U.S.CRC Journal

The journal of ABB's Corporate Research Centers in the United States

Fall 2005

**Welcome to the CRC Journal, a publication of ABB Corporate Research Center locations in the United States featuring news, pertinent articles and conference papers prepared by CRC scientists, technical specialists and others involved in CRC research activities. If you would like to be on the distribution list for this electronic document, please e-mail Robert Loeffler at robert.loeffler@us.abb.com**

# ABB U.S. CRC assists in judging IEEE Computer Society competition in Washington, D.C.

Ten teams from universities around the world competed in the IEEE Computer Society International Design Competition World Finals held in Washington, D.C. The theme for the event sponsored by ABB and Microsoft was "Going Beyond the Boundaries." Teams were challenged to create computer-based systems that perform a socially useful function.

The winning team was from North Carolina State University with its project titled "NEAT"– Networks for Endangered Animal Tracking – presented as an improved method to use devices and systems to track endangered wildlife. The goal is to use NEAT in efforts to repopulate areas of North Carolina with the endangered Red Wolf.

Brian Robinson, a Senior Research Engineer with the ABB Corporate Research Center located in Wickliffe, Ohio, was one of 10 industry experts selected to judge the finals. This is the second year that U.S.CRC has participated in the finals judging and the fifth year that ABB has sponsored this international competition.

Brian noted, "This was the best undergraduate contest I've been to in a long time. The students were very enthusiastic. "I feel the students were really excited to be developing products that can help people. Judging was very difficult due to the wonderful assortment of projects and the very promising applications that were presented.

"Engineering students worked in concert with an array of people from such diverse fields as health rehabilitation, and wildlife conservation as in the case of North Carolina State. It was an excellent exercise having the teams developing solutions for areas they would never have been exposed to."

Teams competed for cash awards for their universities and recognition of their academic achievement. This was the sixth year the event was held. Themes change from year to year making each contest different from the previous one.

# TABLE OF CONTENTS

# Large Scale Storm Outage Management

David Lubkeman, *Senior Member, IEEE*, and Danny E. Julian, *Senior Member, IEEE*

***Abstract*-- This paper describes a process for improving information used by electric utilities for managing restoration of distribution facilities damaged by large-scale storms such as large-scale ice storms and hurricanes. The process is realized in the form of a decision support tool that utilizes the model of the distribution circuit layout, the placement of protective and switching devices and the location of customers to determine how crew resources should be allocated and managed for cost-effective storm outage management. Equipment damage predictions based on the severity and location of the weather are also used to predict where and to what extent circuit damage is likely to occur. Given crew availability and the maintenance crew requirement for each damage type, a storm outage algorithm predicts damage per feeder, area and region and calculates estimated time to restoration for each customer. After the storm hits, additional information such as customer calls, switch status and field damage assessments are utilized by the storm outage algorithm to provide updated predictions of damage and estimated time to restore.**

***Index Terms* — Decision support systems, Ice, Power distribution maintenance, Power distribution meteorological factors, Power distribution reliability, Power system restoration, Resource management, Storms**

## I. Introduction

SEVERE weather conditions such as hurricanes, ice storms, lightning storms, etc. have the potential to wreak havoc with electric utility transmission and distribution systems. High winds and/or icing can knock trees into overhead lines, cause poles to snap, and dislodge overhead lines from crossarms. This type of damage will disrupt electrical service until the physical facilities can be replaced or repaired. Large storms can result in a massive number of customer outages, sometimes taking from days to weeks to repair. In the meantime, customers can be left without electric service during difficult conditions such as extreme cold or extreme heat.

To restore service after a major storm, utilities send a large number of maintenance crews into the field. Often crews are borrowed from neighboring utilities during these emergency conditions. To effectively plan for and respond to major storm-caused disturbances is a challenging task for utility management. Before the storm hits, the utility needs to perform an assessment of what amount of damage will occur and where. Then crews are dispatched to staging areas to be in position to make repairs once the storm hits. Once the storm impacts the service areas, then the utility gets data from a number of sources corresponding to distribution circuit damage. This information includes calls from outaged customers, data points on feeder protection switching operations, and damage assessment reports from survey teams in the field. This large amount of information must be efficiently used to get a picture of what portions of the circuit have been damaged and then used to dispatch maintenance crews.

Utilities typically have what is called an Outage Management System (OMS) to help log customer calls and dispatch crews to the site of the disturbance. However, these applications utilize engines that normally assume that a given set of localized customer outages is associated with a single damage location. Most of the current outage management systems are not optimized to handle the large amount of facility damage associated with major storm events.

This paper describes enhancements that can be added to Outage Management Systems to more effectively handle large-scale storm conditions. Following this section, a summary of the processes already utilized by many utilities to deal with the massive damage caused by severe storms is presented. Section III describes the models, algorithms and information systems for computing the various indices needed by planners and operators for handling storm conditions. Finally, an example shows the types of calculations that can be made on a sample feeder along with conclusions.

## II. Storm Outage Management Background

### A. Overview

Distribution outage situations can be classified regarding the scope of the damage and the number of customers impacted. At the lowest level are the normal day-to-day outages due to minor storms, animal contact, broken tree limbs, etc., that are handled locally through conventional outage management system processes. At the next higher level are outages due to localized storms impacting a small area and causing damage to poles or primary equipment within an operating area. This would necessitate the need for more line crews than normally scheduled. For severe storms impacting a wider area, but still within the same operating area, multiple feeders having mechanical damage could necessitate resources from outside the operating area being called in, although still using available regional resources. At the next higher outage level for storms impacting multiple operating areas, more than 10% of customers could be without power, and in these situations internal resources from other utility operating regions need to be called in. At the highest impact level are severe storms, where centralized storm management is needed to coordinate

both internal crews and extra resources called in from other utilities or contracting agencies.

Distribution outage management systems on the market share much of the same functionality. Inputs to the system include customer and non-customer trouble calls, reports from line patrols, system status transferred from SCADA devices, 911 emergency agencies and possibly fault locations computed from feeder currents. Based on inputs related to a disturbance, a network model is used to infer possible locations/sources of the disturbance event. Note this calculation normally assumes that there is only one cause for the customer outage. An available crew is then dispatched or scheduled to respond to each outage. Estimated time to restore (ETR) is computed based on a number of factors such as time of year, time of day, device type, operations area and the use of an outage to crew ratio multiplier. This information can then be accessed by or communicated to the effected customers. Using the outage management system, the operator can track outages, customers effected, locations of crews and coordinate activity related to the restoration process. The outage information and restoration times are also used to compute network reliability indices.

Storm Outage Management differs from conventional Distribution Outage Management in that large-scale storms cause massive amounts of physical damage, requiring a large number of repair crews for restoring customer service. Major storms usually impact a large part of a utility's service territory and timely restoration involves borrowing crews from nearby utilities and hiring contractors. Conventional Outage Management Systems typically do not operate in a proactive mode as far as predicting damage ahead of time and supporting the staging of a large number of resources. Also conventional outage management systems are not designed to process the large-scale destructive damage that occurs. As a result, utilities have difficulty utilizing conventional outage management systems during and after large storm events and often utilize their own customized tools instead.

Storm outage management is composed of four steps as shown in Figure 1:


**Figure 1: Storm Outage Management process**

The first step, Damage Prediction, involves utilizing a weather forecast and converting this information into a prediction of the damage that will likely occur in the distribution system. The outcome of this stage is a forecast of the number of crews that will be needed in order to restore customer service in a reasonable period of time for the type of storm. The Crew Staging process involves calling in extra crews if needed and locating available crews at various locations in anticipation of where the storm damage will likely occur. Staging is a difficult logistically effort since it involves making arrangements for housing and feeding a large workforce. Once the storm has hit, the Damage Assessment process involves sending teams out to survey the damage and file reports. The damage information is used to dispatch crews and to compute an estimate for customers as to how long they will be without power. The final stage, although the one that lasts the longest, is the Restoration Management process, in which crews are dispatched to the damaged areas to restore service.

