Paper 134-25

Data Warehousing for Manufacturing Yield Improvement

Dr. Robert A. Rutledge, IBM Storage Technology Division, San Jose, CA

ABSTRACT

This paper will describe the data warehouse/decision support/data mining system that we have developed to improve disk drive manufacturing yields in the IBM Storage Technology Division.

The data warehouse combines operational data from our disk, wafer and head stack processing plants. Data is extracted weekly, cleansed and correlated with the test results for the disk drives assembled from the corresponding components. The decision support system, based on SAS/IntrNet®, provides OLAP capability for yield analysis. This system enables the user to compare yield loss rates, by failure code, component source, EC level, date of manufacture, etc. The system also shows the sensitivity of drive yield to each continuous in-line measurement, and predicts the yield improvement that would result from component process improvements. The yield sensitivity studies are also used to identify critical parameters for input to the data mining algorithms of SAS/Enterprise Miner®. We use decision trees to identify combinations of parameters with unusually high or low yield, and neural networks to create multi-dimensional models of yield as a function of in-line measurements.

This system allows us to quickly identify the factors causing yield loss, evaluate the cost/benefit of proposed changes, and operate our plants with optimal processes and specifications.

INTRODUCTION

The disk drive industry is characterized by the rapid introduction of advanced technology that delivers ever-higher capacity and performance for the end users. The industry is also highly competitive, which leads to continuous erosion of prices. In this environment there is a significant advantage to being able to ramp new products to high volume quickly and take advantage of the higher profit margins on leading-edge products. Fast ramp to volume requires fast yield learning to detect, diagnose and correct the inevitable problems that arise in new technology. The IBM Storage Technology Division operates advanced manufacturing facilities for fabricating, assembling and testing all of the major components of IBM data storage products, from disks and magneto-resistive heads through finished disk drives and storage subsystems. This vertical integration leads to close cooperation among the component and drive development and manufacturing teams. We can access the various operational databases to quantify the relationship of drive performance to component measurements. Then we can use the resulting information to discover the root cause of problems, and optimize the specifications on in-line component measurements to maximize the performance and quality of the finished product.

Local and Global SPC

Statistical Process Control provides an efficient structured approach for obtaining and maintaining high manufacturing yield and quality. Process Capability studies are used to determine if the manufacturing process is capable of running at high yields, and statistical control charts are used to quickly detect and correct process changes, which might cause yield problems.

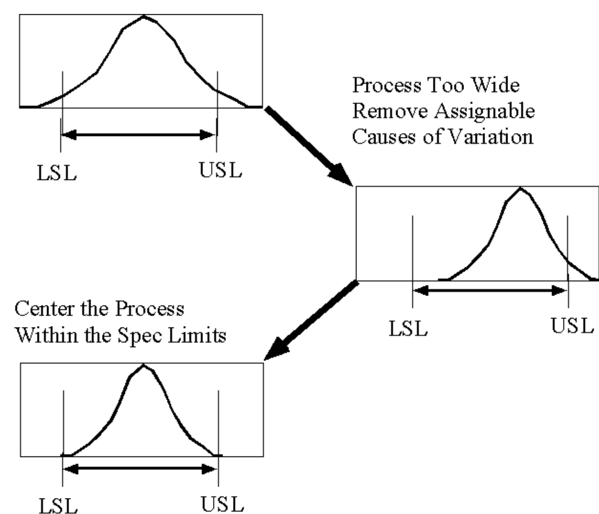

Figure 1: Process Capability Study

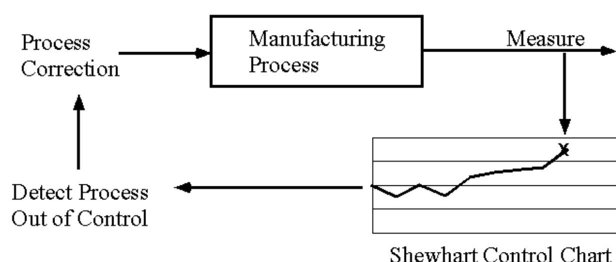

Figure 2: Statistical Process Control Chart

However there are some serious difficulties in applying these techniques to the manufacture of complex systems. The problem is that the total process, including the manufacture of components, subassemblies and systems, may be spread out in factories all over the world. A component level process capability study has limited value without knowledge of the system level measurement data, and the system level operator cannot correct an out-of-control situation caused by a change which occurred weeks or months earlier in a component process. What is required a global SPC system for analyzing and controlling the entire process.

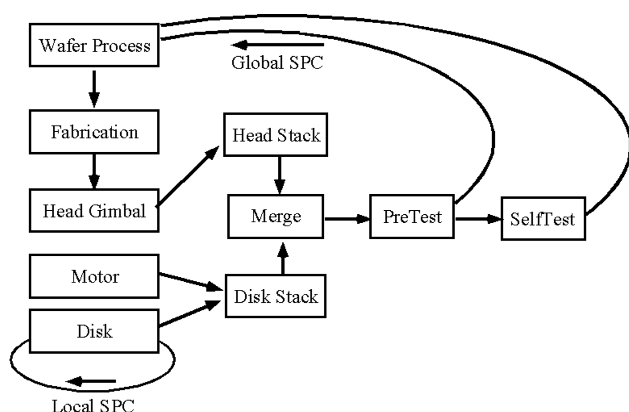

Figure 3: Global SPC

JED: A Quality Data Warehouse

We have developed a system called JED (Just Enough Database) to support global SPC for manufacturing yield improvement. JED is an example of a Quality Data Warehouse as described by Klenz and Fulenwider (1999). The JED system consists of a data warehouse and a decision support/data mining system. The data warehouse combines and correlates data from our Wafer, Disk, Head Gimbal Assembly, Head Stack Assembly and Hard Disk Drive factories. The web-based decision support system, implemented with SAS/IntrNet®, provides automated yield analysis, including OLAP-like analysis of root cause, global process capability studies, and what-if studies to evaluate the effect of process changes. The data mining system includes interactive decision trees and neural network modeling capability bases on SAS/Enterprise Miner®.

