Component-Derived Manufacturing Yield Prediction in Circuit Card Design and Assembly

by

Stephen Trinh

B.S. Mechanical Engineering, Tulane University, 2007

Submitted to the MIT Department of Mechanical Engineering and the MIT Sloan School of Management in Partial Fulfillment of the Requirements for the Degrees of

Master of Science in Mechanical Engineering
and
Master of Business Administration

In conjunction with the Leaders for Global Operations Program at the
Massachusetts Institute of Technology

September 2013

Signature of Author ______________________________
Department of Mechanical Engineering, MIT Sloan School of Management
August 9, 2013

Certified by ______________________________
Martin Culpepper, Thesis Supervisor
Professor, Department of Mechanical Engineering

Certified by ______________________________
Roy Welsch, Thesis Supervisor
Professor of Statistics and Management Science, MIT Sloan School of Management

Accepted by ______________________________
David E. Hardt, Chairman, Committee on Graduate Students
Professor, Department of Mechanical Engineering

Accepted by ______________________________
Maura Herson, Director of MIT Sloan MBA Program
MIT Sloan School of Management

This page intentionally left blank.

Component-Derived Manufacturing Yield Prediction in Circuit Card Design and Assembly

by

Stephen Trinh

Submitted to the MIT Department of Mechanical Engineering and the MIT Sloan School of Management on August 9, 2013 in Partial Fulfillment of the Requirements for the Degrees of Master of Science in Mechanical Engineering and Master of Business Administration

Abstract

Circuit card manufacturing can be a highly risky and volatile proposition due to the placement of hundreds of small, high value components. Operator mistakes, design errors, and defective parts lead to thousands of dollars in troubleshooting and rework costs per product. Raytheon Integrated Defense Systems (IDS) Circuit Card Assembly (CCA) manufactures highly complex circuit cards at a high mix / low volume scale for various purposes. Due to the high input variability and small production lot sizes of this level of circuit card manufacturing, historical trending and defect mitigation is difficult, causing a significant portion of CCA's manufacturing costs to be attributed to troubleshooting defects and rework.

To mitigate these costs, yield prediction analysis software is utilized to predict potential manufacturing defect rates and first pass yields of new designs. This thesis describes the creation and testing of a new data analysis model for yield prediction. By gathering and processing data at an individual component level, the model can predict defect rates of designs at an assembly level. Collecting data at the individual component level drives more comprehensive component-based calculations, greatly improving yield prediction accuracy and thereby allowing cost effective circuit card designs to be created. The increase in prediction accuracy translates to a potential $250,000 saved annually for Raytheon CCA from early defect identification and removal. Updated data retrieval and calculation methods also allow for much easier model maintenance, thereby increasing the relevance of yield prediction. This model can be easily incorporated into other design software as a next step in creating comprehensive concurrent engineering tools.

Thesis Supervisor: Martin Culpepper
Title: Professor, Department of Mechanical Engineering

Thesis Supervisor: Roy Welsch
Title: Professor of Statistics and Management Science, MIT Sloan School of Management

This page intentionally left blank.

Acknowledgments

I would like to thank all the wonderful folks at Raytheon for providing me the opportunity to work on this project as well as all the support and assistance I received throughout. I want to thank Charlie Mullins and Jeff Shubrooks for supporting LGO and the project, Tom Gustafson for all the analysis and ground support, Steve Tribuna for putting up with an unending stream of questions about database structure and data retrieval, Mike Shimazu and Dawn Gilbert for getting me started in SQL properly; and the Test Engineering group for providing different perspectives on the issues at hand.

I would like to thank my advisors, Roy Welsch and Martin Culpepper, for their incredible insights and assistance in completing my project.

The author wishes to acknowledge the Leaders for Global Operations Program for its support of this work. I also would like to thank all of my classmates, especially my summer team, for teaching me leaps and bounds more than I could ever figure out on my own. I will never forget to use the proper frameworks when tackling daunting situations, to recognize and utilize the bullwhip effect efficiently, and to apply myself 110% at every situation, no matter the volume.

Finally, I would like to thank my family for the many years of support and dedication to my future and without whom this would not be possible.

Regul8tors, mount up.

This page intentionally left blank.

Table of Contents

This page intentionally left blank.

List of Figures

This page intentionally left blank.

1 Introduction

Yield prediction analysis has been used in many design-to-manufacturing settings to mitigate the acceptance of potentially costly or defective designs. Prediction software is generally implemented during the design phase, where adjustments can be made with minimal impact to the bottom line. This is traditionally exercised in circuit card manufacturing with the use of component attributes as the driving factors to calculate a "Defects per Million Opportunities" (DPMO) score. Yield prediction is achieved, but with inconsistent accuracy levels. This project will explore the use of individual component data, rather than aggregated component attribute data, in conjunction with component modeling, to create a more robust and accurate method of yield prediction.

This thesis focuses on the execution and results of a six-month research project performed in collaboration with Raytheon Integrated Defense Systems (IDS) Circuit Card Assembly (CCA) at Raytheon's Andover site. These methods will be tested in Raytheon's data rich environment with the ultimate goal of improving current yield prediction methods. The first section of this thesis describes the historical aspects of yield prediction and the current landscape of yield prediction as it pertains to Raytheon. The second section describes the data mapping, retrieval, and analysis used to create an alternate prediction method. The final section describes the result of the analysis its implications on Raytheon and future research. An overview of Raytheon within the context of project management will also be given.

1.1 Project Drivers

Circuit card manufacturing can be a highly risky and volatile proposition due to the placement of hundreds of small, high value components. Operator mistakes, design errors, and defective parts lead to thousands of dollars in troubleshooting and rework costs per product.

Raytheon CCA manufactures highly complex circuit cards at a high mix / low volume scale for various purposes. Due to the high input variability and small production lot sizes of this level of circuit card manufacturing, historical trending and defect mitigation is difficult, causing a significant portion of CCA's manufacturing costs to be attributed to troubleshooting defects and rework. In an attempt to lower these costs, CCA utilizes yield prediction software called PCAT to predict the producibility, or first-pass yield, and DPMO of new designs. Historically, yield prediction for circuit card manufacturing relied on models based off of individual attributes of components, as evidenced in PCAT. In these models, yield is calculated by assigning a weight to each identified component attribute within a derived DPMO scale. Attribute DPMO scores are then combined to generate a DPMO score for a complete assembly. This method, while fairly accurate and the best choice in the absence of manufacturing data, does not take into account manufacturing data and knowledge when data is present. For example, if a single component or a set of components are known to be problematic through manufacturing knowledge, the high risk of using said components would be lost in attribute aggregation if components with similar attributes are not problematic.

CCA currently spends over a million dollars a year in troubleshoot and rework costs. Increased prediction accuracy will reduce the total spent in rework by better identifying known high-risk components and design features from being introduced during the component selection and placement process. Even a defect reduction as small as 5% would translate to $50,000 saved annually, as well as reduced man-hours and less inventory loss.

1.2 Research and Results

Research was performed to identify possible analysis methods using component-based defect data. The results indicate that some inherent flaws in the current prediction method are addressed by component-based calculations, allowing a potential 25% increase in prediction accuracy. This accuracy increase, converted to potential defects averted from final circuit card designs, would save Raytheon over $250,000 annually in rework costs.

Some vulnerabilities and potential improvement opportunities such as input data integrity and outlier adjustments were also identified during the model development and testing process. The details of the analysis performed and the improvements identified are explained in greater detail in Sections 5 and 6.

2 Background Information

2.1 The Printed Circuit Card and Circuit Card Assembly

Early development of the printed circuit card began in the early 20^{th} century, with various inventors scribing conductive material in deliberate patterns onto insulating surfaces. Some notable inventors include Albert Hanson, Thomas Edison, and Charles Durcase. In 1936, Paul Eisler invented the soon-to-be modern printed circuit, which then became commonplace by the mid 1950's. Printing circuits involved a conductive metal, such as copper, to be etched or laminated onto a non-conductive surface, such as silicon. Up until the 1980's through-hole fabrication techniques were used in conjunction with various soldering techniques such as dip-soldering and wave soldering to assemble boards and components. Surface-mount technology (SMT) then replaced through-hole as the common fabrication technology [1]. The most common fabrication methods used today are surface-mount components soldered onto printed circuit cards using reflow soldering techniques.