### *B. Damage Prediction*

The goal of the Damage Prediction phase is to forecast the amount of damage a storm will produce, the resources required for restoration and the resulting time needed to restore service to customers. This is a rough estimate, but it allows the utility to get resources into place or on stand-by, speeding up the restoration process after the storm hits.

Storm Damage Prediction is based on an accurate weather forecast of variables related to distribution circuit damage. For ice storms, this could be inches of ice accumulation on trees surrounding overhead lines or on overhead equipment itself. For hurricanes, this could be peak wind speeds and gust durations. Based on historical information and susceptibility models, expected damage can be predicted with crew requirements being directly computed from estimated damage. This would require knowing of the amount of repair work a crew could achieve on a daily basis.

The functionality of a storm damage prediction module would be given a weather forecast of high wind speeds, lightning, icing, etc., it would predict the location and extent of damage to electric utility circuits. Modeling would be based on past major disturbances and data shared from other utilities. Survey or historical data would also be used to assign a sensitivity code to feeder sections for computing damage due to severe weather conditions.

The outputs from the storm damage prediction for a substation or feeder could include:

1. Expected tree damage per unit length of line
2. Poles broken
3. Primary spans down
4. Secondary spans down
5. Customer services down
6. Transformers damaged

To estimate crew requirements per substation during the planning process, one could take several approaches. The first approach could be to use historical data on system restoration to directly compute crew requirements based on the severity of the weather conditions. A second approach could be to estimate damage, and then apply crew requirements to fix a given unit of damage to estimate total crew requirements.

### *C. Crew Staging*

Crew Staging involves determining the number of crews needed for restoring customer service in a reasonable amount of time, and if a gap exists, to make arrangements for external crews to be called in. Crew Staging also includes determining where crews should be staged and where they should be

lodged and fed. Typically, crews would be staged on the outskirts of the projected problem areas ahead of time.

A crew (and material) staging function would take the damage prediction results and make initial assignments for internal crews, identify need and locations for external crews and identify need and locations for materials. The approach would be to use historical data to identify the total number of resources needed to repair predicted damage per unit time. An ETR target would be used to determine total resources needed. Inputs to the staging process would be the ETR target, damage prediction, crew task statistics, availability of internal crews and sources for external crews. The output to this function would include the types of resources needed, amount of material required, and locations where resources should be deployed.

One option that could be considered for a crew-staging tool is an interface for outside contractors. However, there are firewall and security issues in giving non-utility personnel access to a secure internal network.

### *D. Damage Assessment*

Storm Damage assessment for large scale events involves specially trained teams that ride the circuit to record the types and location of damage on the circuit. Several hundred assessment teams could be involved in a more detailed survey that might take several days to complete. The collected information is then forwarded to the various operations center offices.

The damage assessment serves three functions. First, it is used to guide the appropriate maintenance crews (tree, 2-man, 4-man) to locations where repair work is required. Second, it is used to compute ETRs that can be made available to outaged customers (through systems such as IVR). A third use for the damage assessment information is to update and verify outages in the OMS.

The damage assessment primarily is based on field damage assessments, but could also include trouble calls and SCADA inputs as well.

### *E. Restoration Management*

The restoration process occurs immediately after the storm passes through the utility's service territory. Repair crews are initially dispatched to areas based on storm damage and are assigned to work specific substations and feeders. Crews can also be reallocated as deemed necessary by the Storm Outage Coordinator. It is a rather difficult task to track crew progress in real time, since there are many crews operating at the same time, with many of these crews coming in externally from other utilities or contracting firms.

Note that the damage to unpatrolled sections can be simply estimated by applying a multiplier to damage observed already on the patrolled sections. Also, the model on which these calculations are based could be the same or similar to that used in the Damage Prediction phase to compute the number of crews required. Progress on the restoration process is also accumulated locally and then passed on to the next higher level of management.

## III. STORM OUTAGE MODELING AND IT TOOLS

### *A. Interaction with other IT systems*

An IT system for supporting the storm outage management process can be built around a storm outage application that interacts with the users and combines data from other IT systems and data sources as shown in Figure 2. The supporting IT systems include:

1. Distribution Circuit Configuration Database, which includes the layout of the distribution facilities, location of protection devices, placement and type of sectionalizing switches, and the customer connectivity.
2. Asset Information Database, which contains data about the individual components in the system such as pole types, overhead line construction, construction age, susceptibility of these components to severe weather conditions.
3. Field Observations and Measurements Interface, that includes customer outage calls, data related to protection device operation, and reports from damage survey crews.
4. Damage Prediction Module, that predicts for a given set of weather conditions and asset characteristics, what will be the probable damage in terms of poles broken, downed conductors, damaged transformers on a per mile basis.
5. Crew Requirements Module, that provides the man-hours needed for each type of repair.

Another key IT system driving this process is storm intensity information. This is assumed to come from a weather service and provides data on the severity of the storm such as expected inches of icing, wind speeds and durations, and accumulated rainfall. Both forecasts and actual observed weather quantities are applied. Historical storm data is used to calibrate the models utilized for the Damage Prediction Module and the Crew Staging Module.

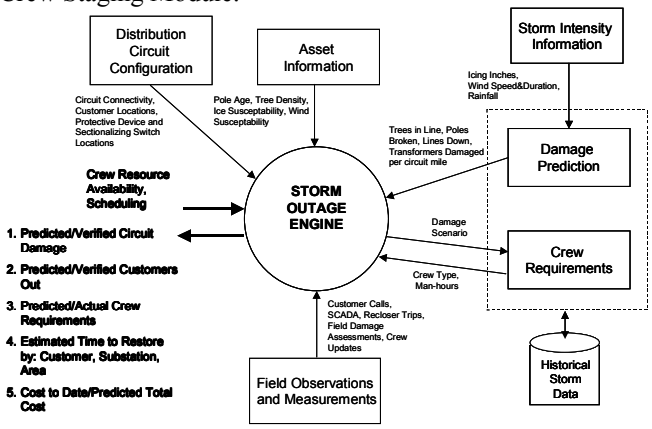


**Figure 2: Interaction of Storm Outage Engine with Other Data Sources**

Based on the IT systems and modules defined above, the storm outage application performs a number of calculations related

to storm management. Given crew resource availability, crew cost, and crew scheduling constraints, the application calculates the following quantities:

1. Predictions of where damage will likely to have occurred and to what extent on the distribution system,
2. Predictions of where customers will be without service,
3. Crew requirements (service by crew type, crew man-hours) for making the needed repairs,
4. Estimated time to restore customer and specified area, and
5. Total accumulated cost of outage and predicted cost for total restoration.

*B. Storm Outage Susceptibility Indices*

The four storm types typically encountered and key features include:

1. Hurricane: Maximum wind speed and duration
2. Ice Storm: Accumulated Icing
3. Snow Storm: Inches of Snow, Temperature, Wind Speed
4. Lightning: Location of strokes and Intensity

Other input information needed pertaining to the distribution circuit would be areas impacted, asset characteristics, tree type and density. The asset characteristics would be in the form of equipment susceptibility indices, somewhat similar to the "failure rates" used for distribution circuit reliability analysis. Unlike reliability analysis though, one would have to account for multiple failures occurring simultaneously. There would be different susceptibility indices depending on weather type such as hurricane high wind, tornado, icing, lightning, etc. The susceptibility indices should be scalable so that a piece of equipment that has an index of 3 times that of another similar piece of equipment would be 3 times more likely to fail during a given storm event.

