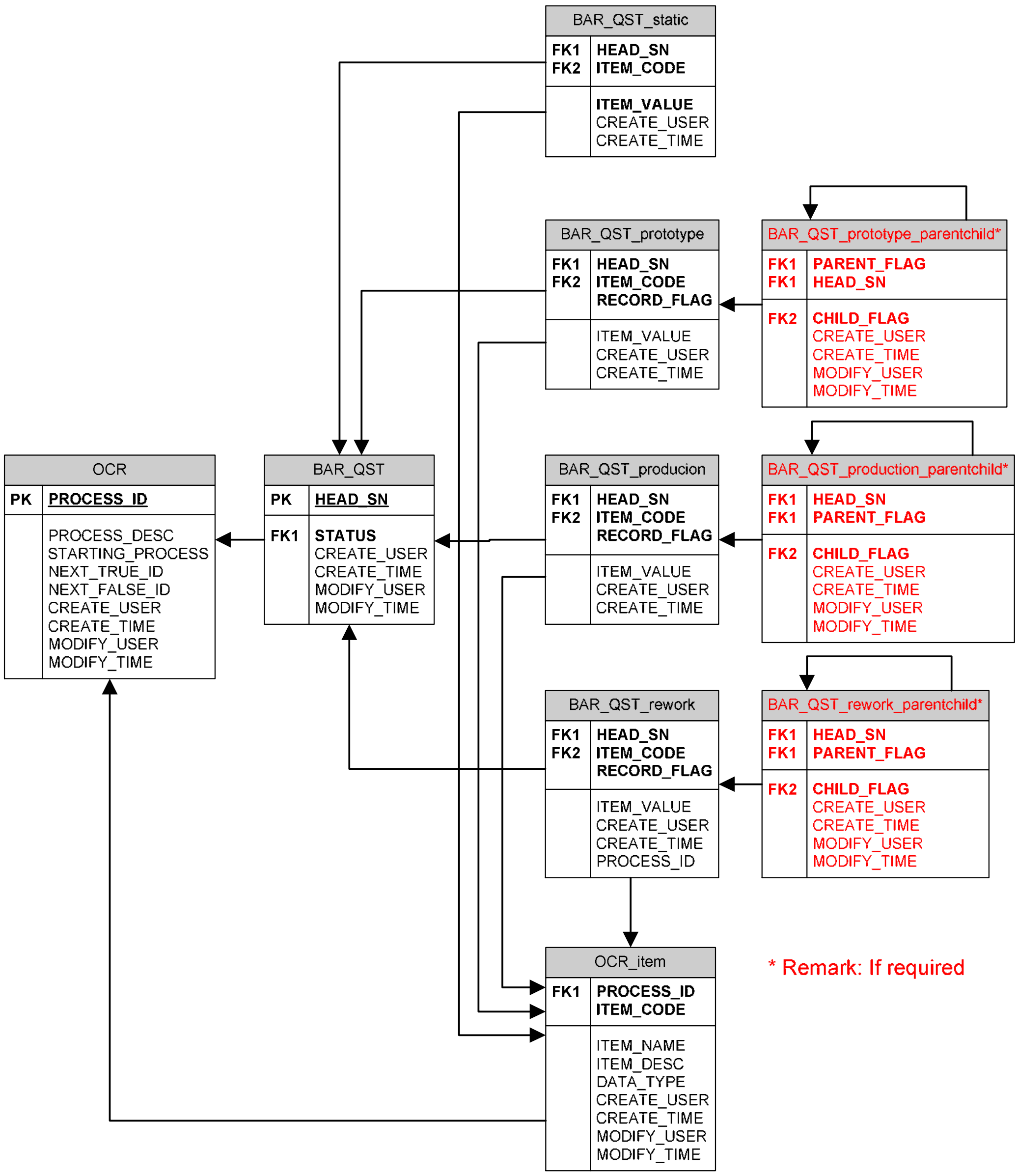



Figure 4.5 Database Model Diagram (DMD) of BAR_QST

### Domain Tables

The domain tables are used to store all the pre-defined information for the system, including the one that was discussed before, ITEM domain. There are also others in the system (Table 4.4). The advantage of using domain tables is the flexibility they give. Developers can write a generic program with variables and then call the required parameters / settings in the domain tables, without “hard code” in the program. Therefore, it is not required to rewrite programs for each new table and all the parameters can simply be modified.

Table 4.4 Domain Tables

| **Domain Table** | **Phase** | **Table Name** | **Description** |
|---|---|---|---|
| Process Domain | 1 | SYSTEM<br>e.g. OCR | Process flow & information |
| Item Domain | 1 | PROCESS_item<br>e.g. BAR_QST_item | Data input item information |
| Key Domain | 2 | SYSTEM_key<br>e.g. OCR_key | Primary keys / composite keys of each table in the system, building table relationship |
| Parameter Domain | 2 | SYSTEM_parameter<br>e.g. OCR_parameter | System parameters |

In fact, domain tables are stored inside the database. They are copied to the buffer at the processing computers with the process and data directors. This design can reduce the query time and increase the overall performance. Synchronization should be made frequently (once a day) to ensure setting data integrity.

**Data Director**

The data director is a program written to manage and control the input information flow (Figure 4.6). By using the pre-set parameters and item information from the two domain tables, the data can be directed and stored in the right section in the database. The corresponding fields and records, such as RECORD_FLAG or field in the PARENTCHILD table will be updated automatically. Normally, the data director only locates in the server side, but an individual server can be put in rather than the database server. The features of the data director are listed below:

- Directs and stores data to corresponding table in particular field under the pre-set information and condition resulting from Parameter Domain and Business Rule System;
- Checks, matches and compares the input data with the pre-defined value from Item Domain;
- Records and reports the history, especially the errors, by creating a log document and / or through the alert system.

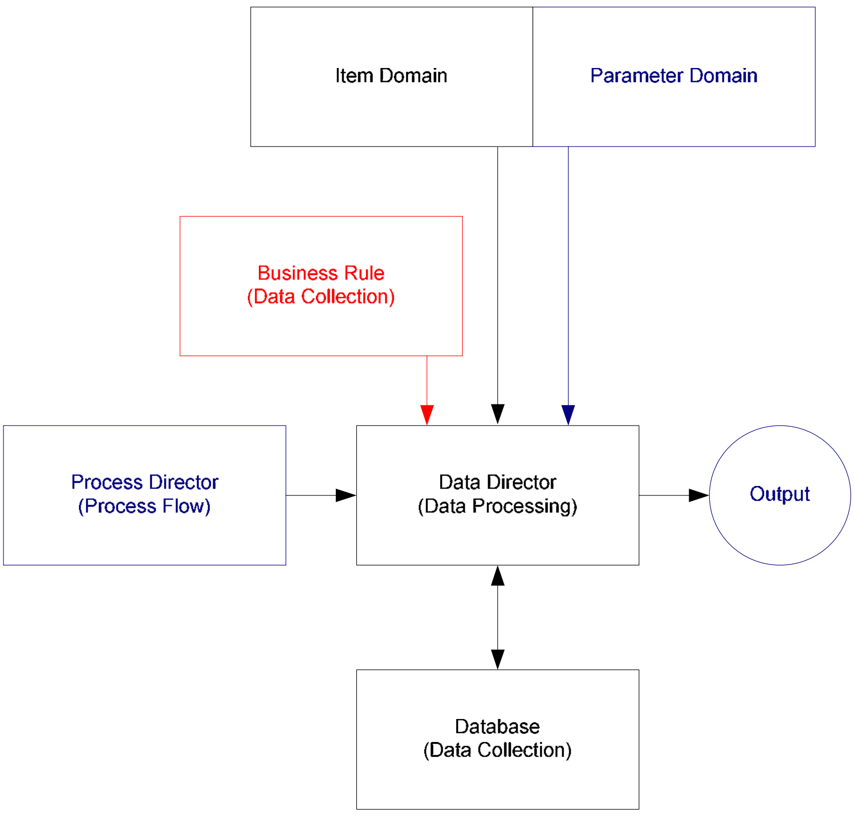



Figure 4.6 Components related to Data Director

PL/SQL was suggested as the language tool for the data director replacing the complicated SQL statement. The program was written as flexible enough to handle the parameter and variable manner, eliminate the needs for modifying or rewriting the program if any new table or module is added.

### 4.2.3 Applying Three-Tier Architecture

Both OTDB and MDDB modules were implemented to nearly all the critical areas on the production shop floors in SAE Magnetics (H.K.) Ltd. The production data is located in different locations including plants in China – Dongguan, China – Changan, the Philippines, Japan and the United States. The standard production data protocol was applied to all the parties. No matter what the formats are they in Access, Excel, SQL Server, Oracle, SAP or CSV, the data is converted to this standard protocol and transferred to the central database through DCOM. For those data that are inside the same factory, the data collection is in real-time. As for other data from remote factories or overseas, it is done by File Transfer Protocol (FTP) with DCOM, DB Link (for the case by using Oracle) and RosettaNet (for the case of sensitive data in SAP between partners including suppliers and customers) in batch mode. The role of DCOM is to distribute the data conversion tasks to the idle data adaptor for processing and uploading data to the OTDB. Production data could be found and forecasted in the IDS within a few seconds after the process had just been completed.

Table 4.5 Production Data Collected to OTDB

| **Product Level** | **Process** | **Transfer Techniques** |
|---|---|---|
| Wafer | Probe | DCOM |
| | Wafer Summary | FTP + DCOM |
| Slider | Bar QST | CSV + DCOM |
| | RIE | DCOM |
| | Slider DP | DCOM |
| HGA | HGA DP | DCOM |
| | HGA D2ET | DCOM |
| | HGA Fly Height | DCOM |
| HSA | Slider-HSA Linkage | DCOM |
| | HSA DP | DCOM |
| | HSA QST | DCOM |
| | HSA D2ET | DCOM |
| HDD | HSA-HDD Linkage | CSV + DCOM |
| | Drive Summary | FTP + DCOM |
| Others | DUV | DCOM |
| | SEM | DCOM |
| | Assembly | DCOM |
| | Bearing Height / Bearing Parallel | DCOM |
| | Pitch / Roll | DCOM |
| | Arm Height | DCOM |
| | Load Gram | DCOM |
| | Alignment | DCOM |
| | Vacuum | DCOM |
| | ATE | DCOM |
| | WYKO | DCOM |

The first phase of implementation was mainly in China – at the Dongguan factory. It was then applied to the Changan plant after the success of OTDB and MDDB implementation. The system design is now being introduced to the Hong Kong research centre and the Philippines production plant. The case study focuses on the first phase of implementation in Dongguan. The IDS with 4 x86-based single core CPU servers was tested so as to be able to handle up to 1,200 connections and more than 200 transactions per second from the wafer fabrication to the HSA production lines. An active-passive server strategy was employed with the Storage Area Network (SAN) for the High Availability (HA) purpose in the central database. The central database size was approximately

1TB for six-month data. Several techniques were applied in the database to eliminate the input and output (I/O) loading of both servers and the storage space required, such as the Redundant Arrays of Independence Disks (RAID) and compress feature in the database storage. Half year production data from the RIE as well as the "before" and "after" process measurement data was collected. Details of the process of data collection are listed in Table 4.5.

## 4.3 Quality Analysis in Neural-OLAP

After the data collecting into the OTDB, the next steps are the ad-hoc analysis by multi-dimensional data structure and the intelligent prediction through classification-based ANN models. The case study shows the implementation steps with examples and results under the RIE processes.

### 4.3.1 Applying OLAP with Examples

Before building the OLAP, it is essential to have a process to identify which kinds of data should be included in the cube and which are related. JMP (Figure 4.7) was suggested to use for the first step to "explore" the data and find out some of the related relationships for setting rules. The engineers can easily use the related data in the data warehouse to find the underlying rules and patterns and thus help make a better decision in IDS. This of course should be done by collaboration between the domain experts who have many years of knowledge and experience in these. JMP acts to help them to prove their ideas on the way to create the OLAP cubes.

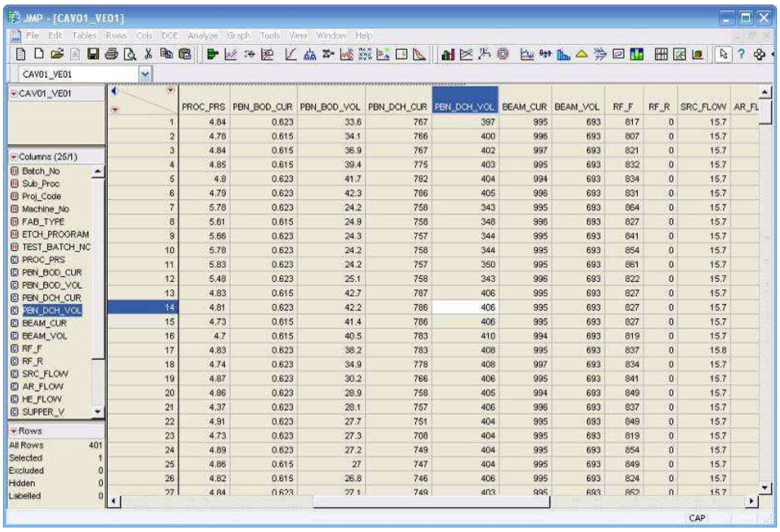


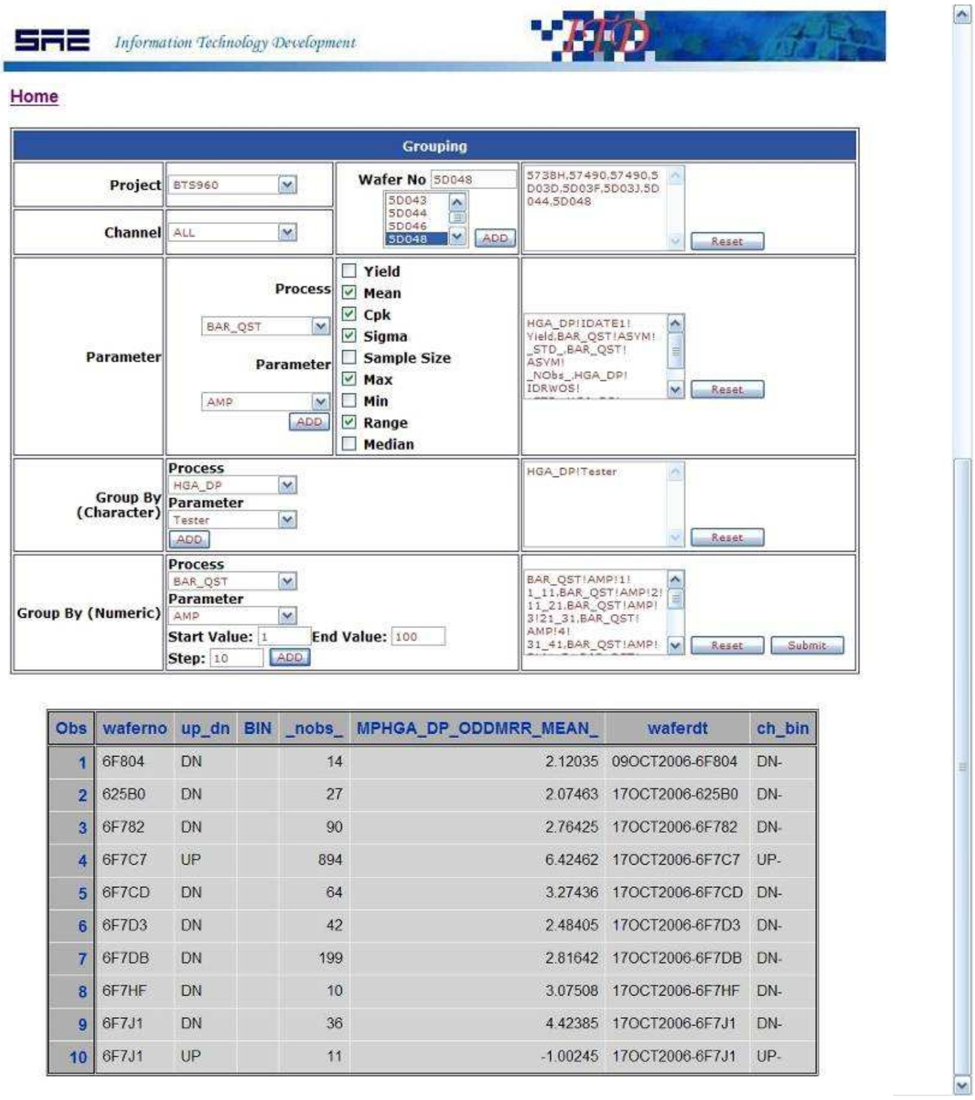

Figure 4.7 Using JMP for Exploring the Data

Some sources of quality problem are actually difficult to identify by a normal table-style report format. Ad-hoc OLAP analysis was found to be the best way for engineers to dig through all the related raw and summarized data by changing different dimensions to compare and check interactively. Once the data and cubes are identified, the OLAP can be constructed. The overall structure of the data warehouse is shown in Table 4.6. Each row of the table is a subject, while the columns represent the dimensions. The subjects of the MDDB are the data groups that can be used to support the analytical and reporting needs of management area within the technical division. Several subjects of the data warehouse can also be combined into larger data groups to create the necessary views of product and process quality data for cross-subject queries and analyses.

Table 4.6 Objects and Dimensions in OLAP

| | Dimensions | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Subjects** | Wafer | Time | Process | Parameter | Machine | Material | Operator | Project | Others |
| Wafer Probe | √ | √ | √ | √ | √ | √ | √ | √ | |
| Wafer Summary | √ | √ | √ | √ | √ | √ | | √ | Supplier, Customer |
| Bar QST | √ | √ | | √ | √ | | √ | √ | |
| RIE | √ | √ | √ | √ | √ | √ | √ | √ | |
| Slider DP | √ | √ | | √ | √ | | √ | √ | |
| HGA DP | √ | √ | | √ | √ | | √ | √ | |
| HGA D2ET | √ | √ | | √ | √ | | √ | √ | |
| HGA Fly Height | √ | √ | | √ | √ | | √ | √ | |
| Slider-HSA Linkage | √ | √ | | | √ | √ | √ | √ | Supplier |
| HSA DP | √ | √ | | √ | √ | | √ | √ | |
| HSA QST | √ | √ | | √ | √ | | √ | √ | |
| HSA D2ET | √ | √ | | √ | √ | | √ | √ | |
| HSA-HDD Linkage | √ | √ | | | √ | √ | √ | √ | Supplier |
| HDD Summary | | √ | √ | √ | √ | √ | | √ | Customer |

In the slider fabrication, process monitoring is not only done by the slider. Data is analyzed by grouping in time and other dimensions. The OLAP acts as a multi-dimensional data model which organizes the measured result and related information from the testers into a hierarchy that represents the levels of details of the data. This powerful query engine assists engineers to find and retrieve the defected statistics or performance measurement with the dimensions of wafer, time, process, parameter, tester (machine), setting, operator and project. Figure 4.8 shows the interface for users to select the data from different dimensions. Cubes are generated in schedule with the predefined summarization rules of standard statistical elements, e.g. the mean, standard deviation / sigma, yield, process capability (Cpk), uniformity, etc. This OLAP engine supports multi-dimensional views of data through array-based multi-dimensional storage engines. They map multidimensional views directly to data cube array structures. If a problem appears, then an engineer can go through OLAP and query different dimension and hierarchies, like whether the problem is coming from the wafer, block or slider level, the source is related to material, machine or man, etc. According to the example listed in Figure 4.9, the OLAP cube was created by the dimensions of date, slider hierarchy and machine. The same defect appeared in certain region of the cube – many row-bars from half-wafer X processed by STS machines on 15th had the same defect. A fast identification could be recommended that there were some problems under the above conditions. Operators and engineers could then take action to continue monitoring and analyzing the issue to achieve quality enhancement.




| Obs | waferno | up_dn | BIN | _nobs_ | MPHGA_DP_ODDMRR_MEAN_ | waferdt | ch_bin |
|---|---|---|---|---|---|---|---|
| 1 | 6F804 | DN | | 14 | 2.12035 | 09OCT2006-6F804 | DN- |
| 2 | 625B0 | DN | | 27 | 2.07463 | 17OCT2006-625B0 | DN- |
| 3 | 6F782 | DN | | 90 | 2.76425 | 17OCT2006-6F782 | DN- |
| 4 | 6F7C7 | UP | | 894 | 6.42462 | 17OCT2006-6F7C7 | UP- |
| 5 | 6F7CD | DN | | 64 | 3.27436 | 17OCT2006-6F7CD | DN- |
| 6 | 6F7D3 | DN | | 42 | 2.48405 | 17OCT2006-6F7D3 | DN- |
| 7 | 6F7DB | DN | | 199 | 2.81642 | 17OCT2006-6F7DB | DN- |
| 8 | 6F7HF | DN | | 10 | 3.07508 | 17OCT2006-6F7HF | DN- |
| 9 | 6F7J1 | DN | | 36 | 4.42385 | 17OCT2006-6F7J1 | DN- |
| 10 | 6F7J1 | UP | | 11 | -1.00245 | 17OCT2006-6F7J1 | UP- |

Figure 4.8 User Interface for the MDDB Query

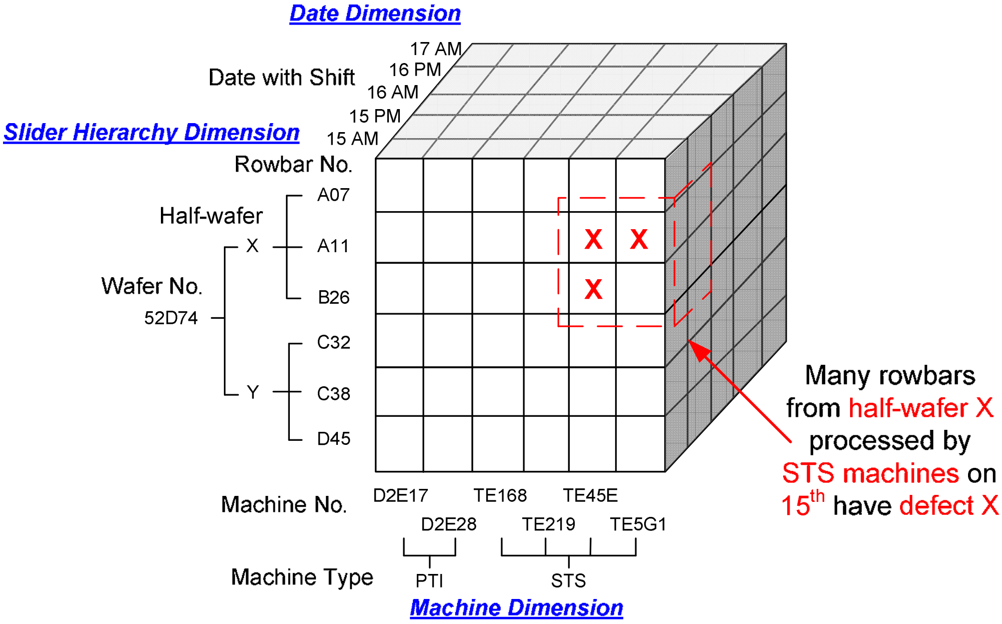



Figure 4.9 An Example of OLAP Application in RIE

Compared with the traditional relationship database, it has the ability to handle changing requirements and model common business situations by means of the predictable query processing. Pre-aggregation, dimensional hierarchy, and sparse data management can significantly reduce the size of the database and the need to calculate values. These approaches remove the need for multi-table joins and provide quick and direct access to the array of data, thus significantly speeding up execution of the multi-dimensional queries. The functionalities of OLAP and data mining can be viewed as disjointed: OLAP is a data summarization / aggregation tool that helps simplify data analysis, while data mining allows the automated discovery of implicit patterns and interesting knowledge hidden in large amounts of data.