JED DATA WAREHOUSE

The JED Data Warehouse is a DB2® database of approximately 50 Gigabytes residing on an IBM RS/6000®. It is updated weekly by extracting the drive test results and parametric measurements for the previous week's production from the operational databases for each of our drive assembly plants. After the drive serial numbers and process time stamps are stored, we search our subassembly and component databases to find the key in-line measurements, which are then stored in JED for easy access. Before the data is stored, the drive and component data are pre-correlated so that one simple query will retrieve drive test results along with measurements taken in the component factory on the same parts.

As the term "Just Enough" suggests, we do not aggregate all of the available data, but only just enough to satisfy our objective of fast and effective yield analysis. We select the drive test results that are directly related to our yield calculations, and the in-line measurements that are likely to be informative for yield and failure analysis. We further reduce the data by capturing data on only a few thousand passing and failing drives of each model each week. For a new product, this is a 100% sample, but as the product volumes increase, the sampling rate is gradually reduced to perhaps 5% of passing drives and 50% of failing drives at each test station. The reduced samples are more than enough for valid data analysis, but much more economical to update, store and analyze.

The JED Data Warehouse has been in operation for over two years now, and has greatly increased our ability to analyze yield detractors and improve yields in our disk drive factories. Before JED, the task of assembling and correlating data across factories was very difficult and time consuming. With the JED data warehouse, an engineer can obtain correctly correlated data in a few minutes with a single SQL query. This simplicity has led to a significant increase in the amount of time spent analyzing data, instead of simply collecting and cleansing it.

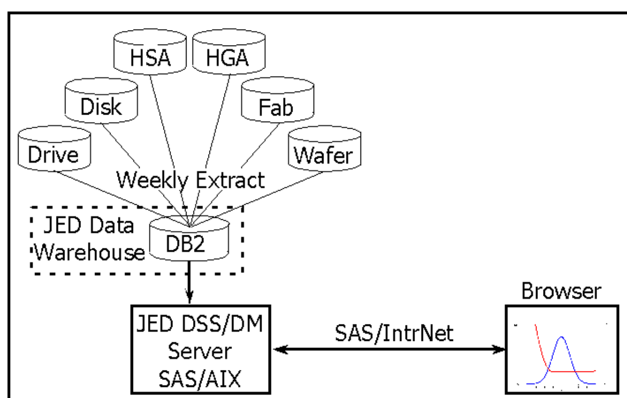

Figure 4: JED Data Warehouse and Decision Support

JED DECISION SUPPORT SYSTEM

The purpose of the JED decision support system is to exploit the value of the data warehouse by automating the search for assignable causes of variation in yield. We have done this by programming a set of SAS/GRAPH® macros to produce diagnostic plots for each of the likely sources of variation. These diagnostic plots are deployed over our Intranet using SAS/IntrNet® running on an IBM RS6000, so that they can be quickly and easily examined by engineers at our factories around the world. We use the charts to determine whether yield loss is related to a particular component source, EC level, manufacturing site or tool, or processing time period.

This is an OLAP system in the sense that it uses pre-summarized data sets to facilitate rapid point-and-click analysis. The summary data sets are not the typical multi-dimensional cubes, but have been designed specifically to support the kinds of analysis we need to do. Figure A-1 (at the end of the paper) shows an example of a wafer map that plots the relative location on the wafer of heads, which failed for selected error codes. The list box at the top allows the user to select a product family for analysis, and the list boxes at the right allow selection of manufacturing sites and dates, and error codes to be plotted. The list box at top right is used for selecting from dozens of different plot types available on this page. Other pages give access to a variety of other analyses.

Global SPC

One of the most categories of analysis is support for yield improvement by modifying the distribution of component parameters, which is the global extension of the traditional (local) process capability study. We begin with a set of charts that quantify the relationship between in-line parametric measurements in our component factories and manufacturing yield at our disk drive factories. By examining such plots we can easily see which component parameters we should concentrate on. Figures 5 and 6 show two typical examples of yield sensitivity curves. Figure A-2, at the end of the paper, shows an example with real data from the JED web site.

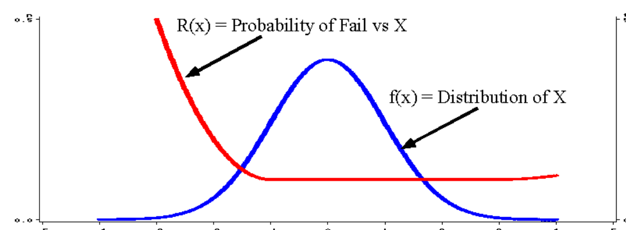

Figure 5: 1-Sided Yield Sensitivity Curve

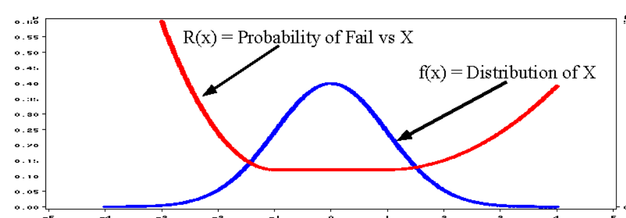

Figure 6: 2-Sided Yield Sensitivity Curve

After such sensitivities have been identified, we can improve yields by shifting the mean, reducing the standard deviation, or screening out the tails of the component distribution, as shown below in Figures 7-9.

