Information management and Process Improvement Using Data Mining Techniques

Gibbons, WM, University of Ulster, N.Ireland, *mw.gibbons@ulst.ac.uk*
Ranta, M, Helsinki, University of Technology, Finland, *Mervi.Ranta@hut.fi*
Scott, TM, University of Ulster, N.Ireland, *tm.scott@ulst.ac.uk*
Mantyla, M, Helsinki University of Technology, Finland,

Abstract

This paper describes a computer component manufacturing scenario which concentrates on the application of data mining techniques to improve information management and process improvement within a manufacturing scenario. The case-study involved, relates to an engineering component manufacturing company with a consortium of several plant outlets in various geographical locations world-wide. Currently, data, information and knowledge management, transparency and communication between plants within the scenario is limited. As a result, important information and best practices are not collectively pooled or conversed with the objective of improving information management and process development procedures on a global basis. This paper explores the possible enhancement of these consortium relationships and it also investigates the use of data mining techniques with the potential to advance information management and process improvement within the manufacturing environment concerned. In brief, the current information management and process improvement situation within the manufacturing consortium is examined and an improved position is recommended. Additionally, work currently in progress within this study is presented and discussed.

Key words : Information management, Transparency, Communication**,** Engineering ontologies, Data mining.

1 Introduction

Contemporary organisations are inundated with data, however, they have little information, even less knowledge, and perhaps no wisdom. Many companies simply hold data in record or archive format, therefore, potentially, valuable information hidden within these databases remains untapped [1]. The sheer volume of data held in corporate databases, in particular, is already too great for manual analysis and understanding, and as the information within them grow, the problems are similarly compounded. Previously, organisations failed to solve business problems due to the deficit of available data, however, to date, the problem has been reversed, as there is a plethora of obtainable data in many of today's modern organisations. Manufacturing environments are extremely costly and time-critical locations in general, as a result of the difficulties in maintaining process control and identifying parameters responsible for variance. Similarly, a vast amount of data, information and knowledge is collected on each individual manufacturing operation, generating a complex information management situation within the scenario. Accordingly, the challenge is to find ways of distilling and managing these large volumes of data and transforming them into valuable information and additionally exploring the use of techniques to improve information and process management to maximise manufacturing benefit. Data mining techniques possess the potential to enhance process improvement, information management and communication within the manufacturing environment of this case-study, on a global basis.

Improving process control to enhance performance, ameliorate product quality and increase productivity is an important consideration within any manufacturing industry. Accordingly it is critical to find efficient methods of both performing and achieving this acumen. The fabrication and production of components carried out within the manufacturing facility in this case-study require a vast number of complex and meticulous processes. Such extensive manufacturing operations are often carried out in clean-room environments and employ continuous quality and precision controls. A major difficulty with process control in the manufacturing industries in general is the extensive quantity and complexity of data and procedures involved within the fabrication process [8]. This particular case study involves the collection of large quantities of data on various composite operations within the production process. Often, process data is collected from more than one database, hence problems analysis and decomposition are laborious tasks. The complexity and magnitude of operations performed in the fabrication process, inevitably impacts upon information management and process control, making them cumbersome procedures. Current methods of process management employed within the manufacturing consortium, such as SPC (statistical process control) and feedback control models, have proved to be inefficient in providing adequate information management and process control improvement. This has caused a reduction in product quality and productivity within the individual plants in this case-study, as engineers are unable to exercise controlled changes due to the limited real-time framework involved. As a result, temporal and monetary detriments have increased in this high cost, time critical environment. The current situation within this domain area highlights the need for new process improvement and information management methods. In the next section we explain the case scenario. This is followed by a discussion on the problems identified within the consortium relationships, including the proposal of possible data mining solutions. Section 4 presents research currently in progress within the consortium.

2 Computer component manufacturing scenario

Our case study is based on a large computer component manufacturer that operates in several geographical areas. In order to keep our discussion focused and to emphasise the characteristics relevant to this research the consortium has been simplified to cover just two geographical areas. The case-study described in this section involves one of the largest global magnetic recording manufacturing industries. This particular organisation is engaged in the time critical fabrication of data storage products world-wide. The problem domain area at the centre of the scenario consortium is outlined succinctly, as the intricacy of the process operations are not within the scope of this paper. As an alternative, the focus is directed to the critical areas of information management and process improvement and how the application of data mining techniques and communication within all plants concerned in the consortium can incur long-term strategic benefits. The problem domain area is discussed with this perspective in mind.

The company as a whole includes many different divisions of which we focus on only one, the recording head division. This consists of a headquarters and individual plants that are located at different areas. At each area the operation is divided into a manufacturing and assembly plant.

- The Headquarters (HQ), that manages and co-ordinates the operation.
- Manufacturing Plant A (MP-A), which manufactures the basic components for recording heads for market area A.
- Assembly Plant A (AP-A), which is responsible for assembling components into final products for market area A.
- Manufacturing Plant B (MP-B), that is similar to MP-A except to its location at area B.
- Assembly Plant B (AP-B), that is similar to AP-A except to its location at area B.