A circuit card assembly is made up of many different electronic components such as diodes, transistors, chips, resistors, and capacitors. Each component has two or more electrical terminals, or leads, that are soldered onto specific connection points on a printed circuit card. For through-hole technology, components have long leads that are fed through holes in the circuit board, soldered, and then trimmed. This has led to waste in the form of holes and trimmed leads. With SMT, components either have short or flat leads that are soldered onto spaces called pads on the circuit card. Figure 1 demonstrates the difference between the two technologies. Since SMT components require smaller or no leads as opposed to their though-hole counterparts, increasingly smaller components have been made with surface-mount in mind (Figure 2). [2]

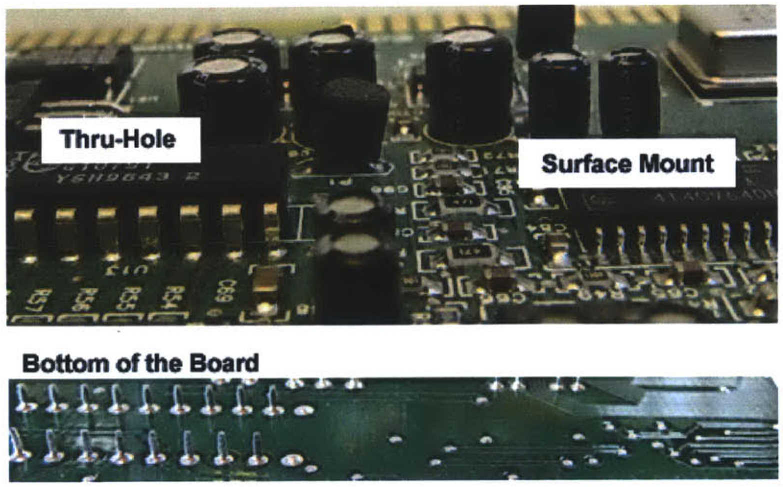

Figure 1: Through-Hole vs. Surface Mount Technology

Figure 2: SMT Component Size

Currently, automated printed circuit assembly uses SMT components. These components are precisely placed onto a printed circuit by pick-and-place machines. They are then either bulk wave soldered or, as in the case of ball grid array (BGA) packages, soldered in reflow ovens depending on component package type. In some cases, components only come in through-hole form, in which case, through-hole and SMT construction is combined. After initial construction, the completed card is tested both mechanically and electrically. If it fails, then diagnosis and possible component replacement may be required, also known as rework.

2.2 Testing and Standard Measurements – DPMO

Defects Per Million Opportunities (DPMO) is a common metric used in process improvement efforts such as Six Sigma. DPMO is used to measure process performance and is meant to be a comparable standard of measure across product complexities. As the name suggests, DPMO counts the number of defects, or predicted defects, and compares it to the number of total defect opportunities. The ratio is then multiplied by 1,000,000 to get DPMO. Defining what constitutes a defect or an opportunity, though, lies in a relatively gray area. Industry standards have provided general guidelines, but the full definitions vary by company.

A defect is generally defined as a deviation or imperfection that causes inadequacy or failure. In circuit card assembly, defects can manifest in many ways, such as bad components, skewed placement, and disconnected leads. Figure 3 below lists the some common defects found. How a defect is counted involves a little bit of creativity, and varies by manufacturer. Some defects are counted simply. For example, a broken component that is replaced counts as one defect. Other defects, on the other hand, may be counted by either the cause or effect of the defect. A slightly skewed component placement that causes two leads to be unsoldered, for example, can be counted as one defect for the skew, two or four defects for the unsoldered leads, or more for a combination of these defects.

Adhesive Adhesion	Component Alignment	Solder Excess
Adhesive Contaminated	Component Lifted	Solder Fracture
Adhesive Crack	Component Defective	Solder Insufficient
Adhesive Missing	Component Orientation	Solder Open
Adhesive Void	Component Lead Damage	Solder Short

Figure 3: List of Sample Manufacturing Defects

An opportunity can be defined as a possibility for a defect to occur. In circuit card assembly, opportunity count can be complex due to the range of failure possibilities. Opportunities generally include the components, the number of leads per component, connection possibilities, and location possibilities. IPC-

7912 has become the standard for DPMO calculation in circuit card assembly [3]. IPC-7912 defines DPMO as

Figure 4 shows.

$$DPMO = \left(\frac{\sum d_x}{\sum o_x}\right) \times 10^6$$

$$DPMO = \left(\frac{\sum d + \sum d_p + \sum d_t}{\sum o_c + \sum o_p + \sum o_t}\right) \times 10^6$$

Component Defect (d)	Damage to a component exceeding the limits of component specification
Component Opportunity (o_c)	Each device or piece of hardware that may be assembled onto a circuit card.
Placement Defect (d_p)	Component presence or positioning error during assembly that violates specified dimensional criteria.
Placement Opportunity (o_p)	Presence or positioning of a component on a circuit card.
Termination Defect (d_t)	Any electrically joined termination that violates specified requirements.
Termination Opportunity (o_t)	A hole, land, or other surface to which a component is electrically terminated.

Figure 4: IPC-7912 DPMO Formula

Raytheon CCA manufactures thousands of products yearly across approximately 1,500 designs, with up to 300 new designs being introduced each year. This involves tens of thousands of individual components of several different types. To simplify obtaining the opportunity count for each component, and subsequent assembly, Raytheon uses the following formula:

$$DPMO = \left(\frac{\sum d + \sum d_p + \sum d_t}{\sum((2 \times o_t) + o_c)}\right) \times 10^6$$

Figure 5: Raytheon DPMO Formula

The main variation from current industry standard is the identification of opportunity. Placement opportunities and termination opportunities are combined and then simplified to two times the terminal count. This calculation method may have been left over from through-hole manufacturing, where placing leads into holes automatically aligned the components, and defects could originate from both the hole and the component terminal. This opportunity count variation has an effect on prediction accuracy, as will be discussed in Section 5.3.

2.3 Prior Research

With the introduction of SMT and Six Sigma practices in the 1980's, researchers have focused on improving circuit card manufacturing through concurrent engineering and yield prediction. Various papers have been published describing the importance of accurate DPMO and yield prediction. These academic papers develop circuit card yield prediction using defect data as a base to calculating DPMO. Given enough data, defect rates can be tied to component package types, which then can be used in turn to calculate DPMO in a weighted algorithm. Analysis programs can then be developed to provide metrics at each stage of manufacturing. [4] [5] [6]

3 The Potential for Improvement

3.1 PCAT and PCAT Express

PCAT is a capability predictor designed by Raytheon engineers about 20 years ago, and has been widely used at Raytheon for the past 10 years. On the back end, process capability data such as costs, quality, and cycle time are used to create a model of various design characteristics. The user then selects specific design characteristics with which PCAT then calculates the prediction in terms of DPMO and other derived metrics.

To create PCAT's internal model, manufacturing data, cost data, and previous design data are combined to generate DPMO ranges for each component package type. Within each package type, component attributes such as lead type, lead pitch, locating features, and component size are given calculated weight scores. For a new design, individual components are weighted based on their attributes to give a DPMO within their component package range. The final design DPMO is calculated using these component DPMOs. Figure 6 and Figure 7 provide visual representations of PCAT's process. [7]