*C. Storm Outage Damage Modeling*

By examining various crew availability scenarios before and after the storm, the storm outage application serves as a decision support tool for optimally managing crew resources with respect to total outage costs.

The underlying methodologies behind the storm outage engine are described using Figure 3 through Figure 6. Figure 3 shows an example circuit consisting of two distribution feeders fed from two different substations. Protective devices are indicated by the circles that include station breakers (B), feeder reclosers (R) and tap fuses (F); Customer locations are indicated by the triangles. Note that fuses isolate lateral circuits off the main primary feeder during faults on the laterals and switches can be opened or closed to either isolate faulted sections of the circuit from healthy sections or switch load to adjacent substations.

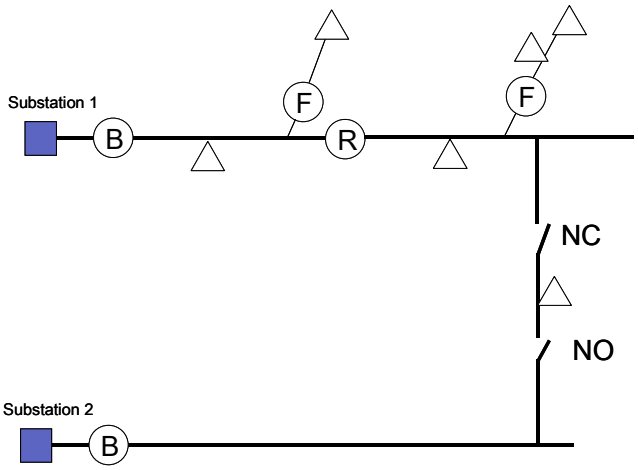


**Figure 3: Power Distribution System Layout**

For pre-storm analysis, the storm outage application is used to predict where in the circuit the damage is likely to occur, the manpower required to repair the damage and estimated times to restore for outaged customers. Based on weather forecast and damage statistics computed by the Damage Prediction Module, components in the circuit are assigned tentative failure rates. These failure rates will be in terms of trees down per mile, poles down per mile, downed conductors per mile, transformers damaged per mile, etc. These failure rates will also be a function of a susceptibility multiplier determined by using information from the Asset Information Database such as pole age, wire size, overhead construction, tree density, and other asset properties. Figure 4 shows the results obtained by the storm outage engine in which the sections having the highest susceptibility to failure are shown as dotted lines, medium level of susceptibility as dashed lines and the sections with the lowest susceptibility are solid. Potential locations and incidences of various types of circuit damage are indicated by the crosses (x).

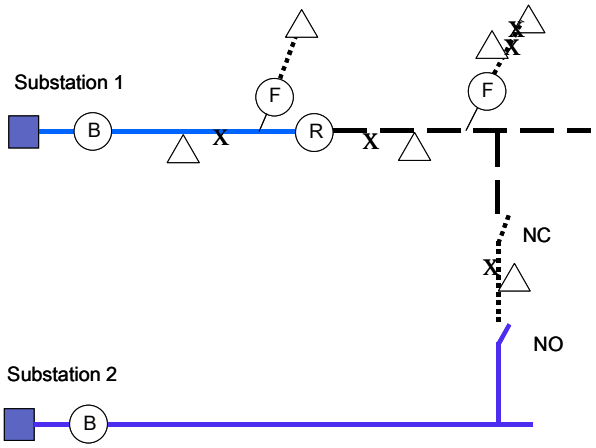


**Figure 4: Pre-storm Damage Prediction**

A number of quantities are calculated based on this prediction. First, the total predicted damage for each feeder could be summed up to provide total expected damage by feeder. Using statistics from the Crew Staging Module, estimates can also be made regarding the required crew types and resources to repair damage on a per feeder basis. Given the number of crews allocated per feeder, estimated time to restore power to each

customer is calculated. This is based on predefined guidelines for how restoration is to take place. This could involve a business restoration process assuming the main primary feeder will be repaired first, that sectionalizers will be employed, priority loads will take preference, and that laterals with small numbers of customers will be fixed last. Total cost estimates for making the repairs based on the crew allocation can also be made. Once the above quantities have been calculated on a per feeder basis, then they can be summed up by substation, by area, by region and by entire service territory.

Once the storm has hit, various sources of information can be integrated together to update the prediction and convert predicted damage estimates to verified damage. Similarly, predicted customer outages can also be converted to verified customer outages. This is illustrated in Figure 5, where sections shown as single solid lines are deemed non-damaged since customers on those sections have not reported an outage. As field survey assessments become available (sections with lines highlighted), then predicted damage is updated to verified damage. Damage to unpatrolled sections of the circuit is still based on the pre-storm predictions. However, as damage reports come in, the damage predictions are scaled up or down, depending on whether the damage estimates to the patrolled sections of the system are higher or lower than anticipated. As the damage assessment improves, other estimates made during the pre-storm stage are updated as well. This includes updates regarding circuit damage, manpower requirements for repair, remaining restoration costs and customer estimated time to restore.

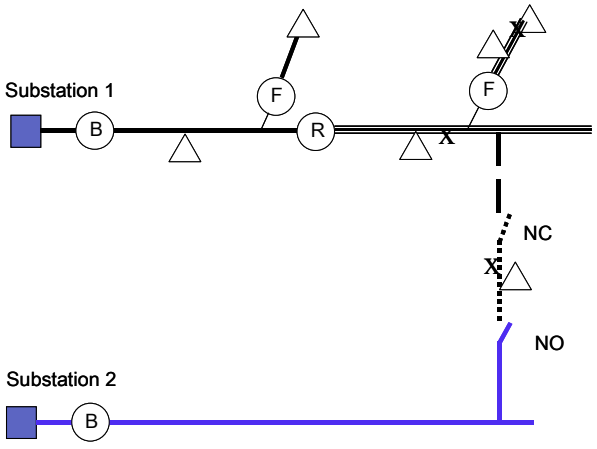


**Figure 5: Storm Outage Restoration with Damage Not Completely Verified**

After all field surveys have been completed, then the storm outage application has a completely verified model of circuit damage. Figure 6 shows that circuit sections are either deemed operational (solid lines) or have verified damage (as shown by highlighted lines). At this stage, the circuit model is still used to allocate crew resources, calculate costs and provide estimated time to restore by customer.

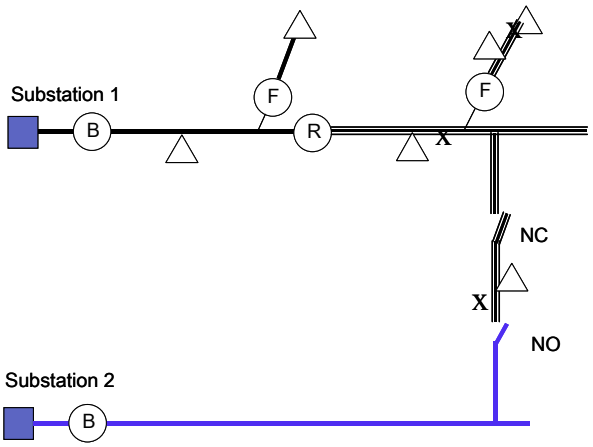


**Figure 6: Outage Restoration after Damage Verified**

## IV. OUTAGE INDICES CALCULATION EXAMPLE

As an example of the types of calculations that could be performed by the storm outage application described in Section III, consider the single feeder illustrated earlier with the customers distributed as shown in Figure 7. Note that customer locations have been greatly simplified since an actual circuit would be difficult to view in the context of this paper.