The customers of the recording head division may be either internal customers or external customers, i.e., they may be manufacturers, distributors, and system integrators coming either from the same company or from outside it. In addition the customers may be served by the headquarters as global customers or by either of the areas as local customers.

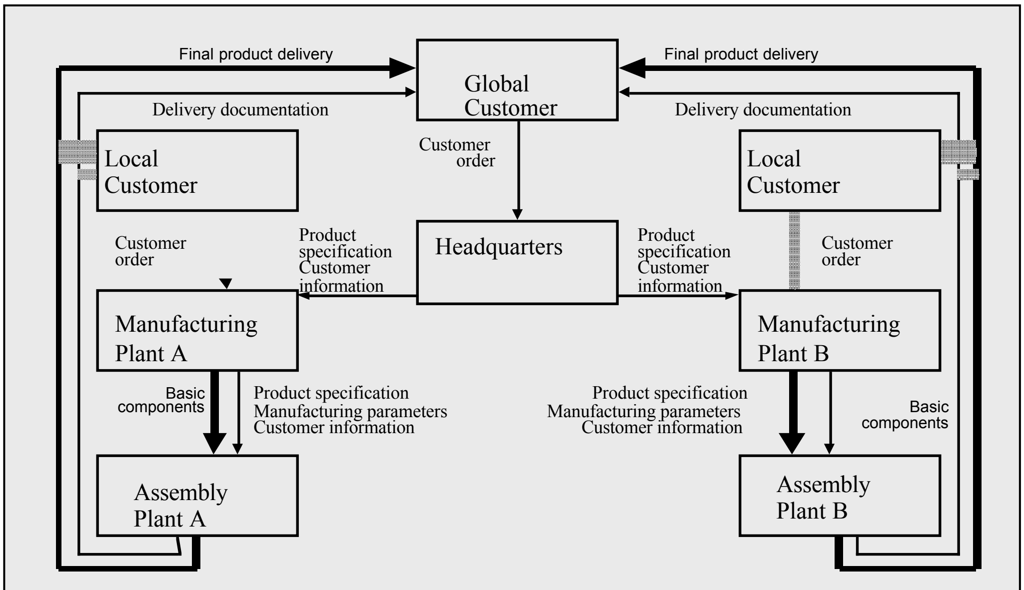


Figure 1: The customer order process

All outlets, HQ, MP and AP's A and B possess customer, Product development and Research and Development teams that deal with product and process development. Data individual to the plant is controlled and managed by the specific plant teams of that area alone. Product and process development teams decide upon the introduction and improvement of new products and processes for each plant. Assembly and shipment of the product to the customer occurs at the AP nearest to the customer in question. Figure 1 illustrates the operations within the recording head division during a customer order satisfaction process. The thin arrows illustrate the information flow and the thick arrows are used to display the flow of material between plants.

Processing a customer order received by the HQ and the local customer

The HQ includes a sales office that maintains the customer relationships and handles orders from global customers. When HQ receives a product request from the customer, it processes the customer order and forwards it to the most appropriate manufacturer, according to geographical proximity. MP-A has the facilities for the manufacturing of basic components for the product specification received from HQ. The manufacturing process is a complicated process that is very difficult to control precisely and the resulting components inherently include variation from the normal process dimensions. MP-A forwards the components including some information regarding the process operations conducted and parameters identified to the AP-A, as these will act as guidelines for adjusting the assembly stage. It is here that the MP-A's responsibility for the product ends. AP-A possess the facilities for assembling the basic components into a final product that fulfils the customer requirements. The AP-A receives the basic components and specification of the product that is to be assembled from the MP. It also receives certain information on the specific manufacturing parameters that the MP-A used in the manufacture of the components, as the manufacturing process has an immense impact on the assembly of the product. According to this information the AP-A experts are able to predict the possible variations on the components and take this into account during the assembly process. When the AP-A has finished the assembly process it delivers the final product to the customer that originally ordered it.

The MP's and AP's A and B have their own internal and external sales customers. Figure 1 shows examples of a customer order resulting from either MP-A or MP-B due to the geographical location of the customer. The difference is that the customer order is sent directly from the local customer to the MP-A or B which processes the product specification. The material and information flows between MP-A/B and AP-A/B during a local customer order are similar as in the case of a global customer order. Finally AP-A/B delivers the final product to the local customer. In addition, local plants may serve their own area as the MP's and AP's also possess their own customers according to the product specialisation of the plant, customer data on local customers is managed at these particular plants only.

3 Challenges of the scenario

The most important challenges found in the investigation can be categorised as follows:

1. The limited co-ordination and management of information, data and knowledge among the consortium partners.
2. The deficient transparency and communication of data, information and knowledge across areas and between MP and AP's.
3. The difficulty of discovering new and relevant manufacturing information and knowledge and the problems of managing process control within the manufacturing process are quite predominant.

On investigation, many problems were identified within the current situation of the manufacturing consortium. These are discussed in detail in this section.