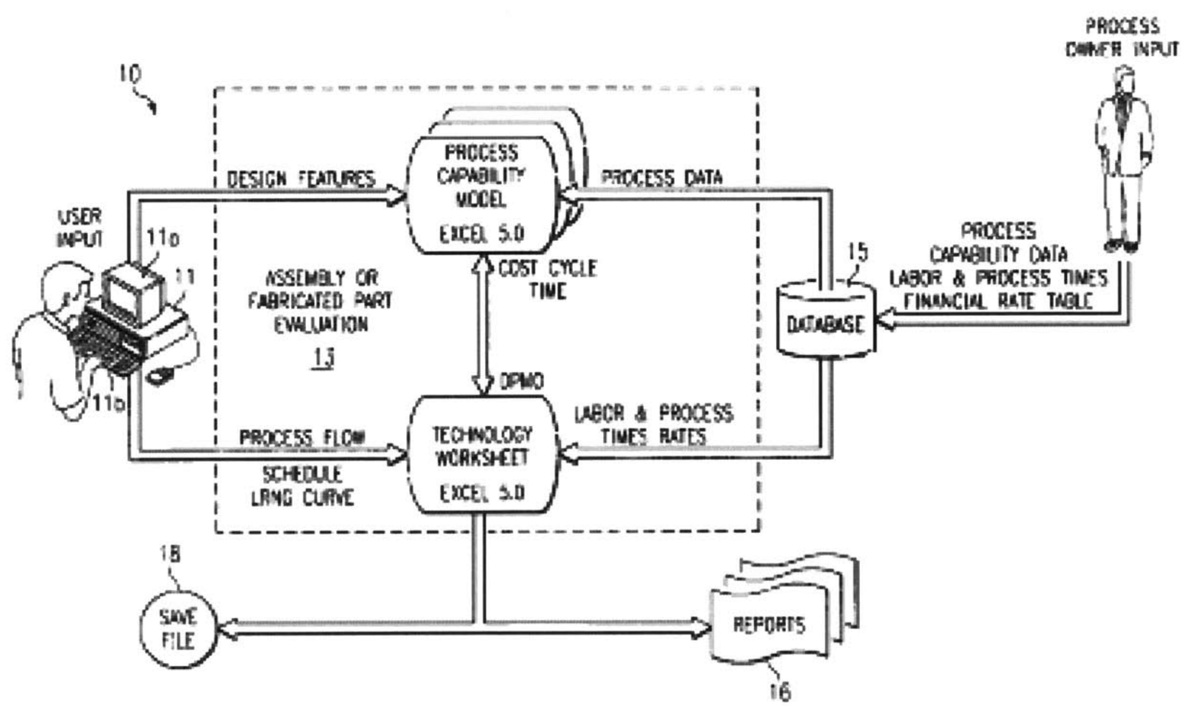

Figure 6: PCAT Process Graphic

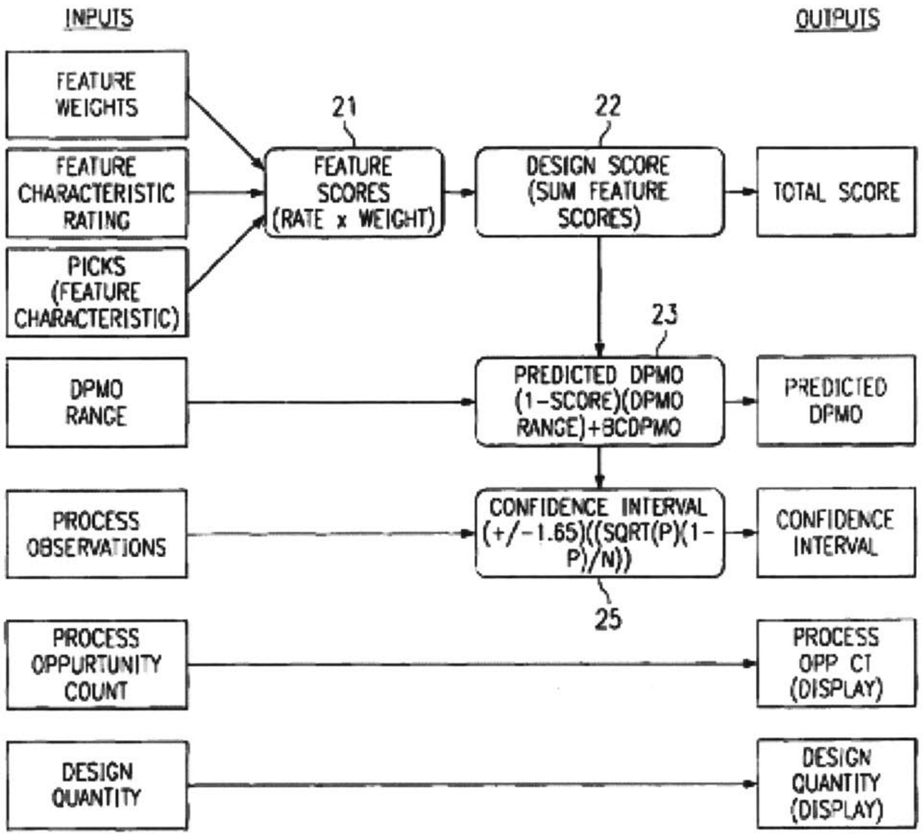

Figure 7: PCAT Process Flow Chart

With newer advances in manufacturing technology and standardization of many components and other design features, many of the analysis details and choices that go into PCAT calculations were becoming obsolete. These items cluttered the user interface and made the process long and arduous. For example, PCAT required the entry of every dimension of a particular component layout, which had become standardized for all related components. Manually inputting standard dimensions became very tedious and required the user to be a subject matter expert. Thus, PCAT Express was created. PCAT Express was a program that was built on top of PCAT. Express automated input of most of a component's attributes, with only a few important component attributes such as lead pitch and lead material requiring manual input from the user. Express also streamlined user input methods to make the overall PCAT experience more efficient. It is necessary to emphasize here that PCAT Express is an overlay on top of PCAT; all original PCAT calculations are used in PCAT Express.

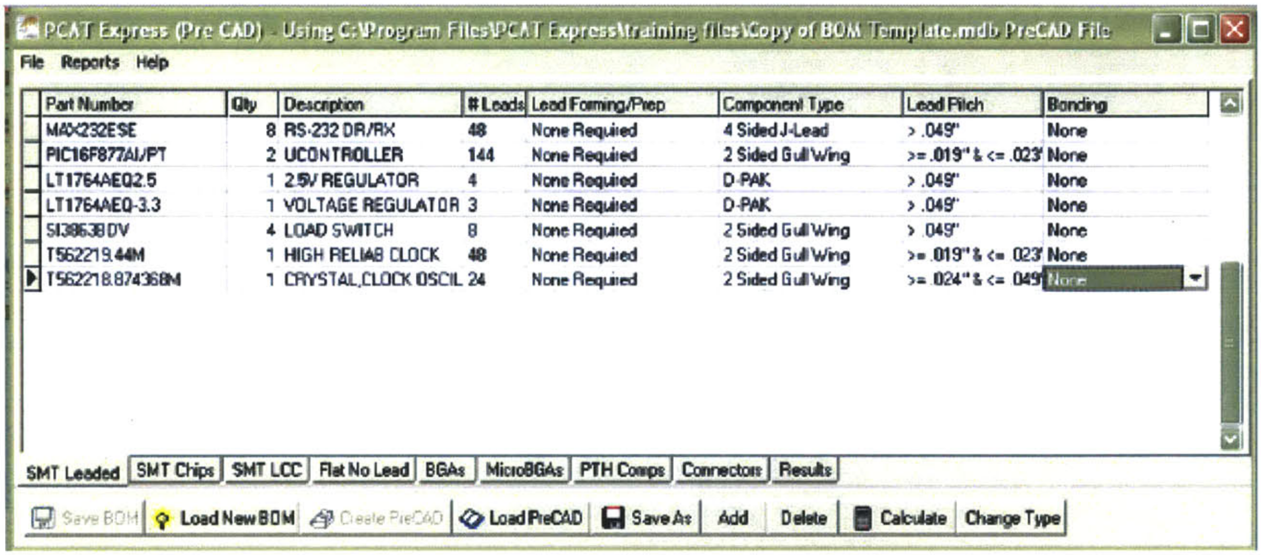

Figure 8: PCAT Express Program

3.2 The Potential for Improvement

Until recently, it was believed that PCAT scores were accurate to within 10% of actual range. PCAT's score rating system can be summarized in Figure 9. This system rated the quality of designs based on PCAT score, and was the basis for testing the accuracy of PCAT in the development of a separate Design-for-Test tool. The subsequent PCAT accuracy plot is provided in Figure 10. [8]

Design Perspective Rating System	CCA DPMO Range	Sigma Range	Color Code
Best Design	Less than 300	Over 5.0	
Average Design	300-700	4.7-5.0	
Below Average	700-900	4.6-4.7	
Difficult Design	Over 900	Less Than 4.6	

Figure 9: PCAT Scoring System

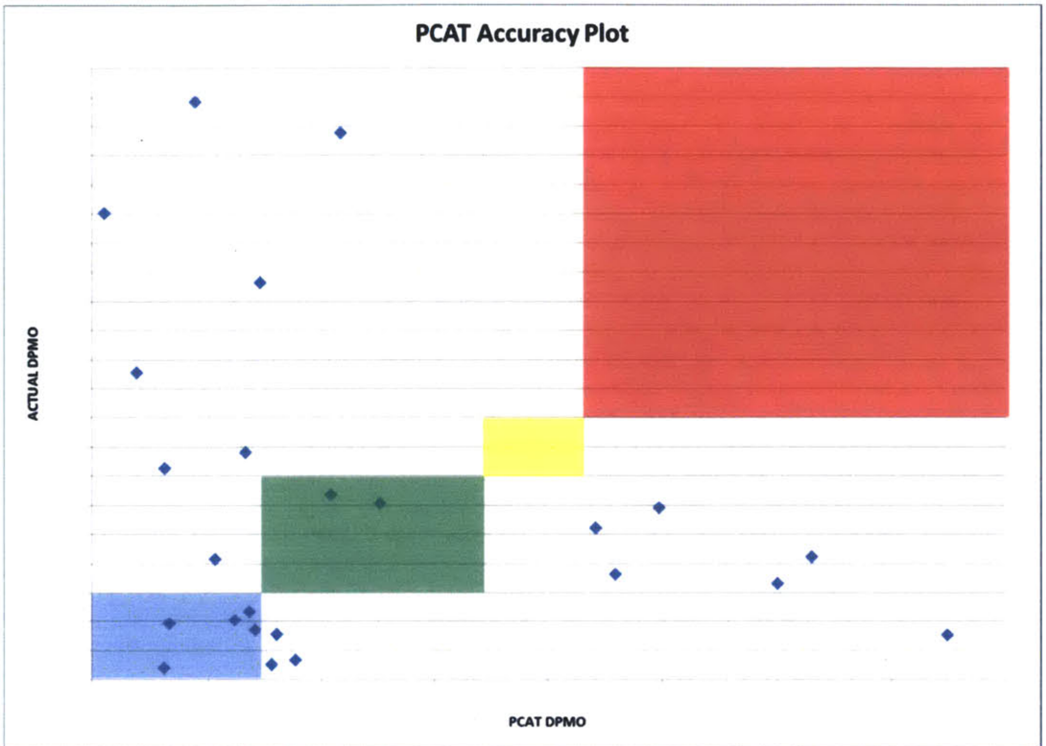

Figure 10: PCAT Accuracy Plot

As can be seen in Figure 10, the predictions perform reasonably well for most low DPMO assemblies. As actual DPMO increases, PCAT accuracy greatly decreases, diverging farther and farther from actual DPMO in both the positive and negative directions.

PCAT's inaccuracy stems from two main sources: the difficulty of updating such a complex model and the inherent inaccuracy of aggregating individual attributes to assess whole components.

In theory, updating PCAT's model is complex. In practice, it is much more difficult. Theoretically, each section and subsection of the model must be updated individually. For each component attribute, a new DPMO range must be determined, followed by the weight given to each possible choice for that attribute. Since the same attribute may be chosen for components with different sets of attributes, the changes must be compatible with all possible permutations involving that attribute. After every attribute has been updated, the overall model is then tested against existing assemblies for accuracy. In reality, great difficulty comes in trying to collect and parse data to update individual attributes. The same attribute may have a large range of defects, from very low in a simple component, to very high in a complex component. This makes determining a proper DPMO range and weight distribution very hard and tedious.