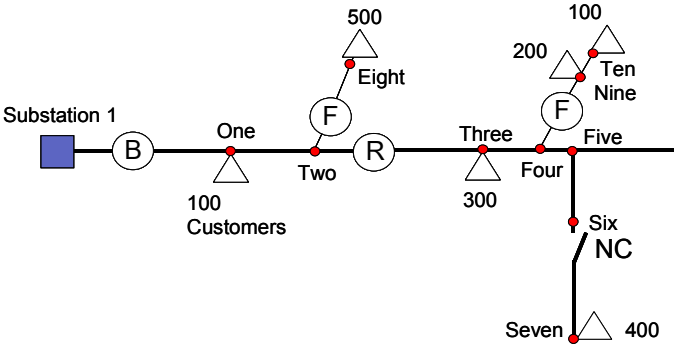


**Figure 7: Example System showing Customer Locations**

Asset susceptibility indices can be used to predict circuit damage and ETR for both pre-storm planning and post-storm analysis with preliminary damage assessment data. Sections of the circuit will have different levels of susceptibility for the various storm types and intensities. In the example shown in Figure 8, sections with low susceptibility are show by solid lines, medium susceptibility as dashed lines, and high susceptibility as dotted lines. In this case the dashed section is assumed to be three times more susceptible than the solid section while the dotted sections are assumed to be ten times more susceptible than the solid sections.

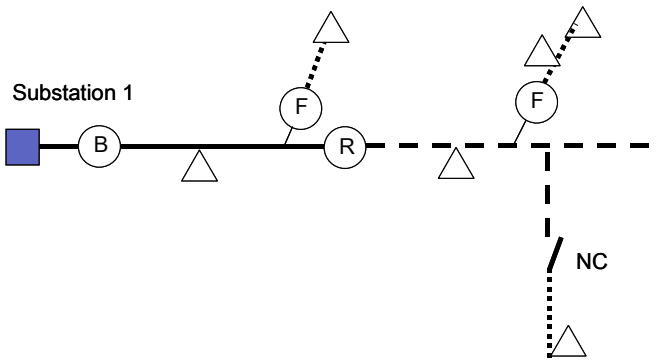


**Figure 8: Circuit Section Susceptibility**

Before the storm hits, the expected distribution of the damage can be used to roughly estimate the number of crew-days needed to restore service and how many crews would be needed per feeder to get all customers restored in a target time frame. For example on the circuit shown in Figure 7, the following type of data could be generated as a planning estimate based on assigning two line and tree crews:

| | |
|---|---|
| Total Customers Out | 1600 |
| Expected Line Crew Manpower | 7.7 days |
| Expected Tree Crew Manpower | 6.2 days |
| Estimated Restoration Cost | $50,938 |
| Estimate Time to Restore (ETR) | 3.86 days |

**Table 1: Initial Damage Estimate Using Asset susceptibility Indices**

Other types of damage could also be included, but only span and tree damage are shown to simplify the example. Obviously, the ETR would be higher or lower depending on the crew resources allocated.

Once the storm hits, then customer calls are logged, and outaged areas and protective device operations are identified. As initial field assessments are completed, then damage can be estimated in unsurveyed sections based on the susceptibility factors highlighted in Figure 8. This is illustrated in Figure 9, where surveyed areas are shown circled, and damage is indicated by an (x). Note that damage shown in unsurveyed areas is predicted.

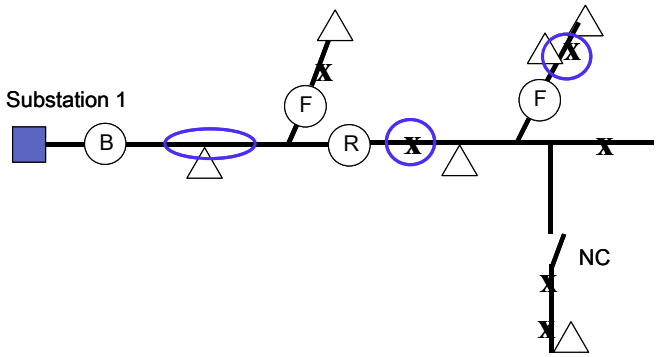


**Figure 9: Initial Damage Assessment**

For the damage assessment shown in Figure 9, the sample report in Table 2 could now be generated:

| | |
|---|---|
| Total Customers Out | 1500 |
| Percent of Customers Out | 94 |
| Percent of System Assessed | 25% |
| Verified Spans Down | 21 |
| Verified Trees Down | 37 |
| Line Crew Days Remaining | 8.3 |
| Tree Crew Days Remaining | 6.0 |
| Line Crews Assigned | 2 |
| Tree Crews Assigned | 2 |
| Estimate Time to Restore (ETR) | 4.16 days |

**Table 2: Updated Damage Estimate with Asset Susceptibility Indices and 25% Damage Assessment Complete**

After the damage assessment is completed, all ETRs are based on verified damage. At this stage, individual customer times to restore can now be estimated with much more accuracy. The estimates are based on crew productivity indices and restoration priority guidelines that were established before the storm. Figure 10 illustrates restoration status, where the planned restoration sequence is shown in the circles, the customer ETRs are shown in the boxes, and remaining damage is denoted by an (x):

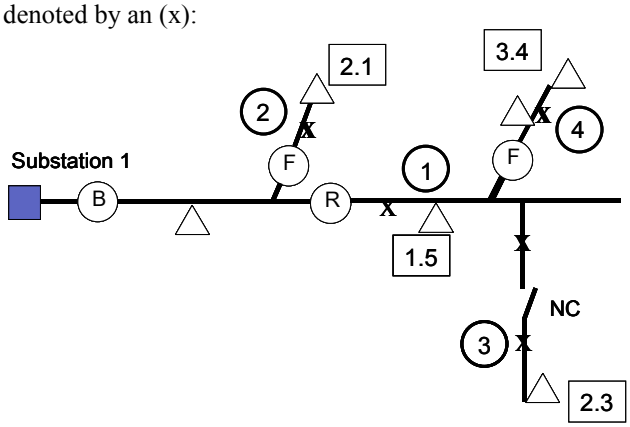


**Figure 10: Initial Restoration Process**

The report in Table 3 shows the type of information that could now be produced:

| | |
|---|---|
| Total Customers Out | 1500 |
| Percent of Customers Out | 94 |
| Percent of System Assessed | 100 |
| Verified Spans Down | 69 |
| Verified Trees Down | 125 |
| Repaired Spans Down | 17 |
| Repaired Trees Down | 33 |
| Estimated Restoration Cost | $55,640 |
| Line Crew Days Remaining | 6.9 |
| Tree Crew Days Remaining | 5.0 |
| Line Crews Assigned | 2 |
| Tree Crews Assigned | 2 |
| Estimate Time to Restore (ETR) | 3.45 days |
| ETR at Transformer Three | 1.50 days |
| ETR at Transformer Seven | 2.30 days |
| ETR at Transformer Eight | 2.10 days |
| ETR at Transformer Nine | 3.45 days |
| ETR at Transformer Ten | 3.45 days |

**Table 3: Updated Damage Estimate 100% Damage Assessment Complete**

As the damage is repaired, then ETR calculations are updated. These ETRs can be made available to customers through the outage management voice response system. A sample schematic showing the ETR status at a later time is shown in Figure 11:

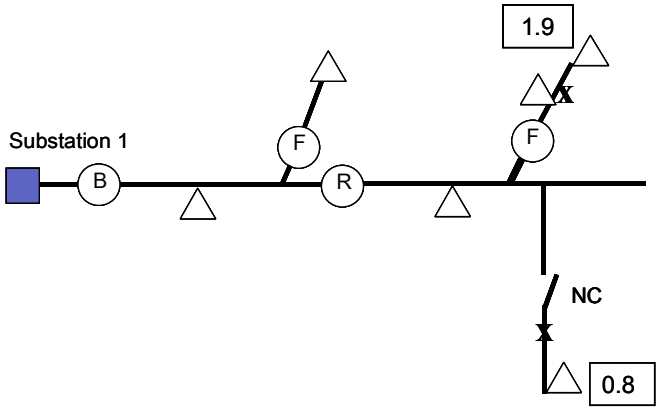


**Figure 11: Tracking of Restoration Process**

## V. Conclusions

This paper describes the information technology tools needed to make decisions regarding the use of crew maintenance resources after a major storm event. The modeling incorporated into the storm outage application makes it possible to analyze the relationship between crew allocations, crew costs, predicted and verified damage, and estimated time to restoration. In the storm-planning phase, this allows managers to determine the number of crews needed, the number of crews to be requested from neighboring utilities, and where these crews should be sent. After the storm hits, the storm outage application makes it possible to dispatch crews before a total damage assessment has been completed, effectively using existing information to predict where damage has occurred. During the restoration process, managers can also use the storm outage engine to look at the trade-offs between adding more resources and the difference in customer restoration times, to most cost effectively use resources. Since massive storm restoration costs large utilities on the order of tens of millions of dollars, being able to dispatch crews more effectively can result in large cost savings.

## VI. Acknowledgements

The authors would like to acknowledge Larry Nunnery of Progress Energy for sharing his experience regarding storm outage management.

## VII. Biographies

**David Lubkeman** is a Senior Scientist with ABB Corporate Research. He received his B.S., M.S., and Ph.D. degrees in Electrical Engineering from Purdue University in 1979, 1980, and 1983 respectively. Prior to that, he was an Associate Professor at Clemson University and an Assistant Professor at North Carolina State University. Dr. Lubkeman is a Senior Member of the IEEE Power Engineering.. His expertise is in power distribution system analysis and automation.

**Danny E. Julian** (S'1988, M'1992, SM'99) is currently working with ABB Corporate Research. In Raleigh, NC. Prior to working at ABB, worked for GE Harris Energy Control Systems. He received his BSEE and MSEE degrees from Clemson University in 1991 and 1992 respectively and is a registered Professional Engineer in the States of Florida and North Carolina with areas of interest including computer applications in power systems, power system planning, operations and control and economics and optimization of power systems. He is also a member of the Eta Kappa Nu and Phi Eta Sigma honor societies, a Senior Member of IEEE and a member of CIGRE.

# Design and Implementation of a PVC Reaction Optimizer

Bruce W. Benjamin, WaiBiu Cheng, Xiangming Hua, Ramasamy Selvaraj

## Introduction

Polyvinyl chloride (PVC) is one of the oldest polymers and the second largest thermoplastic in terms of volume manufactured in the world. This widespread use arises from a high degree of chemical resistance and a truly unique ability to be mixed with additives to give a large number of reproducible PVC compounds with a wide range of physical, chemical, and biological properties. Most PVC resins are produced by batchwise aqueous suspension polymerization. This is an extremely complex process and it is difficult to achieve an optimum operation. Therefore, a rigorous and accurate model representation of this process, focused on optimizing its operation, should provide a basis for significant economic advantages over current operations and control techniques.

An optimizer for suspension PVC reactors has been developed on the *ABB Optimize$^{IT}$ Dynamic Solutions* platform, an advanced general-purpose process modeling and simulation software framework. Dynamic Solutions contains a simulation environment and built-in optimization and estimation capabilities. This allows the user to conduct parameter estimations and optimizations, in addition to dynamic simulation applications, on a single platform. gPROMS (general Process Modelling System from Process Systems Enterprise Ltd.) is the underlying model representation language and solution engine.

## Process Description

Most of the current batch suspension polymerization reactors for PVC production are water-jacketed and lined with glass or stainless steel to minimize polymer buildup on the walls. A Dynamic Solutions flowsheet of a typical batch suspension PVC reactor is shown in Figure 1.

Vinyl chloride monomer (VCM), including recovered vinyl chloride (RVCM), is first finely dispersed in water by vigorous agitation. A small amount of suspension agent(s), typically partially saponified polyvinyl alcohol (PVA), is added to minimize coalescence of the growing polymer grains through the formation of a protective coating. Polymerization is induced by the addition of oil- or monomer-soluble initiators. Peroxides have typically been used for initiation. The reaction takes place in the organic droplets. The reactor's contents are heated to the required temperature where the initiator(s) begin to decompose and polymerization commences. The heat of polymerization is transferred from the monomer droplets to the bulk water phase and then to the reactor wall, which is cooled by water flowing through the reactor's jacket.

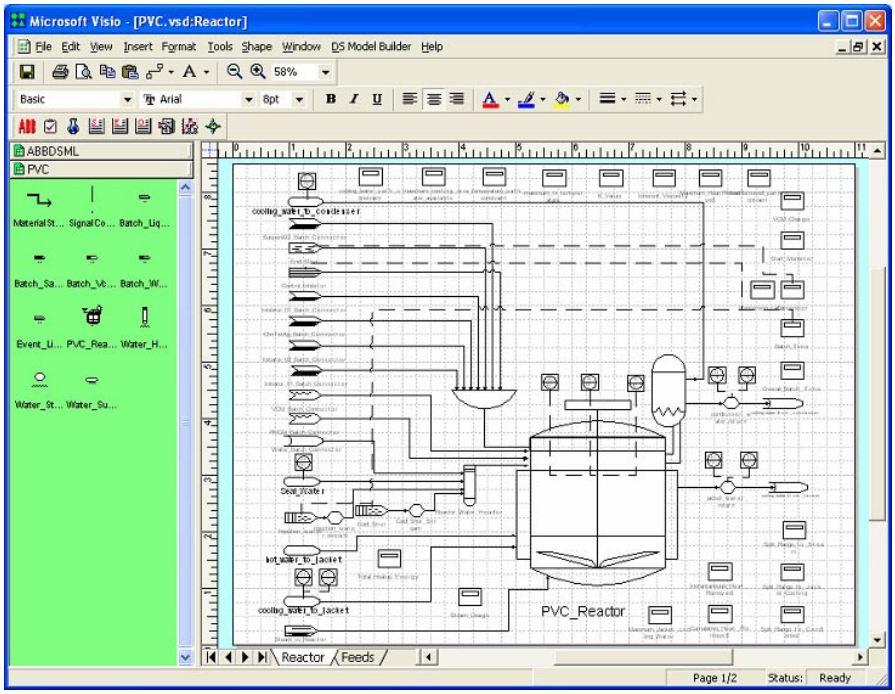


Figure 1. Dynamic Solutions PVC Reactor Flowsheet

Depending on the reactor size and design, a very efficient cooling jacket may provide the means for all the heat removal, often eliminating the need for a reflux condenser or chilled water. If the reactor includes a reflux condenser, it is typically provided as an upper extension to the reactor.