A. Co-ordination and information management within the division

The first problem highlighted on investigation of the manufacturing consortium can be categorised as the co-ordination and management of data, information and knowledge among the consortium members. The computer component company has set very clear goals for the operation of its division: production of smaller, faster and cheaper products with the least possible waste. Each plant within the recording head division works towards these goals in its own area of operation. The strategy of establishing quite autonomous plants at each market area in order to simplify logistic chains and to acquire competence for serving local customers has proven to be successful. The operation of the division is managed centrally by the HQ, but it does not maintain close information exchange connections to all the plants within the consortium. As the division has grown and global competition has become more competitive, therefore, the necessity of co-ordination and information management on the level of the whole division has increased immensely.

B. Transparency and communication of information across areas and between plants

The second problem identified within the manufacturing consortium may be defined as the limited transparency and communication of information across areas and between plants. The critical control dependency and static relationship between MP's and AP's within the scenario and the unexplored relationship between MP's and AP's A and B in different geographical areas are crucial factors in making information transparent for the communication across areas and between MP's and AP's. The distinction between the plants of area A and area B lies in their geographical locations in different areas and continents, even as parts of the same company. The plants, however, act as autonomic entities with hardly any co-ordination of operation or exchange of knowledge. The predominant point of contact involves HQ, although this is not essential as contact with HQ does not affect the daily operations of any of the MP's or AP's A or B.

The product manufactured by the company within the consortium specifies particular requirements on the tight relation of the processes of the MP's and AP pairs since the occurrences at the MP have a direct affect on the AP outcome. Since it is not possible to measure certain crucial dimensions of the already manufactured components, the variation from the nominal specifications must be forwarded to the AP in terms of the description of the process parameters and the steps that are applied during manufacturing. The AP utilises these descriptions in order to compute the corresponding parameters for the assembly-based processes that complete the product. The character of the product and production technology as well as the work division between the MP and AP highlighted the need for information transparency, communication and management between the two. In order to diminish the amount of scrap, the processes have to be controlled as accurately as possible, considering both as separate entities and as an integrated wholeness. It is crucial to recognise which information is critical and which is non-critical, i.e., which manufacturing parameters or steps can be freely changed without affecting the assembly. The critical pieces and parameters of the manufacturing process must always be communicated to the AP as assembly is affected by it, whereas best practice information and knowledge tends to remain a competence solely of the MP.

The most important unaccustomed relationship that the consortium has the possibility to exploit involves information sharing between plants on a global basis. Currently, on the one hand the AP's and on the other hand the MPs are not realistically sharing information properly. For example overlapping product or technology development may be conducted in parallel resulting in redundant work. Furthermore, the autonomy may cause variation on the quality of the consortium's output due to the independent practices.

The co-ordination and information problem has already emphasised the need for the HQ to co-ordinate the sharing of best practices and discovery of new information and knowledge by analysing the data that is collected throughout the division, for example on customers. The requirement of the availability of information considers all of the parties in the consortium that are operating as an entity with common goals and objectives, to achieve optimum results. In order to make the shared information usable within other plants it is necessary to not only exchange results of manufacturing or development processes, but also to share descriptions on the processes, methods and decisions that lead to the results, applied background information and various kinds of rationale of products, processes and strategies. Obviously this level of transparency requires extensive effort among the parties. As the plants of the Recording Head Manufacturer belong to a common company they have the basis and need for global effort that is a necessity for a plant to make its data, information, knowledge and operations visible to other plants [2,3,4,5].

C. Discovery of relevant manufacturing information

The third problem demonstrated within the manufacturing scenario involves the discovery of new and relevant manufacturing process data information and knowledge. At present, there is great difficulty in managing manufacturing process control, identifying process control parameters and introducing new manufacturing products and processes. Data mining techniques can be used to improve process control and discover new manufacturing process knowledge in various manufacturing sectors. It is important to discover new manufacturing information in order to improve the manufacturing process. The manufacturing process involves a very intricate and delicate set of operations, therefore, difficulties are immense. One problem involves maintaining the control of the manufacturing process operations due to the effect of process variance, which is a major problem inherent in manufacturing processes in general. The identification of second order process control parameters currently responsible for causing process variance, are very difficult to discover due to the inherent complexity of manufacturing procedures. The discovery of new manufacturing process information and knowledge on process and customer data across areas and between plants to enhance overall information management and process improvement is essential to the improvement of the consortium.

Technology push and the new demands appearing in the customer orders mainly initiate product development initiatives within the consortium. Manufacturing and assembly plants are constantly developing their facilities and methodologies according to the development and manufacture of new technologies and according to the need be competitively viable. Customer orders are received by the sales office of the HQ that manages local customer relations or the individual MPs. Headquarters do not co-ordinate much of the centralised development efforts but they encourage continuous information and process development efforts to meet customer demands. As the MPs maintain a close internal connection to local customers via marketing and sales, they are able to follow more closely the needs among potential customers. A number of issues arise where the possibilities for enhancing product development and predicting future customer needs are concerned. The tough competition requires headquarters to take efforts to define and execute a customer and market driven product strategy that is based on the systematic gathering of information, as a basis for future needs. Currently the HQ does not document or analyse the local marketing and sales happenings that are handled directly by the local MP's and AP's. Thus the central sales office controls only the customer data related to customers dealing through HQ and not the individual manufacturing and assembly plants.