PCAT places weights on a component's attributes through a combination of historical data and engineering theory to determine its effects on DPMO score. While this method allows for fairly quick and easy modeling techniques, it may not be capable of capturing all possible attribute interactions. For example, two components with nearly identical attributes may have similar predicted defect rates but vastly different actual defect rates whether from a quirk in individual design, manufacturer, etc.

3.3 Test Theory

To address these improvement opportunities, data analysis and statistical prediction calculations are tested using component-based defect data as the base. The test methods will be referred to as Version 2. Prior research has indicated the viability of this method. [4] [5] [6] Selection of this particular method for this situation is based on the following reasons.

- Designers and engineers are habitual in their design and component selection process. They have a tendency to incorporate the same design features and choose the same or similar components in different designs. If a component has been used previously, it will likely be used again.
- Component placement is assumed to be discrete and independent events on an assembly. This means the placement success or failure of one component does not affect the placement success or failure of another component. This can be reasonably assumed due to the method of construction and testing for circuit cards.
- In statistical analysis and modeling, the contribution of discrete parts such as component attributes can be easily determined. The contribution of part interactions, however, can be incredibly difficult to determine. Assuming that attribute interactions exist and components are

discrete and independent of each other, the effects of individual attributes and their possible interactions can be captured by defining components as the base unit to measure.

- Defect data based on components can be further used beyond yield prediction. The data can easily be further analyzed to create easier routes for troubleshooting, reducing overall inventory by eliminating visibly error-prone and obsolete parts, and other uses.

4 Method

4.1 Understanding the Environment

To be able to effectively navigate and utilize stored data, one must understand the structure in which the data is stored. Many types of data structures exist, each with its own specific purpose. Examples are hierarchal, network, and object-based models. Raytheon uses a "data warehouse" structure to capture its manufacturing data. From that data warehouse, Raytheon employs a star schema model to access the data in an organized manner.

The star schema model separates data into two distinct table types with links between them. Fact tables store the metrics of specific events. They generally contain certain numeric values and foreign keys to dimensional data, used to link to dimension tables. Examples of fact tables are customer orders, manufacturing completion events, and non-conformance events. Dimension tables store specific characteristics of fact data attributes. These tables are linked to fact tables through a primary key to fully define fact data. Examples of dimension tables are item descriptions, time dimensions, and non-conformance types. Figure 11 gives an example of a visual representation of a star schema database. [9]

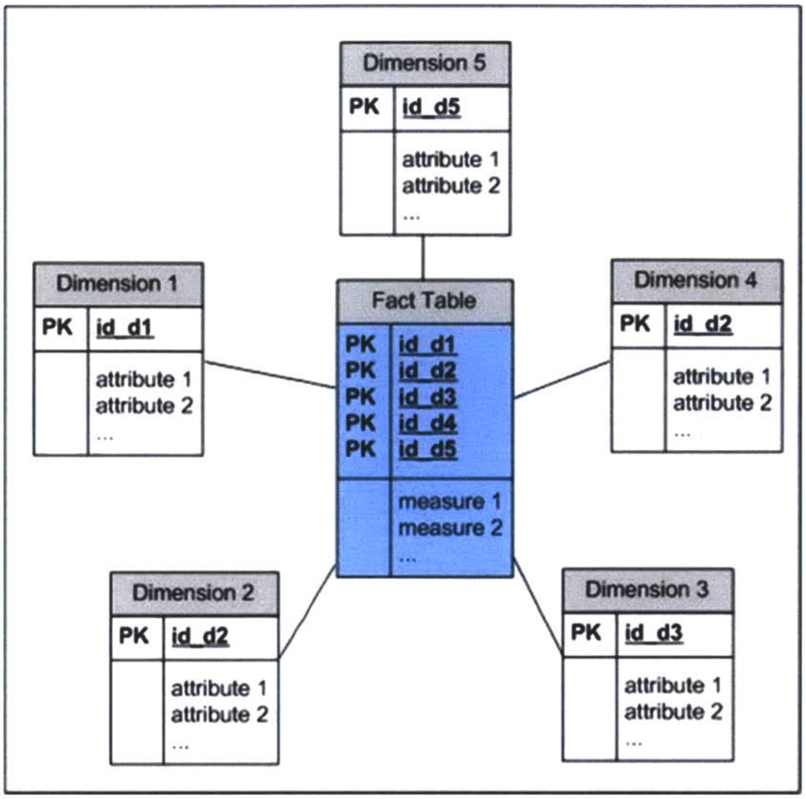

Figure 11: Star Schema Representation

A star schema structure is a denormalized data model, giving it several advantages and disadvantages. Denormalization allows for simpler queries when using join logic, query performance gains by limiting search scope, and faster data aggregations. The primary disadvantage comes in the form of data integrity. Denormalization relaxes data entry restrictions found in other models, opening the possibility of redundant data and erroneous data loading. This disadvantage is usually mitigated with strict controls during data entry.

4.2 Samples

To test Version 2, a sample of circuit card assemblies was selected for comparison. For comparison ability, the sample set was chosen with PCAT analysis as a criterion, and therefore not a random population sample. It is not entirely clear where the data behind PCAT's algorithms comes from, whether comprehensive or targeted manufacturing data. As such, no guarantees can be made that PCAT tested assemblies are representative of the population. While this set can be used as representative of CCAs that will be affected by the new test method, using fully comprehensive component data in the test method may or may not affect the accuracy of the comparison. Within this caveat, these assemblies were selected with the widest range of qualities available to increase testing integrity. The assemblies:

- All have reported PCAT scores,
- Range from ~50 to ~2600 installed components per assembly,
- Range from ~20 to ~20,000 assemblies built in the past two years, and
- Have predicted and actual DPMO scores ranging from ~20 to ~2,000.

To provide better modeling effects, assemblies with large external factors affecting actual DPMO were not considered in the sample. Figure 12 provides an example of the magnitude of assembly outliers created by known external factors.

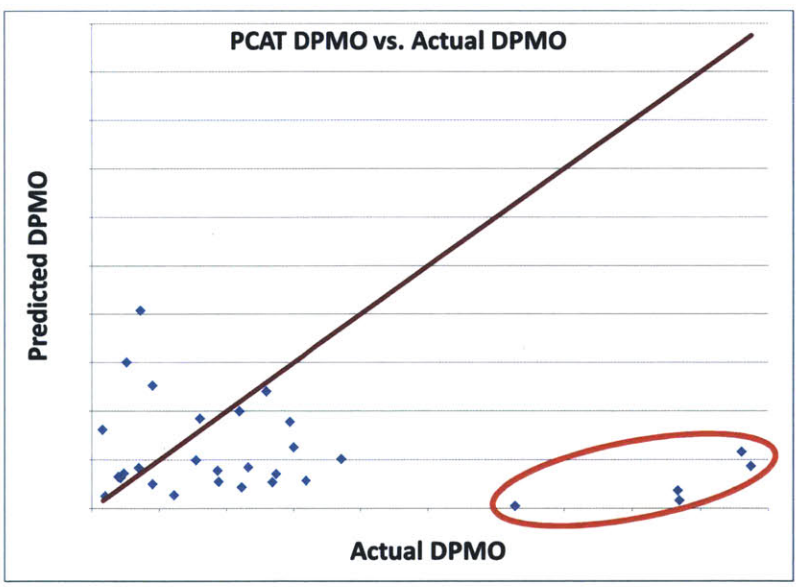

Figure 12: Example of External Factor Effects on DPMO

Figure 13 plots PCAT predicted DPMO of the sample set against the actual DPMO found from manufacturing data. In the figure, an accurate prediction would fall close to the red reference line. PCAT difference from actual DPMO in this sample set ranges from a few defects to over 1500, with deviations up to 800%.

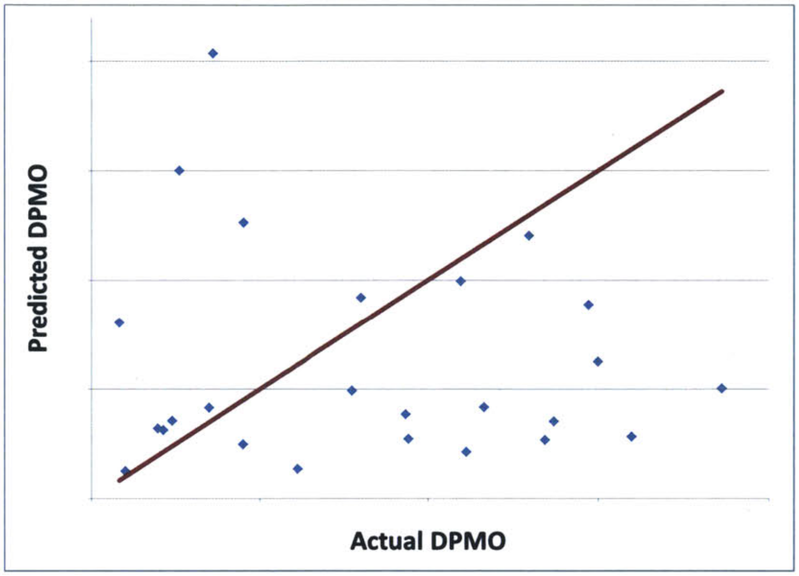

Figure 13: PCAT vs. Actual DPMO

4.3 Test Methodology and Algorithm Structure

Calculating Component Opportunity Fail Rate

Prior to assembly DPMO prediction, component fail rate must be calculated. Component fail rate is defined as the following:

$$P(Fail)_{comp} = \frac{\sum defects_{comp}}{\sum opportunities_{comp}}$$

Total opportunities for a component can be found by multiplying the total number of a single component built by the defined opportunity count of the component.

$$\sum opportunities_{comp} = \sum built_{comp} * opportunities\ per\ component$$

Total number of a single component built can be found by determining the number of a single component used in one assembly multiplied by the total number of that assembly built. The number of a single component built by assembly is then aggregated for all assemblies.

$$\sum built_{comp} = \sum \#\ used\ in\ assembly * \#\ assemblies\ built$$

As defined in Section 2.2, Raytheon calculates opportunities per component as twice the terminal count plus one for component.

$$opportunities\ per\ component = 2 * terminals_{comp} + 1$$

Incorporating these formulas gives the probability of failure per component opportunity. This is the probability per component opportunity, not per component.

$$P(Fail)_{comp} = \frac{\sum defects_{comp}}{\sum(\#\ used\ in\ assembly * \#\ assemblies\ built) * (2 * terminals_{comp} + 1)}$$

Defects, number used in an assembly, number of assemblies built, and number of terminals were all drawn from reported manufacturing data. An important decision to note here is the segregation of the data into monthly production groups. Section 5.3 provides the logic behind this method of data segregation.