### 4.3.2 Investigation of RIE and Related Process Flow

Theoretically, the etch rate equation (8) below can predict the corresponding etch rate in the slider fabrication. This method is found accurate, however, only in a laboratory standard environment. Equation (9) is used to calculate the required etch time based on the target etch rate and the target thickness.

$$E_R = Ae^{-\frac{E_a}{kT}} \tag{8}$$

Where $E_R$ = Etch Rate, $A$ = Rate Constant, $E_a$ = Activation Energy, $k$ = Boltzmann Constant, $T$ = Temperature

$$t_{ideal} = t_{obs} + (T_{des} - T_{obs}) / E_R \tag{9}$$

Where $t_{ideal}$ = Ideal Etch Time, $t_{obs}$ = Observed Etch Time, $T_{des}$ = Designed Target Thickness, $T_{obs}$ = Observed Target Thickness

Seven additional factors were included in the consideration of decision tree-based ANNs model development, these inclusions were taken from the engineering knowledge domains and from experiments:

- Rate constant ($A$) dependence with the reactant density
- Changing temperature ($T$) during process
- Project various – requirements of the finished product
- Sub-process and program various – the process specifications
- Dummy / empty jig in the batch affecting the total processing surface area
- Machine idle time – the machine warm-up effect
- Chamber life cycle including the side effect of residents

In addition, there are two thickness measurements, one before, the other after the RIE. The first one is used to measure the actual thickness that is used to forecast the estimated length of etching time required. The second one is used to measure the final thickness and check how the etching process has behaved. A comparison, therefore, can be done by these two measurements.

### 4.3.3 Determination of RIE Critical Factors

The major objective of employing the ANN model is to model the RIE process as well as predict the etching rate and etching time. For building the ANN model, the architecture of the ANN, which includes a number of nodes in the input and output layers, a number of hidden layers, and a number of hidden nodes, needs to be defined first. The critical input factors of the ANN are required to be determined first. Generally speaking, these critical factor selections are based on the literature and also the experience of experts in the field as mentioned in Chapter 4.3.2. To ensure all the important elements were considered in this case study, all the points of data collection were studied and audited several times by means of a flow chart, Pareto analysis and a cause and effect diagram.

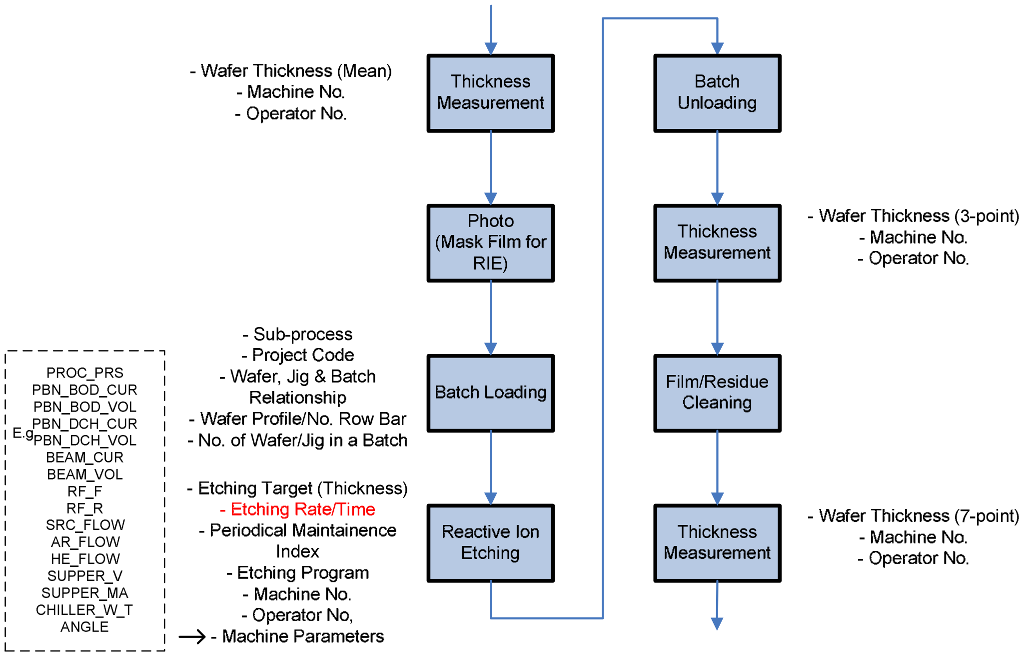



Figure 4.10 RIE Procedures and Related Data Collection

Figure 4.10 shows all the important points in the RIE procedure for data collection. Other than the photo mask film procedure, batch unloading, film / residue cleaning, data collection was suggested to be included in all other steps. Normally, those data collected directly from the RIE machine should be included, such as the voltage, current, gas flow ($CHF_3$ and He / Ar), etc. For mass production, however, other information including the machine number, the operator number, the project code, sub-process, the jig and batch relationship, wafer profile, the number of jigs in the batch, the periodical maintenance index, etching program, etc are also important factors that the previous literature does not include. By using the cause and effect diagram, factors related to Man, Machine, Material, Method and Environment (4M1E) can also be determined.

### 4.3.4 Comparison of Decision Tree and Naïve Bayesian Approaches

Since the classification-based ANN approach is suggested as the core for the prediction module in IDS, two popular approaches are used and compared so as to find out which is the most appropriate and accurate one for the classification. Naïve Bayesian is a powerful tool for decision making and reasoning under uncertainty, but it is based on a very strong independence assumption (Ben Amor *et al.*, 2004). They are robust to isolated noise points because such points will be averaged out when estimating conditional probabilities from data, which can also be used for missing values by ignoring the example during model building and classification (Tan *et al.*, 2006). Decision trees, which are among the most well-known of machine learning techniques (Wang *et al.*, 2005), are used for comparison with Naïve Bayesian in this section.

From the previous experiments in SAE Magnetics (H.K.) Ltd., it was first found that the ANN model behavior was not so accurate by using one single ANN model. It is not suitable for some of the cases, for example, the design of the ABS in the slider is very different in different project codes, so a genetic ANN model for all project codes becomes meaningless. As one single neural network may not be able to fulfill the needs and conditions of process prediction, certain numbers of neural networks are therefore built based on the specific situation and environment for that prediction. By finding the most accurate and reliable network model, the classification is involved in classifying the case to the most suitable predictive network. Classification is used to identify whether the new batch of sliders is similar to the previous experimented ones by using the same settings and the same ANN model for prediction.

Four of the attribute types of data are defined based on the trial-and-error method conducted by experts and engineers. By avoiding the over-fitting in classification, noise and outliners were first filtered (For example, those typing mistakes or special cases such as experimental RIE samples and prototype). They are:

1. Project code
2. Sub-process
3. Etching program
4. Etching machine number (different group of machine number)

It is required to first identify the actual ANN models needed for the prediction and then find out the relationship one-by-one with enough amounts of data for the training sets. Five sets of ANN models were found. Details of which models were discovered and constructed are included in the next section. The evaluation of classification efficiency is based on the Percent of Correct Classification (PCC) of the instances belonging to the testing set (Ben Amor *et al.*, 2004). Since the above four parameters are all in character / text type, so there are no continuous variables that may be required for consideration of the normal distribution or kernel density estimation.

Results of the classification with training and testing data between the decision tree and Naïve Bayesian-based approaches are listed in Table 4.7. It shows that both decision tree and Naïve Bayesian are completely in accordance with the training set which means that this latter is coherent, i.e., almost all training instances characterized by the same attributes' values belong to the same class. This behavior is also kept with the classification phase with a little bit of an advantage for the decision trees, but this never exceeds 1.5% of the difference between the different approaches. It is clear that the decision tree gives slightly better results. However, from a computational point of view, the construction of Naïve Bayesian is much faster then the decision tree. According to Table 4.8, both learning and classifying in Naïve Bayesian is approximately 5 times faster than using the decision tree. Since the objective of IDS is accuracy and the computer processing power is not a major concern nowadays, as most of the minimum level servers are having dual core Central Processing Units (CPU), such as Pentium D or Xeon from Intel and Athlon 64 X2 or Opteron from AMD, otherwise Naïve Bayesian will be the choice. On the other hand, the more samples used in the decision tree construction, the higher the accuracy achieved. This is a trade-off between accuracy and information / computation burden.

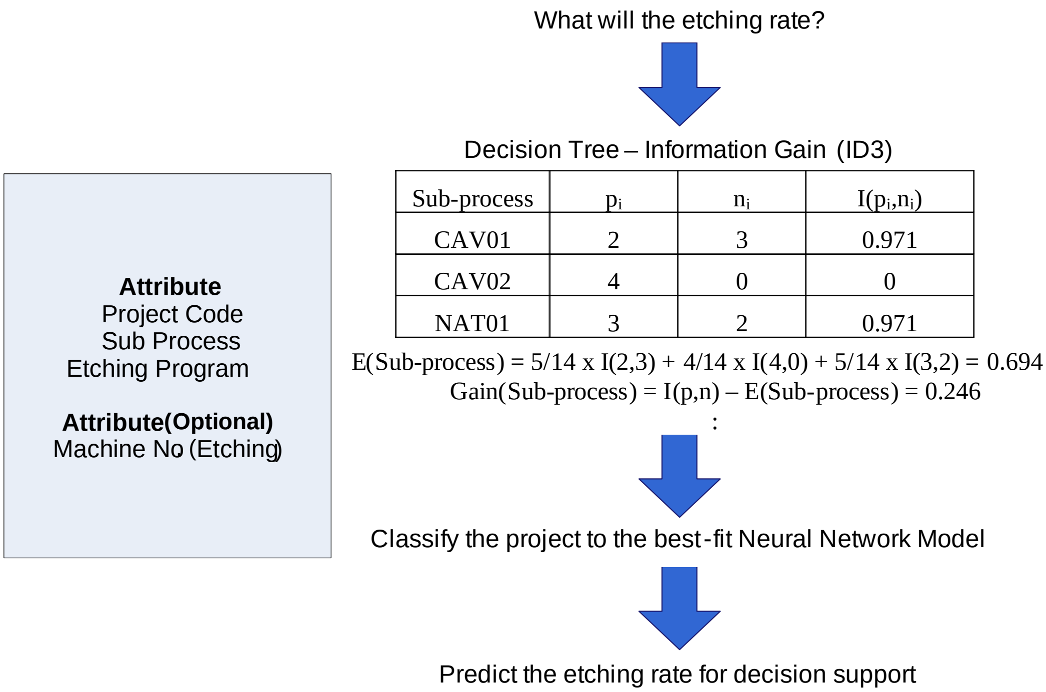
Table 4.7 PCC's in Decision Tree & Naïve Bayesian Classification

| *Training Set* | | *Testing Set* | |
|---|---|---|---|
| *Sample Size* | *PCC* | *Sample Size* | *PCC* |
| 25,000 | 100% | 10,000 | 100% |
| **Decision Tree** | | | |
| 24,951 | 99.80% | 9,218 | 92.18% |
| **Naïve Bayesian** | | | |
| 24,786 | 99.14% | 9,063 | 90.63% |

Table 4.8 Speed Performance between Decision Tree & Naïve Bayesian

| | Decision Tree | Naïve Bayesian | Decision Tree / Naïve Bayesian |
|---|---|---|---|
| *Learning* | 7m 12.004s | 1m 34.068s | 4.60 times |
| *Classifying* | 3m 21.962s | 49.187s | 4.11 times |

From the above comparison of the results, the decision tree-based approach was chosen as the classification algorithm integrated with the ANN for the corresponding model selection. Figure 4.11 shows the flow of the decision tree model classifier. The entropy and information gain were calculated based on the training data, then the testing data were inputted to validate the model as mentioned above.




| Sub-process | $p_i$ | $n_i$ | $I(p_i,n_i)$ |
|---|---|---|---|
| CAV01 | 2 | 3 | 0.971 |
| CAV02 | 4 | 0 | 0 |
| NAT01 | 3 | 2 | 0.971 |

Figure 4.11 Flow of Decision Tree Model Classifier

Figure 4.12 shows the decision tree and the rules discovered for the five ANN models. Model A is the generic model for most of RIE in case the sub-process code is NATxx which means no empty jig is inside. The sliders with a large area of air groove are mostly under the project TMK8xx and SSV9xx, the best suitable RIE model for ANN is model B. Model C is the one with PRO-xx1, xx2 and xx4 which represent the fact that the machines have just warmed-up.

Finally, the rest of the classification is based on the machine types A and B to identifying the belonging of model D and E respectively.

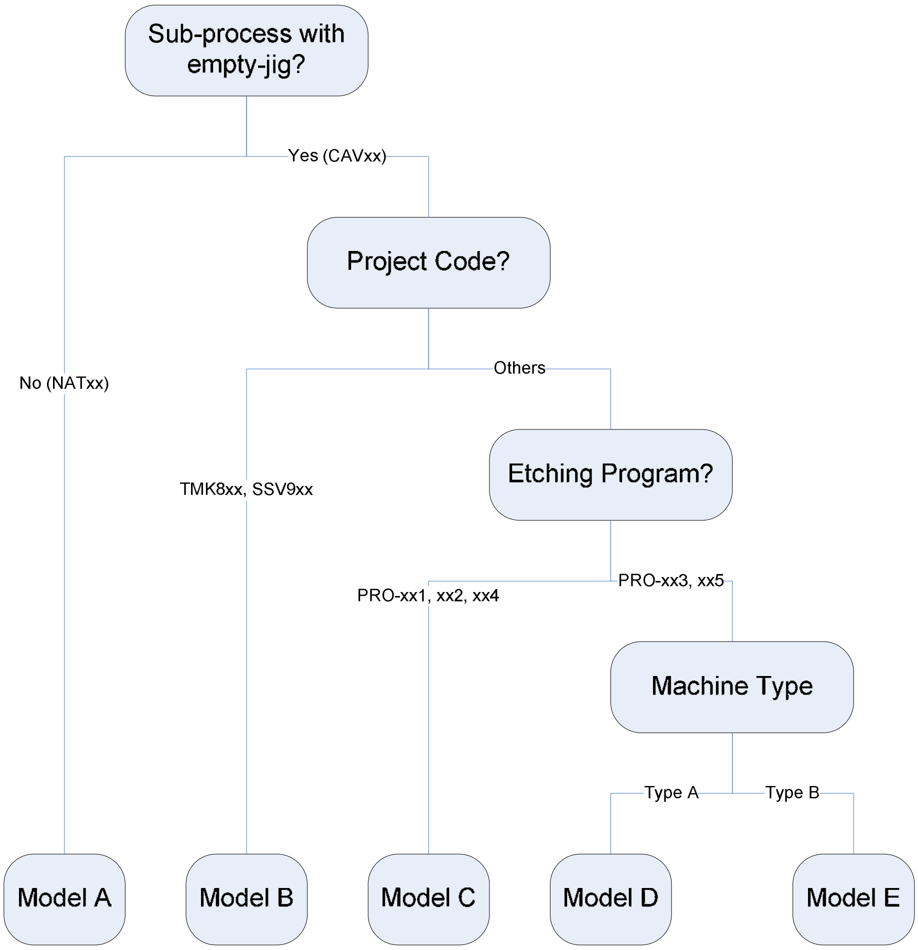



Figure 4.12 Decision Tree Diagram for the Five ANN Models

The hybrid intelligent system, by combining the decision tree and ANNs, was used in the data mining engine in IDS. This system helps and provides a more accurate and flexible solution for some complex problems under the dynamic situations. More attributes for classification might be required in case the model of ANN is found inaccurate when new factors appear.

### 4.3.5 Creation of ANN Models

As the back-propagation ANN was chosen in this research, the first step in modeling is to create the ANN model. Back-propagation ANNs (Haykin, 1999) have multi-layered architecture as shown in Figure 4.13. The first layer is called the input layer, and does not have computing activity. It is simply used to input independent variables such as the various significant formulations of RIE and process parameters, to the hidden layer. The last layer is called the output layer, and is used to process the outcome for the dependent variables such as the etching rate. Hidden layers stay in between the input and output layers, and provide interconnection between the layers. The connection could be fully connected or partially connected. For a fully connected ANN model, each node on the first layers is connected to every node on the second layer. The direction of the connection can be fed forward and bi-directional. For feed forward connection, the nodes on the first layer send their output to the nodes on the second layer, but they do not receive any input back from these nodes. As for a bi-directional connection, there is another set of connections carrying the output of the nodes of the second layer into the nodes of the first layer.

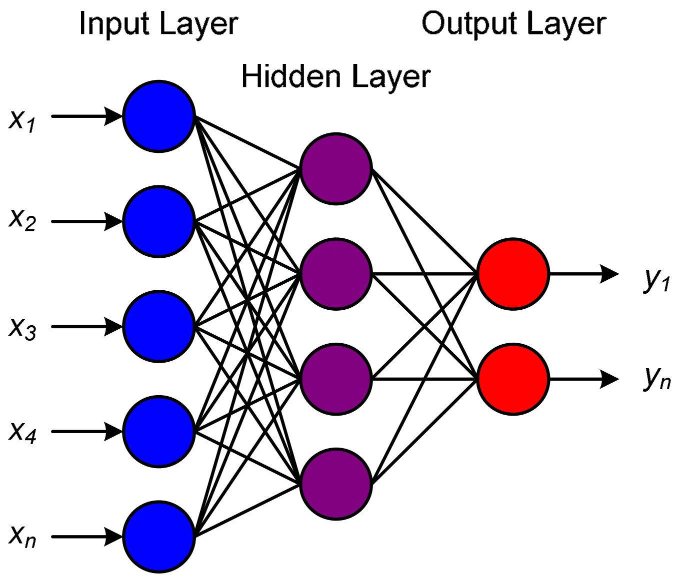



Figure 4.13 Back-propagation ANN Architecture

Artificial neurons or nodes are the core component in ANN, which is an information processing unit that is fundamental to the operation of an ANN (Haykin, 1999). There are three basic elements in each neuronal model:

1. A set of synapses or connecting links, each of which is characterized by a weight or strength of its own
2. An adder for summing the input signals, weighted by the respective synapses of the neuron; the operations described here constitute a linear combiner
3. An activation function for limiting the amplitude of the output of a neuron

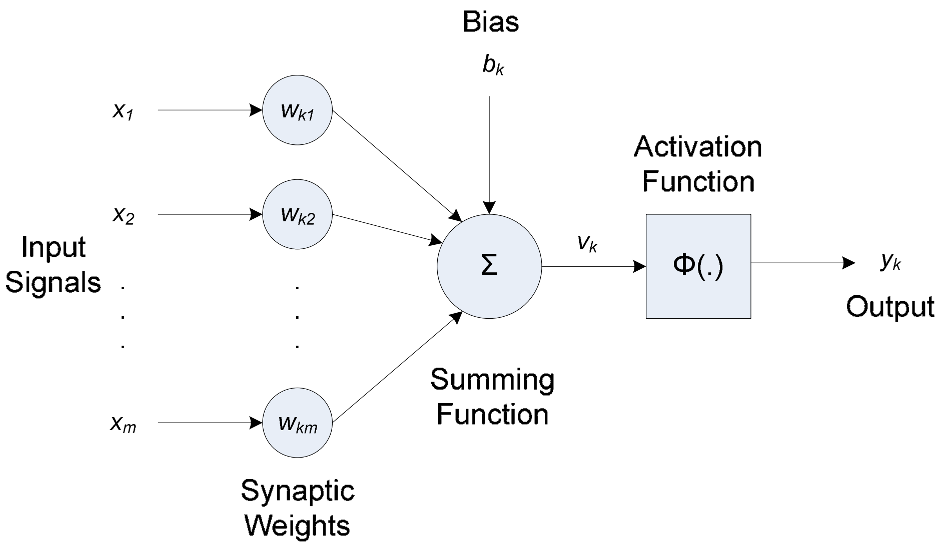



Figure 4.14 Nonlinear Model of a Neuron (Haykin, 1999)

The neuronal model of Figure 4.14 shows the above three elements have an externally applied bias, which has the effect of increasing or lowering the net input of the activation function. A neuron *k* can be described by the following equations (10) and (11):

$$u_k = \sum_{j=1}^{m} w_{kj} x_j \tag{10}$$

$$y_k = \varphi(u_k + b_k) \tag{11}$$

Where $x_1$, $x_2$, …, $x_m$ are the input signals; $w_{k1}$, $w_{k2}$, …, $w_{km}$ are the synaptic weights; $u_k$ is the linear combiner output due to the input signals; $b_k$ is the basis; $\varphi(\cdot)$ is the activation function; and $y_k$ is the output signal of the neuron

The synaptic weight links between the processing nodes serve in a critical role during the learning process. They are part of the memory capacity of the ANN. Since the numerical values of the weight factors change according to the training data sets in order to minimize the differences between the actual outputs and model predicted outputs. Thus, the relationship between causal factors and response is mapped during the learning process.