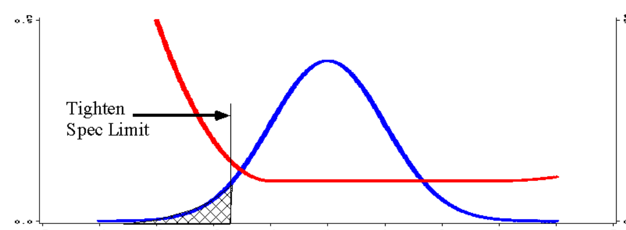

Figure 7: Yield Improvement by Spec Tightening

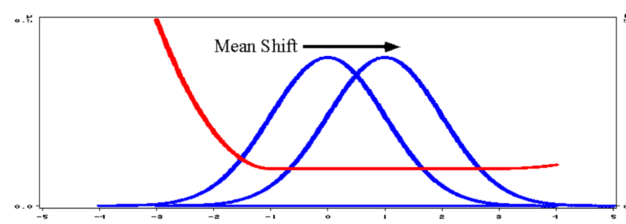

Figure 8: Yield Improvement by Mean Shift

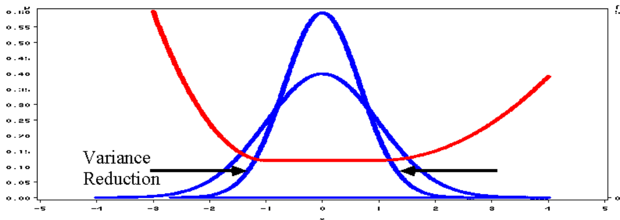

Figure 9: Yield Improvement by Variance Reduction

Each yield improvement method has advantages and disadvantages as shown in Table 1. Spec tightening can be implemented immediately, but incurs the cost of scrapping components. A mean shift can generally be achieved at little or no cost, but takes longer to implement, and is less effective when the yield is sensitive to both low and high parametric values, as in Figure 6. Variance reduction is generally slow and expensive as it might require invention or investment in new tools.

Yield Improvement Method	Time to Implement	Cost of Implementation
Component Spec Change	Immediate	Moderate
Mean Shift	Moderate	Low
Variance Reduction	Slow	High

Table 1: Yield Improvement Methods

The JED Decision Support System includes functions to quantify the cost and benefit of each of these methods. Expected yield can be computed from the empirical component distribution, f(x), and the yield sensitivity function R(x), using the formula:

$$Y = 1 - \Sigma R(x) f(x)$$

The improved yield for each method is easily computed as

$$Y^* = 1 - \Sigma R(x) f^*(x)$$

Where f*(x) is the distribution obtained by truncating, or changing the mean or variance of f(x), respectively,

dY/dS

For the particular case of yield improvement by spec tightening, it is easy to quantify the cost and benefit of various levels of tightening and arrive at an economically optimal decision. We define:

$$dY = \text{Yield Gain} = Y^* - Y$$
$$dS = \text{Scrap Loss} = \Sigma f(x),$$

where the summation is over the values of x that would be rejected under the tightened specification. We also define:

Cs = cost of scrapping one component
Cf = cost of one failure at assembly test

In general the cost of failure increases sharply as a component is used in higher levels of assembly, so that Cf >> Cs.

The financial benefit of a tighter component spec is

$$\text{Savings} = (Cf)(dY) - (Cs)(dS)$$
$$= (dS)(Cf)[(dY/dS) - (Cs/Cf)]$$

Therefore we can realize a benefit by spec tightening as long as (dY/dS) > (Cs/Cf), and, for a given value of (dY/dS), we get maximum benefit by making dS as large as possible. Figure 10 shows a typical plot of dY vs. dS. The relative benefit of truncation, dY/dS, is greatest for small values of dS and gradually declines to the point where further cutting would actually result in a loss. We adjust the spec to maximize dY subject to the constraint dY/dS > K, for some suitably chosen value of K. In the actual charts (e.g. Figure A-3) we plot dY vs. dS for both right and left side tightening, and we leave off the financial information (K, Cs/Cf) because this may vary with time, site, etc. and is difficult to estimate precisely. An experienced engineer can use these charts together with his knowledge of Cs and Cf to arrive at an optimal trade-off.

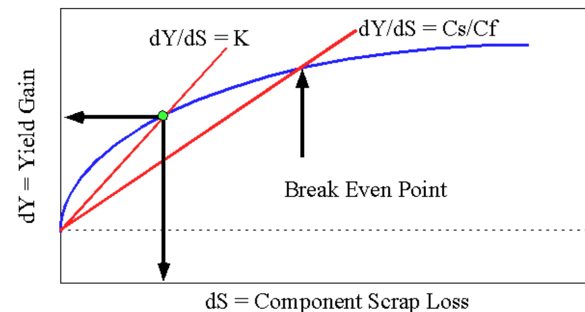

Figure 10: dY vs. dS

Mean Shift and Variance Reduction

For the case of yield improvement by component process mean shift or variance reduction, the costs are more difficult to quantify, but the JED system can quantify the benefits by computing the yield gain expected for various amounts of change, as illustrated in Figures 11-12.