Due to the dependent relationship between the manufacturing and assembly plants across areas, it is imperative that new manufacturing process information on process, product, customer and information management improvement, are discovered, made transparent and communicated across the consortium. This solution may be accomplished via the application of data mining techniques to model the manufacturing product and process domain areas in order to discover new and relevant process and customer information and knowledge, these are discussed further in section 4.

3.2 Improved situation

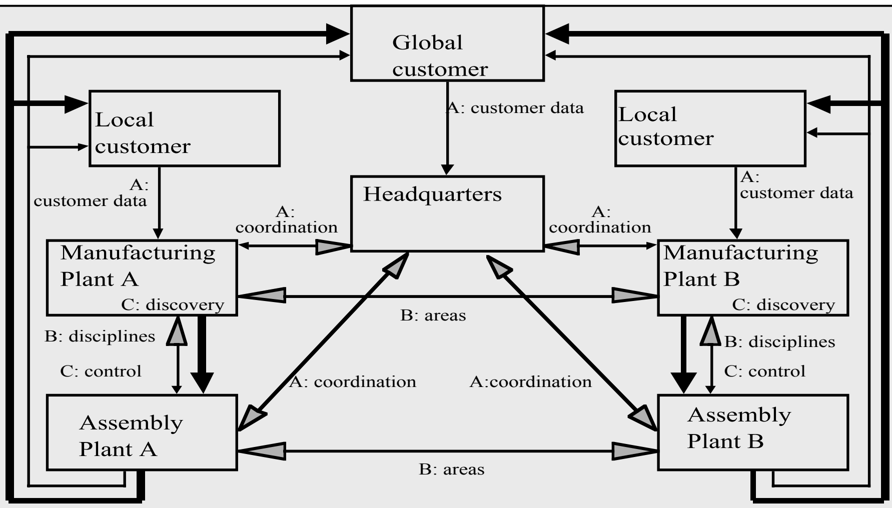


Figure 2: Improved customer order process

Solutions to improve the current situation within the manufacturing consortium are based on the following factors: improved information management techniques, engineering ontologies, application of data mining techniques. Each of these proposed approaches for improvement are discussed in the following and figure 2 illustrates the recommendations of an improved situation.

A. Improved information management techniques

The development of new relationships between plants and across areas within the consortium are an essential factor in improving the current information management situation. Sharing information between all areas and plants are required and the co-ordination and management of both information communication lines and relationships between areas and across plants must be maintained to ensure efficient co-ordination and management of information among the consortium partners in all areas. New consortium relationships must be enforced and managed by HQ and the management of information can be achieved by implementing a central data warehouse and by using intelligent data analysis, management and retrieval tools in all plants.

Data mining offers the means for obtaining control of the dependency of the manufacturing and assembly processes and also in identifying critical and non-critical information, however, it can not pass this information on to the subsequent AP as only the MP has the knowledge and means to do this. Critical process knowledge and related information is currently stored in each individual plant. Only the necessary information is distributed to the AP's, trial and error experiments are not documented for the purposes of assisting the other plants within the division. The issue of communicating and managing this information and knowledge gained via experiences is a matter of defining explicitly all relevant best practices with the contents of the product documentation that is exchanged between the plants and the co-ordination of these procedures.

It appears that both the MP's and AP's have a strong responsibility for this among themselves, as they are both familiar with the entire manufacturing process. In reality, this does not happen since these measures are not enforced by the divisions HQ. Present strategies must be revised and all lines of communication within this consortium must be exploited for mutual benefit. HQ must enforce distribution and transmission of best practices between MP's and AP's in similar and different geographical areas. Adoption and deployment of this strategy would improve data, information and knowledge management within this consortium on a long-term basis, as would the constant exchange of ideas and regular plant interaction. These practices would undoubtedly impact upon manufacturing process improvement as a whole.

The solutions to the problems involving plants in across areas are quite similar to those solutions to the problems between MP's and AP's. Information management, involvement and sharing between HQ division and MP's and AP's A and B have the potential to resolve the problems recognised within the relationships among the consortium collaborators. Subsequently, determining a strategy that allows plants to specialise, combining the expertise to discover new information and knowledge and emphasising collaboration and not competition will also contribute to the exchange and management of ideas within the consortium globally.

Figure 2 shows four two-directional grey arrows with the label "A: co-ordination" connecting the HQ to all of the MP's and AP's. They demonstrate the need for the HQ carry out information management on the division level in order to be able to establish and implement efficient division strategies. division wide information management, or even company wide, allows defining and sharing best practices among the plants. Furthermore, product development efforts can be conducted collectively by assigning certain process improvement tasks to a certain plants. Thus, the plants are able to learn from each others experiences and repetition research or development work is avoided. Such an approach also allows the plants to specialise and to offer different kinds of product mix. As a result, the static approach of assigning customer orders on a geographic basis is developed into a more dynamic strategy of selecting parties for each project according to the competencies of the available plants. An important example of the need for information management by the HQ is the gathering of customer data which is illustrated in figure 2 by the three labels "A: customer data" besides the information flow arrows from customers. In order to plan and successfully realise a customer and market driven product strategy it is necessary to maintain and analyse a database on the potential customers and their requests. The obstacle in the current efforts of analysing customer data lies in the fact that information on the local customers is currently not attainable by the HQ, as the local MP manages local customers autonomously.