The activation function, denoted by $\varphi(v)$, defines the output of a neuron in terms of the induced local field $v$. The sigmoid function is used in this research as the activation function. It is defined as a strictly increasing function that exhibits a graceful balance between linear and nonlinear behavior (Haykin, 1999). Equation (12) is the logistic sigmoid function:

$$\varphi(v) = \frac{1}{1 + \exp(-av)} \tag{12}$$

Where $a$ is the slope parameter of the sigmoid function

By varying the parameter $a$, the sigmoid function of different slopes is obtained. It is sometimes desirable to have the activation function range from -1 to +1, in which case the activation function assumes an anti-symmetric form with respect to the origin; that is, the activation function is an odd function of the induced local field. Specially, the threshold function is defined as (13):

$$\varphi(v) = \begin{cases} 1 & if \quad v > 0 \\ 0 & if \quad v = 1 \\ -1 & if \quad v < 0 \end{cases} \tag{13}$$

For the corresponding form of a sigmoid function, the hyperbolic function (14) is used:

$$\varphi(v) = \tanh(v) \tag{14}$$

Back-propagation algorithm has an additional feature compared with other algorithms as it learns by iteratively processing a set of training samples, and comparing the network's prediction for each sample with the actual known class label (Han and Kamber, 2001). For each training sample, the weights are modified so as to minimize the mean squared error between the network's prediction and the actual class.

The number of hidden layers and number of hidden nodes in each layer is strongly dependent on the complexity of the problems such as the number of input and output variables, the number of training sets and required prediction accuracy, which need to be studied. The optimal number of hidden nodes depends on the following factors (Ning, 2003):

- Number of input and output nodes
- Number of training data sets
- Amount of noise in the targets
- Complexity of the function or classification to be learned
- Architecture
- Type of hidden node activation function
- Training algorithm

According to the case of RIE, the ANNs inputs can be classified into three catalogues, numeric (nominal), numeric (ordinal) and attribute (Figure 4.15). The nominal inputs such as thickness, and the machine parameters such as the voltage, current, gas flow rate and etch angle are used directly for the ANNs, while the ordinal and attribute inputs are required to be transformed into a column structure in advance. Each individual unique "value" in the inputs is seen as a new item combined with the original input's name as a new column in binary format. For example, the five RIE machines in the attribute input, D2E#1, D2E#2, D2E#3, D2E#4 and D2E#5, will be transformed into five column inputs as Machine_No_ D2E#1 to Machine_No_ D2E#5 in the ANNs. It is also known as a "dummy value" in some of the programming language. If the new input record is processed in machine D2E#1, the input of column Machine_No_ D2E#1 will be set to "1" and other machine input will be set to "0". After the input data transformation methodology is constructed, the next step is initialing the training, testing and validation process.

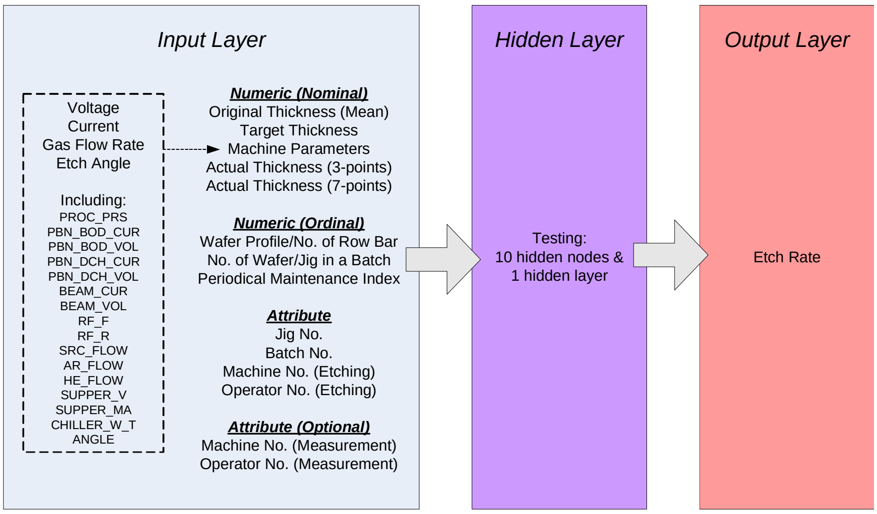



Figure 4.15 ANNs Input and Output for RIE

### 4.3.6 Training, Testing and Validation of ANNs

Before performing the prediction, the ANNs need to be trained. The training process involves the fitting of the data to an ANN model. Training is supervised for an associating ANN model in which the model is presented with input and output datasets. This process adjusts the weighting factors of links among the nodes when the training data, including input and output data, presents to the ANN model. The weights are updated after processing the entire training set for the batch training process through the training algorithm.

In a back-propagation approach, the system is taught by a set of training data where known solutions are supplied to the system. The delta rule back-propagation of errors is used as the training algorithm. During the training process, the network's response at the output layer is compared with the given set of known answers – training targets, which are equivalent to standard deviation that was presented to the system. In order to train a network, the structure of the ANN must be determined with past history of records as training datasets. After an ANN model has been designed and the initial weight factors assigned random small vales, the ANN model can be trained through an iterative process of presenting training datasets to the model and adjusting weight factors. All iterations include two steps, a feed-forward step and a back-propagation step. During the feed-forward step, the training dataset is presented to the model. The processing nodes in the hidden layer sum the inputs based on the randomly assigned weight values and then apply the sigmoid transfer function to compute the outputs. The predicted outputs for the inputs can be obtained at the output layer. During the back-propagation step, the error for the output is calculated first. This is accomplished by comparing the actual output values to the predicted output values. Errors for all the processing nodes are calculated and weight adjustments are then computed for all interconnections. The weight adjustment is then sent back to the model for slight weight correction. This iterative training process keeps on going until the error has reached the criteria. The network therefore would make a better prediction if it sees similar data in the future. The training algorithm adjusts the weights in an attempt to drive the network's response error to minimal. The iterative procedure of processing inputs, determining the errors, and back-propagating the errors through the network to adjust the weights constitutes the learning process.

The training data are separated into three sets: the including training set, testing set and validation set. The training set is used to train the ANN by adjusting the linked weights of the network. This data set should be large enough to contain all the required information and must include a wide variety of data from different experimental conditions. Another proportion of data set is reserved and not to be used in the training process, they are the independent testing data. It is presented to the ANN periodically and is used to check the training process of ANN. There is another important function that is used to keep an independent check on the progress of the algorithm. If the test error stops decreasing or alternatively starts to rise, it indicates that the ANN model is starting to over fit the data. The process should be stopped accordingly. The validation set is then used as the tool to ensure the relationship between inputs and outputs are read and not artifacts of the training process comparing to the training and test sets. The final model usually is tested with the validation data set to confirm its accuracy.

Six months data of RIE data, (from 1$^{st}$ October 2005 to 31$^{st}$ March 2006), was selected as the data set. This is around 43,200 records of RIE process from two types of RIE machines. Approximately 12% of the data was filtered out due to poor quality, such as including missing parameters, incorrect data format, noisy data, human / operation mistakes and abnormal machine data which might affect the training of the model and its accuracy. The abnormal machine parameters and their specifications are:

- PBN_BOD_CUR < 1
- BEAM_CUR < 1000
- BEAM_VOL > 640
- RF_F > 500
- AR_FLOW < 30

Therefore, over 37,500 records were then used for building the model. Before doing the training process, data was extracted, transformed and loaded (Extract, Transform and Load - ETL) in the OLAP database according to the needs of prediction and requirement of ANN. During the data preparation study, there are some more new inputs defined during data transformations which are listed below:

- STOP_TIME is one of the important factors affecting the performance of the machine after several ANN model trainings:
    - STOP_TIME = LOAD_TIME – UNLOAD_TIME
    - If STOP_TIME > 4 hrs, then it is assumed that the machine is idle
    - If 0.75 hr < STOP_TIME =< 4 hrs, then it is assumed that the machine ran dummy in the previous batch
- ACCU_TIME is the accumulating run time which represents the conditions of the machine (cool / warm)
- GRID: A / B represents different materials in RIE

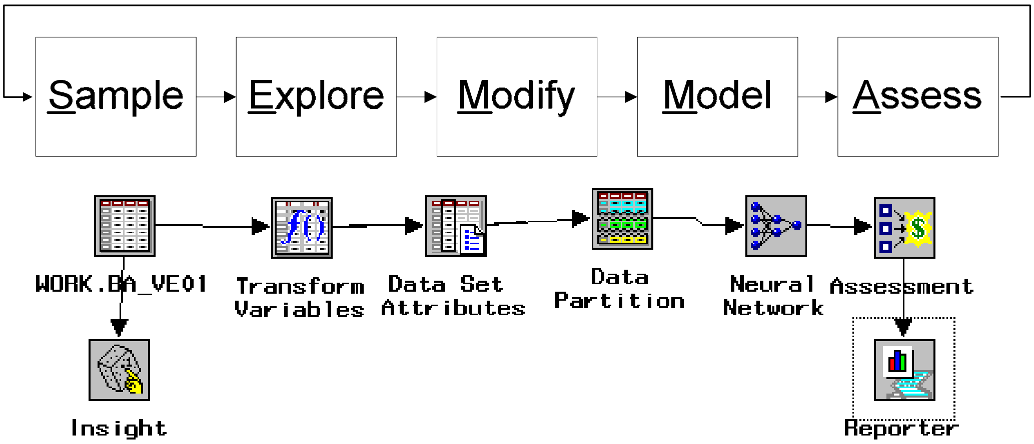



Figure 4.16 SEMMA Approach in SAS Enterprise Miner

The ANN model was created through the SAS Enterprise Miner by the SEMMA approach as shown in Figure 4.16. SEMMA stands for the five important steps of model building (Sample, Explore, Modify, Model and Assess) recommended by the SAS Institute. The model was first built through the user-friendly interface by drag and drop (Figure 4.17). Once the ANN model has been defined, trained and tested, the related structure, design and weights were then transferred to the programming of IDS.

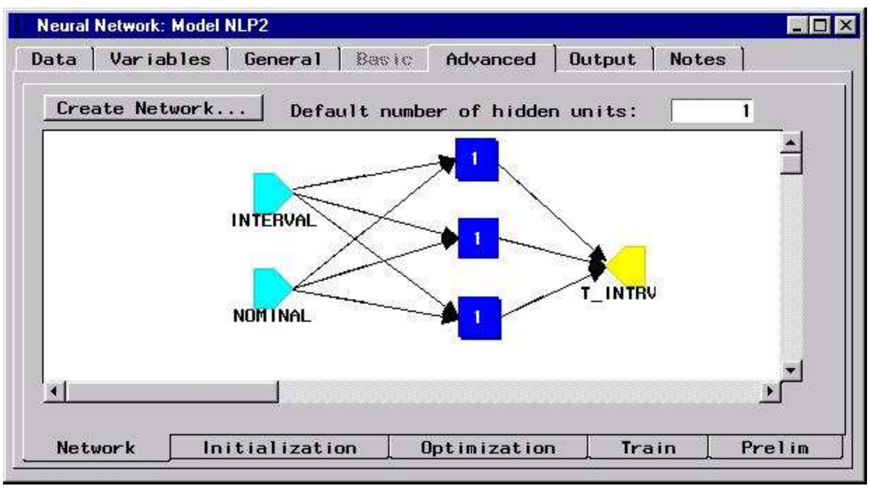



Figure 4.17 User Interface of SAS Enterprise Miner

A simple random approach was used to separate the data, the ratio was set 70% : 30% for the training to testing data sets. The Quickprop in Enterprise Miner was selected as the training algorithm. It usually requires more iterations than conjugate gradient methods, but each single iteration is very fast. Quickprop is a Newton-like method but uses a diagonal approximation to the Hessian matrix. It is much faster and more reliable than other algorithms and requires little tuning. An additional 1,000 records were used to validate the training result and get the score under the assessment node in Enterprise Miner.

Both the learning rate and the momentum coefficient are two parameters that need to be defined for the training. The learning rate is an adjustable factor that controls the speed of the learning process, and its use results in the ANN model being learnt much faster. The oscillation of weight change can impede the convergence of the error surface and may lead to the overshooting of a near-optimal weight factor if the learning rate is too high. In contrast, the ANN may get caught in a local error minimum instead of the global minimum if the learning rate is too slow. Constant learning rates of 0.1 – 10 and 0.4 – 0.6 throughout the training process were proposed by Wythoff (1993) and Zupan & Gasteiger (1993) respectively. The learning process can be facilitated by starting with a high learning rate initially, followed by its gradual reduction. A momentum coefficient is used in weight updating for back-propagation ANN to avoid local minima and to reduce the oscillation in weight change. To obtain faster learning without oscillation, the weight change is related to the previous weight change to provide a smooth effect. The momentum coefficient determines the proportion of the last weight change that is added into the new weight change.

Convergence is the process of searching a set of weight factors for the ANN model so that the prediction errors can be reduced to a minimum. The most common criterion of convergence is based on the sum of squared errors. Supervised ANN measures the difference or error between the predicted output value and the actual output value during the training process. The Sum of Squared Errors (SSE) for the training and test subsets can be calculated using the following equation (15) (Basheer & Hajmeer, 2000):

$$SSE = \frac{1}{N}\sum_{p=1}^{N}\sum_{i=1}^{M}\left(t_{pi} - o_{pi}\right)^2 \tag{15}$$

Where $o_{pi}$ is the actual output of the *i*th output node from the *p*th sample; $t_{pi}$ is the target output of the *i*th output node from the *p*th sample; *N* is the number of training samples; and *M* is the number of output nodes

Training error such as SSE for the training dataset decreases indefinitely with increasing number of hidden nodes or training iterations. The initial quick drop of SSE is done to learning. However, the subsequent slow reduction of SSE could be attributed to memorization or over-fitting because of the excessively large number of training cycles or excessive number of hidden nodes. On the other hand, the test error decreases initially, but subsequently increases due to memorization and over-fitting of the ANN model. Thus, the training would be stopped when the test error starts to increase, and the optimal number of hidden nodes would be picked when the test error is at its minimum.

During the building of several ANN models at the beginning, it was discovered that there were different groups of data included, and the characteristic and behavior of input and output data were quite different. Although model training and network tuning were performed, no significant improvement could be drawn. According to the advice from the experienced engineers in this area and the data pattern discovered and the ad-hoc analysis with clustering, five different groups of model were suggested to be classified (Table 4.9).

Table 4.9 Several Groups of ANN Models

| ANN Group | Characteristic | Percentage |
|---|---|---|
| A | Generic model for most RIE (except those with empty jig in the same batch) | 67.3% |
| B | Model for particular slider design with large area of air groove | 10.1% |
| C | Model for process after machine just warm-up | 14.8% |
| D | Model for machine type A when the machine maintained index is relatively small | 3.1% |
| E | Model for machine type B in case there is empty jig | 4.7% |

Groups of RIE were classified by the decision tree-based approach which was mentioned at Chapter 4.3.4. The input data was then transformed to the format and loaded to the ANN model from MDDB for the training, testing and validation. The four steps of the back-propagation training algorithm were then started by using the approach suggested by Negnevitsky (2004):

- Step 1: Initialization

  Set all the weights and threshold levels to random numbers uniformly distributed inside a small range:

$$\left(-\frac{2.4}{F_i},+\frac{2.4}{F_i}\right)$$

Where $F_i$ is the total number of inputs of neuron *i* in the network. The weight initialization is done on a neuron-by-neuron basis

- Step 2: Activation

  Activate the back-propagation neural network by applying inputs $x_1(p), x_2(p), \ldots, x_n(P)$ and desired outputs $y_{d,1}(p), y_{d,2}(p), \ldots, y_{d,n}(p)$

  a. Calculate the actual outputs of the neurons in the hidden layer:

  $$y_i = sigmoid\left[\sum_{i=1}^{n} x_i(p)w_{ij}(p) - \theta_j\right] \quad (16)$$

  Where *n* is the number of inputs of neuron *j* in the hidden layer, and *sigmoid* is the sigmoid activation function

  b. Calculate the actual outputs of the neurons in the output layer:

  $$y_k(p) = sigmoid\left[\sum_{j=1}^{m} x_{jk}(p)w_{jk}(p) - \theta_k\right] \quad (17)$$

  Where *m* is the number of inputs of neuron *k* in the output layer

- Step 3: Weight Training

  Update the weights in the back-propagation network propagating backward the errors associated with output neurons

  a. Calculate the error gradient for the neurons in the output layer:

  $$\delta_k(p) = y_k(P)\times[1-y_k(p)]\times e_k(p) \quad (18)$$

  Where $e_k(p) = y_{d,k}(p) - y_k(p)$

  Calculate the weight correction:

  $$\Delta w_{jk}(p) = \alpha \times y_j(p) \times \delta_k(p) \quad (19)$$

  Update the weights at the output neurons:

$$w_{jk}(p+1) = w_{jk}(p) + \Delta w_{jk}(p) \tag{20}$$

b. Calculate the error gradient for the neurons in the hidden layer:

$$\delta(p) = y_i(p) \times [1 - y_i(p)] \times \sum_{k=1}^{l} \delta_k(p) \times w_{jk}(p) \tag{21}$$

Calculate the weight corrections:

$$\Delta w_{ij}(p) = \alpha \times x_i(p) \times \delta_j(p) \tag{22}$$

Update the weights at the hidden neurons:

$$w_{ij}(p+1) = w_{ij}(p) + \Delta w_{ij}(p) \tag{23}$$

- Step 4: Iteration

  Increase iteration $p$ by one, go back to Step 2 and repeat the process until the selected error criterion is satisfied

Once the data is inputted to ANN during training, it is also essential to check the model and define the numbers of both hidden layers and hidden nodes. The Gaussian hill diagram was used for the hidden layers and hidden nodes selection. For example, a multi-layer perceptron with an identity output activation function can easily fit the hill by surrounding it with a few sigmoid hidden units, but there will be spurious ridges and valleys where the plane should be flat, as shown in Figure 4.18. It is required to take dozens of hidden units to flatten out the plane accurately as shown in Figure 4.19.

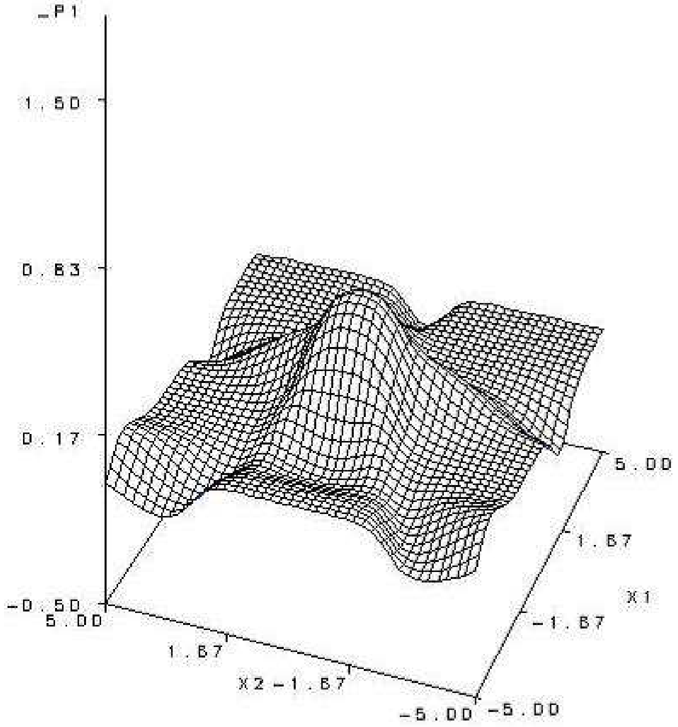



Figure 4.18 Gaussian Hill Diagram with a Few Sigmoid Hidden Units

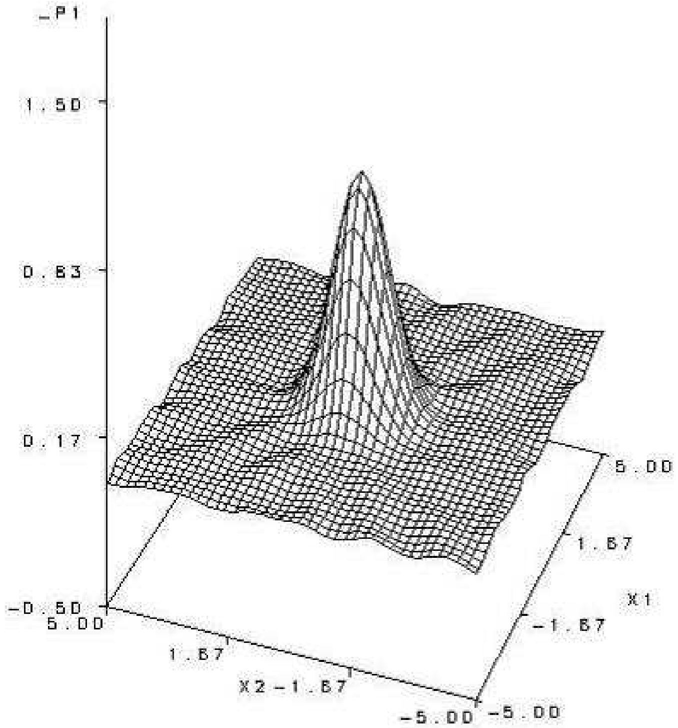



Figure 4.19 Gaussian Hill Diagram with Flatten Plane

After using the approach by hill diagram, one hidden layer was chosen for all five models but with different hidden nodes. The ideal learning constant and momentum were set to 0.1 and 0.5 in order to reduce the possibility of any weight oscillation and controls over how much iteration an error adjustment persists. About three quarters of 37,500 sample datasets were used for the training. Experimental outcomes were evaluated using Root Mean Square (RMS) error in which the less is the better. The normalized root mean square error should be less than 0.01. Typical iteration to fulfill this criterion ranged from 30,000 to 70,000 (Wang, 2003). A multiple correlation coefficient is used to quantitatively evaluate how well the model predicts the physical model (Kleinbaum and Kupper, 1978), which is defined as (24):

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_p - y_i)^2}{\sum_{i=1}^{N}(\bar{y} - y_i)^2} \tag{24}$$

Where $y_p$ is the ANN predicted value, $y_i$ is the testing data point and $\bar{y}$ is the sample mean of the $N$ testing data of $y_i$

By using model A as an example, less than 0.0094 of root mean square error (Figure 4.20) was measured between the set targets and the network output with the error range from 0 to 1 and a correlation coefficient of 0.709 by using 40,000 iterations. The first 250 records comparing the predicted and actual results are listed in Figure 4.21.