Since PVC is mostly insoluble in its own monomer, once the polymer chains are first generated, they precipitate immediately to form a separate phase in the polymerization droplet (a polymer-rich phase with entrapped monomer). Reactions continue in both the free liquid monomer-rich phase and the entrapped monomer in the polymer-rich (gel) phase dispersed about the formed polymer. When polymerization is complete, the polymer is in the form of colloidal, spherical particles dispersed in water. If the polymerization conditions are properly controled through the course of the batch, a polymer product having an extremely narrow particle-size distribution can be obtained.

The reaction temperature is used for molecular weight control. Sometimes, a chain transfer agent may be added to control molecular weight in the free radical polymerization. A polymerization inhibitor (e.g. hydroquinone) may also be used in this system to terminate the reaction. Typical polymerization times can vary between 3.5 to 6 hours, depending on the molecular weight of the polymer resin being prepared, as well as the heat-removal capacity of the reactor system.

When the free liquid monomer has been consumed, the pressure in the reactor begins to fall as a result of increased monomer mass transfer from the vapor phase to the polymer-rich phase due to sub-saturation conditions. In industrial PVC production, the reaction is usually stopped when a certain pressure drop has been reached. At this point, the conversion is typically in the range of 84 to 88%.

After completion of the batch, the mixture (polymer slurry) is transferred to a blow-down vessel where unreacted vinyl chloride is recovered. The PVC slurry is then stripped, dried, and stored.

## Model Description

The major reactions that are included in the optimization model are given below:

Initiation:

$$I(i) \rightarrow 2\ I^{*}(i)$$
$$I^{*}(i) + M \rightarrow R^{*}(1) + I^{inactive}(i) \qquad i=1,..,N_d$$

Propagation:

$$R^{*}(x) + M \rightarrow R^{*}(x+1)$$

Chain transfer to monomer:

$$R^{*}(x) + M \rightarrow P(x) + R^{*}(1)$$

Chain transfer to polymer:

$$R^{*}(x) + P(y) \rightarrow P(x) + R^{*}(y)$$

Termination by disproportionation:

$$R^{*}(x) + R^{*}(y) \rightarrow P(x) + P(y)$$

Termination by combination:

$$R^{*}(x) + R^{*}(y) \rightarrow P(x+y)$$

Inhibition:

$$R^{*}(x) + Z(i) \rightarrow P(x) + Z^{*}(i)$$
$$Z^{*}(i) + Z^{*}(i) \rightarrow 2\ Z\ (i) \qquad i=1,..,N_z$$

where:

- I : unreacted initiator
- $I^{*}$ : reacted initiator, a free radical
- $I^{inactive}$ : inactive free radical initiator
- M : monomer
- $N_d$ : number of initiator species
- $N_z$ : number of inhibitor species
- P(x) : polymer chain of length x
- $R^{*}(x)$ : free radical polymer chain of length x
- Z : unreacted inhibitor
- $Z^{*}$ : free radical inhibitor

The basic kinetic relationships for the above reactions, with one exception, are set up as first order with respect to reactants and scaled with an Arrhenius expression. The one exception is the

reaction of initiator with monomer which is modeled with the quasi-steady-state approximation. All of the above reactions, except for the chain transfer to polymer, are modeled as occurring in both the monomer and the gel phases. Chain transfer to polymer occurs only in the gel phase. Additional reaction rate scaling, representing the Trommsdorff, or gel, effect, is included in the gel phase reaction kinetics.

A total of four phases are modeled – monomer rich phase, polymer rich (gel) phase, water phase, and vapor phase. The assumptions are as follow:

- All components are assumed to be at equilibrium across all phases in which they are present
    - Dead polymer chains exist only in the polymer rich phase
    - Distribution of monomer across the phases follow the assumptions of Xie [2]:
        - A modified ideal gas equation describes the distribution of the vinyl chloride monomer in the vapor phase
        - Henry's law describes the solubility of the monomer in the water phase
        - Flory-Huggins equation describes the solubility of the monomer in the polymer-rich phase
    - Water exists only in the vapor phase and water phase. Distribution is based on the water vapor pressure.
    - Distribution of all other components including radical polymer chains follow these rules:
        - Concentrations in monomer-rich and polymer-rich phases are equal
        - Concentration in water is zero
        - Concentration in the vapor space is dependent upon the component vapor pressures
        - Vapor pressure of radical polymer chains is zero
- Reactor pressure is balanced in all phases
- The method of moments approach [1] is adopted to determine the average molecular weights of the radical and dead polymer chains
- Provisions are made to include impurities, according to the work of Penlidis [3] and of Zegel'man [4], and chain transfer agents in the model
- No air is assumed to be present. Reactions involving oxygen are not modeled

## Parameter Estimation

The tuning of the PVC reactor model to represent real process behavior is crucial for the optimizer to implement a successful optimization of the reactor performance. Parameters such as reaction kinetics and heat transfer coefficients may need to be tuned depending on changes in the plant operating conditions such as the use of different raw materials, specific charging recipes, and product grade requirements. Moreover, for a specific plant, even if the model has been well tuned based on historical plant data, it may be necessary to re-tune the model from batch to batch, due to unknown process disturbances and the time-varying behavior of the process. For example,

the heat transfer coefficients of the cooling jacket and the reflux condenser may tend to drift from batch to batch because of the build-up of deposits on the reactor walls.

In practice, it is very important to ensure the validity of the data used for model tuning. Proper audit of the instrumentation, and data collection systems should be conducted prior to collecting data for model tuning and validation.

In Dynamic Solutions, parameter estimation is internally constructed as an optimization problem. A maximum likelihood objective function is used to perform parameter estimation. In other words, the objective of the parameter estimation problem is to maximize the probability that the model will predict the measured values as closely as feasible. An on-line test should be executed early in the batch to confirm the validity of the parameters obtained from the previous batch.

In the estimation, the selected model parameters are the reaction rate coefficients and the heat transfer coefficients for heat transfer through the reactor wall. Other parameters such as vinyl chloride feed impurities can also be included if sufficient data are available to support such estimation. Sample plots of estimation results are given in Figures 2, 3, and 4.