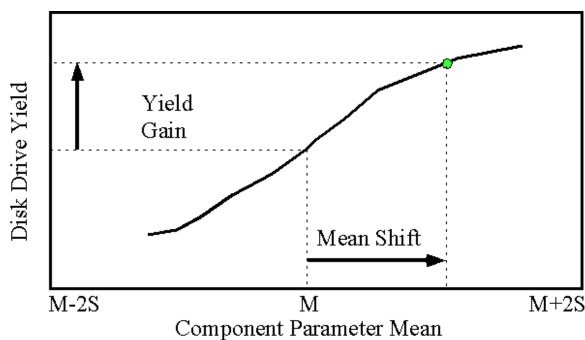

Figure 11: Yield Improvement by Process Shift

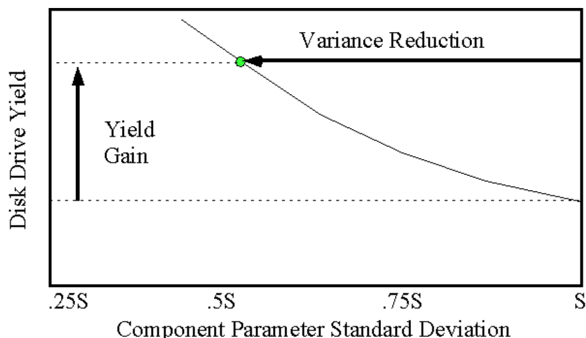

Figure 12: Yield Improvement by Variance Reduction

This yield improvement process can be thought of as a global process capability study, analogous to the conventional process capability shown in Figure 1, but adapted to the requirements of a complex distributed manufacturing system

DATA MINING FOR YIELD IMPROVEMENT

The methods described above are very effective for evaluating the relationship of each component parameter to disk drive yield. However we have hundreds of potentially important parameters, and several different disk drive models manufactured at four different sites around the world. We need methods to quickly identify the important parameters and relationships.

dY/dS for Multiple Parameters

For the case of yield improvement by spec tightening, we can easily pick out the important parameters simply by overlaying the dY vs. dS plots for all parameters, as shown in Figures 13. It is evident from this plot (shown more clearly in Figure A-4) that the greatest benefit will result from cutting parameter 11, and a substantial but lesser benefit from cutting parameters 2, 3 or 4. The remaining parameters are of less interest and can be ignored.

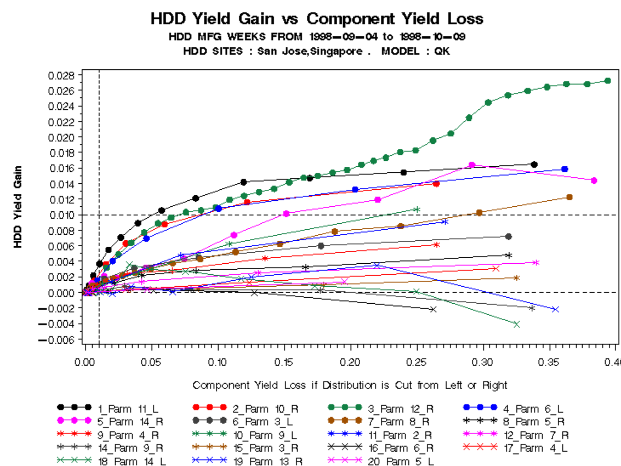

Figure 13: dY vs. dS for Multiple Parameters

A weakness of this chart is that it does not allow us to see the benefits of multiple spec changes. If we want to evaluate the cumulative effect of two or more spec changes, we can't tell whether the effects on dS and dY will be additive. For that analysis we use a decision tree customized for the requirements of the problem at hand.

Decision Trees for Spec Optimization

The splitting criteria most commonly used for building decision trees are CART, C4.5 and CHAID, which are implemented in SAS/Enterprise Miner as the Gini, entropy and logworth criteria respectively. These are generic criteria, which are appropriate for implementation in a general-purpose data-mining tool set, but they are not necessarily the best choice for a particular practical problem. Our problem is to maximize dY subject to the constraint $(dY)/(dS) > K$, where K is some suitably chosen constant greater than (Cs/Cf). Therefore we have implemented a decision tree capability based on this criterion.

The tree is constructed interactively on the JED web site. Figure 14 shows an example of the chart used to determine the first split. (The full web page is shown in Figure A-5.) It is clear that we want to split on parameter 3, and the program recommends splitting at the point of greatest dY for which dS<5%. All other possible splits with dS<5% are available for user selection in the list box above the graph (See Fig. A-5.)

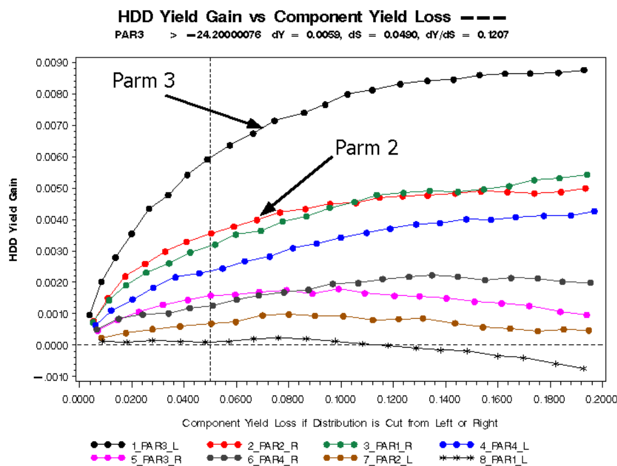

Figure 14: Graph for Choosing First Split

After the user selects the first split, the choices for the next split are presented in a similar graph (Figure 15), which shows the additional yield gain possible after the original spec change for parameter 3 is made. The user continues until no further economically advantageous splits are available.