Computing Predicted Assembly DPMO

Given the probabilities of an individual piece, statistical aggregation can be used to compute the probabilities of the whole. Failure probabilities of component opportunities can be used to determine the potential failure probability and subsequent DPMO of an assembly through statistical methods.

Assuming that each component installed is a discrete and independent event, that is, the installation of one component does not affect the installation of another component (Section 3.3), component probabilities can be combined as follows to find assembly DPMO.

The probability of failure per component opportunity is inverted to produce the probability of success per component opportunity.

$$P(Success) = 1 - P(Fail)$$

Given the above assumption, success probabilities are combined to give the probability that ALL opportunities will be successful.

$$P(Success)_{All} = P(Success)_{x_1} P(Success)_{x_2} P(Success)_{x_3} \dots P(Success)_{x_n}$$

The probability of total success is inverted to give the probability of ANY opportunity failure. It is important to note that this probability states that any one opportunity failure is an assembly failure. This can also be defined as First Pass Yield (FPY).

$$P(Fail)_{Any} = 1 - P(Success)_{All}$$

To find the average probability of failure for each opportunity, the probability of any opportunity failing is divided by the number of opportunities per assembly. This gives the average fail rate for each opportunity.

$$P(Fail)_{avg} = \frac{P(Fail)_{Any}}{n}$$

Multiplying by 1,000,000 gives the final predicted assembly DPMO.

$$DPMO_{Assy} = P(Fail)_{avg} * 1{,}000{,}000$$

The combined full formula is as follows.

$$DPMO_{Assy} = \frac{1 - \left(\left(1 - P(Fail)_{x_1}\right)\left(1 - P(Fail)_{x_2}\right)\left(1 - P(Fail)_{x_3}\right) \dots \left(1 - P(Fail)_{x_n}\right)\right)}{n} * 1{,}000{,}000$$

$$x = components$$

$$n = number\ of\ components\ in\ the\ assembly$$

5 Results

5.1 Algorithm

As described in the above section, the Version 2 converts individual failure probabilities into an aggregate DPMO for an assembly.

$$DPMO_{Assy} = \frac{1 - \left(\left(1 - P(Fail)_{x_1}\right)\left(1 - P(Fail)_{x_2}\right)\left(1 - P(Fail)_{x_3}\right) \dots \left(1 - P(Fail)_{x_n}\right)\right)}{n} * 1{,}000{,}000$$

$$P(Fail)_{comp} = \frac{\sum defects_{comp}}{\sum(\#\ used\ in\ assembly * \#\ assemblies\ built) * \left(2 * terminals_{comp} + 1\right)}$$

To put Version 2 into effect, data was retrieved from Raytheon's VM databases. Number of defects, opportunities per component, components per assembly, and assemblies built were tabulated and inputted to obtain final predictions. These predictions were compared against actual defect rates and PCAT predictions to determine potential improvements.

5.2 Improvements Achieved

To determine the accuracy of Version 2 and demonstrate any improvements over the existing methods, the test results were plotted against PCAT scores and actual results. Figure 14 plots PCAT scores vs. actual DPMO and Version 2 scores vs. actual DPMO. Actual DPMO was plotted along the x-axis with the predicted scores plotted along the y-axis. A red line was added to indicate the perfect prediction line. Prediction accuracy increases as a plotted point gets closer to the red line. Initial impressions lean towards Version 2 scores being more accurate than PCAT scores, with accuracy decreasing as actual DPMO increases.

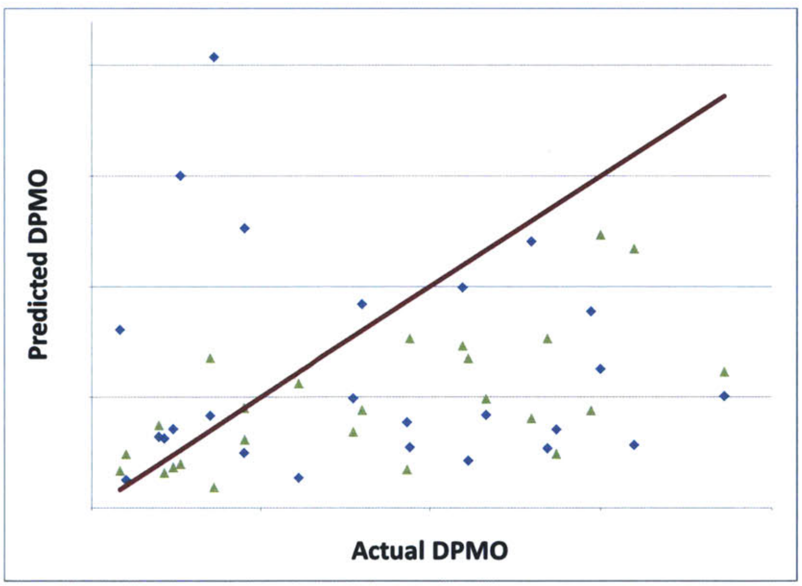

Figure 14: PCAT and Test Predictions Plotted vs. Actual

Examining PCAT scores plotted against actual data separately in Figure 15 shows a trend of DPMO underestimation and overestimation. The majority of PCAT scores are in the good to fair design range at the lower end of the graph. This indicates two things:

1. PCAT's model may not be tied to or related to real manufacturing data and design implications.

2. Designers have been designing to PCAT prediction specifications that are creating unforeseen defects.

Figure 15 also shows that designs with very high predicted DPMO scores have been passed to manufacturing. What this implies cannot be determined without further investigation.

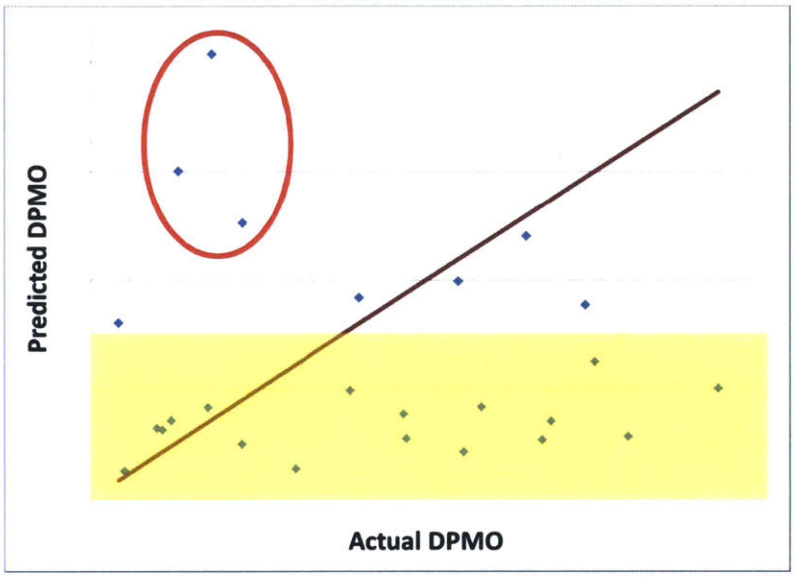

Figure 15: PCAT Predicted vs. Actual DPMO

Figure 16 tells a different story for Version 2 predictions. Most of the predictions underestimate DPMO, with a few exceptions due to known accuracy issues. The predictions also fall in a more conical cross-section, creating the implication that with increasingly problematic boards, component-based defect data becomes less relevant. It indicates that DPMO is somewhat correlated to components. This also shows that Version 2 predictions more closely follow actual events.