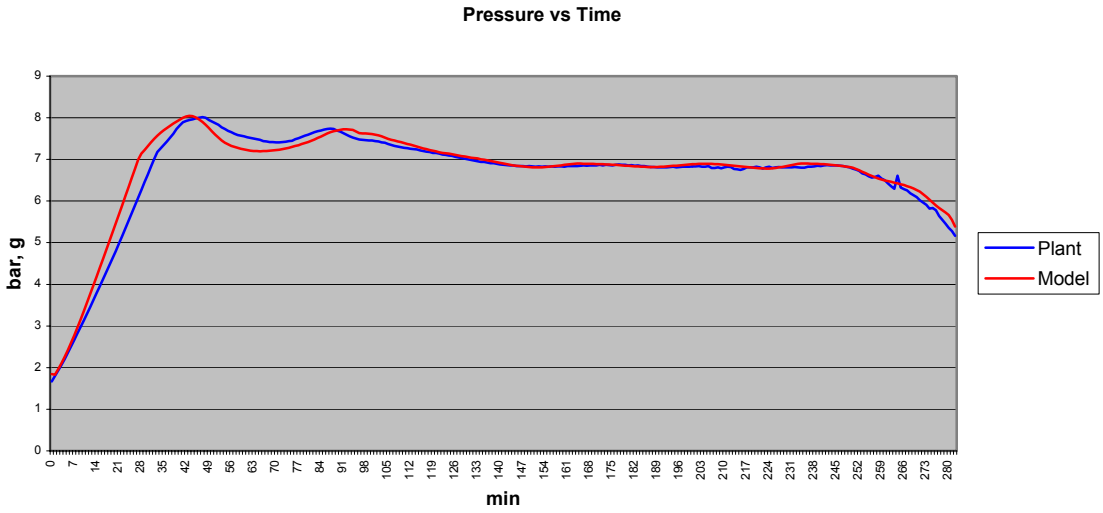


Figure 2: PVC Reactor Pressure Profile Results from Parameter Estimation Run

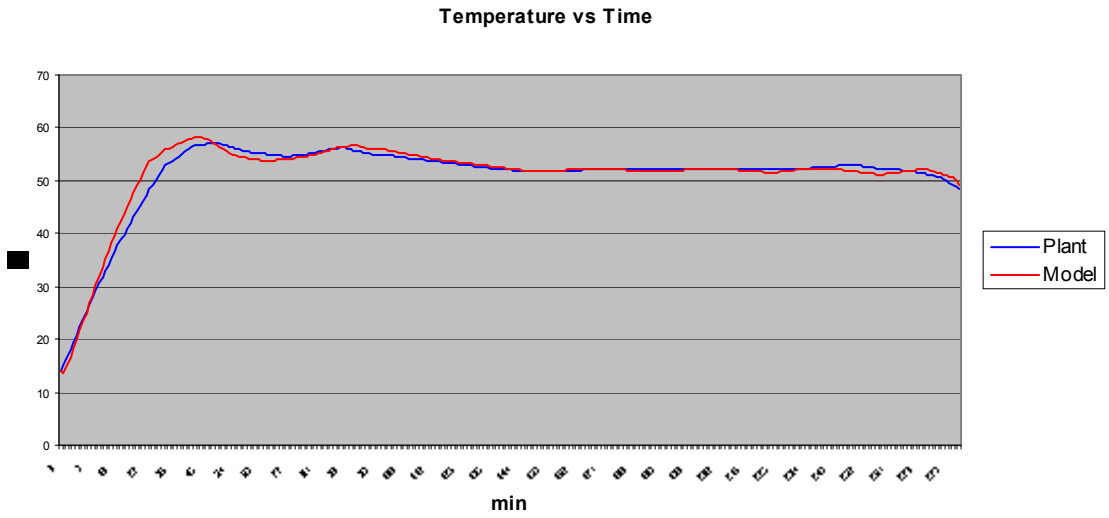


Figure 3: PVC Reactor Temperature Profile Results from Parameter Estimation Run

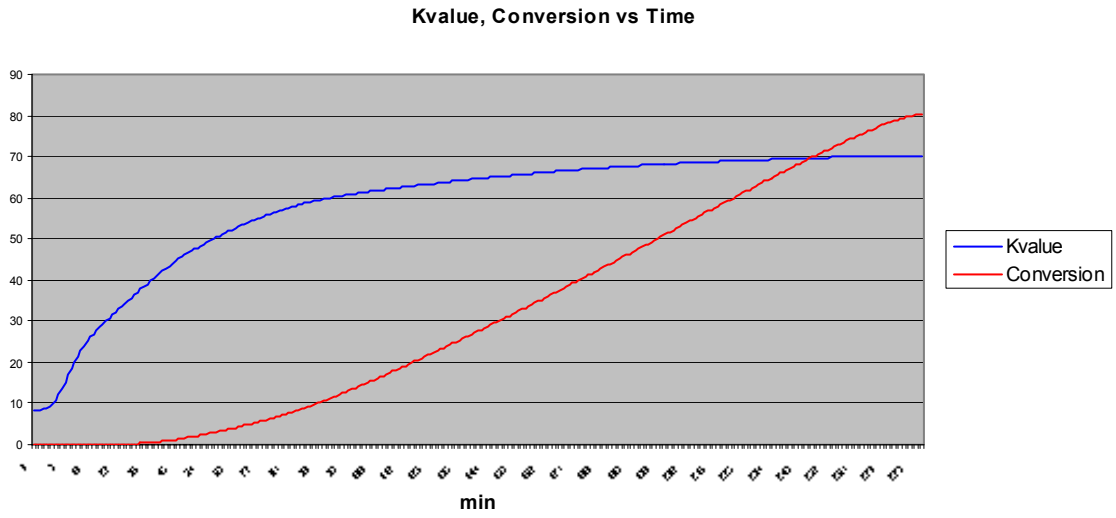


Figure 4: PVC Reactor Kvalue and Conversion Profile Results from Parameter Estimation Run

## Dynamic Optimization

Once properly tuned, the model can be used to optimize the PVC reaction operation. The optimization of the PVC process is tied, primarily, to reducing the total time for a batch run to complete. In order to achieve this, the batch recipe must be set so that the profile of heat removal over time is square and the heat removal rate is at the maximum constraint throughout. Industrially, attempts to establish this square profile are often centered on the use of dual initiators, one fast and one slow. The fast initiator starts off the reaction, while the slow initiator replenishes the free radical population of the system, which would otherwise decrease as a result of the various termination reactions.

In this development effort, the goal of the batch optimization is to determine the optimal operating policy (recipe and reactor temperature setpoint etc.) to minimize operating cost. One

tactic is to minimize the batch time for a given conversion while operating at the maximum heat removal rate. The batch optimization is formulated as a dynamic optimization problem incorporating an economic objective function, the decision variables, and all the important physical and operation constraints, and solved using the dynamic optimization facilities included in Dynamic Solutions.

**Objective function**

To implement a batch optimization problem, it is important to define an objective function for the PVC reaction process that reflects the underlying economic incentives. Therefore, it should include the following objectives:

- Maximize polymer produced
- Minimize initiator charge
- Minimize batch time
- Minimize the utility cost

These factors may be combined with various weights to form a single optimization objective function. The actual cost information can also be included to formulate a cost-based economic objective function.

**Decision variables**

The variables available to the optimizer to "adjust" in order to find an optimal solution are known as decision variables. Decision variables must be independent, and, ideally, can be directly controlled by an operator. Examples include:

- Initiator 01 charge
- Initiator 02 charge
- Reactor temperature setpoint profile (setpoints and durations)

**Constraints**

In addition to the objective function and decision variables, the optimizer must consider constraints, which identify certain physical or operational limits that the optimizer must observe as it establishes targets for the decision variables. Examples are:

- Maximum available cooling water flow rate to jacket
- Maximum available cooling water flow rate to condenser
- Maximum heat removal rate
- Maximum and minimum coolant temperature
- Reactor temperature
- Final K_value should be on target

## Application Scenario

A scenario in which the PVC optimizer may be applied on line for each cycle of the PVC batch operation is outlined as follows:

1. Batch set-up / Reactor Charging
   Optimizer on standby

2. Reactor heat-up and normal operation
   Collect on-line data periodically, e.g., once every five minutes
   Predict and display remaining batch profile periodically, e.g., once every five minutes

Display available cooling capacity and display and alert the operators of possible run-away conditions based on simulation results (ideally at least 30 minutes ahead of the predicted happening)
Determine the optimal reaction end point

3. Reactor Run Down
Re-estimate model parameters with the data of the last batch
Determine optimal recipe and reaction conditions for the next batch

The optimizer may also be used to collect long term reaction trend data and determine and advise on optimal maintenance frequency.

## Potential Benefits

The PVC reaction optimizer can potentially offer significant benefits in the following areas:

1. Shortening of batch cycle times

This is achieved by matching initiator charge to current heat removal capacity, by optimizing ratios of blended initiators to achieve squarer heat generation profiles, and by stopping the reaction at the optimum time.

2. Reduction of overall product loses

This is achieved by reducing frequency of aborted batches and by reducing the occurrence of batches outside product quality limits.