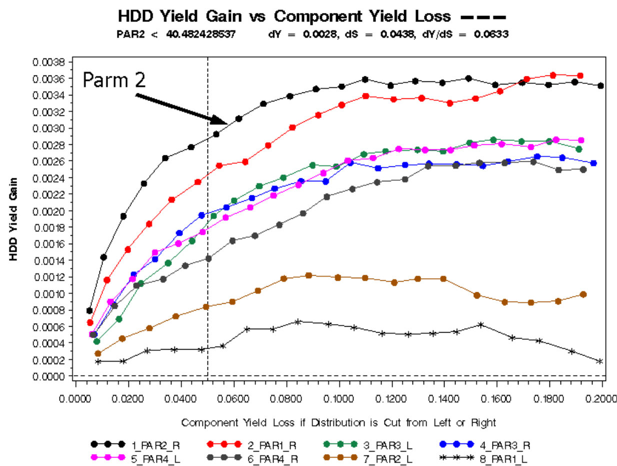

Figure 15: Graph for Choosing Second Split

The tree resulting from the first two splits is shown in Figure 16. This shows that we can reduce yield loss by 20%, from 4.2% to 3.3%, if we scrap 9.1% of the components. This is the best we can do by changing each spec separately, but we might do better by defining a joint spec involving multiple parameters, e.g. Parm 2 + Parm 3 < N. We can search for such opportunities by continuing to grow the tree in both directions, but it is more efficient to use a neural network model to represent multi-dimensional regions in the parameter space that have poor yield.

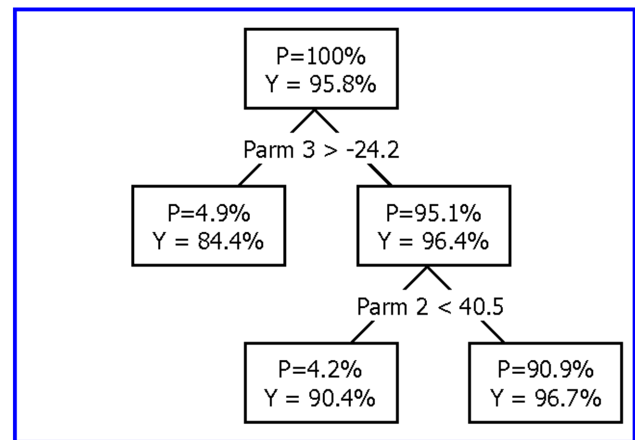

Figure 16: Decision Tree for Yield Improvement

Neural Network Models for Yield

We use Generalized Additive Neural Networks (GANN) as described by Potts (1999) to model yield as a function of several component parameters. The GANN models have the advantage of being more interpretable than general neural networks, because the relative effect of variation in one parameter is independent of the values of the other parameters, and can be shown in a set of partial residual plots as in Figure 17.

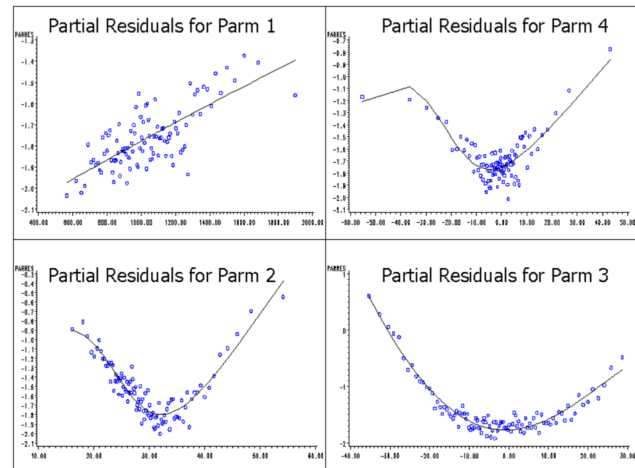

Figure 17: Partial Residuals of GANN Model

The residuals in Figure 17 are from the GANN fit of the same four parameters used in Figures 14 and 15. If we add this GANN yield prediction as a new parameter we see (Fig 18) that the new parameter is significantly better than Parm 3 as a yield predictor. A multi-parameter spec of the form

GANN Prediction < M,

which is approximately elliptical in Parm 2 and Parm 3 (Fig. 19), will give a better value of dY/dS than the best individual spec.