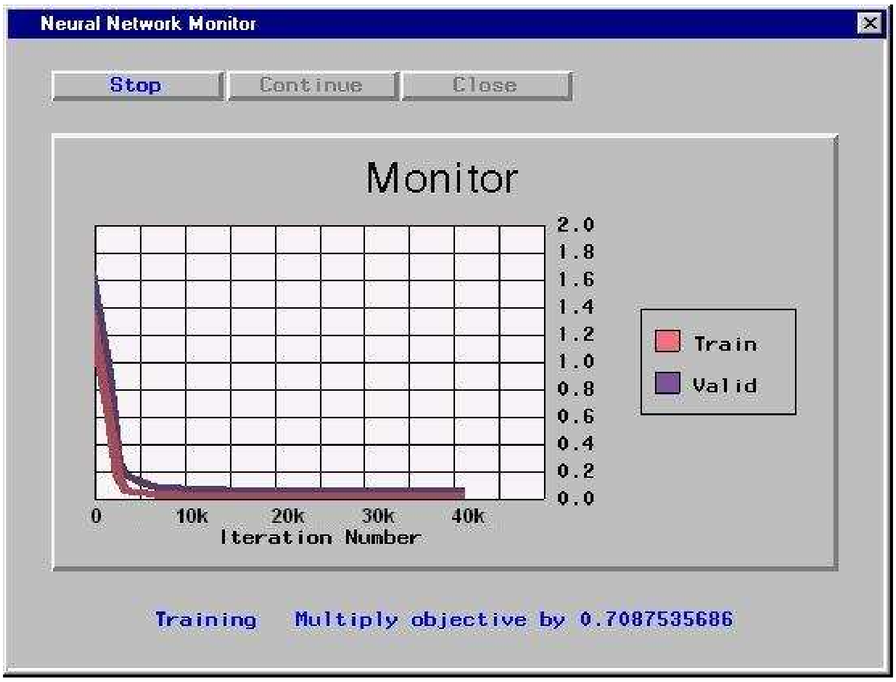



Figure 4.20 Neural Network Monitor of RMS Error

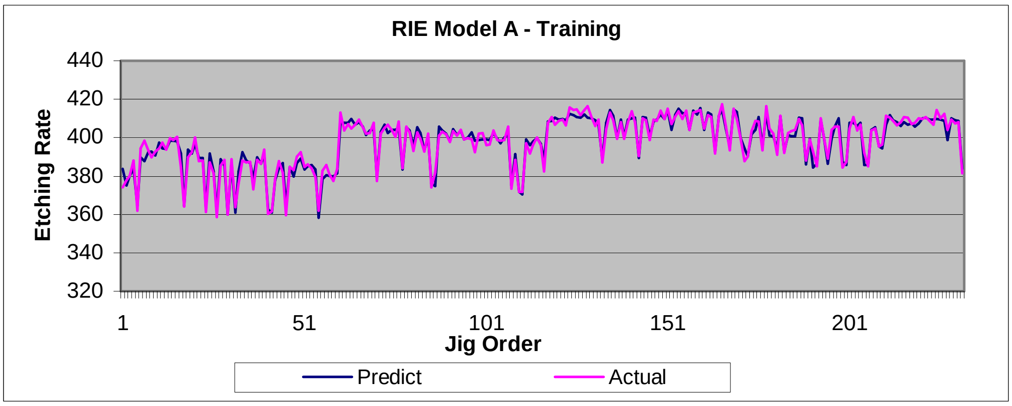



Figure 4.21 Training Result of Model A

The corresponding models, B, C, D and E were using the same methodology to train and validate the ANN models. Detailed summaries are listed in Table 4.10. The validation data was then used to validate the performance of the models construction.

Table 4.10 ANN Models Training and Testing Result

| Model | Training Data Set | Testing Data Set | Correlation Coefficient | RMS Error |
|---|---|---|---|---|
| A | 12366 | 5300 | 0.708 | 0.0094 |
| B | 1856 | 795 | 0.815 | 0.0069 |
| C | 2720 | 1166 | 0.782 | 0.0078 |
| D | 570 | 244 | 0.778 | 0.0072 |
| E | 864 | 370 | 0.802 | 0.0080 |

Comparisons between the actual values and the predicted values from ANNs are shown in Figure 4.22 and Table 4.11. It should be noted that there was only a little data available for model D and E during the experiment due to the limitation of changing production schedule. Approximately 500 and 250 records, however, were used to validate the model (i) A, B, C and (ii) D, E respectively.

Table 4.11 Maximum Deviation of all ANN Models

| Model | Max (Positive) | Max (Negative) | Max (Absolute) |
|---|---|---|---|
| A | 2.55% | -2.67% | 2.67% |
| B | 0.61% | -0.57% | 0.61% |
| C | 1.31% | -1.23% | 1.31% |
| D | 1.08% | -0.82% | 1.08% |
| E | 1.24% | -1.04% | 1.24% |

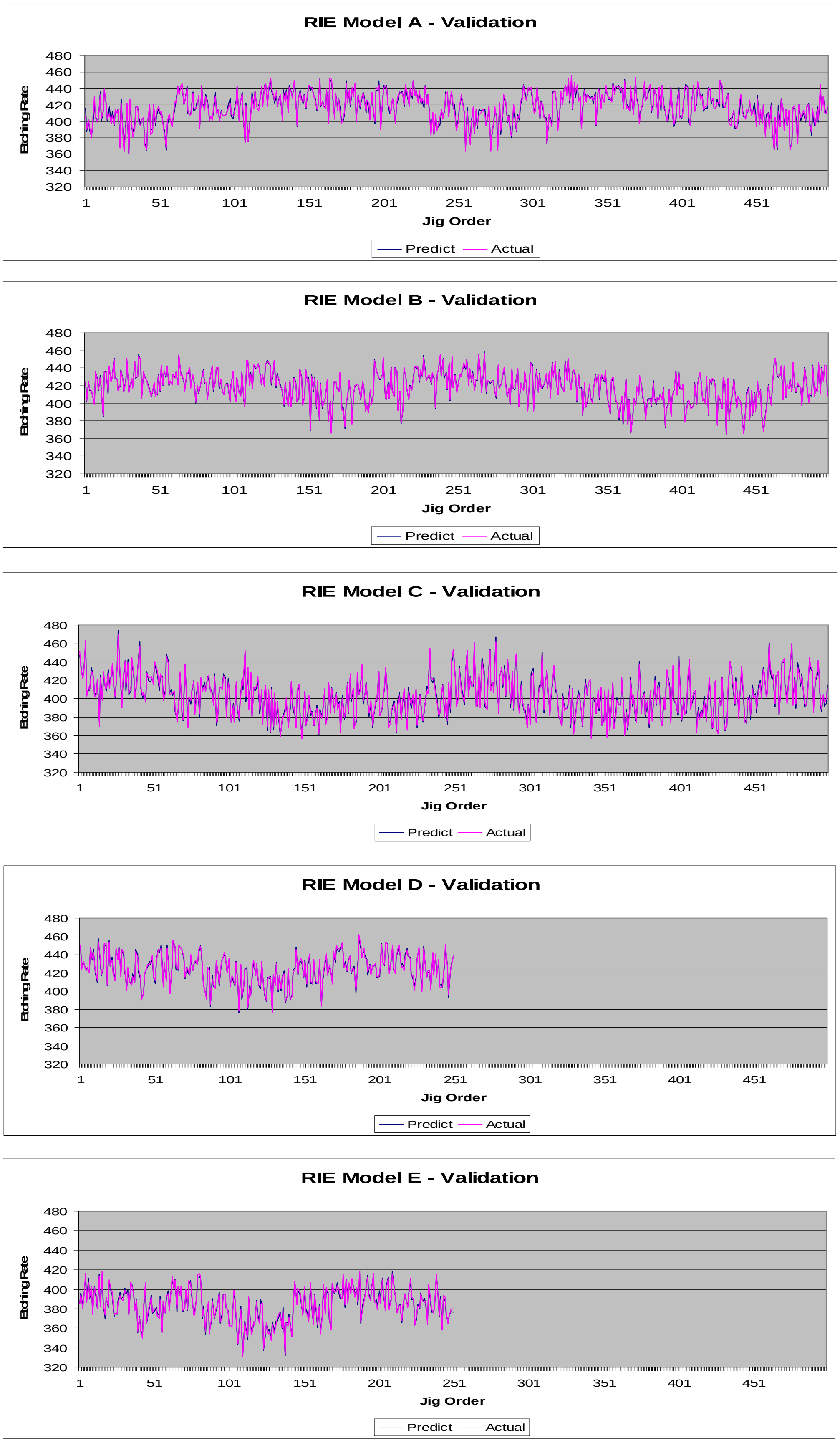



Figure 4.22 Validation Result of Model A, B, C, D & E

The best model is model B with 0.61% maximum deviation between the predicted value and the actual value. Model A has a relatively larger deviation of 2.67% than the other four models. As the maximum deviation is less than ±10% of the actual value, the developed ANN models can be regarded as a suitable tool to predict the etching rate required for the RIE of slider fabrication under the SAE Magnetics (H.K.) Ltd. mass production environment. This decision tree-based ANN approach could also be applied in other process predictions with different sets of sample data. Delighted results were obtained through the visual presentations of web-based information portal.

## 4.4 Evaluation and Summary of Neural-OLAP

The IDS prototype was used in the SAE Magnetics (H.K.) Ltd. for a continuous three-month period to evaluate its performance, to help the engineers to discover the quality problem and to support their decision making for process optimization through the processed data and related graphs. Both OTDB and MDDB architecture were implemented into the whole manufacturing environment from wafer fabrication, slider fabrication, HGA / HSA assembly to HDD production. The predictive module of the decision tree-based ANN focused on the RIE in slider fabrication. After implementing the prototype, the result indicates that significant improvements can be achieved after applying the optimum process parameters in compliance with the adoption of the Neural-OLAP approach. One of the examples in slider fabrication is attached. Testing machine D2E#14 contains one of the testing parameters, MAXSLOPE, this indicator’s abnormal reject rate 3.5% was higher than the other testers. The mode

D auto alert was then detected by the predefined SPC rules in MDDB of IDS (Figure 4.23), which meant that the tester could then be repaired in time and the work process returned to normal, resulting in a saving of sliders 3.5% × 57600 sliders = 2016 sliders per day. Appendix D includes the data processing algorithm for this case.

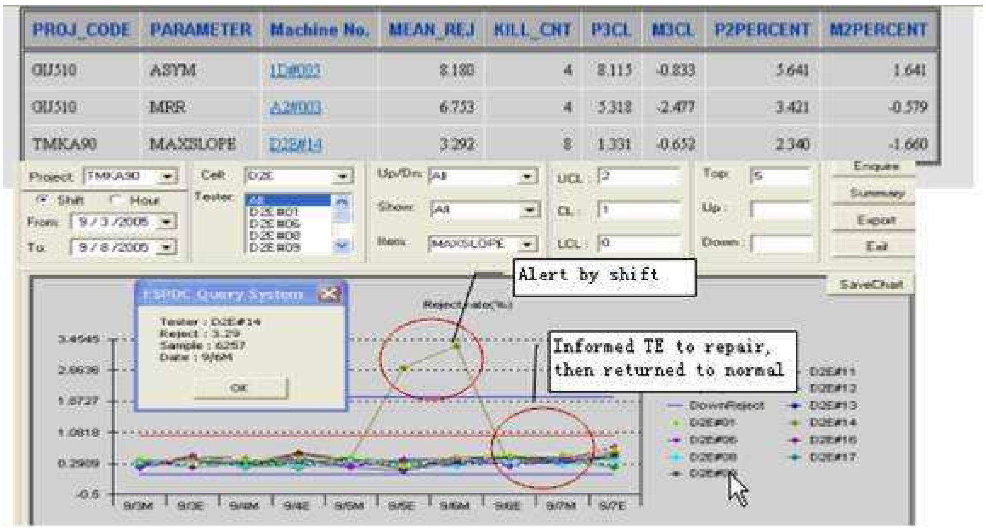



Figure 4.23 Successful Examples in MDDB of IDS

The overall yield of the RIE process has been improved from 87.3% to 94.6%. The quality performance of a finished slider is compared with that using the traditional SPC approach. It is obvious that there is a significant decrease in the total quality costs. The frequency of the need for reworking and the production of scrap are also reduced resulting in better customer satisfaction and a shorter manufacturing time. Screenshots of the system interface with the related graph and report are attached (Figure 4.24a & 4.24b).

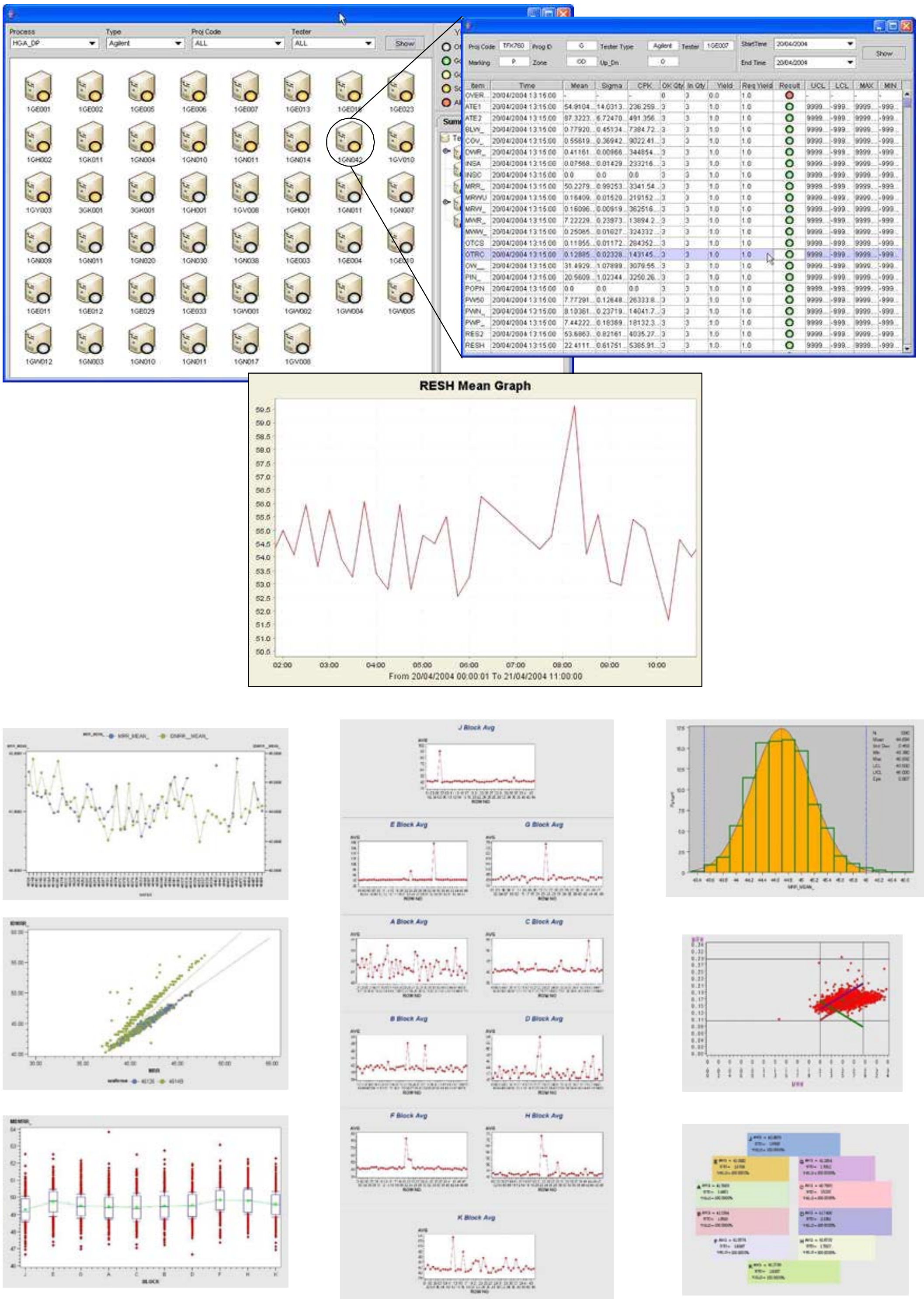



Figure 4.24a Screenshots of IDS

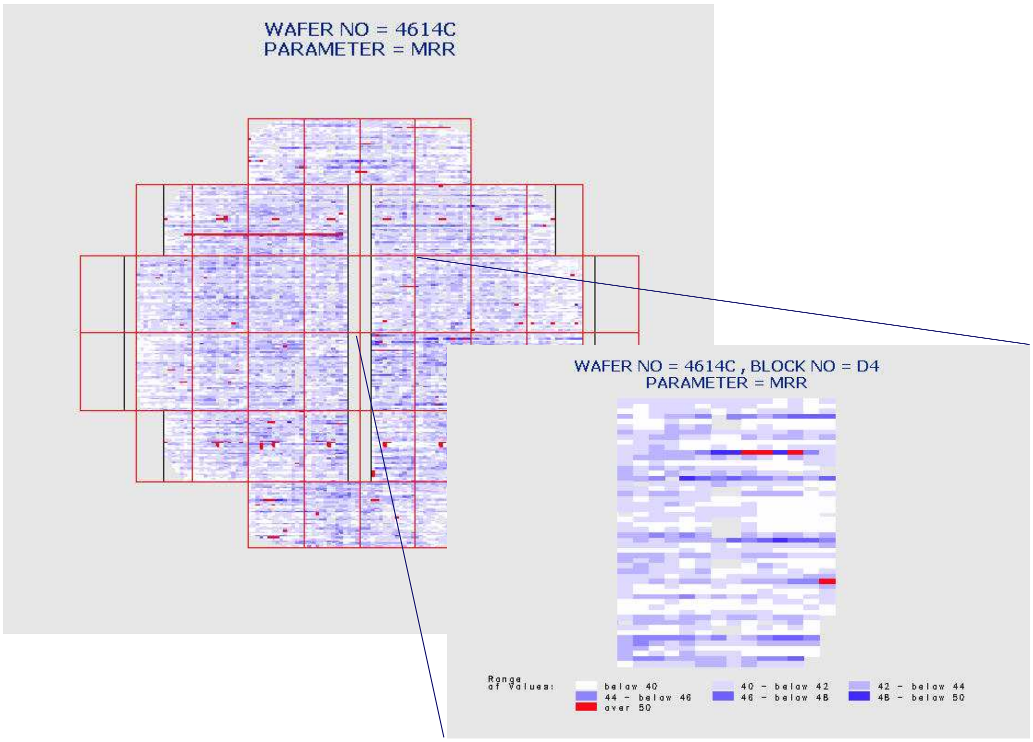



Figure 4.24b Screenshots of IDS

There are, however, also some limitations that IDS cannot handle. Factory shop floor conditions are usually not well-controlled and noise is always present because not all variables in a manufacturing process can be monitored. The datasets collected from production systems are often nonlinear and noisy so that some complex models may not get any creditable results. Further refinements of the proposed methodology are, therefore, needed.

Since IDS was implemented and was considered as a company-based project, there were several evaluations taken to measure the actual performance of IDS, they are:

**Return-Of-Investment (ROI)**

For the point of view of the top management, the Return-Of-Investment (ROI) was used to evaluate the actual performance of IDS in term of values. Total investment cost, including hardware, software and manpower, was compared with the saved direct cost, indirect cost and extra value gained. The lead time of ROI was found around 6 to 12 months time in different implementation sites in SAE Magnetics (H.K.) Ltd.