B. Engineering ontologies

Ontologies, product and process modelling and the communication of these will improve the communication and transparency of information across areas and between plants. Ontologies has the capacity to assist in modelling the manufacturing product and process to enable the development of engineering ontologies with which to facilitate the communication process between plants and across areas to be more comprehensive. Communication of these via the new relationships described in the co-ordination and information management problem will permit the transparency and communication of information among the manufacturing consortium partners. This may be accomplished by newly developed engineering ontologies and terminology's. Engineering ontolgies can also provide the raw data in an acceptable format with which to conduct data mining. Communication can again be enforced by HQ and made easy by a data warehouse and suitable data management, analysis, retrieval and transfer tools within all plants**.** However, this is the subject of further research and discussion [2,3,4,5].

In figure 2, all the grey arrows with the labels "A: co-ordination", are needed to establish the transparency of information and operations. Furthermore, information also must be made transparent among the plants across both area and discipline divides. The two horizontal grey arrows in the middle labelled "B: areas", illustrate the need to share information across areas A and B by establishing a connection on the one hand between the MPs MP-A and MP-B and on the other hand between the AP's AP-A and AP-B. In this relationship, the challenge lies in the area dependent differences in cultures, legislation and standards etc. When information is shared, these differences must be taken into account and translation or mapping must be applied to ensure correct interpretation of the exchanged information. Another obstacle involves the sharing of information, as the plants have operated as autonomous entities, they have developed their own methods, concepts and naming practices for describing their products and processes, i.e., not only the products and processes have become differentiated but also the language and concepts used in describing them have developed to separate directions. In figure 2, the two vertical arrows between each pair of MP's and AP's are labelled with the term, "B: disciplines", these denote the establishment of a higher degree of transparency between the plants within same area but with different disciplines. The challenge of different "languages" and "concepts" is more difficult as the understanding of the information differs according to discipline. However, the requirement for transparency is greater as the manufacturing process operations have a strong impact on the assembly results and the pair of plants must be toned to work together.

C. Discovery of manufacturing process information and knowledge

Within the manufacturing process, an engineer monitors and analyses the critical control parameters within the manufacturing processes that are crucial in determining the quality of the final product. Engineers employ these methods to exercise control over the process in order to improve it. The process is monitored and analysed on a daily basis as all manufacturing processes are subject to variation changes, causing manufactured parts to vary in shape and size. An additional activity performed by the engineers involves the search to discover new control parameters to assist in the continuos improvement of the manufacturing process. Engineers currently discover additional control parameters via time-consuming processes such as statistical random searches, by setting up organised experiments or in the event of process failure. Due to the dynamic nature of magnetic recording manufacturing, engineers rarely have sufficient time to uncover all control parameters before the process changes. Data mining techniques have the potential to improve this situation via automating the search for second order process control parameters, provide predictive process control and assist in the discovery of new and relevant process information and knowledge.

Data mining also offers tools with the potential for analysing the customer order and data in order to discover or predict future customer profiles as a basis for marketing and maintaining customer relations. Furthermore, the large database of customer data and orders can be used for predicting the directions of future customer needs. Additionally, the basis of the market need prediction can be strengthened by either collecting all the customer data from all the manufacturing and assembly plants and the headquarters as one target base for data mining or by at least utilising similar mining operations and gathering the results for central analysis and strategy definition by the headquarters. According to the data mining based market predictions the HQ have the potential to initiate developments of either completely new products as a centralised effort or alternatively, encourage concentration on certain existing products and their enhancement of marketing efforts.

Section 3 identified and discussed a number of problems among the relationships regarding information and process management operations within the consortium. However, it is evident that co-ordination, information management, transparency and communication of information is needed primarily in order to discover new manufacturing knowledge properly. Ontologies, product and process modelling can facilitate this transparency and aid information communication for improved data mining within the consortium. This is, however, the basis for further research considering the use of engineering ontologies

4 Discovery of manufacturing process information and knowledge

This section discusses research currently being conducted onsite within the manufacturing consortium involving the implementation of data mining techniques to improve the discovery of new and relevant manufacturing process knowledge. The manufacturing processes of both MP-A and MP-B have been developed through continuous development that has meant introducing new steps to the process in order to gain the ability to address new process and customer requirements. New features have been added to the process according to the new possibilities offered by improved technology and according to the pressures of industrial competitors. The expertise of the MPs have been gathered as a storage of various manufacturing processes that have been developed to match different specifications. This pool of processes are utilised to fulfil the product specifications.