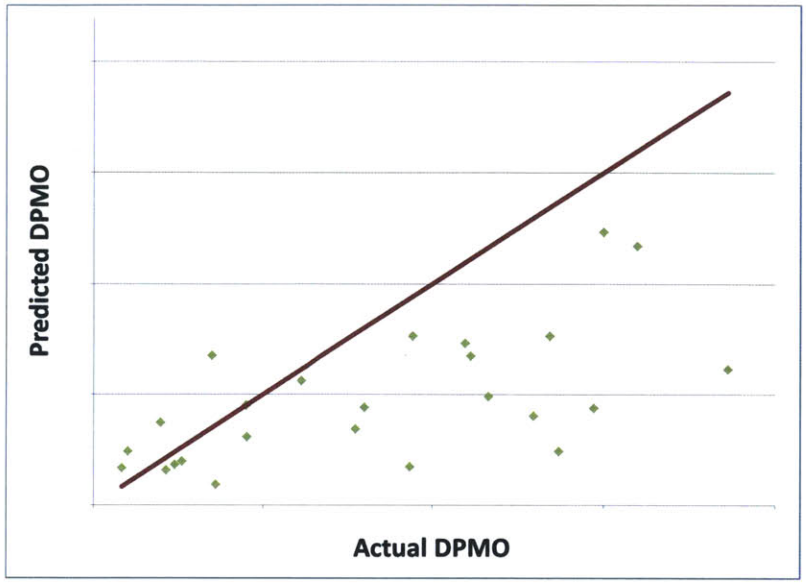

Figure 16: Test Predicted vs. Actual DPMO

Trend lines for both prediction sets were plotted to more visibly indicate how closely each prediction method follows actual data in Figure 17.

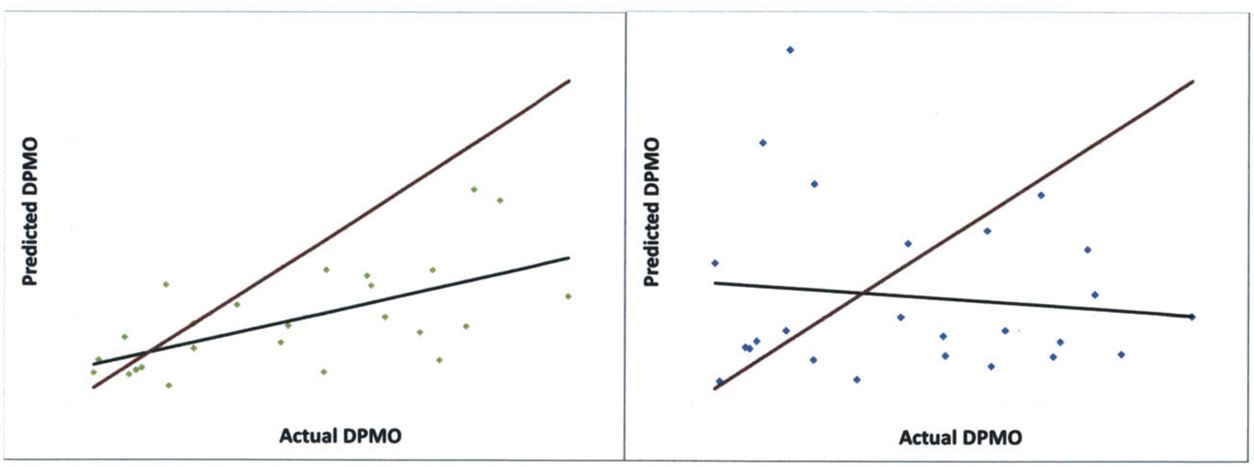

Figure 17: Prediction Score Trend lines

Figure 18 plots the absolute difference, in DPMO, of the predicted scores from actual data by sample assembly. The assemblies are ordered from smallest actual DPMO to largest actual DPMO to better distinguish data trends. The main trend indication is that the more problematic the board, the less accurate the prediction is. This again implies that, increasingly, defects are somewhat tied to components, and other factors are affecting defect rate.

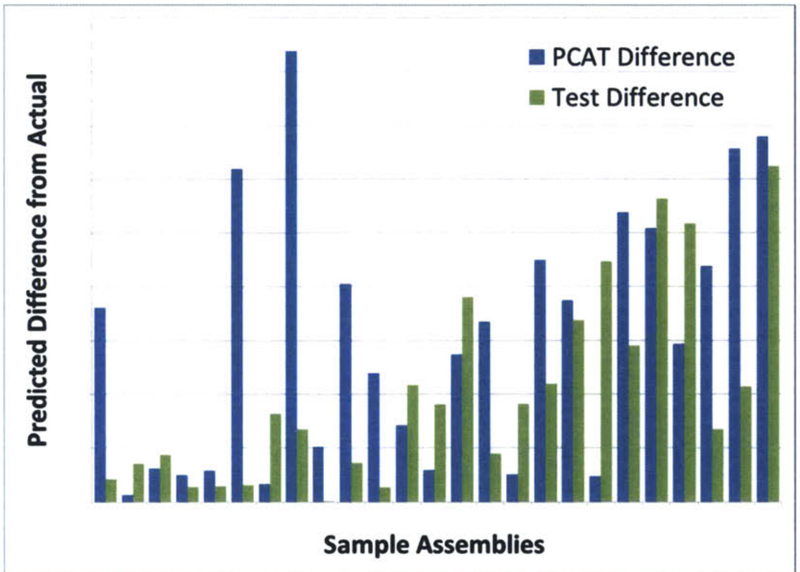

Figure 18: Prediction Score Absolute Difference from Actual DPMO

The above error data was further analyzed to better represent prediction differences. Figure 19 reports the range, average, and median deviations for the prediction models within the sample set.

	PCAT	Version 2
Range of Differences from Actual DPMO	66 – 1675	2 – 1249
Average Difference from Actual DPMO	610	407
Median Difference from Actual DPMO	587	326
Range of Deviation from Actual	7% - 881%	0.4% - 139%
Average Deviation from Actual	123%	53%
Median Deviation from Actual	63%	45%

Figure 19: Numeric Deviation from Actual

Figure 19 clearly shows that Version 2 produces more accurate predictions than current methods. Represented in a different manner in Figure 20, the accuracy difference can clearly be seen. The positive bars represent how often and to what magnitude Version 2 is closer to actual data. This is compared to the negative bars, which represent how often and to what magnitude PCAT is closer to actual data.

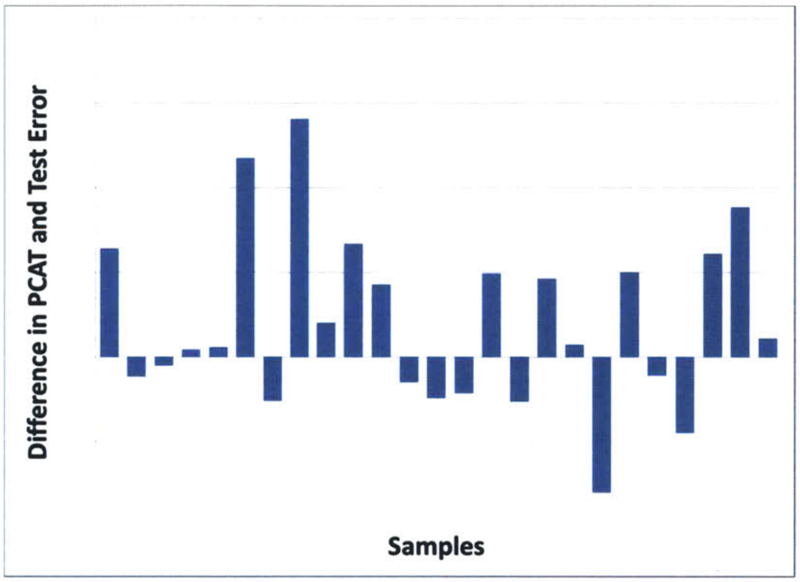

Figure 20: Prediction Score Error Comparison

As reported in Figure 19, Version 2 reduces average deviation from actual DPMO by 70%. This value was obtained by calculating the percent deviation for each sample assembly for each prediction method. These individual deviations were then averaged to find the difference in accuracy. The 70% value actually places an emphasis on the impact of individual sample assemblies, with the value increasing due to a few very high PCAT deviations. To get a result that is more closely tied to model behavior, the total deviations for both prediction methods were compared against the total actual DPMO. This gives a value of approximately 25%. This value also coincides with Figure 20. On average, Version 2 identifies 25%

more potential DPMO than PCAT. Within this test sample of assemblies, that translates to ~200 DPMO identified.

5.3 Accuracy

One of the most important features of any modeling analysis, and the one subject to the most scrutiny, is the accuracy relative to the actual data. Accuracy in Version 2 calculations varies with different choices and assumptions placed on the input data. Ideally, the prediction should be close to 100% accurate at all times. That level of accuracy requires perfect, complete data and a comprehensive understanding of manufacturing events and interactions. In the absence of perfect data and an incomplete understanding of event cause and effect, certain assumptions are incorporated to try to bridge the gap. These assumptions involve handling outlier data, filtering bad data, filling in missing data, and accounting for external and non-recorded events.

Component Data Outliers

The choices made when defining, gathering, and presenting raw data affects the visibility of data trends and outliers. The level of granularity chosen to group the data must be fine enough to see data anomalies, but not so fine as to hide the anomalies in a sea of data points. The easiest data collection method would be to count total defects and opportunities found for all time, but this aggregate data does not provide any trend data in any way. Breaking data down to individual defect events would create difficulty in data processing and trend spotting as well. Segregating component defect data by month provides an adequate level of visibility for outliers and some trending data as well as providing a natural database update frequency in line with existing Raytheon practices and norms.

Monthly data separation revealed outliers and awkward trends in some component data. For example, one component averaged 15 DPMO for over 20 months before spiking to over 12,000 in January 2013. Discovering these anomalies, while good for issue investigation, is not good for modeling. Unrelated outliers should be removed prior to modeling to provide nominal results. Also, separation by other metrics like assembly, work center, or manufacturer could allow better data analysis and more significant outlier visibility. Unfortunately, with the data given, separating data using more significant metrics than time was not viable.