**Lead-time for New IT Project Development**

Another measurement commonly used in IT project management is the lead-time for the new project development. This study was done by the IT management team. It measured and compared IDS with previous cases for new requirement for IT systems. The lead time of development using IDS approach was reduced from 3 - 6 months to 2 - 4 weeks. It proved the efficient and effective of IDS implementation with high customer satisfaction.

**Cost Productivity, Quality / Time Measurement, and Cause-effect Analysis**

Production department also studied and evaluated the performance of IDS by three different methods. The first one is the cost productivity by the cost / output ratio (25) in monthly time period, comparing with traditional production line:

(i) cost/output ratio = total cost/output value 25

The second method used the failure rate, scrap rate, number of Work-In-Process (WIP), machine setup time and process cycle times to measure the result after IDS implementation.

The cause-effect analysis of using fishbone or Ishikawa diagram is the latest method used to study how the IDS affected the cause-effect of quality failure and make improvement. Since the above studies and related data were considered as confidential commercial information, they could not be shown. Only the methodologies are discussed.

# CHAPTER 5.  DISCUSSION

## 5.1 Introduction

This chapter on the discussion of the results includes the comparison between the traditional approach and IDS, the implication for different groups of people and the limitations of the system.

## 5.2 Comparison of the Traditional Approach with IDS

Before using IDS, most of the tasks for planning the process optimization were done manually. Although some software tools were used, such as Excel and in-house built SPC software (Figure 5.1), the process optimizations were still not so satisfactory. Because the SPC tool was designed mainly for process monitoring and control, the system cannot discover the problems nor suggest related corrective methods. It requires knowledgeable engineers to analyze the results and devise a plan for the appropriate solutions.

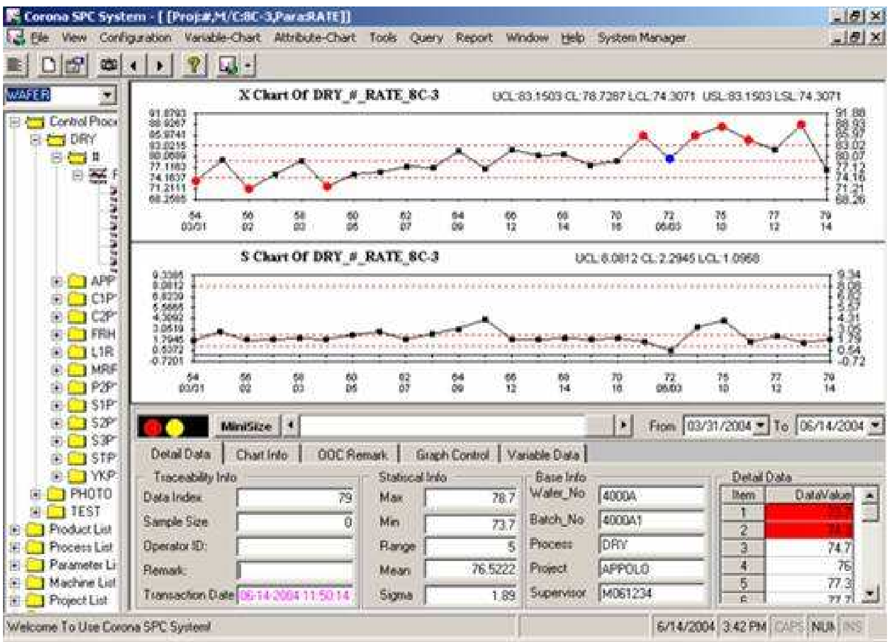



Figure 5.1  In-house Developed SPC Tool

The focus of IDS on a parameter-based manufacturing process, however, is a great breakthrough. Comparatively, previous works have not paid enough attention to the influences of changes to the parameters change which actually are a crucial component contributing to the smoothness of a successful manufacturing process. The newly-proposed method not only provides a theoretical yet scientific way to examine the parameter-based manufacturing process, but also solves the real-life manufacturing industry issues. The marriage of a parameter-based industrial process and a mathematical approach directly leads to a more scientific examination, observation and improvement of the manufacturing operation to achieve process optimization. By adjusting the correlations involved, the whole set of mindset could be used in a generic way. That is, for nearly all kinds of process in the manufacturing environment.

From the IT perspective, whole system architecture with a three-tier real-time data collection structure, a central data warehouse, a data mining engine and a web-based information portal are more reliable and easier to manage than the traditional stand-alone SPC program. This fundamental framework provides a powerful and reusable platform for the continuous process optimization, not only in a single machine or process, but is also designed for the implementation of the whole mass production environment no matter which processes or machines they are. The application software or solution, such as SPC, data traceability, tester monitoring, and all other analysis and data modeling application, can be built into this foundation as a module (Figure 5.2). Some of them can even be integrated together, similar to IDS. It is the enterprise solution for intelligent data analysis by turning the strategies into action, maximizing the customer profitability and proactively managing internal processes.

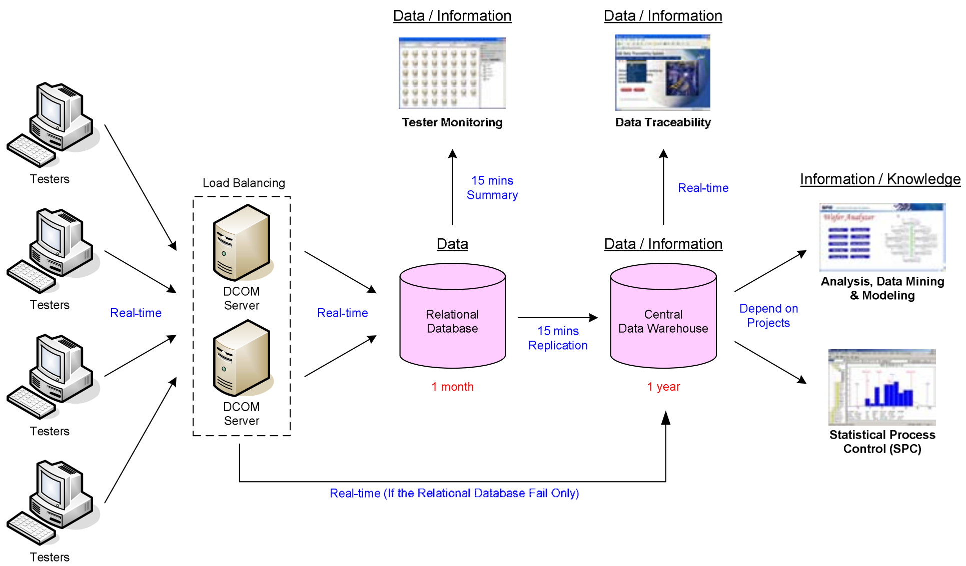



Figure 5.2 Implementation of Architecture Framework in SAE

As for the hybrid Neural-OLAP approach, the model can handle the complex and non-linear relationships among the mass production environment of slider fabrication, especially the RIE, where a more accurate solution could be obtained compared with the conventional approaches. The proper setting of process parameters can shorten both the system development and model tuning times. Determination of process parameters for RIE processing is a highly skilled job and based on a skilled engineer's "know-how" and intuitive sense acquired through long-term experience rather than a theoretical and analytical approach. Faced with the global competition whose emphasis is now on the high quality products and short time-to-market, the previous traditional practices in the determination of the process parameters for RIE seems to be inadequate. The

IDS is therefore one of the best solutions for the process optimization of the complex manufacturing problems nowadays.

## 5.3 Implications of IDS

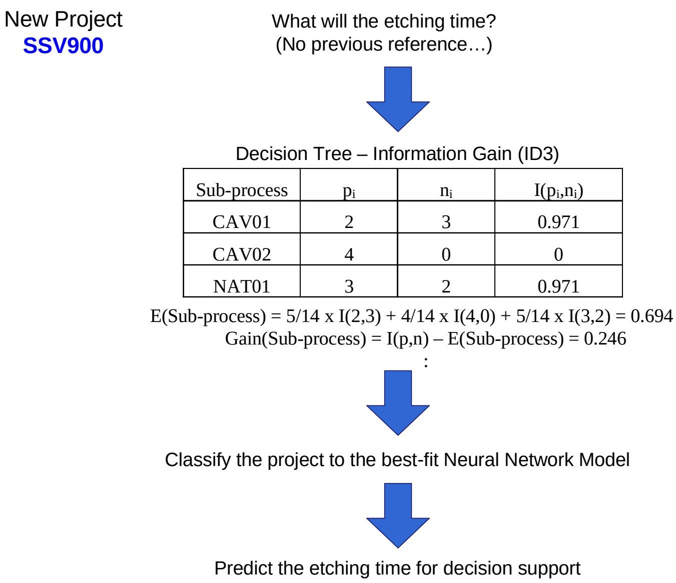
Selecting the most appropriate methodology for use in process optimization under the mass production environment within the semiconductor industry in order to achieve the high manufacturing quality performance is highly challenging. Enhancing product quality and process performance in such an industry is severely constrained by resource conflicts, and limitations and requirements from customers and plant managers. IDS gives an immense implication from operation to top management.

### 5.3.1 Implication for Operation

Operators usually are the people who work on the production shop floor for operating the machines. They are the first people affected by the implementation of IDS. According to the experience and records of the company, it takes a few weeks and up-to a month for the team to become familiar with the new settings or new projects in production. This lead time is relatively too long for the manufacturing of a short product life cycle product, such as the magnetic heads and other electronic components in the competitive market. After the IDS implementation, the lead time is reduced to a few working days according to the results and feedback from the operation. It is believed that the information found and solutions provided by IDS can largely reduce the time and risks in the real world environment. The initial preparation time and quality defects at the beginning of production can be reduced in order to lessen the cost and defects in manufacturing related to human mistakes and errors.

### 5.3.2 Implication for Engineering

Similar to the above, engineers are also benefited by the IDS implementation. Before having IDS, it is required to have a team of engineers to investigate the process parameters when a new project or changing project specification is introduced. The solutions from IDS can largely reduce the time and effort required for the engineers to do those tasks after system implementation. For example, the model classifier in Neural-OLAP helps engineers to define the new project in a systematic way by relating it to the previous model that has the matched characteristics according to the historical records (Figure 5.3). Risks of having quality failure and production defects are thereby reduced.




| Sub-process | $p_i$ | $n_i$ | $I(p_i,n_i)$ |
|---|---|---|---|
| CAV01 | 2 | 3 | 0.971 |
| CAV02 | 4 | 0 | 0 |
| NAT01 | 3 | 2 | 0.971 |

Figure 5.3 New Project Classification through Model Classifier

Relatively more appropriate and accurate solutions with strong supporting evidence can be provided to engineers for them to analyze the models and take action. They can, therefore, handle more projects in a shorter period of time with less workload on each project. Furthermore, engineers can know more about the industrial process in depth by the ad-hoc analysis from MDDB. They can discuss the knowledge with experts and modify the inputs and outputs in the system in order to increase its efficiency and accuracy for the continuous process optimization, which includes knowledge sharing and systemic knowledge reuse. Combining the innovation with the company culture, engineers can create their own knowledge and embed the new knowledge in the process so as to increase the smartness of employees and the competitive edge of the company.

### 5.3.3 Implication for Management

At management level, IDS provides them with high-level information for the production line's monitoring, planning and control. Managers can keep track of the quality of the process by plotting presentable graphs with the process data store in OLAP and then they compare the data before and after applying the ANN. Figure 5.4 shows an example of using trend chat and distribution chart in IDS to monitor the production line and product quality performance while different parameters. Normally, different machine data, different process data and even different department data are stored in the same data warehouse which can be accessed by different department managers, of course including the top management. Managers can monitor the process status and compare its capacity with other processes through comparing the optimal process data.

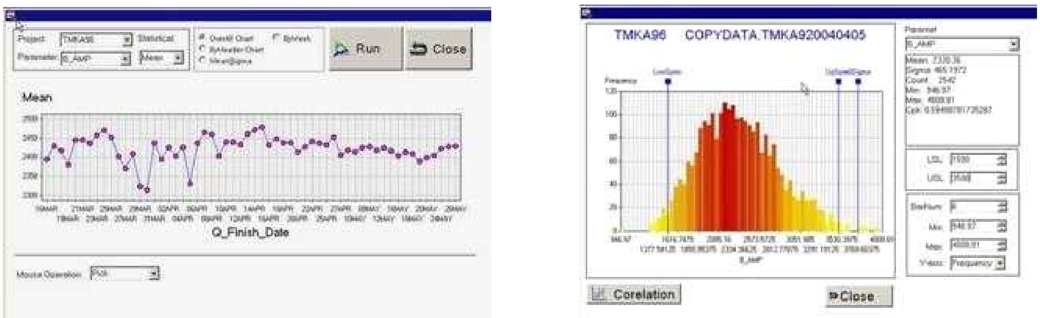



Figure 5.4 Using IDS to Monitor the Production Line

Although IDS only includes the hybrid decision tree-based ANN approach at this moment, managers can export the data into other analysis packages for doing the classification, clustering and further analysis of data. Once the analysis methodology is defined, it can be developed as the module in IDS for future reuse. The IDS foundation of OTDB and MDDB can also be applied in other areas.

### 5.3.4 Implication for SMEs

With a view to the short term, the production costs can also be decreased once after the processes have been optimized. Since the number of trial runs for investigating the optimal parameters is also reduced, the scrap is reduced and the material costs can be reduced at the same time. As a result, the risks to production, especially for a new product or project with new specifications, are also eliminated. Company investment can be reduced or re-allocated to other areas and the competitive advantage can be extended. Consequently, the Return on Investment (RIO) will grow and the company market share will be expanded.

From a long term perspective, the organization culture has also been changed by using this innovative analysis tool. According to Leung (2004), the optimization experts expect that engineers ask them to do optimization initiatively, rather than spreading the technology throughout the company. Thus, they are comfortable doing things in the way they have always done them (Parmee and Hejela, 2002). Insufficient accuracy of mathematical models, too little time for optimization, difficulty in iterating, and difficulty in setting objectives are also some of the major reasons that the process enhancement cannot be implemented successfully (Pardalos and Korotkikh, 2003). By introducing the IDS into the SMEs, the organization culture has been changed due to the de-structuraling. Companies now set up a process optimization team which includes different kinds of employees such as project engineers, manufacturing engineers, quality engineers, production and material planner supervisors, industrial engineers, testing equipment engineers, IT engineers, etc. to work together. In order to optimize all the company processes, not only the industrial processes, but also the designing process, delivering process, engineering change process were also included. It is believed that a person with a manufacturing and industrial systems background is the most appropriate to lead the optimization project within such an environment and situation. Finally, the support from top management, of course, is the most important factor in ensuring the project's success.

A side effect was also discovered during the system's development and implementation that had not been considered before. The importance of the IT infrastructure and architecture affect the possibility of the system and its final results. The ideas of having a well-designed system framework including the data warehousing and system application architecture (such as three-tier data collection and web-based information portal) proved to be the essential element as the production system is different from the other financial, marketing or supply chain management systems. The industry is dynamic, such that related parameters and requirements are changed from day to day. The “hard-coded” design for rapid programming development cannot handle the changes anymore. The latest idea of Object Oriented Architecture (OOA) and even Service Oriented Architecture (SOA) gives a remarkable insight into IDS. Further development by using the SOA approach and Extendable Markup Language (XML) for data exchange will be the future trend as more powerful and flexible techniques lead to further production system development. Also, IDS, based on the SOA design, can be extended by using the modules with web service enabled design.

## 5.4 Limitations of IDS

There are several limitations of IDS, including the ANN algorithm, process implementation, time and cost, and system implementation.

### 5.4.1 Limitation of ANNs

The accuracy of a neural network depends heavily on the desired deviation value and the design of network patterns, that is, the number of neurons and the number of hidden layers. The determination of these parameters requires multiple training in trial and error. First of all, a correlation analysis between all possibilities should be performed, such as the example shown in Figure 5.5. The correlations between all the nodes in RIE for ANN development were tested in JMP. ANNs require experienced engineers or ANN experts to analyze the data and design the predictive models. Technical training is required for juniors in case that they are responsible for developing the models themselves.

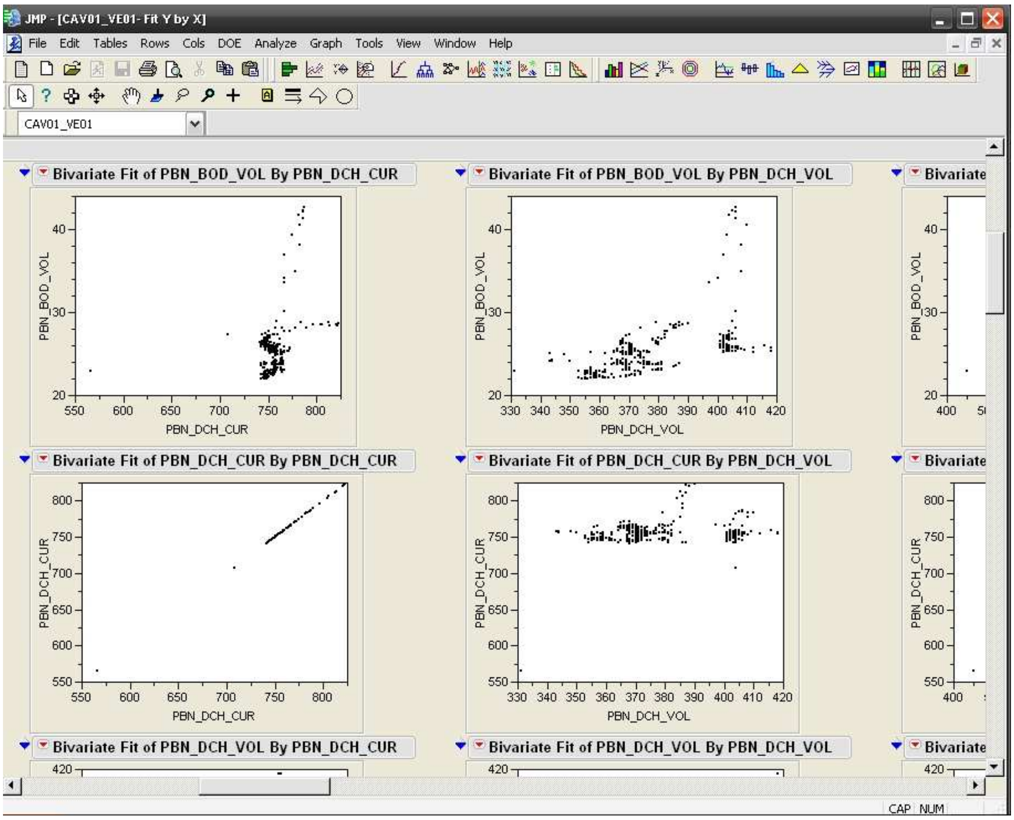



Figure 5.5 Correlations of Nodes in RIE

The maintenance and amendment of the system requires a thorough understanding of the system and the relationships among the parameters. The developer (can either be a business analyst or a system analyst with both IT and engineering background) is responsible for learning the needs of the enterprise, and for being able to define requirements in order to design enterprise wide information systems.

As for the ANN processing, the larger the number of hidden layers and nodes, the more time is needed to perform the training due to the fact that each node has a set of weights associated with it and each set of weights must be adjusted during the training process. The larger number of nodes, the more data that will be required to prevent over fitting or memorizing the data. In addition, the more complicated the network becomes, the less likely it is to find good solutions. When designing a neural network, collecting, analyzing, and preprocessing of data would also consume a lot of time. Not only that, the model needs to be re-trained, re-tested and re-validated regularly to ensure its accuracy.

### 5.4.2 Limitation of Process Implementation

There are also some limitations that IDS cannot handle. Factory shop floor conditions are usually not well-controlled and noise is always present because not all variables of a manufacturing process can be monitored. Therefore, not all manufacturing processes can be implemented into IDS. Data audit should performed to investigate the data availability and its quality before applying IDS. Otherwise it will become “garbage in, garbage out” and nothing can be benefited, while a lot of effort and resources, however, have been spent.

### 5.4.3 Limitation of Time and Cost

Due to the long setup time for the RIE, it takes nearly an hour for an RIE process and up-to three hours for the related machine calibration, process preparation and thickness measurement tasks, and still only a small amount of good quality example data was available for validation. “Clean” historical data were therefore used for building the ANN models instead. Another major reason is that one batch of sliders are costly, approximately HK$ 300,000 to 400,000. It will be pricey if whole batches of sliders are over-etched due to wrong settings for etching rate and time. As time goes on, more process information could be accumulated, thus a more reliable ANN model could be developed which providing a more accurate source for engineers on which to make decisions.

### 5.4.4 Limitation of System Implementation

It is not recommended to apply the system to any process without an in-depth investigation has been carried out. The quality of the optimal result will be affected if exact numbers of input data are unknown or missed. The model will be meaningless and cannot reflect the actual performance and prediction needs. Similarly, it is not recommended to apply the system to the process if the industrial process has only a few parameters which contain few choices. Using AI techniques of course can solve these problems through the intelligent models. The predicted result is more or less similar to those using the traditional methods or fundamental theories. The time, cost, efforts and resources are relatively too "expensive" compared with the traditional approaches.