As a result, the manufacturing processes have grown to be complicated and hard to manage and the manufacturing engineer has lost the track on the wholeness and on the rationale and affect on the parts of the manufacturing process. Thus, the development work to match new specifications has not necessarily resulted in continuous improvement, but may have caused overwhelming complexity that the engineers involved in the manufacturing and assembly processes have to cope with. It is difficult to comprehend the huge amount of process data in order to maintain it properly. Support is needed for the manufacturing plants to keep track of the side affects and rationale of steps or pieces of process data and their relationships with the manufacturing process. The extreme of these problems is that the manufacturing process may include void manufacturing process steps that no longer have any affect on the final results. Such a situation may result when the process includes new later stages that override the affect of previous steps. This happens very easily when MPs do not want to touch a tested and working process description, so instead a few new steps are added after it to modify the result according to the new requirements.

A further problem is caused by the close dependency of the manufacturing and assembly processes, as mis-processing during manufacturing can have a destructive effect on the actual assembly of the final product. The computer component that is being manufactured requires two steps of first manufacturing the basic components and then assembling them into the product requested by the customer. These steps are separated into distinct plants with specialised facilities and know-how. As there is no practice for communicating the documentation of changes, problems and best practices that occurred in the MP, to the AP, the product assembly is performed without the chance of taking advantage of the changes or discoveries of the manufacturing process. A particular characteristic of the product is that the manufacturing process requires high precision but it is still unpredictable and the results vary from the expected nominal dimensions. Moreover, it is not possible to measure the final dimensions, as they can be only predicted and indicated by tolerances that are derived according to the parameters of the manufacturing process that were used for the particular component. Thus, the AP's need the information on the manufacturing process in order to be prepared for the possible accuracy deviations and keep them under control when possible.

4.1 Solutions

Compared to most discrete manufacturing industries, semiconductor wafer fabrication encounters composite process, quality and performance control difficulties. Additionally, as the semiconductor device design evolves, new process constraints are continuously added, making the planning process more complex. To date the manufacturing process has become so complex that the number of constraints, their interrelations, and the need to consider numerous scenarios is more than humans can handle without sophisticated computer-based tools to assist them [8].

The problem of discovering void operations within the manufacturing process operations and also parameters that cause the manufacturing process to go out of control, were addressed as the first stages of process improvement. Data mining was an obvious choice for the task of processing large amounts of data in order to discover some structure of the process steps and finally those pieces that become overridden by following steps. In order to improve manufacturing process control, the application of product and process modelling methodologies must be implemented to build a structured model of the manufacturing process domain. This means that all process steps must be described and monitored with the ultimate goal being be a model that simulates the manufacturing operations in order to validate the process and even discover possible variance causes (second order process control parameters). This solution has been implemented and applied successfully within one of the MP's alone, however, application to all MP's within the consortium is required.

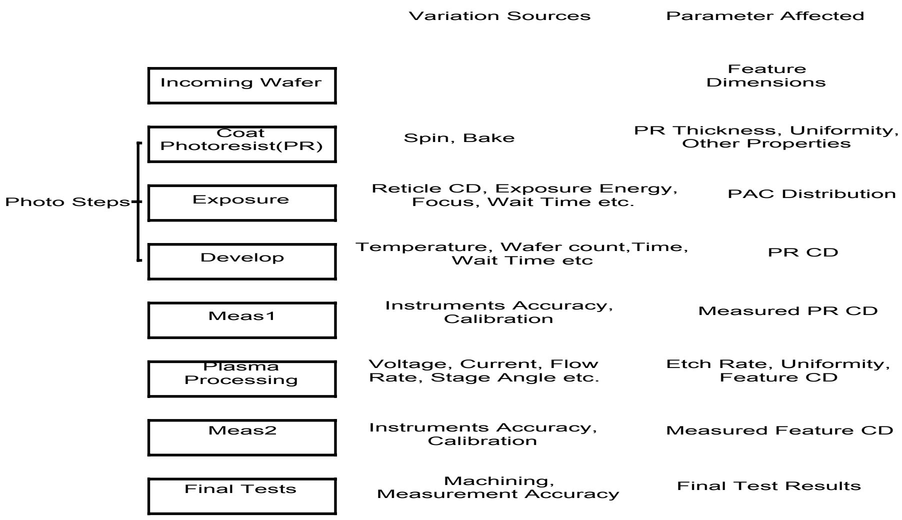


Figure 3: Illustrates the process steps & parameters involved in the Ftest operation

The first aspect of current research in this area involves the identification of second order process control parameters. The case study discussed in section two outlines the major process control, time delay and yield problems prominent in the magnetic recording industry in particular, including the intricacy of the process operations. One such operation involves plasma machines, which are widely utilised during critical steps within the recording head and semiconductor manufacturing processes. The complexity of the physics involved and inability to monitor the plasma characteristics on a real-time basis makes the control of this process equipment very difficult. Work is currently being carried out, in accordance with the manufacturing case study, on one of the most critical process steps during the manufacture of recording heads. Figure 3 illustrates the processes and parameters involved. Statistical process control is applied to measured results before and after the plasma processing step (Meas1 and Meas2) of the manufacturing process and at the final test results (Ftest), respectively.