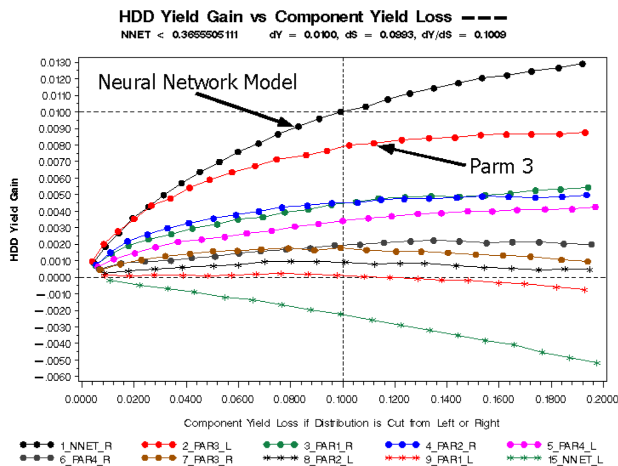

Figure 18: dY vs. dS for GANN Model Spec

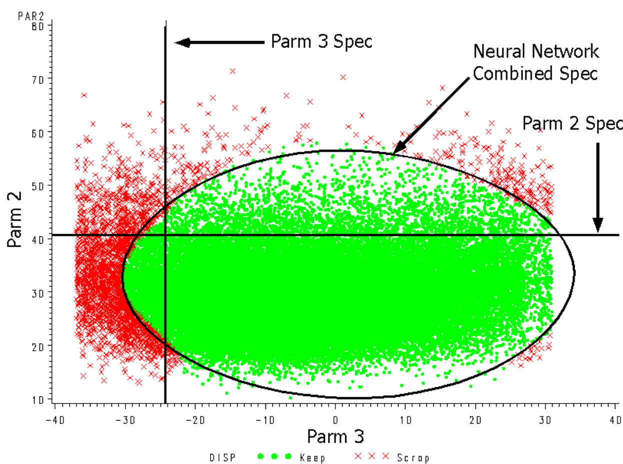

Figure 19: Acceptance Region for GANN Spec

CONCLUSION

Complex manufacturing operations, especially in a vertically integrated enterprise, require a global approach to SPC. This requires a quality data warehouse for assembling and correlating cross-functional data, and analysis tools for detecting change and diagnosing root cause. The JED system, which we developed to support these requirements, is proving very effective in improving and maintaining the yields in the disk drive factories of the IBM Storage Technology Division. The system makes extensive use of SAS software, including SAS/ACCESS for data extracts, SAS macros for data manipulation and plotting, SAS/IntrNet for web deployment, and SAS Enterprise Miner for in depth analysis and modeling.

REFERENCES

Klenz, B. W. and Fulenwider, D. O. (1999), *The Quality Data Warehouse: Solving Problems for the Enterprise*, Proceedings of SUGI 24, SAS Institute Inc.

Potts, W. J. E., (1999) *Decision Tree Modeling*, Course Note, SAS Institute Inc.

Potts, W. J. E., (1999) *Neural Network Modeling*, Course Notes, SAS Institute Inc.

Rutledge, R. A., Tai, L. H. and Helf, G. W. (1999), *Just Enough Database for Manufacturing Yield Analysis*, Proceedings of SUGI 24, SAS Institute Inc.

ACKNOWLEDGMENTS

The author would like to express his appreciation to the team at IBM who have worked on the JED project: Linda Tai, Garth Helf, Roger Buland, Barbara Okimoto, Heidi Leung, Sysouk Sananikone and Catherine Turnbull.

And to the many SAS employees who helped me to understand and use the SAS software products:
- Kim Fields and Greg Lloyd (SAS San Jose Office)
- Jim Georges and Will Potts (SAS Enterprise Miner)
- Rena Soffir and Aaron Hill (SAS/INTRNET)
- Sam Atassi and Margo Scharer (SAS/INTRNET)
- Jean Moorefield (WEBAF)
- Regina Louie and Jim Luther (SAS/CONNECT)
- Bob Rodriguez, Brad Klenz and Donna Fulenwider

CONTACT INFORMATION

Robert A. Rutledge
IBM Storage Technology Division
5600 Cottle Road, San Jose, CA 95193
Work Phone: (408) 256-0804
Email: brutledg@us.ibm.com

SAS, SAS/INTRNET, SAS/GRAPH and SAS Enterprise Miner are registered trademarks of SAS Institute Inc. in the USA and other countries. IBM, DB2 and RS/6000 are registered trademarks of International Business Machines Corporation. ® indicates USA registration.