## 5.5 Software Implementation Aspects and Major Challenges

The whole IDS can be considered as a software implementation project for SAE Magnetics (H.K.) Ltd. for solving the production quality problem and achieving process optimization. It started with production and process study, then followed with user requirement collection, overall system analysis, architectural framework development, building of mining model and algorithm, prototype programming, system quality testing, system pilot-run, user acceptance test, system implementation, quality and performance evaluation, etc. There are some major challenges faced during the whole software implementation:

- Since the scope and the size of the project were comparatively large, it was one of the largest IT systems, which could be compared with the Enterprise Resource Planning (ERP). The system development took so long and required quite a large of manpower to get them done on time.
- The business logics and user requirements were changed from time to time; a lot of efforts were required for the re-programming for those changes. A more flexible design therefore is quite important for development such system in the future. The Service Oriented Architecture (SOA) is recommended.
- Another major challenge of IDS was many departments involved during implementation, the research student and other developers were required to work with and interview quite a large amount of users, analyzed and consolidated their requirements.
- For the whole IDS, a lot computer technical skills were required, including Visual Basic (VB), Java, Oracle, SAS, Windows Server, Linux, Apache, etc. Research student and developers needed to learn and covered most of the techniques to get the project done.

# CHAPTER 6. CONCLUSION AND FURTHER WORKS

## 6.1 Conclusion

In this research project, a hybrid intelligent approach of using OLAP and decision tree-based ANNs for optimizing manufacturing processes is discussed. The implementation of the IDS was successful and the university and the cooperated company were very satisfied. The results when comparing IDS's use with the use of the traditional approach showed the effectiveness and efficiency of the proposed hybrid methodology. The following is a brief summary of the main contributions and findings of this research:

1. Based on the findings of traditional quality control and management, a new approach is then formulated for the step-by-step development of the infrastructure. The key elements and techniques for building this quality enhancement system were discovered. The infrastructure framework – the Neural-OLAP engine – was designed for the actual environment of a magnetic head manufacturer and the broad review on the methodology of recent artificial intelligence and data mining technologies. It is believed that this infrastructural framework can be used in other electronic component manufacturing industries for quality prediction, quality control and production process optimization.
2. A new data warehouse, by using the object technology, was implemented which collects all the production data from the testing machines on the shop floors and delivers it to the core database in real-time manner. It contains six-months of production data from more than 20 processes for around1 TG size, including all the major

information from wafer fabrication to hard disk drive manufacturing. An engineer can easily identify the sources of process failure by multi-dimensional ad-hoc analysis. The evaluated result is delighted providing a well-centralized and good data quality environment for future analysis and development.

3. By applying the idea of a model classifier, the decision tree-based algorithm is collaborated with the Neural-OLAP engine thereby gaining the benefit of handling new conditions without the training data set. It can help to classify the new input to the corresponding best suitable ANN model with the largest information gain based on certain input characteristics. A more accurate and reliable result can thus be predicted.
4. The decision tree-based Neural-OLAP algorithm was tested & five specific models for Reactive Ion Etching (RIE) were built based on the fine-tuned number of hidden nodes. A reliable and systematic etching time prediction can then be made. The preliminary results were highly satisfactory.
5. The research project gave a new perspective and idea to the company for the process optimization and quality enhancement in a systematic and efficient way. It also encouraged the further research cooperation in research between the academia & industry as a result of the successful implementation of this project.

The case study in slider fabrication proves the effectiveness of IDS prototype implementing in an actual industrial site. It is expected that this proposed infrastructural framework will contribute to other similar industries, especially Small and Medium Enterprises (SMEs), providing a key and an introduction to the use of such techniques for gaining benefits through process optimization. By using the developed IDS, substantial improvements and long-term benefits can be obtained in the following aspects:

- Process operations are optimized based on data mining results
- The process associated with the cause of defects can be positively identified, thus greatly improving the future trend of related quality problems
- Production lead time can be shortened due to more efficient assembly operations
- Substantial cost saving will be realized through the reduction in the defect percentage of the magnetic components

The system was implemented into the industrial process of a slider manufacturer. The overall performance of the system was evaluated with the use of ANN theory. Results of the validation indicate that Neural-OLAP can determine a set of optimal process parameters quickly without relying on personal experience while the end products of the process can reach the company’s acceptable defect level. Engineers can find the optimal process settings and machine tuning for the new project or product within several days thus reducing the processing time compared with using the traditional trial-and-error methods. Engineers can also organize the knowledge captured from the system for knowledge management. Mangers can monitor the process status and capability with other processes through comparing the optimal process data. The data mining and data exploring can be done by the Neural-OLAP that the managers can investigate the process trends and business strategy. Finally, it is believed that the ROI can be increased and benefit the business. Other SMEs can use the same methodology of IDS framework for building their own system modeling to achieve process optimization and sharpen their competitive edge.

## 6.2 Further Works

Based on the current study and the above conclusions, the following further works is suggested:

1. The results of this research study further confirmed that a statistical approach combined with an artificial intelligence approach is a highly efficient and reliable way of providing credible models of the process. The next step in this research study will be to apply these models to identify the economical conditions for optimizing the RIE.
2. The current hidden layer and hidden node selection was based on the Gaussian hill diagram, the effective approach of hidden layer and hidden node selection is required, especially the reliable error-tuned automatic selection approach will be the future trend.
3. The classification-based hybrid would be able to help other AI or data mining techniques to enhance the predictive capability, further research on the model classification of new conditions without historical training data in ANN is recommended.
4. The development and further investigation of the Service Oriented Architecture (SOA) as the framework of the intelligent systems is

essential, as it is believed to eliminate the system development time in the future and the resources can be reused.

5. Apart from ANN, other AI techniques, such as fuzzy logic, Genetic Algorithm (GA), Case-based Reasoning (CBR) can also be applied to other industrial processes for gaining more benefits for process optimization.

## REFERENCE


Basheer, I.A. and Hajmeer, M. Artificial Neural Networks: Fundamentals, Computing, Design and Application. *Journal of Microbiology Methods*, 2000, Vol. 43, 3-31.

Bayes, T. *Philos. Trans.* London, Royal Society, 1763, Vol. 53, 370-418.

Ben Amor, N., Benferhat, S. and Elouedi, Z. Naïve Bayes vs Decision Trees in Intrusion Detection Systems. *ACM Symposium on Applied Computing*, 2004 420-420.

Bequette, B.W. *Process Dynamics: Modeling, Analysis and Simulation*. New Jersey: Prentice Hall, 1998.

Besterfield, D.H. *Quality Control*, 7th Ed. New York: Pearson Prentice-Hall, 2004.

Bob, P and Coronel, C. *Database System: Design, Implementation and Management*, 5th Edition. Boston, Mass.: Course Technology, 2002.

Boillat, E., Kolossov, S., Glardon, R., Loher, M., Saladin, D. and Levy, G. Finite Element and Neural Network Models for Process Optimization in Selective Laser Sintering. *Part B of the Proceedings of Intuition of Mechanical Engineers, Journal of Engineering Manufacture, 2004*, Vol. 218, 607- 614.

Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.J. *Classification and Regression Trees*. Belmont: Wadsworth, 1984.

Chaudhuri, S. and Dayal, U. An Overview of Data Warehousing and OLAP Technology. *ACM SIGMIN Rec.*, 1997, Vol. 26, 65-74.

Churchland, P.S. and Sejnowski, T.J. *The Computational Brain*. Cambridge, MA: MIT Press, 1992.

Cios, K., Pedrycz, W. and Swinarshi, R. *Data Mining Methods for Knowledge Discovery*. Massachusetts: Kluwer, 1998.

Connoly, T. and Begg, C. *Database Systems: A Practical Approach to Design, Implementation, and Management*, 4th Edition. Essex, England: Pearson Education, 2005.

Dahr, V. and Stein, R. *Seven Methods for Transforming Corporate Data into Business Intelligence*. Englewood Cliffs, NJ: Prentice Hall, 1997.

DARPA Neural Network Study. *A Connectionist / Neural Network Bibliography*. MIT Lincoln Lab, 1989.

Elmasri, R. *Fundamentals of Database Systems*, 3rd Ed. Reading, Mass.: Addison-Wesley, 2000.

Fahn, C.S, Lan, K.T. and Chern, Z.B. Fuzzy Rules Generation Using New Evolutionary Algorithm Combined with Multilayer Perceptrons. *IEEE Transition on Industrial Electronics*, 1999, Vol. 46, No. 6, 1103-1113.

Fan, H.T. and Wu, S.M. Case Studies on Modeling Manufacturing Process Using Artificial Neural Networks. *American Society of Mechanical Engineers (ASME) Winner Annual Meeting: Neural Networks in Manufacturing and Robotics*, Nov. 1992, PED-Vol. 57, 127-141.

Fayyad, U. Thesis, *On the Induction of Decision Trees for Multiple Concept Learning*. Ann Arbor: The University of Michigan, 1991.

Fayyad, U., Piatetsky-Shapiro, G. and Smyth, P. The KDD Process for Extracting Useful Knowledge from Volumes of Data. *Communication of ACM*, 1996, Vol, 39, 27-34

Forcht, K.A and Cochran, K. Using Data Mining and Data Warehousing Techniques. *Industrial Management and Data Systems*, 1999, Vol. 99, No. 5, 189-196.

Green, M. *Knowledge Aided Design*. London: Academic Press, 1993.

Grochowski, E. Presentation, *HDD Technology Overview 2003*. San Jose: Hitachi Global Storage Technologies, San Jose Research Center. 2003. (Available from: http://www.hitachigst.com/hdd/hddpdf/tech/hdd_technology2003.pdf)

Gruska, G. and Kymal, C. Use SPC for Everyday Work Process. *Quality Progress*, Jun. 2006, Vol. 39, No. 6, 25-32.

Han, J. and Kamber, M. *Data Mining: Concepts and Techniques*. San Diego, CA: Academic Press, 2001.

Han, S.S., Ceiler, M., Bidstrup, S.A., Kohl, P., and May, G. Modeling the Properties of PECVD Silicon Dioxide Films Using Optimized Back-Propagation Neural Networks. *IEEE Trans. Component, Packaging and Manufacturing Technology*, Part A, 1994, Vol. 17, No. 2, 174-182.

Haouani, M., Lefebvre, D., Zerhouni, N. and Elmoudni, A. Neural Networks Implementation for Modeling and Control Design of Manufacturing Systems. *Journal of Intelligent Manufacturing*, 2000, Vol. 11, 29-40.

Haykin, S. *Neural Network: A Comprehensive Foundation*, 2nd Ed. Upper Saddle River, New Jersey: Prentice Hall, 1999.

Himmel, C.D. and May, G.S. Advantages of Plasma Etch Modeling Using Neural Network Based Modeling of the Plasma Enhanced Chemical Vapor Deposition of Silicon Dioxide. *Proceedings of IEEE International Electronic Manufacturing Technology Symposium*, 1993, 458 -463.

Ho, G.T.S., Lau, H.C.W., Ip, W.H. and Ning, A. An Intelligent Information Infrastructure to Support the Streamlining of Integrated Logistics Workflow. *Expert Systems*, Jul. 2004, Vol. 21, No. 3, 123-137.

Ho, G.T.S., Lau, H.C.W., Lee, C.K.M. and Ip, A.W.H. An Intelligent Forward Quality Enhancement System to Achieve Product Customization. *Industrial Management and Data System*, 2005, Vol. 105, No. 3, 384-406.

Hong, S.J. and May, G.S. Neural Network-based Real-time Malfunction Diagnosis of Reactive Ion Etching Using in situ Metrology Data. *IEEE Transactions on Semiconductor Manufacturing*, Aug. 2004, Vol. 17, No. 3, 408-421.

Inmon W.H. *Building the Data Warehouse*. New York: Jon Wiley and Sons, 1996.

Jenkins, M.W., Mocella, M.T., Allen, K.D. and Swain H.H. The Modeling of Plasma Etching Processes Using Response Surface Methodology. *Solid State Technology*, Apr. 1986, 175-182.

Joseph, B. Wang F.H. and Shieh, D.S.S. Exploratory Data Analysis: A Comparison of Statistical Methods with Artificial Neural Networks. *Computers Chem. Eng.*, 1992, Vol. 16, No.4, 413-423.

Kaya, M. and Alhajj, R. Fuzzy OLAP Association Rules Mining-based Modular Reinforcement Learning Approach for Multiagent Systems. IEEE *Transactions on Systems, Man and Cybernetics – Part B: Cybernetics*, Apr. 2005, Vol. 35, No. 2, 326-338.

Kagami, T., Kuwashima, T., Miura, S., Uesugi, T., Barada, K., Ohta, N., Kasahara, N.' Sato, K., Kanaya, T., Kiyono, H., Hachisuka, N., Saroki, S., Inage, K., Takahashi, N. and Terunuma, K. A Performance Study of Next Generation's TMR Heads Beyond 200 Gb/in$^2$. *IEEE Transactions on Magnetics*, Feb 2006, Vol. 42, No. 2, 93-96.

Khizroev, S. *Perpendicular Magnetic Recording*. Boston: Kluwer Academic, 2004.

Kim, B., Kim, D. W. and Park, G. T. Prediction of plasma etching using a polynomial neural network. *IEEE Transactions on Plasma Science*, Dec. 2003, Vol. 31, No. 6, 1330-1336.

Kim, B. and Kim, S. GA-optimized Backpropagation Neural Network with Multi-parameterized Gradients and Applications for predicting Plasma Etch Data. *Chemometrics and Intelligent Laboratory Systems*, 2005, Vol. 79, 123-128.

Kleinbaum, D.G. and Kupper L.L. *Applied Regression Analysis and Other Multi-variable Methods*. North Scituate: Duxbury Press, 1978.

Krishnamoorthy, C.S. and Rajeev, S. *Artificial Intelligence and Expert System for Engineers*. Florida: CRC Press, 1996.

Lee, S.J. and Siau, K. A Review of Data Mining Techniques. *Industrial Management and Data Systems*, 2001, Vol. 101, No. 1, 41-46.

Leung W.K. Thesis, *Generic Methodology for the Design and Development of an Intelligent Optimization System*. Hong Kong: The Hong Kong Polytechnic University, 2004.

Leung, R.W.K., Lau, H.C.W. and Kwong, C.K. An Expert System to Support the Optimization of Ion Plating Process: an OLAP-based Fuzzy-cum-GA Approach. *Expert Systems with Applications*, 2003, Vol. 25, 313-330.

Liu, T.C., Li, R.K. and Chen, M.C. Development of an Artificial Neural Network to Predict Lead Frame Dimensions in an Etching Process. *International Journal of Advanced Manufacturing Technology*, 2006, Vol. 27, 1211-1216.

Mendel, J.M. and McLaren, R.W. Reinforcement-learning Control and Pattern Recognition Systems. In Mendel, J.M. and Fu, K.S. eds. *Adaptive Learning, and Pattern Recognition Systems: Theory and Applications*, New York: Academic Press, 1970.

Microsoft Corporation. Database Migration Nets Manufacturer 100% Uptime and ROI in Six Months. *Microsoft Windows Server System Customer Solution: Case Study*. USA: Microsoft Corporation, 2003.
(Available from: http://download.microsoft.com/documents/customerevidence/6856_final_western_digitalGCRP.doc)

Minsky, M.L. and Papert, S.A. *Perceptrons*. Cambridge, MA: MIT Press, 1969.

Mitchie, D., Spiegelhalter, D.J., and Taylor, C.C. *Machine Learning, Neural and Statistical Classification*. New York: Ellis Horwood, 1994.

Mitchell, T.M. *Machine Learning*. Singapore: McGraw-Hill, 1997.

Monson, J.E. Capturing Data Magnetically. In Daniel, E.D. Mee, C.D. and Clark, M.H. *Magnetic Recording, The First 100 Years*. New York: IEEE Press, 1999.

Narendra, K.S. and Parthasarathy K. Identification and Control of Dynamical System Using Neural Network. *Neural Networks, Trans. IEEE*, 1990, Vol. 1, No. 1, 4-27.

Negnevitsky, M. *Artificial Intelligence: A Guide to Intelligent Systems*, 2$^{nd}$ Edition. Essex, England: Pearson, 2004.

Ning, A. Thesis, *Hard Chrome Plating for Ting Grooves of Piston Crown*. Hong Kong: The Hong Kong Polytechnic University, 2003.

Nirmalakhandan, N. *Modeling Tools for Environmental Engineers and Scientists*. London: CRC Press, 2000.

Nguyen, D.H. and Widrow, B. Neural Networks for Self-Learning Control Systems. *IEEE Control Systems Magazine*, 1990, 19-23.

Oakland, J. S. *Statistical Process Control*, 3rd Edition. Boston: Butterworth-Heinemann, 1996.

Pang, C.C. Dissertation, *The Studies of Wall Profile Effect on Taper-less Air Bearing in Magnetic Hard Disk Drive*. Hong Kong, The Hong Kong Polytechnic University, 1998.

Paplambros, P.Y. and Wilde, D.J. *Principles of Optimal Design: Modeling and Computation*. Cambridge: Cambridge University Press, 2000.

Pardalos, P.M. and Korotkikh, V. *Optimization and Industry: New Frontiers*. Netherlands: Kluwer Academic Publishers, 2003.

Parmee, I.C. and Hajela, P. *Optimization in Industry*. London: Springer, 2002.

Paul, B.D. *Database System*, 3rd Edition. Basingstoke: Palgrave Macillan, 2004.

Quinlan, J.R. Induction of Decision Trees. *Machine Learning*, 1986, Vol. 1, 81-106.

Rao, S.S. *Engineering Optimization Theory and Practice*. New York: Wilet-Interscience Publication, 1996.

Rastogi, R. and Shim, K. PUBLIC: A Decision Tree Classifier that Integrates Building and Pruning. *Data Mining and Knowledge Discovery*, 2000, Vol. 4, 315-344.

Rietman, E.A., Patel, S.H. and Lory, E.R. Modeling and Control of a Semiconductor Manufacturing Process with an Automata Network: An Example in Plasma Etch Processing. *Computers Operation Research*, 1996, Vol. 23, No.6, 573-585.

Rutledge, R.A. Data Warehousing for Manufacturing Yield Improvement. *SAS and IBM White Paper*, 2000, Paper 134-25, SAS Institute.
(Available from: http://www.sas.com/partners/directory/ibm/wp_222.pdf)

Samarskii, A.A. and Mikhailov, A.P. *Principles of Mathematical Modeling: Ideas, Methods, Examples*. London: Taylor and Francis, 2002.

Shaffer, R.E. Dissertations, *Optimization Methods for the Multivariate Analysis of Infrared Spectral and Interferogram Data*. Athens: Ohio University, 1996.

Spanos, C.J., Guo, H.F., Millar, A. and Levine-Parrill, J. Real-time Statistical Process Control Using Tool Data. *IEEE Transaction on Semiconductor Manufacturing*, 1992, Vol. 5, No. 4, 308-318.

Stevens, L.D. Data Storage on Hard Magnetic Disk. In Daniel, E.D. Mee, C.D. and Clark, M.H. *Magnetic Recording, The First 100 Years*. New York: IEEE Press, 1999.

Takeda, H., Veerkamp, P., Tomiyama, T. and Yoshikawa, H. Modeling Design Process. *AI Magazine*, Winter 1990, 37-48.

Tan, P.N., Steinbach, M. and Kumar V. *Introduction to Data Mining*, 1st Ed. Boston, Mass: Pearson Addison Wesley, 2006.

Thompson, J.R. and Koronacki, J. *Statistical Process Control, The Deming Paradigm and Beyond*, 2nd Ed. Florida: Chapman and Hall, 2002.

Two Cows. *Introduction to Data Mining and Knowledge Discovery*, 3rd Ed. Potomac, MD: Two Cows Corporation, 1999.

Utgoff, P.E. and Brodley, C.E. An Incremental Method for Finding Multivariate Splits for Decision Trees. *Machine Learning: Proceedings of the Seventh International Conference*, 1990.

Veloeo, M. and Borrajo, D. Learning Strategy Knowledge Incrementally. *Proceedings of Sixth International Conference on Tools with Artificial Intelligence*, 1994, 484-490.

Wang, K.J., Chen, J.C. and Lin, Y.S. A Hybrid Knowledge Discovery Model Using Decision Tree and Neural Network for Selecting Dispatching Rules of a Semiconductor Final Testing Factory. *Production Planning and Control*, Oct. 2005, Vol. 16, No. 7, 665-680

Wang, S.X. and Taratorin, A.M. *Magnetic Information Storage Technology*. USA: Academic Press, 1999.

Wang, X.A. *Process Modeling, Optimization and Control Using Artificial Neural Networks*. Michigan: UMI, 2003.

Westkämper, E. and Schmidt, T. Computer-assisted Manufacturing Process Optimization with Neural Networks. *Journal of Intelligent Manufacturing*, 1998, Vol. 9, 289-294.

Whitney, M.C. and Fowler, L. Motorola's Engineering Data Analysis System: 10 Years of Analytical Excellence. *Proceedings of the 25th Annual SAS User Group International Conference*, 2000, Paper 44-25.
(Available from: http://www.units.muohio.edu/doc/sassystem/sugi25/25/ad/25p044.pdf)

Wu, S.M., Ni, J. and Hu, S. Next Generation Quality Control in Manufacturing – Real-time Defect Prevention. *Proceedings of the 1989 American Society of Mechanical Engineers (ASME) Winner Annual Meeting*, 1989.

Wythoff, B.J. Backpropagation Neural Network: A Tutorial. *Chemometr. Intelligence Laboratory System*, 1993, Vol. 18, 115-155.