Figure 4 depicts the various control loops involved in this critical process operation. The objective of these controls are to achieve the final test results (Ftest) within the product specification or target measurement. The drawbacks of the current method are immense, primarily because SPC is applied off-line which means that the corrective actions are taken after the out-of-control actions have occurred. Secondly, the feedback time between the point at which the SPC charts are analysed and the process step is controlled, constitutes an extremely lengthy control loop. Similarly, the feedback time between Meas2 and the processing step concerned may range from two to four days. Feedback between the final test results (Ftest) and the processing step can range from anything up to five weeks. When a fault is discovered, substantial quantities of wafers may have already been processed on these machines, resulting in these wafers becoming redundant. As part of this research project, real-time equipment information is collected and sent to a central database. The real-time equipment parameters involved are characteristically auto-correlated and cross-correlated, rendering traditional SPC approach no longer valid. As a result of the vast amount of data generated for each individual run (over a hundred parameters) and also the inherent non-linear nature of the plasma process, difficulties arise and it becomes extremely laborious for an engineer to analyse such volumes of data using basic statistical software. Consequently, a more effective and systematic approach is necessary [9].

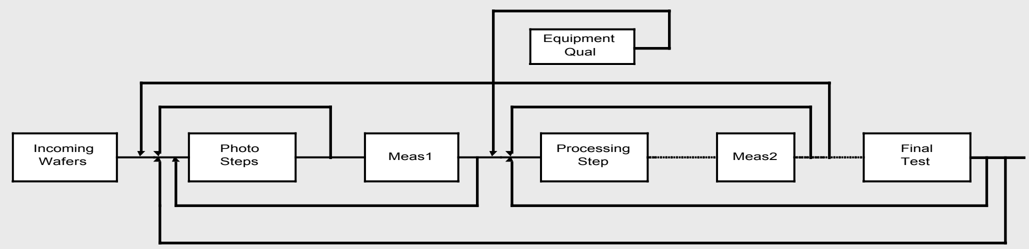
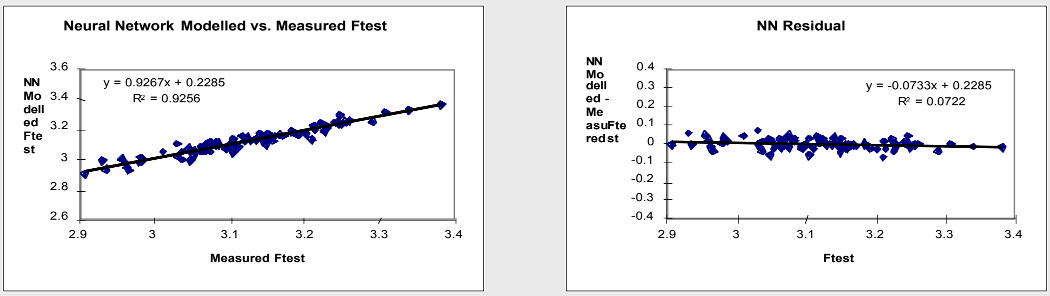


Figure 4: Illustrates the control loops involved in obtaining the final test results

Current research has illustrated that a greater level of process control can be achieved using a predictive model. This may be accomplished by applying multivariate or neural network analysis techniques on process data, including production lot history data, in-line physical measurement data, real-time equipment data, equipment maintenance data and end-of-line product test and final yield data. This model may be integrated within a real-time knowledge-based-system to provide engineers with real-time run-to-run control in a user-friendly format using feedback and feed-forward control predictive models. Due to the complexity of the equipment parameters involved and the highly correlated nature between them, it was necessary to employ multivariate analysis techniques such as Principal Component Analysis (PCA) [10]. This is a powerful technique which applies data reduction techniques on highly correlated data variables, whilst simplifying the analysis process [8]. An in-depth analysis was carried out on over one hundred parameters using the Unscrambler software's (by Camo) PCA technique.

A predictive model was built using The Unscramblers multivariate analysis technique Partial Least Square Regression (PLS) and a feed-forward backpropagation three layer neural network [10]. PLS was used to primarily reduce the dimension of the data, the results of which were used as input for the neural network. This approach succeeded in accelerating the training time of the neural network and also reduced the chances of over-training. Three different models were built and explored individually. In the first model process parameters included wafer feature dimensions, photo step process parameters, plasma machine data, Meas1 and Meas2, were used to model and predict the final test results (Ftest). Promising results were obtained as illustrated by figures 5 (A & B).



Figures 5 A & B : Illustrates the Ftest results of the neural network predictive model

The accurate prediction of the final test results at this stage has incurred two significant benefits. Primarily, the neural network modelled results (using ISL's Clementine software) proved to be a superior final predictor to the final test results in comparison to Meas2 results, providing a better level of control. Secondly, as a result of accurately predicted final test results at Meas2, it has been established that the lengthy control loop prior to this step is redundant.