Input Data Integrity

As noted above, the star schema data warehouse model suffers from a potentially significant vulnerability when it comes to data integrity. Since data restrictions that are present in other data management models are relaxed to accommodate the star schema model, rigorous controls must be in place during data entry to maintain data integrity. Despite the conversion to a new data warehouse model, Raytheon has maintained the legacy methods of data entry. These methods were optimized for quick changeovers and ad hoc repairs on the manufacturing floor by having low restrictions on the data entered. This, combined with changing organization structures and non-standard data entry practices, has created databases with both redundant tables and incomplete tables. For these reasons, use of the prediction model built on this data must be tempered with outside verification until more vigorous data entry methods are implemented and model accuracy improves.

Missing Opportunity Count Data

Specifically for this iteration of Version 2, some data must be assumed. Not all opportunity counts for each component were found by the time of this analysis. Therefore, two extremes were used to replace those numbers. First, the missing counts were assumed to be zero. This gives component success rates of 100%, ultimately providing the lowest possible predicted DPMO for each assembly. Second, the missing counts were assumed to be five. This gives lowest possible component success rates, ultimately providing the highest possible DPMO for each assembly. For all missing opportunity count data, final DPMO for an assembly will fall within this range. To stay conservative, the analysis performed above assumed worst case DPMO with missing opportunity counts set at five.

Overall Increase due to External Factors

Version 2 contains defect data based on assembly and component error. Certain other factors that may affect DPMO such as human error cannot always be captured due to the lack of concrete data. These factors are currently not incorporated into the final model. For a more complete model, studies should be performed to properly quantify the effects of these factors for incorporation into a later version of the model.

Opportunity-based Defect Reporting Considerations

Opportunity-based defect reporting can be imprecise at times. This stems from the fact that the opportunities counted are not independent of each other, and these dependencies must be taken into consideration. For example, a defective component is considered one defect. A missing component is also considered one defect. Two terminals on a single component unattached are considered either two or four defects because the terminals have to be reworked individually. Although the defective and missing component defects require more work to fix and involving more points of manual labor, the two unsoldered terminals count for more defects. This type of inaccuracy is unavoidable, but must be accounted for.

Sample Set Size

As with any test environment, the derived accuracy of Version 2 depends on the sample size used. Although the data pool is large, the test sample size is small. This brings into question the validity of accuracy claims placed on the test model. For a more reliable accuracy statement, Version 2 must be tested with many more test assemblies.

6 Discussion

6.1 Raytheon Implications

The potential impact to Raytheon, both in tangible and intangible effects, from implementing the improved Version 2 is sizable. The 25% increase in prediction accuracy can potentially translate to $250,000 saved annually. In addition, this data analysis establishes a foundation for more extensive analyses and improved tool development. PCAT can also be improved using the data gathered from Version 2.

Identifying improvements gained from using Version 2 requires careful consideration of test assumptions as well as recognition that, due to the nature of any prediction model, the benefits gained are potential benefits. Static characterization of direct benefits is difficult as best. Version 2 can improve prediction accuracy by 25%. Since most predictions are underestimating actual DPMO, this can be seen as 25% more potential defects revealed. Using sample set average actual cost set at 100% and PCAT predicted value calculated at 25% (masked to hide sensitive data), 25% lower defects would equate to a potential savings of $250,000 yearly.

In addition to potential monetary value, the improved data model can be used as a springboard for improving current and future troubleshooting and analysis systems. Identifying defect data in this way provides better visibility for meaningful outliers. These outliers can be investigated to find hidden causes of defects in both manufacturing and design settings. Systems can be developed to automatically flag and manage these outliers as they are identified.

In building Version 2, attempts to verify and compile defect data revealed vulnerabilities in CCA's current data input and management process. With these vulnerabilities identified, as discussed in Section 5.3, CCA can improve the affected data systems and create a more robust data environment.

Raytheon is currently in the process of creating and implementing a comprehensive design environment that will incorporate an array of design tools for engineers to use. Some of these tools are designed to assist in the design process, while others create design gates that must be passed through for design acceptance. Yield prediction will be integrated to mitigate defects during the design phase. Using the model defined here will help grade new designs, reducing defects at the most cost effective point of the process. This will also provide a meaningful way of passing manufacturing knowledge to designers.

6.2 Broader Implications

This analysis simultaneously proves the efficacy of component-based yield prediction while revealing that component-based data provides only part of the whole picture. As described in past research as well, breaking prediction models down to logical base units and developing data at that level will provide a decent prediction model. There is a deficiency, however, in the model, as evidenced by predicted and actual values diverging at higher defect rates. More research must be done to create a complete model.

6.3 Future Research

From here many possibilities exist for future research and improvement. The most prominent ones are listed below.

- For ease of collection and due to time constraints, defect data was grouped by month. Other groupings should be analyzed to reveal more meaningful trends. For example, categorization by assembly or supplier may show that certain assemblies or manufacturers are more problematic than others. This information can be used to streamline manufacturing efforts.
- Similar to the PCAT model, linking defect data with component attributes will allow the creation of predictive component modeling when direct data is scarce or unreliable. This line of research can be taken further than current models by introducing design features as a model variable.
- Other factors that span many components during manufacturing, such as specific manufacturing step or the labor intensity of a step, may also have an effect. This is evidenced by the divergence of the model from actual data as defect rates increased. Properly identifying non-component categories and analyzing their possibilities would create better model accuracy.
- Identifying possible links between components, component attributes, and other as yet unknown factors will also generate a more robust and accurate prediction model.

7 Raytheon

This section provides an overview of Raytheon in the context of project management and this project.

7.1 Company Background

Raytheon, headquartered in Waltham, Massachusetts, designs and manufactures defense systems for the U.S. Department of Defense and other customers around the globe. Founded in 1922, Raytheon started as a machinery manufacturing company, with notable products such as gas rectifiers and vacuum tubes. Raytheon, through a series of acquisitions of companies such as Hughes Aircraft Company and Texas Instruments Defense Systems, and with the divestiture of most nondefense business units, became a prominent player in the defense system industry. Today, Raytheon is the fourth largest defense contractor in the U.S., with $24 billion in revenue and employing approximately 68,000 people in 2012.

As a consequence of their growth by acquisition, Raytheon operates largely as multiple independent business units under the governance of a large umbrella company. These business units are headquartered around the U.S. with discrete presidents, goals, products, and revenues. A quick summary of the 6 Raytheon business units is shown in Figure 21.

Business Unit	Abbr.	HQ	Emp.	Rev.	Primary
Integrated Defense Systems	IDS	Tewksbury, MA	13,900	$5.0B	Missile defense systems, radar, naval electronics
Intelligence and Information Systems	IIS	Garland, TX	8,300	$3.0B	Information and intelligence products for military and government
Missile Systems	RMS	Tucson, AZ	11,900	$5.7B	Missile systems for the US and allies
Network Centric Systems	NCS	McKinney, TX	11,900	$4.1B	Network systems for sensing, command and control, air traffic control
Space and Airborne Systems	SAS	El Segundo, CA	12,300	$5.3B	Radar, intelligence and surveillance systems
Technical Services	RTS	Dulles, VA	10,400	$3.2B	Training, engineering services

Figure 21: Raytheon Business Units

Within IDS, Circuit Card Assembly (CCA) Center of Excellence (COE) sits almost as an independent "business within a business." CCA is the de facto producer of circuit cards for parts of IDS. For all other Raytheon business units, CCA COE must bid on and win contracts just the same with third party suppliers. This drives CCA to remain competitive and to continually seek out better and more efficient methods of circuit design and manufacture. It is within CCA that this project takes place.

CCA operates as an independent manufacturing entity. It currently produces ~1500 active assemblies, with approximately 300 new assemblies introduced every year. To do so it employs its own development and support groups such as quality, testing, finance, human resources, IT, and business development.

To better understand Raytheon, IDS, and CCA as a whole, the perspective of MIT Sloan's Three Lenses will be taken. These lenses consist of:

Strategic Lens: "organizations are designed (engineered) to achieve agreed-upon goals"

Cultural Lens: "organizations are shared mental maps, identities, assumptions"

Political Lens: "organizations are contests for power and autonomy among internal stakeholders"

7.2 Strategic Design

Organizations are machines. An organization is a mechanical system crafted to achieve a defined goal. Parts must fit well together and match the demands of the environment. Action comes through planning. Raytheon's growths through acquisition and ever changing plethora of products have greatly influenced the formal structure in which work is done.

7.2.1 Matrix Organization

Raytheon employs a matrix style of organization and structure, as demonstrated in Figure 22. In this structure, project-based teams such as specific products and task forces are populated with individuals from functional groups like operations, research, and engineering. An advantage of this structure is the dynamic and diversified teams with strong functional support, while the disadvantages are the increase in hierarchal complexity and split loyalties between functional managers and project managers.