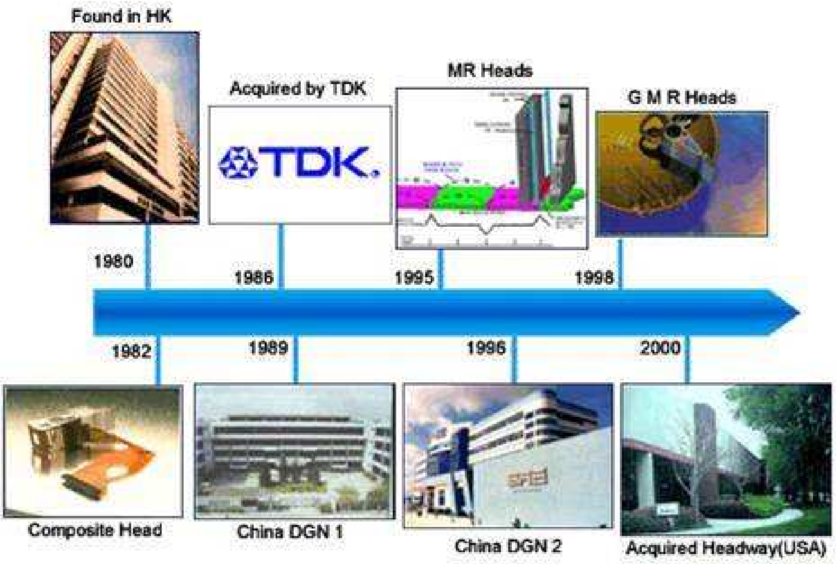
Zant, P. V. *Microchip Fabrication: A Practical Guide to Semiconductor Processing*, 4th Ed. USA: McGraw-Hill, 2000.

Zupan, J. and Gasteiger, J. *Neural Networks for Chemists: An Introduction*. New York: VCH, 1993.

# APPENDIX A. RESEARCH COMPANY BACKGROUND

**SAE Magnetics (H.K.) Ltd.**, a wholly owned TDK subsidiary, is one of the world's leading independent manufacturers of magnetic recording heads, Head Gimbals Assemblies (HGAs) and Head Stack Assemblies (HSAs) for computer disk drivers. SAE also supplies high performance recording heads for video tape recorders.

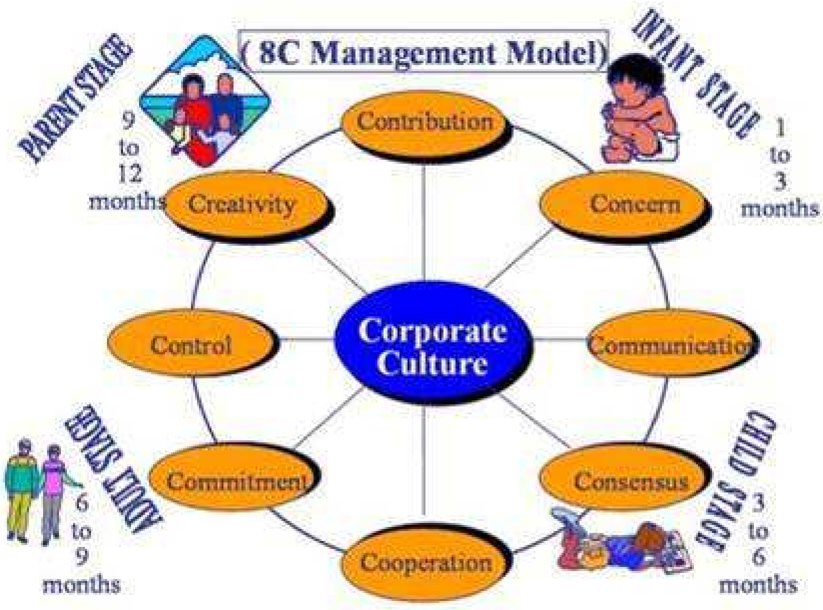
SAE Magnetics (H.K.) Ltd. was established in October 1980 to meet the increasing demand for magnetic recording heads. In August 1986, SAE joined the TDK Corporation Group, the world's largest and most advanced ferrite manufacturer. Since then, the combination of SAE's technologies in head manufacturing and TDK's expertise in magnetic materials has brought SAE to a leading position in head manufacturing.




SAE continues to provide the best support to customers by developing and producing the best magnetic recording heads, including the advanced Magneto-resistive (MR) technologies and the Giant MR (GMR) technologies. SAE is headquartered in Hong Kong and has multiple manufacturing operations in China.

## Corporate Culture

SAE's Corporate Culture is 8C Culture (8C Management Model). They are:




- Concern – love and care for team members and help achieve team performance;
- Communication – ability to express own ideas and have the proper skills open minded team objectives;
- Consensus – ability to execute general agreement in undertaking task, responsibility and target;
- Cooperation – ability to cooperate with others, willing to be a contributing team player;
- Commitment – make convincing promise while taking and go all out reach the target;
- Control – ability to well-arrange in unfavorable situation, willing to summarize lessons into values;
- Creativity – have power to make new or original ideas or things happen, can stimulate innovation in the undertaking field;
- Contribution – willing to take responsibility, and have consistent passion to work for the company.

## Quality Assurance

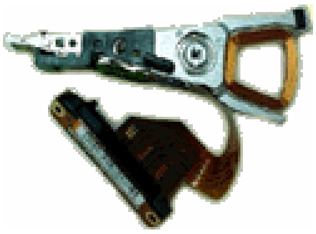
SAE's philosophy is one of continual improvement in all areas of the business. Our Quality Assurance group supports each department to help maintain our policy of quality products delivered on time to achieve total customer satisfaction.

We believe that quality and reliability are "Built-In" by careful attention to details at each process step. This is achieved through strict adherence to our Total Quality Management (TQM) System at all stages of manufacturing. Suppliers are monitored and supported technically to provide SAE with only top quality components. Comprehensive training and certification programs ensure operators possess the up-to-data skills necessary to meet our quality requirements. Line audits backed by rigorous inspection and testing programs provide the feedback necessary for continuous improvements as well as guaranteeing our outgoing quality levels. SAE's dedication in building quality into the products enables us to supply high quality HGA and HSA unmatched in the industry.


SAE's quality system covers all ISO9000 and ISO14000 requirements. We care for environment. SAE has the obligation to support environmental protection and prevention of pollution in balance with socio-economic needs. To comply with international as well as local legal environmental protection requirements, we already established the Environmental Management System (EMS) at our facilities. We always believe that "QUALITY IS THE ONLY WAY TO SURVIVE AND PROSPER IN BUSINESS".

## Company Products

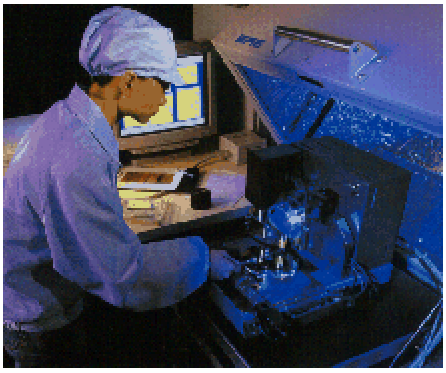
SAE offers a wide range of HGA, HSA and spindle motors to the hard disk industry. We also manufacture various types of video tape recorder Heads used in Video Cassette Recorder (VCR). SAE has built its reputation on supplying both mainstream and specialized state-of-the-art products to the data storage industry with the most favorable combination of quality, cost, and delivery.



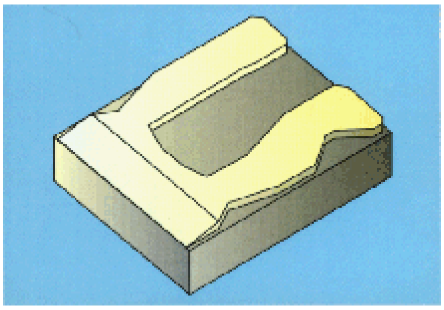
SAE scientists and engineers in our research and development teams have been continuously pushing the boundaries of magnetic for the increasing demand of even greater recording density. The company's extensive research and development efforts ensure that we will be there to supply the industry with future generations of products. From MIG composite, advanced thin film inductive to the latest in MR and GMR technologies, SAE is able to supply turn-key solution to specific customer requirement.

## Product Development

In the dynamic, fast changing world of the HDD business - for instance, 60% annual increase in areal density; rotational speed from 5,400 rpm, 7,200 rpm to 10,000 rpm and beyond, SAE is ready to offer complete services of customers supports for our products from initial product design, prototyping, pre-mass production to mass production until product end of life. Our experienced team of product design engineers utilizes state-of-the-art computer simulations for both magnetic transducer and air bearing design in order to meet the most stringent performance requirements. These simulation models are calibrated and verified with data obtained from our prototype lines in Hong Kong and are constantly refined to provide the most advanced and reliable technologies to our customers.



During the pre-mass production and mass production phases of a project, SAE is able to perform timely failure analysis and conduct implementation of process and product improvements that further enhance functionality and reliability. Backed by our overseas service center in U.S.A., Japan, Singapore, and Korea, our Project Engineering Team is able to respond quickly to provide professional supports necessary to ensure total customer satisfactions. With state-of-the-art advanced analitical equipment such as SEM, AFM, FIB, FTIR, LC, etc., our reliability laboratory is unmatched in its ability to analyze, identify and resolve micro-analitical issues before they become problems.

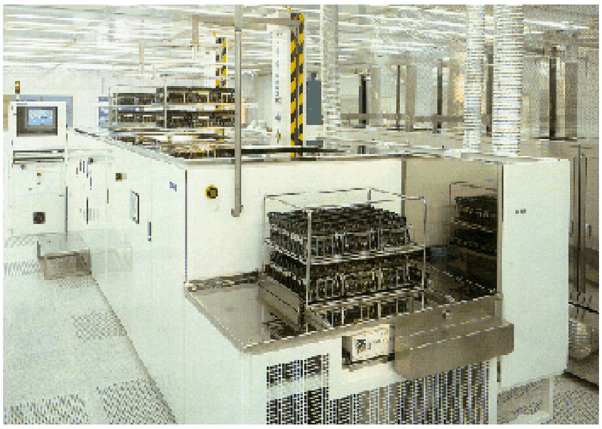


Finally, at the end of a project's life, SAE's Materials Team will ensure that inventories are managed to the benefits of the customers, and will take competent measures to minimize inventory balances while at the same time ensure the flexibility to match the extension of any schedule.

## Slider Fabrication

SAE has over 18-year's experience of building various sliders e.g. MIG composite, thin film inductive and magneto-resistive (MR) sliders. Wafer slicing, one of the first steps performed in the manufacture of MR sliders, demands extreme accuracy. Wafers are sliced into row bars utilizing highly stable precision CNC slicing systems. SAE routinely holds production tolerances to less than plus or minus 1 micron. Air Bearing Surface (ABS) lapping defines the final surface finish and flatness of the slider air bearing surface. Row bars consisting of a specific number of magnetic head transducers are lapped dynamically using real-time measurement feedback to achieve sub-micron MR stripe height tolerance at high volume. Air bearings are formed using semi-conductor type photolithography processes that assure accuracy and uniformity in subsequent etching and coating operations. Combined stepper and photo-mask, the tolerances are held better than plus or minus 0.2 microns. Ion milling and Reactive Ion Etching (RIE) are used to form the complex patterns of the air bearing that ensures correct head flying height profiles. SAE has invested heavily in these areas to ensure process capabilities well into the future. Together with our highly skilled and experienced workforce, this commitment enables SAE to deliver one of the most uniform and stable negative-pressure air bearing sliders available in the market today.

## HGA and HSA Assembly and Testing



Producing HGA and HSA in volume with high yields requires skilled workers using the most sophisticated equipment. At SAE, new methods and the latest equipment are assembled using both fully automated and semi-automated process equipment and then tested and inspected with the latest metrology systems. Using a linear manufacturing approach, SAE builds in the quality at every step of the assembly process while minimizing production inventories and providing quick product turnaround.

# APPENDIX B. TERMINOLOGY IN RESEARCH

| Terminology | Description |
|---|---|
| 4M1E | Man, Machine, Material, Method and Environment |
| AAB | Advanced Air Bearing |
| ABS | Air Bearing Surface |
| AI | Artificial Intelligence |
| ANN | Artificial Neural Network |
| ANOVA | Analysis of Variance |
| BPN | Back-propagation Network |
| CAPE | Computer-Aided Production Engineering |
| CBR | Case-based Reasoning |
| Cpk | Process Capability |
| CPU | Central Processing Unit |
| CSV | Common Separated Variable |
| DBMS | Database Management System |
| DCOM | Distributed Component Object Model |
| DMD | Database Model Diagram |
| EDAS | Engineering Data Analysis System |
| ERD | Entity Relationship Diagram |
| ETL | Extract, Transform and Load |
| FIS | Factory Information System |
| FTP | File Transfer Protocol |
| GA | Genetic Algorithm |
| GMR | Giant Magneto-resistive |
| HA | High Availability |
| HAD | Head Disk Assembly |
| HDD | Hard Disk Drive |
| HGA | Head Gimbals Assembly |
| HSA | Head Stack Assembly |
| I/O | Input and Output |
| IDS | Information Discovery System |
| INDIN | International IEEE Conference on Industrial Informatics |
| IT | Information Technology |
| JED | Just Enough Database |
| KDD | Knowledge Discovery in Database |
| LPCVD | Low-Pressure Chemical Vapor Deposition |
| MDDBMS | Multi-dimensional Database Management System |
| MOLAP | Multi-dimensional OLAP |
| MR | Magneto-resistive |
| MTF | Machined Taper Flat |
| NGQC | Next Generation Quality Control |
| OLAP | On-Line Analytical Processing |
| OO | Object Orientation |
| OOA | Object Oriented Architecture |
| OTDB | Object Technology Database |
| PCB | Printed Circuit Board |

| Terminology | Description |
|---|---|
| PCC | Percentage of Correct Classification |
| RAID | Redundant Arrays of Independence Disk |
| RAMAC | Random Access Method of Accounting and Control |
| RBF | Radial Basis Function |
| RDB | Relational Database |
| RIE | Reactive Ion Etching |
| RMR | Perpendicular Magnetic Recording |
| RMS | Root Mean Square |
| ROI | Return on Investment |
| SAN | Storage Area Network |
| SME | Small and Medium Enterprises |
| SOA | Service Oriented Architecture |
| SPC | Statistical Process Control |
| SQC | Statistical Quality Control |
| SQL | Structured Query Language |
| SSE | Sum of Squared Errors |
| TB | Tetra Bytes |
| TMR | Tunneling Magneto-resistive |
| TPC | o Transverse Pressure Contour |
| TQM | Total Quality Management |
| UICP | University-Industry Collaboration Programme |
| WIP | Work-In Process |
| VB | Visual Basic |
| XML | Extendable Markup Language |

# APPENDIX C. RELATED PROESS DOCUMENTS

SAE Magnetics (H.K.) Ltd.
Dongguan Plant 2

## Process Parameter Notice

No.: VID-005-003
Version : H

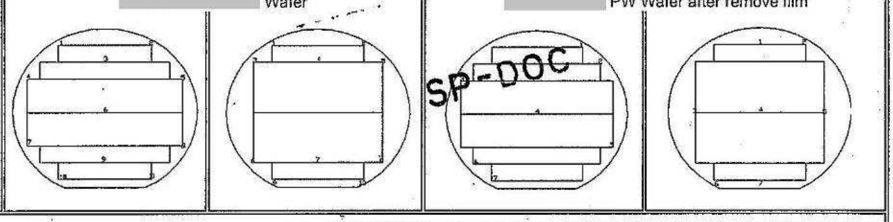
| | | | | | |
|---|---|---|---|---|---|
| Process | RIE DEPTH MEASUREMENT | Ref . Doc. | DAC000-0-M0501/0481/0482/0793 | Effect date: | 07-18-04 |
| Project | ALL | Version : | H | M/C No. | |
| Remark | | | | | |

| parameter | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Machine type | | | | | | | | | | |
| Program name | Master -1 | Master -2 | CAVITY | TTP | FILM | Master -1 | Master -2 | CAVITY | TTP | FILM |
| Scan length[um] | 2000 | 500 | 500 | 200 | 400 | 2000 | 500 | 500 | 200 | 400 |
| Scan speed[um/s] | 100 | 100 | 100 | 20 | 50 | 100 | 100 | 100 | 20 | 50 |
| Sampling rate[HZ] | 200 | 200 | 200 | 100 | 100 | 200 | 200 | 200 | 100 | 100 |
| Multi-scan average | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Scan direction | → | | | | | → | | | | |
| Stylus | | | | | | | | | | |
| Stylus force | 10 | 10 | 10 | 7 | 5 | 10 | 10 | 10 | 7 | 5 |
| Contact speed | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| Vertical ranging | | | | | | | | | | |
| Range/Resolution | 13um/0.0078125A | | | | 65um/0.04A | 26um/0.015625A | | | | 131um/0.357A |
| Profile tyep | | | | | | | | | | |

Measure Map

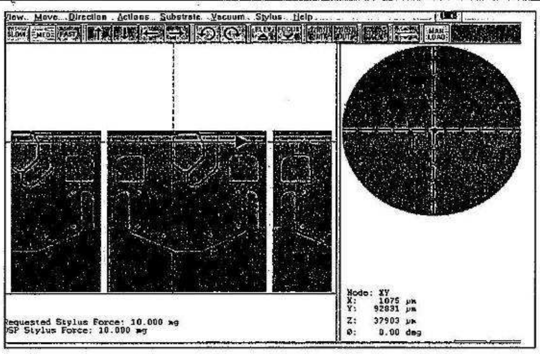
Wafer

PW Wafer after remove film




### 操 作 规 则

1> 将JIG放好后调整Row bar 的水平,请参考右图。
2> 关好送料窗口门。
3> 选择测量位置在pole 区(非玻璃面), 请参考右图。
4> 选择正确程序测量。
5> 在测量时手不可以碰机器和防震台。
6> 每次扫描结束后要先调整扫描线水平,以后
读取和输入正确数据。




Remark :

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Issue by | | Date | 7-17-04 | Check by | | Date | 7/17/04 |

Process Parameters Notes for RIE Depth Measurement

** Sensitive Information was hidden*

SAE Magnetics (H.K.) Ltd.
Dongguan Plant 2

# Process Parameter Notice

No.: VID-005-059
Version: B
Due Date:
M/C D2M&D2E ALL [redacted]

Remark: Shallow recess is measured by [redacted], Main cavity recess by [redacted]
30% all MR&GMR Project

| Project | Shallow Specification USL | Shallow Target | Shallow LSL | Recipe | Cavity 01 Specification USL | Cavity 01 Target | Cavity 01 LSL | Cavity 01 Etch Depth | Recipe | Cavity 02 Specification USL | Cavity 02 Target | Cavity 02 LSL | Cavity 02 Etch Depth | Recipe | Part No. / Ref Doc. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| [redacted] | 1550 | 1450 | 1350 | SWSHX000 | 15400 | 14000 | 12600 | 12550 | CANSX000 | | | | | | 176-7727/7728-01 |
| | 1547 | 1397 | 1247 | SWSHX000 | 16720 | 15200 | 13680 | 13803 | CANSX000 | | | | | | 176-6614/6613-01 176-7744/7743-01 |
| | 1470 | 1330 | 1190 | SWSHX000 | 15675 | 14250 | 12825 | 12920 | CANSX000 | | | | | | 176-7739/7740-01 176-6620/6619-01 |
| | 1550 | 1450 | 1350 | SWSHX000 | 15400 | 14000 | 12600 | 12550 | CANSX000 | | | | | | 176-7739/7740-01 176-6620/6619-01 |
| | | | | SWSHX000 | 14850 | 13500 | 12150 | 13500 | | | | | | | 176-7801/7802-01 |
| | 1634 | 1756 | 1878 | SWSHX000 | 12100 | 11000 | 9900 | 9244 | CANSX000 | | | | | | 176-7773/7774-01 |
| | 1550 | 1450 | 1350 | SWSHX000 | 13805 | 12550 | 11295 | 11100 | CANSX000 | | | | | | 176-6672/6673-01 |
| | 1878 | 1756 | 1634 | SWSHX000 | 15452.8 | 14048 | 12643.2 | 12292 | CANSX000 | | | | | | 176-6666/6667-01 |
| | 1346 | 1270 | 1194 | SWSHN080 | 22352 | 20320 | 18288 | 19050 | CANSX000 | | | | | | 176-6659/6660-01 |
| | 1485 | 1385 | 1285 | SWSHN080 | 23749 | 21590 | 19431 | 20205 | CANSX000 | | | | | | 176-7767/7768-01 |
| | 1117.1 | 1054.1 | 991.1 | SWSHH330 | 17665.7 | 16510 | 15354.3 | 15455.9 | CAHRH330(Veeco) | 34925 | 31750 | 28575 | 15240 | CAHRX000[redacted] | 176-7788/7787-01 176-7790/7789-01 |
| | 1096 | 1016 | 936 | SWSLN025 | 22352 | 20320 | 18288 | 19304 | CANSX000 | | | | | | DENW-00643-03("XN" marking) |
| | 2200 | 2000 | 1800 | SWSHX000 | 17600 | 16000 | 14400 | 14000 | CANSX000 | | | | | | 176-6112/6111-01 |
| | 2600 | 2400 | 2200 | SWSHX000 | 19620 | 18000 | 16380 | 15600 | CANSX000 | | | | | | 176-6114/6113-01 |
| | | | | | 18094 | 16600 | 15106 | 16600 | | | | | | | 176-6652/6653-01 |
| | 2500 | 2300 | 2100 | SWSHX000 | 19620 | 18000 | 16380 | 15700 | CANSX000 | | | | | | 176-5555/5554-01 |
| | 1304.5 | 1219.2 | 1133.9 | SWSHX000 | 12573 | 11430 | 10287 | 10210.8 | CANSX000 | | | | | | 176-6681/6682-01 |
| | 1986 | 1860 | 1734 | SWSHN080 | 20746 | 18860 | 16974 | 17000 | CANSX000 | | | | | | 176-5521/5522-01 |
| | 1980 | 1850 | 1720 | SWSHN080 | 22000 | 20000 | 18000 | 18150 | CANSX000 | | | | | | 176-5537/38-01 |
| | 1986 | 1860 | 1734 | SWSHN080 | 20746 | 18860 | 16974 | 17000 | | | | | | | 176-5690/91-01 |
| | 1382.4 | 1292 | 1201.6 | SWSHN080 | 15991.8 | 14538 | 13084.2 | 13246 | CANSX000 | | | | | | DENW-000384-03("CT" marking) |
| | 1635 | 1545 | 1455 | SWSHX000 | 16888 | 15494 | 14100 | 13994 | CANSX000 | 35249 | 32049 | 28849 | 16555 | CANSX000 | DENW-01492-02 |
| | 1485 | 1391 | 1297 | SWSHX000 | 16888 | 15494 | 14100 | 14103 | CANSX000 | 35249 | 32049 | 28849 | 16555 | CANSX000 | 176-6658-01 |
| | 1500 | 1400 | 1300 | SWSHN080 | 14960 | 13600 | 12240 | 12200 | CANSX000 | | | | | | 176-5531/5532-01 |
| | 2120 | 2000 | 1880 | SWSHN080 | 18700 | 17000 | 15300 | 15000 | CANSX000 | | | | | | 176-6643/6644-01 |
| | 2568 | 2400 | 2232 | SWSHH330 | 19250 | 17500 | 15750 | 15100 | CANSX000 | | | | | | DENW-00349-03("ET" marking) |
| | 2190 | 2050 | 1910 | SWSHH330 | 19800 | 18000 | 16200 | 15950 | CANSX000 | | | | | | 176-7777/7778 |
| | 1590 | 1500 | 1410 | SWSHH330 | 11000 | 10000 | 9000 | 8500 | CANSX000 | | | | | | 176-7779/7780-01 |
| | 1590 | 1500 | 1410 | SWSHX000 | 11000 | 10000 | 9000 | 8500 | CANSX000 | | | | | | 176-7763/7764/01 |
| | 1590 | 1500 | 1410 | SHSHX000 | 11000 | 10000 | 9000 | 8500 | CANSX000 | | | | | | 176-7807/7808-01 |
| | 1306 | 1270 | 1133 | SHSHX000 | 15070 | 13700 | 12330 | 12430 | CANSX000 | | | | | | 176-7855/7856-01 |
| | 1751 | 1651 | 1551 | SWSHHx000 | 13990 | 12954 | 11918 | 11303 | CAHRX000(Veeco) | 31839 | 29210 | 26581 | 16256 | CAHRX000[redacted] | DENW-00980-03("USA3E" marking) |
| | 1481 | 1397 | 1313 | SWSHN080 | 13990.3 | 12954 | 11917.7 | 11557 | CAHRH330(Veeco) | | | | | | 176-6646/47-01 |
| | 1249 | 1168 | 1087 | SWSHH330 | 19202.4 | 17780 | 16357.6 | 16612 | CAHRH330(Veeco) | | | | | | DENW-00926-03 |
| | 1470 | 1330 | 1190 | SWSHX000 | 15675 | 14250 | 12825 | 12920 | CANSX000 | | | | | | 176-7739/7740-01 176-6620/6619-01 |
| | 1420 | 1280 | 1140 | SWSHX000 | 14300 | 13000 | 11700 | 11720 | CANSX000 | | | | | | 176-6635-01 176-7752-01 |
| | 1223 | 1143 | 1063 | SWSHX000 | 16764 | 15240 | 13716 | 14097 | CANSX000 | | | | | | 176-7795-01 |