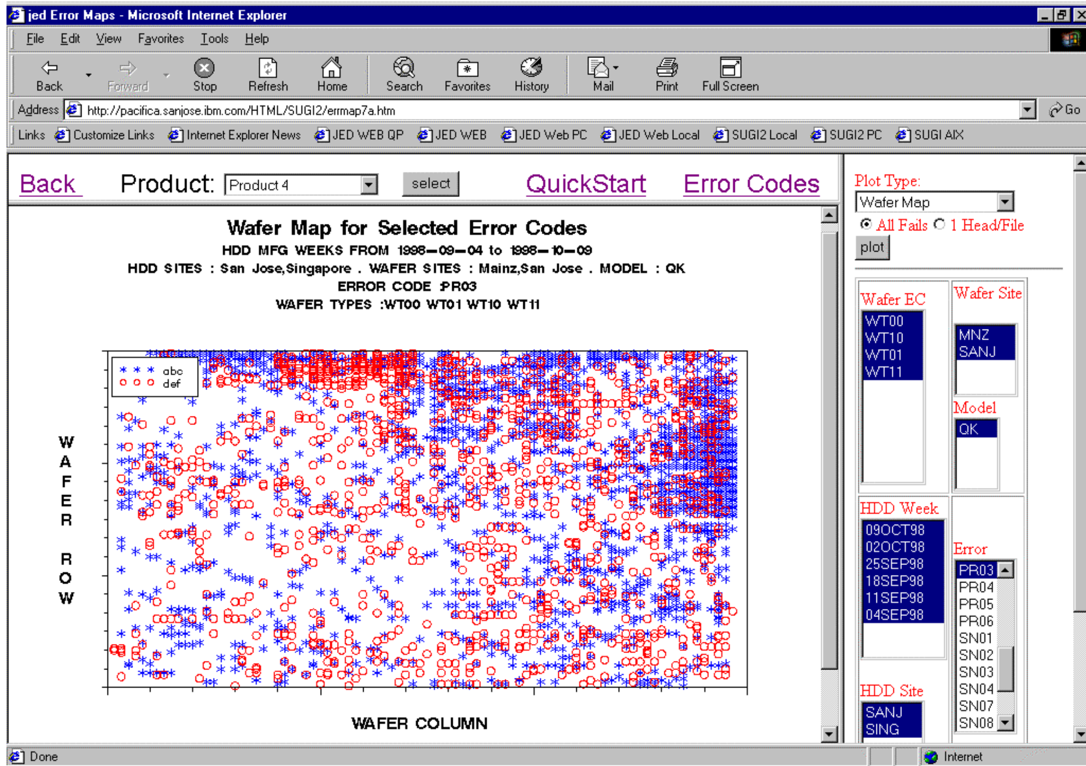

Figure A-1: JED Web Page Showing Wafer Map for Selected Error Codes

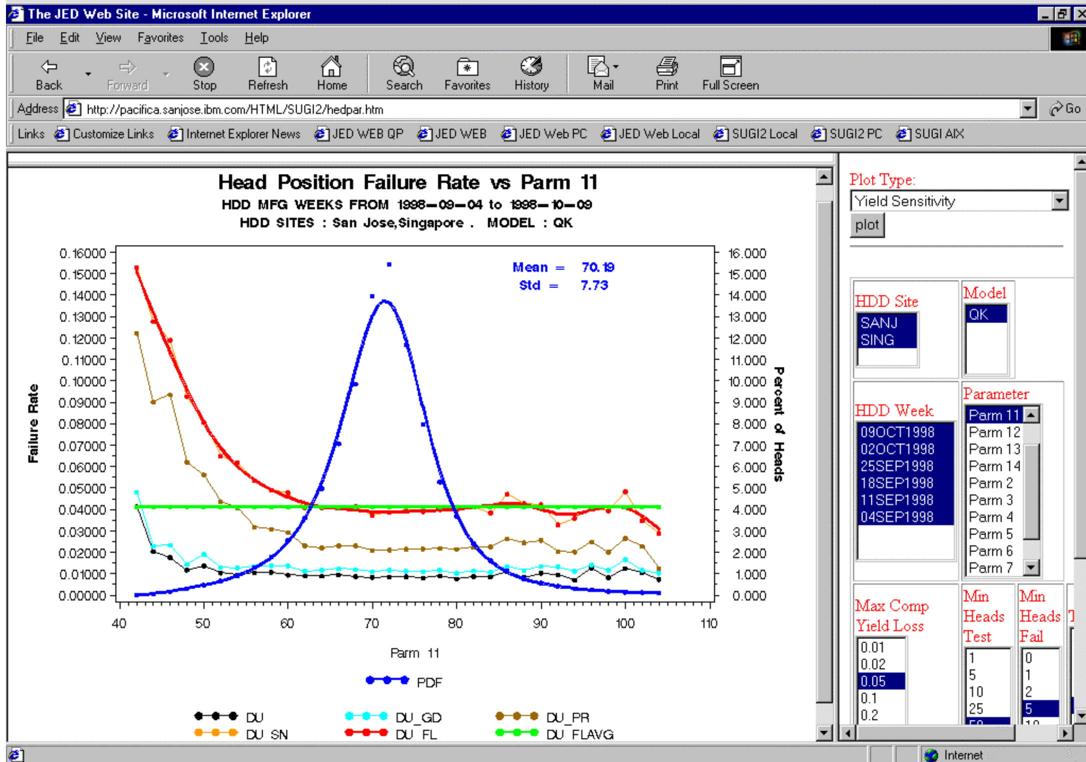

Figure A-2: Yield Sensitivity Curve

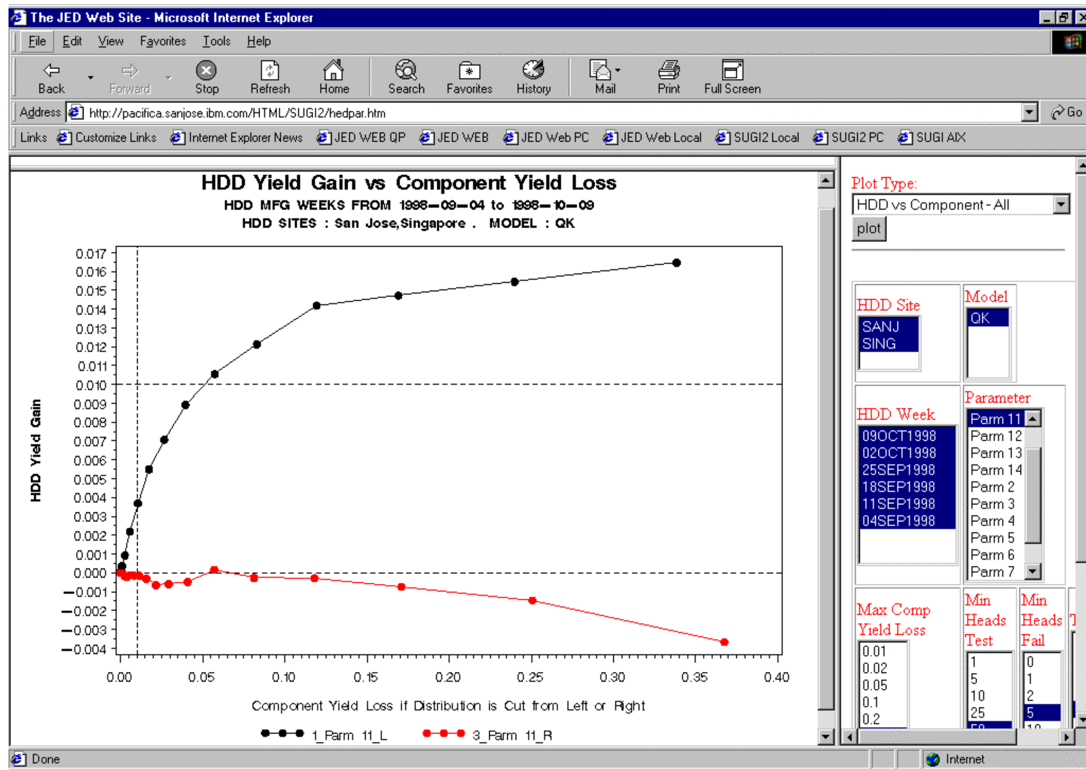

Figure A-3: dY vs. dS for Component Specification Tightened on Left or Right Side