In the second model, all input parameters except Meas2 are used to model the final test results. As a result, as soon as the plasma processing step was over, it was possible to predict the final test results and take corrective action for the next wafer in line if required. To date, it was not possible to implement control action for the next wafer waiting to be finally tested as equipment performance could only be assessed after Meas2 is noted. Results of the second model have also incurred additional benefits as it enables control to be taken immediately after the processing step without waiting for the Meas2 results, thus reducing the control loop by up to four days. Another advantage incurred by this model is that Meas2 and associated process steps may also be eliminated which would shorten the overall process cycle time and reduce line yield losses at these steps.

In the third model, modelling is conducted immediately after Meas1 using previous run equipment data. This model demonstrates that preventative actions can be taken so that the expensive action of scraping or reworking wafers could be avoided, achieving run-to-run control in the process. At the moment, this research is in the process of working towards implementing these models to be part of the final testing (Ftest) control strategy. As part of a previous research project, a real-time knowledge-based-system was specifically developed for the engineers to provide them with automated SPC chart and alarm generation. The predictive models will eventually be integrated into this knowledge-based-system to provide engineers with real-time information to make data analysis and interpretation more efficient.

Implementation of predictive models using data mining techniques has proved that they have the potential to achieve real-time run-to-run control when utilised as a process control mechanism within the manufacturing environment. These models have not only reduced the length of the control loop considerably, but they also provided a better level of control for the critical parameter final test results (Ftest). The integration of these models within a real-time knowledge- based-system will make analysis and interpretation of vast amounts of data feasible on a real-time basis. The possible elimination of certain process steps not only constitutes a more condensed cycle time, it also reduces line yield loses at these particular steps. Since these predictive models are characteristically generic, they may be applied it to other process operations and areas and also across all plants within the manufacturing consortium.

5 Conclusions and work in progress

This paper provided a case-study of a manufacturing consortium involving the customer, HQ and several MP's and AP's in different geographical areas within the specific recording head division of an international company. The current situation within the consortium entails a manufacturing consortium, recording head division, involving HQ, MP and AP's A and B which operate on quite an autonomous basis and only communicate process information as required by their position within the network of consortium relationships. On investigation several inherent problems were identified within the current structure and relationships of the consortium as a whole. Solutions proposed and discussed in this work presented sufficient global communication, exchange of ideas and information management techniques in the form of data mining, ontological, product and process modelling methods which have the potential to inadvertently impact upon manufacturing process control. Data mining procedures currently in progress have also described possible methods of enhancing process control and information management between plants on a global scale, to make co-ordination, management, transparency and communication of process and customer information, more accessible and beneficial to all consortium partners on a long-term basis. Work in progress considers an ontological approach that is currently being developed to assist in the application of the proposed co-ordination, information management, transparency and communication solutions discussed in this paper. Completion of this research will provide a recommended set of solutions in the form of a suite of methods, techniques and tools suitable for consortium use in order to improve the overall manufacturing operational strategy on a global basis.

6 References

[1]. McClean, S & Scotney, B, (1996) "The data mining report", Unicom Seminars Ltd, Middlesex.

[2]. Büchner, A., Ranta, M., Hughes, J., and Mäntylä, M., (1999) Semantic Information Mediation among Multiple Product Ontologies Proc. 4th World Conference on Integrated Design & Process Technology, 1999.

[3]. Ranta, M., Büchner, A., Mäntylä, M., and Hughes, J. (1999) Collaboration in Heterogeneous Environments along Product Processes, Proc. 4th World Conference on Integrated Design & Process Technology, 1999.

[4]. Mäntylä, M., and Ranta, M. (1998) Engineering Process Ontologies for Communication, Co-operation, and Co-ordination in a Virtual Enterprise, CD-Rom of Prolamat '98, Trento, Italy, September 1998.

[5]. M. Ranta, M. Mäntylä, R. Smeds, P. Haho, J. Alvesalo, Improving product development: Towards a framework for integrating artefact and process models, to be published in the proceedings of the Advances in Production Management Systems Conference APMS '99, Berlin, Germany, September 1999.

[6]. Shim, K & Mathon, A, (1993) "Simulation of Manufacturing Models and it's Learning with Artificial Networks", PROC of International Joint Conference on Neural Networks, Vol.1, pp.991 - 994, 1993.

[7]. Fan, H-T & Wu, S, (1995) "Case Studies on Modelling Manufacturing Processes Using Artificial Neural Networks", Journal of Engineering for Industry, Vol.117, No.3, pp.412-417, 1995.

[8]. Esbensen, K Midtgård, T, Schönkopf, S (1994) *Multivariate Analysis in Practice*, Camo AS, Trondheim.

[9]. Everitt, B & Dunn, G, (1994) "Reducing the Dimensionality of Multivariate Data: Principal Components and Correspondence Analysis", in *Applied Multivariate Data Analysis*, Edward Arnold, London.

[10]. Fayyad, U, Piatetsky-Shapiro, Smyth, P & Uthurusamy, R (1996) "*Advances in Knowledge Discovery and* Data *Mining*", AAAI Press.