Unit:Å | Issued by: [redacted] | Date: 6/17-03 | Checked by: [redacted] | Date 6/17/03 | Cross Checked by: [redacted] | Date: 7.1.03

Process Parameters Notes for Shallow and Cavity Specification

** Sensitive Information was hidden*

SAE Magnetics (H.K.) Ltd.
Dongguan Plant 2

## Process Parameter Notice

| | | | | | |
|---|---|---|---|---|---|
| Process | RIE etching | Ref . Doc. | DAC000-0-M0482/ DENW-00639-04 | Effect date: | 6/11/04 |
| Project | All project | Version : | R | M/C No. | |
| Remark | | | | | |

No.: VID-005-007
Version : D

| CANSX | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Gas(Sccm) | | CF4(±15%) | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 |
| | | C2H2(±15%) | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| | | Ar(±20%) | 10 | 10 | 9 | 9 | 10 | 10 | 10 | 10 | 10 |
| Power(W) | Step1 (ESD1) | Coil(±10%) | 500 | 565 | 565 | 565 | 565 | 615 | 615 | 560 | 450 |
| | | Platen(±10%) | 5 | 5 | 5 | 5 | 5 | 0 | 0 | 0 | 0 |
| | | Process Time | 1 Second | | | | | | | | |
| | Step2 (Warm Up)) | Coil(±10%) | | 565 | 565 | | 565 | | | | |
| | | Platen(±10%) | | 200 | 225 | | 200 | | | | |
| | | Process Time | 10 Second | | | | | | | | |
| | Step3 (CANS) | Coil(±10%) | 500 | 565 | 565 | 565 | 565 | 565 | 565 | 565 | 450 |
| | | Platen(±10%) | 400 | 400 | 450 | 450 | 400 | 400 | 400 | 400 | 370 |
| | | Process Time | Set up for every batch | | | | | | | | |
| | Step4 (ESD2) | Coil(±10%) | 500 | 565 | 565 | 565 | 565 | 565 | 565 | 565 | 450 |
| | | Process Time | 1 Second | | | | | | | | |
| Process Pressure(mTorr)(±20%) | | | 3 | 3 | 3.5 | 3.5 | 3 | 3 | 3 | 3 | 3 |
| APC Degree | | | AUTOMATIC | | | | | | | | |

| DLC TRIM-Trailing pattern etching | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Power(W) | Step1 (recipe:SFS-1) | O2(±50%) | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| | | Coil(±10%) | 600 | 600 | 600 | 600 | 600 | 600 | 600 | 600 | 600 |
| | | Process Time | 5 minutes | | | | | | | | |
| | Step2 (recipe:SFS-2) | CF4(±50%) | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| | | Coil(±10%) | 600 | 600 | 600 | 600 | 600 | 600 | 600 | 600 | 600 |
| | | Process Time | 1 minutes | | | | | | | | |
| Process Pressure(mTorr)(±20%) | | | 3 | 3 | 3.5 | 3.5 | 3 | 3 | 3 | 3 | 3 |
| APC Degree | | | AUTOMATIC | | | | | | | | |

| DLC TRIM-PDLC etching | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Power(W) | Step1 (recipe:PDLC-1) | O2(±50%) | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| | | Coil(±10%) | 600 | 600 | 600 | 600 | 600 | 600 | 600 | 600 | 600 |
| | | Process Time | 8minutes | | | | | | | | |
| | Step2 (recipe:PDLC-2) | CF4(±50%) | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| | | Coil(±10%) | 600 | 600 | 600 | 600 | 600 | 600 | 600 | 600 | 600 |
| | | Process Time | 5minutes | | | | | | | | |
| Process Pressure(mTorr)(±20%) | | | 3 | 3 | 3.5 | 3.5 | 3 | 3 | 3 | 3 | 3 |
| APC Degree | | | AUTOMATIC | | | | | | | | |

| PM O2 | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Gas(Sccm) | O2(±10%) | 49 | 49 | 49 | 49 | 49 | 49 | 49 | 49 | 49 |
| Power(W) | Coil(±10%) | 950 | 950 | 950 | 950 | 950 | 950 | 950 | 950 | 950 |
| | Platen(±10%) | 30 | | | | | | | | |
| | Process Time | 45 Minutes | | | | | | | | |
| Process Pressure(mTorr)(±20%) | | AUTOMATIC(use manual APC to keep 40mT) | | | | | | | | |
| APC Degree | | 78° | 78° | 78° | 78° | 78° | 83 | 86 | 78 | 75 |

| Base Setting | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Chiller Temperature(±2°C) | 7 | 10 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| He Flow Rate(sccm) | 2.5±0.2 | | | | | | | | |

**Remark:** **1.He flow: 2.5± 0.2sccm .**
**2. Chiller temperature: 7± 2°C;STS # 3 10± 2°C; : 5± 2°C.**
**3. can not running new recipe at current.**

| | | | | | | |
|---|---|---|---|---|---|---|
| Issue by | date | 6/11-04 | Check by : | | date | Jun-11-04 |

Process Parameters Notes for Machine Type A Setting

** Sensitive Information was hidden*

# Process Parameter Notice

| SAE Magnetics (H.K.) Ltd.Dongguan Plant 2 | | | | No. | VID-004-021 |
|---|---|---|---|---|---|
| | | | | Version : | F |
| Process | | Ref . Doc. | DAC00-0-M0793 | Effect date : | 09-Jun-04 |
| Project | All Project | Version : | N | M/C No. | |

**Recipe Component**

| Recipe Name | Schedule Name | Process Name | Step No.1 | Step No.2 | Step No.3 | Step No.4 | Step No.5 | Step No.6 |
|---|---|---|---|---|---|---|---|---|
| SWSH | SWSH | SWSH | ESDAr | SH | | | | |
| SWSH | SWSH | SWSH | ESDAr | ESDTR2 | H330 | Empty | ESDAr | SH |
| SWSH | SWSH | SWSH | ESDAr | ESDTR1 | N080 | Empty | ESDAr | SH |
| SWSH | SWSH | SWSH | ESDAr | ESDTR1 | N025 | Empty | ESDAr | SH |
| SWSL | SWSL | SWSL | ESDAr | SL | | | | |
| SWSL | SWSL | SWSL | ESDAr | ESDTR2 | H330 | Empty | ESDAr | SL |
| SWSL | SWSL | SWSL | ESDAr | ESDTR1 | N080 | Empty | ESDAr | SL |
| SWSL | SWSL | SWSL | ESDAr | ESDTR1 | N025 | Empty | ESDAr | SL |
| | | | | | | | | |
| TPLR | TPLR | TPLR | ESDAr | LR | | | | |
| | | | | | | | | |
| CAHR | CAHR | CAHR | ESDAr | HR | | | | |
| CAHR | CAHR | CAHR | ESDAr | ESDTR2 | H330 | Empty | ESDAr | HR |

**Step Parameter List**

| Step Name | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Beam Voltage (V) | 350±10 | 400±10 | 500±10 | 400±10 | 350±10 | 175±10 | 700±15 | 400±10 | 500±10 | 0 |
| Beam Current ('mA) | 300±10 | 500±10 | 800±10 | 450±10 | 300±10 | 250±12 | 1000±15 | 500±10 | 800±10 | 0 |
| Supperssor Voltage (V) | -200±10 | -400±10 | -400±10 | -400±10 | -200±10 | -400±10 | -400±10 | -400±10 | -400±10 | 0 |
| RF Forward(W) | 200 | 300 | 500 | 300 | 150 | 300 | 500 | 300 | 500 | 0 |
| RF Ref(Ref .W) | <15 | <15 | <15 | <15 | <15 | <15 | <15 | <15 | <15 | 0 |
| Source Gas Flowrate(Sccm) | 9±2 (Ar) | 9±2(Ar) ,15±2(O2) | 9±2(Ar) ,15±2(O2) | 9±2 (Ar) | 9±2 (Ar) | 9±2 (Ar) | 16±2 (Ar) | 30±2 (O2) | 30±2 (O2) | 0 |
| PBN Ar Flowrate(Sccm) | 4±0.5 | 4±0.5 | 4±0.5 | 4±0.5 | 4±0.5 | 4±0.5 | 4±0.5 | 4±0.5 | 4±0.5 | 0 |
| Flowcool Rate (Sccm) | 20±2 | 20±2 | 30±2 | 20±2 | 20±2 | 15±2 | 30±2 | 20±2 | 30±2 | 0 |
| Milling Angle( Degree) | -65±0.5 | -65±0.5 | -65±0.5 | -45±0.5 | -45±0.5 | -45±0.5 | -45±0.5 | 0±0.5 | 0±0.5 | -65±0.5 |
| Milling Time (Seconds) | 120 | 10 | 10 | | | | | 25/80 | 330 | 30 |
| Chiller Temperature(°C) | 3±1 | 3±1 | 3±1 | 3±1 | 3±1 | 3±1 | 3±1 | 3±1 | 3±1 | 3±1 |
| Shutter Status | Close | Close | Close | Open | Open | Open | Open | Open | Open | Close |
| K factor | 1.5 | 1.2 | 0.8 | 1.2 | 1.5 | 2.5 | 0.65 | 1.2 | 0.8 | 0 |

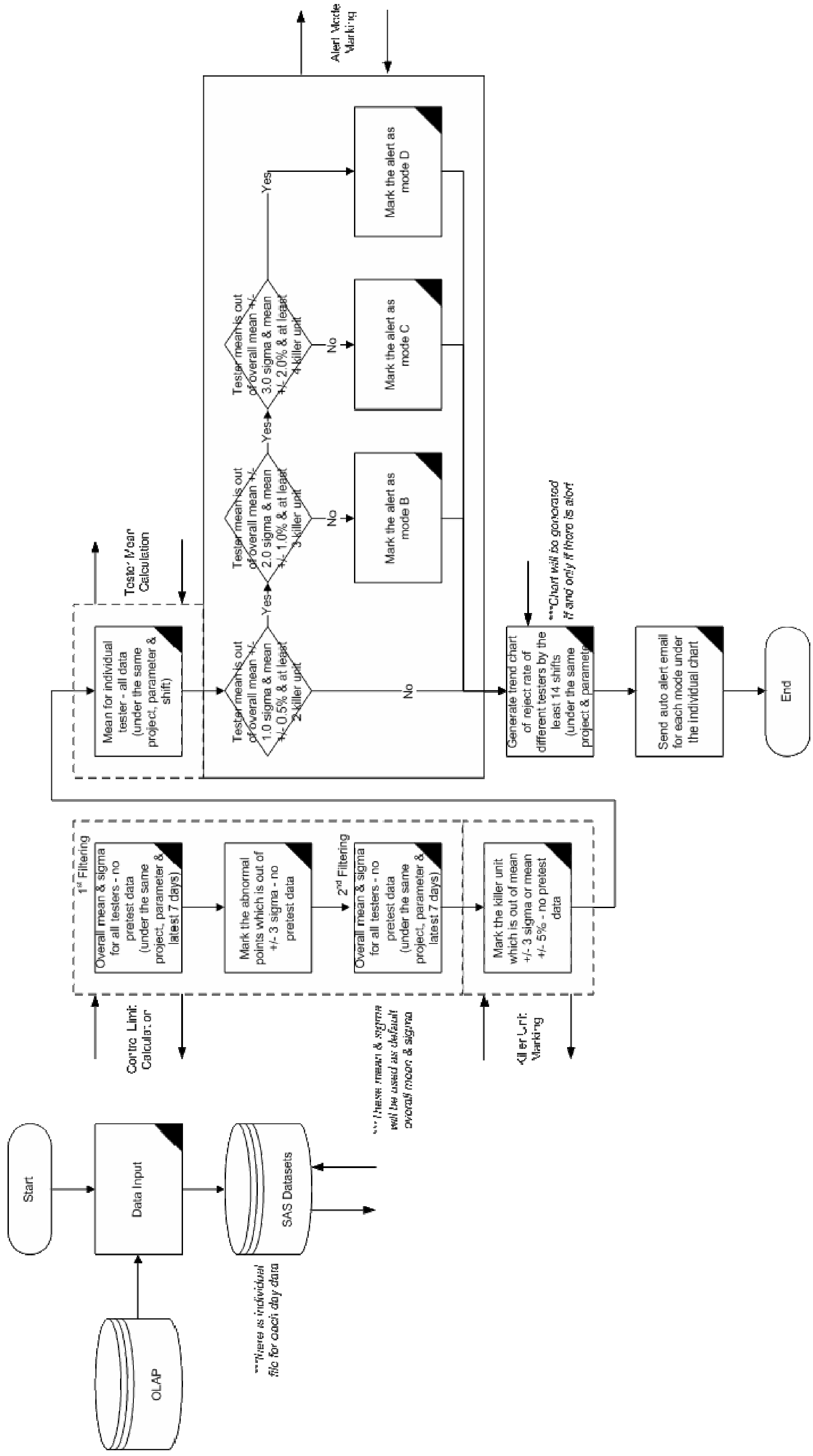
Remark : Veeco#6"SH" Step RF Forward is 400 W.

| Issue by/ Date | 6/8-04 | Checker / Date | 6/8/04 |
|---|---|---|---|

Process Parameters Notes for Machine Type B Setting

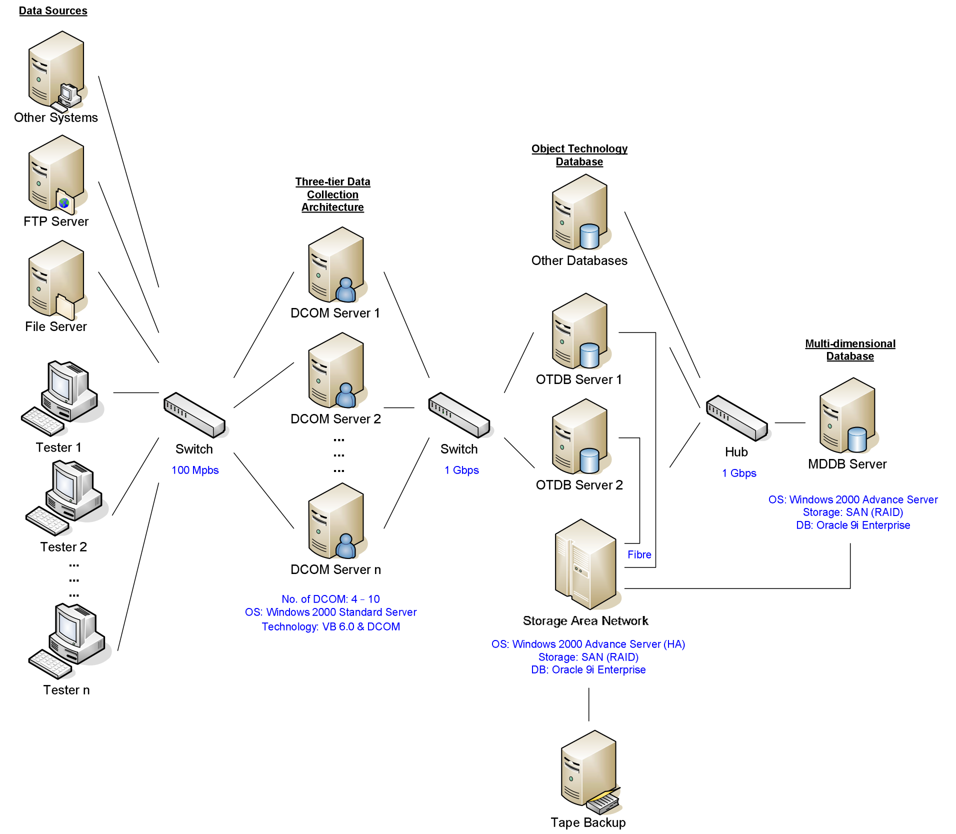
** Sensitive Information was hidden*

# APPENDIX D: EXAMPLE OF DATA PROCESSING

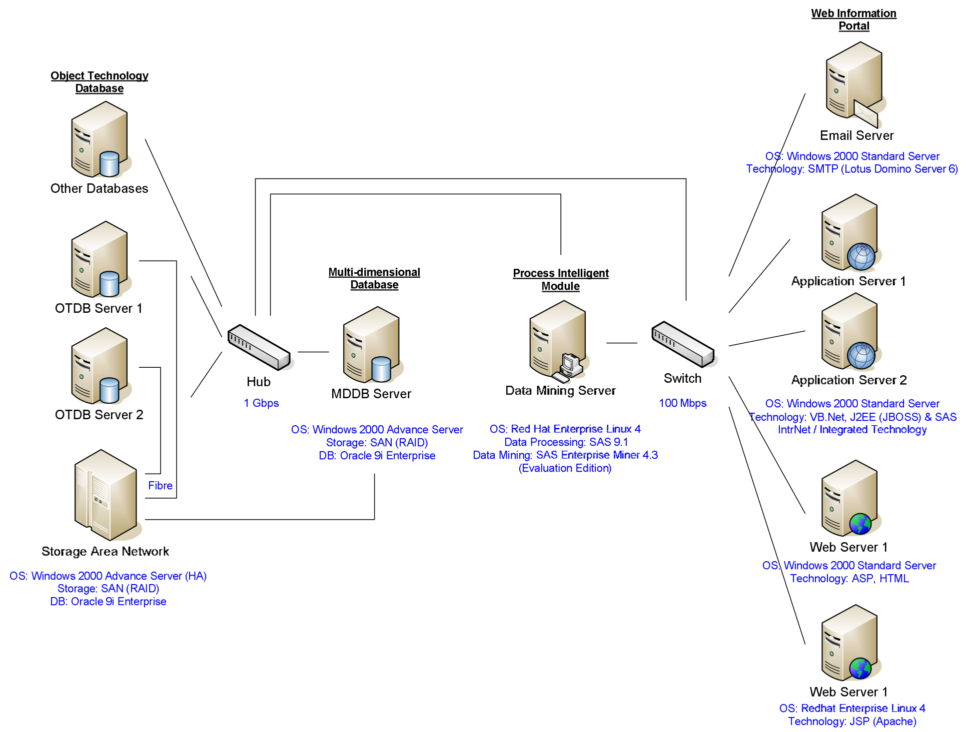


Example of the Data Processing for Data Filtering, Summary and Auto-alert

# APPENDIX E. ARCHITECTURE DIAGRAM OF IDS




Architecture Diagram of IDS: Data Collection Layer




Architecture Diagram of IDS: Application Layer