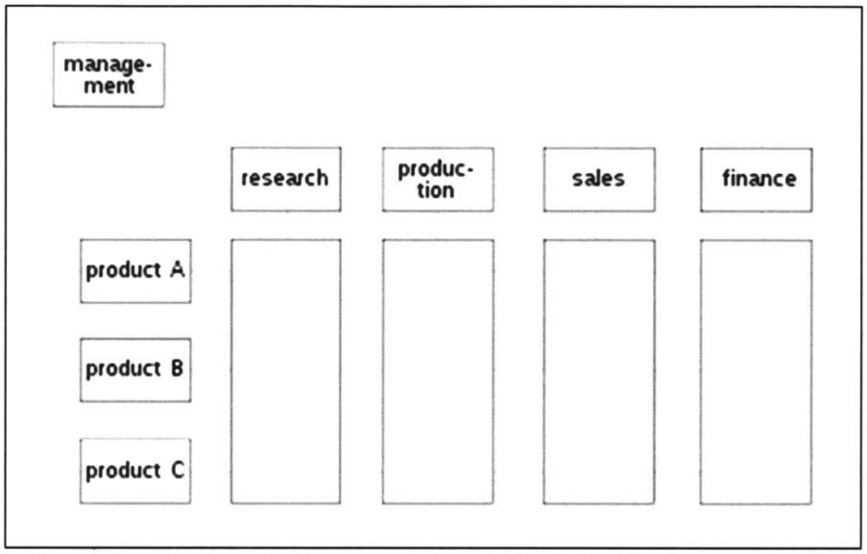

Figure 22: Example Matrix Organization Chart

CCA can be considered as a "micro business unit" with several project-based teams working on various products. Cross Business Teams (CBT) are the functional groups from which the product teams are formed. For example, an engineer working in the Patriot Defense System will report to both the Patriot and the engineering chains-of-command. The net effect is that most personnel resources are over budgeted trying to cover two sets of goals and deadlines. While most personnel are open to the idea of change and very willing to help, finding time and money for them to commit to a third objective is extremely difficult, if not impossible.

7.2.2 Independent Businesses and Business Goals

Raytheon's corporate vision and strategy focuses on existing and evolving customer needs in their 4 core markets to become defense and aerospace industry leaders as described in Figure 23 and Figure 24. To do so, Raytheon builds on a customer-centric mindset while creating and maintaining a competitive advantage. Since Raytheon is such a large company with focuses in many different areas, each business unit interprets and executes this strategy differently. Within CCA, this translates to a few focus items shown in Figure 25. The topic of this thesis falls under focus five (5) of Figure 25 in that the final results will be used to add functionality to the common tool set used by designers.

OUR VISION

To be the most admired defense and aerospace systems company through our world-class people, innovation and technology.

Figure 23: Raytheon's Corporate Vision

OUR STRATEGY

- Focus on key strategic pursuits, Technology, and Mission Assurance, to sustain and grow our position in our four core markets:
  - **Sensing:** Provide the breadth of sensing solutions required to meet our customers' mission needs.
  - **Effects:** Leverage kinetic energy-based expertise into EW, directed energy and cybersolutions.
  - **C3I:** Broaden market presence in communications, C2, networking and knowledge management.
  - **Mission Support:** Expand product support, engineering services and training.
- Leverage our domain knowledge in air, land, sea, space and cyber for all markets.
- Expand international business by building on our relationships and deep market expertise.
- Continue to be a Customer Focused company based on performance, relationships and solutions.
- Deliver innovative supply chain solutions to accelerate growth, create competitive advantage and bring valued, global solutions to our customers.

Figure 24: Raytheon's Corporate Strategy

1. Expand our Work Safe mindset
2. Exceed customer Quality expectations
3. Drive competitiveness through Affordability
4. Accelerate development of a flexible, enabled workforce aligned with the business
5. Implement a common test strategy/tool set to ensure customer success

Figure 25: CCA 2012 Business Focus

To maintain their competitive advantage and uphold Raytheon's vision, CCA products must be delivered with ever increasing quality and decreasing costs. Quality and cost are directly affected by the amount of troubleshooting and rework performed. These in turn are directly tied to the manufacturability of new circuit card designs, which can be predicted to a certain degree of accuracy by CCA. The nature of Raytheon's segregated businesses, however, demonstrates a weak link in manufacturing knowledge transfer. Often not involved until the end of the design process, the lessons learned from the manufacturing floor are not properly transmitted to designers until too late. Then, due to potential adverse effects to the originating group's bottom line, designs not optimized for manufacturability are passed, with CCA taking the brunt of troubleshooting and rework costs. To address this, CCA is currently moving towards manufacturing data based modeling at the early stages of design, with most of the work focused on data input and modeling improvement.

7.3 People and Culture

Organizations are institutions. An organization is a symbolic system of meanings, artifacts, values, and routines. Informal norms and traditions exert a strong influence on behavior. Action comes through habit. Raytheon carries a proud, distinct culture of engineering gumption and customer, the "warfighter," centric focus. These cultural identities are deeply ingrained and underlie many actions and decisions of the people at Raytheon.

7.3.1 Demographics

Culturally, Raytheon is a more traditional engineering company. A generation gap exists, with the older engineers having worked in their area of expertise for 20+ years, and the younger generation generally not staying in any one position for more than 3 years. The majority of engineers involved with this project fit into the former category. They are very focused on completing their assigned tasks and jobs, almost to the exclusion of more global issues. The existence of these work "silos" creates obstacles that must be overcome to work efficiently. While getting assistance from a particular person may not be difficult, finding the right person to ask can be. Most people do not know what the next group over does, so getting references into a new group can be hit or miss. One example is trying to get in contact with a group that maintains a certain database. While many people use, query, input, and rely on that database, not one person asked could provide direct contact to the group that maintains the database.

7.3.2 The Importance of Time

The employees of Raytheon are also quite proud and hard working. There is a true belief in what they are trying to accomplish, and that what they do makes a difference. Phrases like "It has to work, our warfighters depend on it" can be heard around the office. This belief translates into their work schedule and willingness to do what it takes to get the job done. Meetings routinely run well over time, filled with people determined to get the problem resolved before moving on. Raytheon employs a 9/80 schedule, where one can work 9 hours a day and take every other Friday off. Most salaried employees on the schedule, working 9+ hours a day, can still be found at the office on their Friday off, finishing up various tasks. This dedication to the warfighter and strong work ethic are a big part of Raytheon and its employees.

7.4 Political Arena

Organizations are contests. An organization is a social system encompassing diverse, and sometimes contradictory, interests and goals. Competition for resources is expected. Action comes through power.

Raytheon's complex structure directly influences its political landscape and the implications of that political landscape.

7.4.1 Implications of a Matrix Organization

In normal situations, the balance of powers and the underlying routes to getting a task with many stakeholders completed can be easily seen. The complexity of a matrix organization serves to hide these political aspects, or at the very least, makes navigating them more complicated. With multiple managers and multiple projects and deadlines keeping each employee very busy, capturing the attention of a needed employee for any appreciable amount of time can be challenging. This can be done in a few ways.

Money, connections, and coattail riding were all used to become temporarily important on someone's to do list. With the amount of work that needed to be done, most employees were hesitant to put time into non-value-added work or non-budgeted work. Having a charge code tied to a budget helped facilitate discussions by billing them as chargeable work. Utilizing the connections and networks of fellow team members generated an initial level of respect in contacting new people allowed for much quicker network building overall. Finally, fully utilizing the reputation built up and left over from the previous MIT intern opened many doors and paths. Within a complex matrix structure, successfully navigating the political arena depends on the power and reputation generated and maintained.

7.4.2 Manufacturing vs. Design

Like most typical design and manufacturing organizations, CCA continuously deals with the conflict between the design community and the manufacturing community. In quite simplistic terms, when poor results are recorded for any project or product, the design side blames the manufacturing side for being unable to manufacture products properly and cheaply. The manufacturing side blames the design side for creating poor, difficult-to-manufacture designs. Poor communication seems to be the underlying issue. With CCA generally being the last group to be included in the design process, designers do not get the knowledge and lessons learned from the manufacturing floor until very late, if at all. Chronic problematic components, difficult orientations, and testability issues do not get communicated, and then continue to persist. Due to program politics, budget considerations, and reputations, what little input that CCA does have tends to have very little effect on the final design. CCA's evaluation of the manufacturability and potential costs of certain design decisions often do not get introduced into many aspects that have already been "set in stone." Tight deadlines also restrict potential changes to the design from being implemented. This leads to the perception that CCA has little or no impact in the design process.

To strategically improve this issue, a new concurrent engineering initiative is being carried out. The initiative, being backed by many important entities within Raytheon, should foster communication between the design and manufacturing communities and provide dedicated lines to do so. Moving to a model based design process with direct links to manufacturing data will have a large impact on manufacturing affordability.

8 References

[1] Harper, Charles A. Handbook of materials and processes for electronics. New York: McGraw-Hill, 1970.

[2] http://www.pcmag.com/encyclopedia/term/52271/surface-mount

[3] IPC, Assembly Qualification Profile. IPC 7912A Standard, Northbrook, IL, January 2004.

[4] Soukup, Radek. "Optimization of printed circuit board design assembly using yield prediction." Electronics Technology, 2008. ISSE'08. 31st International Spring Seminar on. IEEE, 2008.

[5] Soukup, Radek. "Defect level prediction of printed circuit board assembly manufacturing based on DPMO metric." Electron Devices (CDE), 2011 Spanish Conference on. IEEE, 2011.

[6] Verma, Amit. "Management of DPMO Metrics Reduces the Cost of PCB Assembly." APEX (2003).

[7] Drake Jr, Paul J., et al. "Capability predictor." U.S. Patent No. 7,110,956. 19 Sep. 2006.

[8] Bergevin, Glenn. *A Predictive Troubleshooting Model for Early Engagement*. Diss. Massachusetts Institute of Technology, 2013.

[9] http://datawarehouse4u.info/Data-warehouse-schema-architecture-star-schema.html