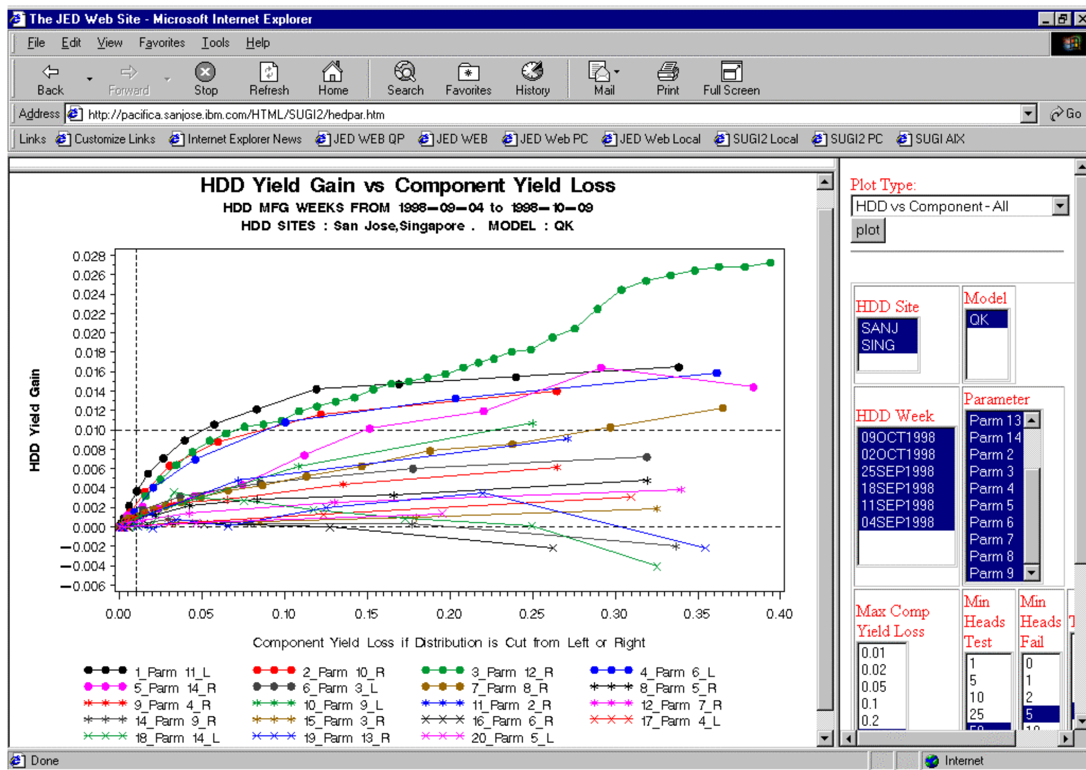

Figure A-4: dY vs. dS for Multiple Parameters

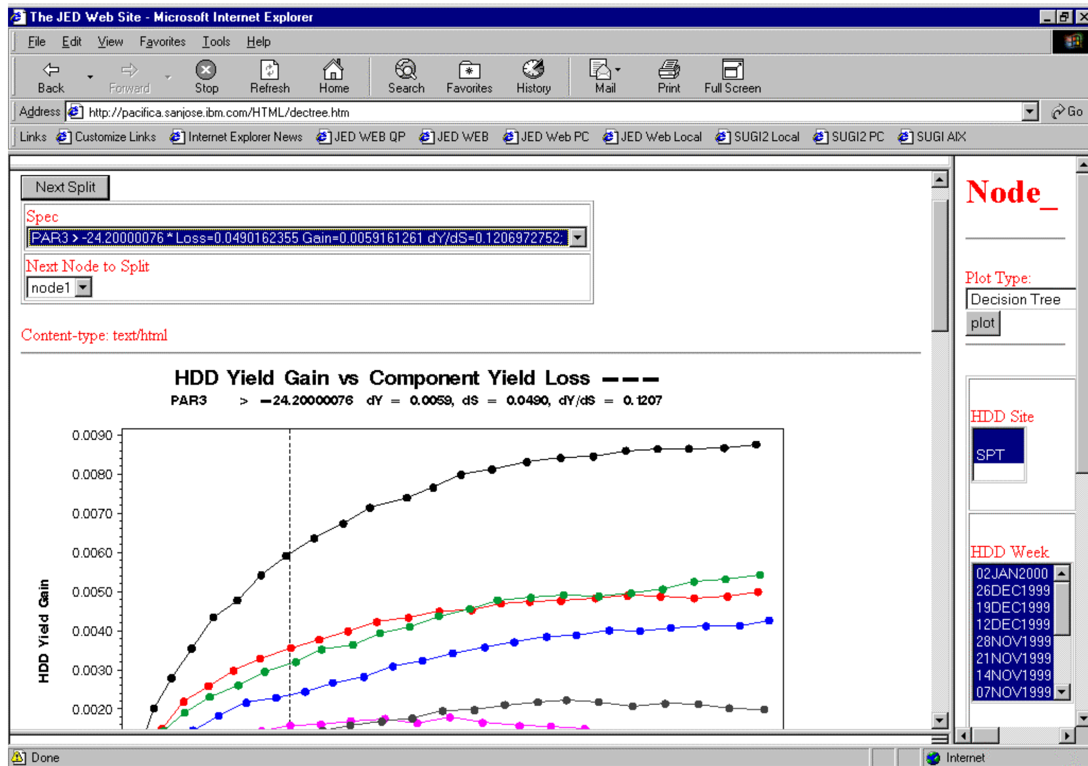

Figure A-5: Plot for Choosing First Decision Tree Split