3. Others

Other benefits include accurate prediction and repeatability of batch cycle times, improved product consistency by eliminating batch to batch variability, and the ability to advise on the optimum timing for reactor overhaul.

In short, the optimizer offers several advantages for PVC operation. When used as an on-line "shadow" for a process run, it can detect conditions, such as higher than expected heat generation requiring a higher than available cooling capacity, that may lead to possible runaway situtations later on in the batch. In these cases, provisions could be made to evaluate the consequences of various actions taken to regain control of the reactor. Through the monitoring of predicted heat transfer coefficients, it could be used to specify the point at which a full cleaning of the reactor and condenser walls would be required. It can also predict the conversion, product quality, and other details for normal runs. As an optimizer, it can be used to evaluate the practicality of run recipes using timed or scaled initiator additions that can match the current production rate and/or raw material cost requirements, while operating within variable constraints or safety limits.

## Future Directions

There are two development items that should be further investigated in order to realize and enhance the potential benefits of the PVC optimizer: (i) model detail and parameter tuning, (ii) modeling platform.

In general, the parameter estimation runs of the current optimizer closely match the available data for a given operating temperature in a given plant. However, developing a single set of tuned parameters capable of representing all likely commercial configurations has been elusive. This is probably due to deficiencies in the quantity and quality of available reactor data as well as deficiencies in the model detail.

Future development of the model should, therefore, address the following outstanding issues:

(i) Ensure availability of operation data (plant or laboratory) that cover a wider range of temperature, heat removal, and product quality such as average molecular weight.
(ii) With improved data availability, expand the parameter estimation scope by including more model parameters such as phase equilibrium parameters.
(iii) Investigate and improve on the modeling of gel effect which currently is based on a simple correlation of conversion
(iv) Investigate and improve on the modeling of chain transfer agents

The process model was originally developed on the Dynamic Solutions platform. However, a tighter integration with the plant control system and data collection facilities would be desirable for online applications. It should also be noted that the batch operation and the associated sequence control logic often have to be implemented in external functions (foreign objects) in Dynamic Solutions. This necessitates frequent re-initialization of the model and thus significantly increases the CPU time requirement. This may prove to be a hindrance in online applications where execution speed is very important.

In a related development, a decision has been made recently within the ABB CRC organization to unify the various modeling platforms across ABB and base the common modeling platform on Modelica, which is an object-oriented and multi-domain modeling language designed for physical system modeling. Modelica offers many significant advantages and features such as symbolic pre-processing, selective suppression of re-initialization, and the inclusion of sequential calculations in an otherwise equation-based, simultaneous environment.

The runtime environment with supporting estimation and optimization engines will be provided by ABB's Dynamic Optimization System Extension (DOSE) which is designed to work closely with ABB's Industrial IT 800xA Extended Automation System. The tight integration with 800xA system also enables one to benefit from the other services such as history logs, trends, and OPC DA data transfers to external systems. When fully developed, a tight and seamless integration of data collection, preprocessing, real time and online optimization, and closed loop control will be available to optimize the operation of complex processes such as PVC production plants (Fig. 5).

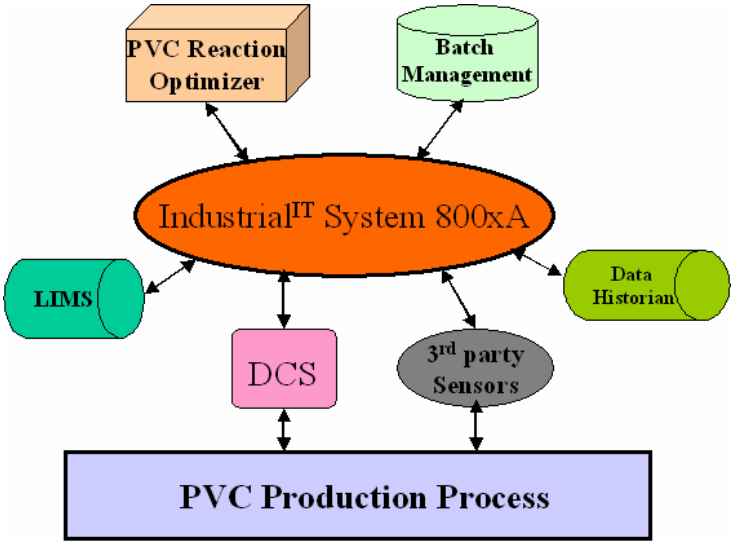


Figure 5. Overview of a PVC Reaction Optimizer package integrated into 800xA

## References

[1] "Dynamic Simulation of Industrial Poly(vinyl chloride) Batch Suspension Polymerization Reactors", C. Kiparissides et al., Ind. Eng. Chem. Res. 1997, 36, 1253 - 1267

[2] "Experimental Investigation of Vinyl Chloride Polymerization at High Conversion – Temperature/Pressure/Conversion and Monomer Phase Distribution Relationships", T. Y. Xie et al, Journal of Applied Polymer Science 1987, 34, 1749-1766

[3] "Effect of Impurities on Emulsion Polymerization: Case I Kinetics", A. Penlidis et al, Journal of Applied Polymer Science 1988, 35, 2023-2038

[4] "Effect of Impurities in Vinyl Chloride on the Kinetics of its Polymerization and on Degradation of PVC", V. I. Zegel'man et al, Polymer Science USSR 1985, 27, 882-890

# On Identifying Deficiencies in a Knowledge Management System

**Prashant Baheti[1], Laurie Williams[1], Aldo Dagnino[2], Andrew Cordes[2]**
[1]North Carolina State University, Department of Computer Science
{ppbaheti, lawilli3}@unity.ncsu.edu
[2]ABB, Inc.
{aldo.dagnino, andrew.cordes}@us.abb.com

**Abstract.** The Abou-Zeid Knowledge Management Reference Model provides a structure for the elements necessary for a holistic knowledge management support system. The model underscores the importance for an effective system to consist of three layers: business aspects and associated knowledge nuggets; processes which enable the manipulation and update of these nuggets; and technology that makes this manipulation and access to the nuggets effective and straightforward. This paper supports the Abou-Zeid model through its application at ABB, Inc. in the implementation of the ABB Software Process Initiative (ASPI) Knowledge Base. The ASPI group has strengthened the process and the technology aspects of its previous experience database. The reengineered system demonstrates strength at all three Abou-Zeid layers, leveraging the utility of the captured knowledge nuggets. As structured, initial indications support the reengineered system as significantly more valuable to ABB employees than had been the earlier system.

## 1. Introduction

Organizations need to identify, manage, and exploit their knowledge assets [1]. Often, a large amount of organizational knowledge leaves the organization when an employee and the knowledge he or she possesses leave the company. Additionally, the knowledge within an organization should be available to whoever needs it so that repetitive work can be avoided. By using a knowledge base, organizations can leverage both the positive and negative experiences of its employees.

A software organization's main asset is its intellectual capital, and Knowledge Management (KM) capitalizes on the retention, impact, and distribution of this intellectual capital [2]. Software engineering projects are often challenged by poorly defined requirements, frequent staff turnover, and volatile hardware and software platforms [3]. Software organizations possess knowledge in different areas; each knowledge area is critical for achieving business goals. Some of these areas include documented knowledge about new technology, the organization's domain, local policies and practices, and the knowledge of the employees [2]. Organizations often face problems identifying the content, location, and use of their knowledge because it exists in different forms in the organization.

In the software engineering (SE) domain, KM can be defined as "a set of activities, techniques, and tools supporting the creation and transfer of SE knowledge throughout the organization" [4]. Many organizations realize that KM is an integral part of their process improvement initiative, which involves facilitating access to process-related artifacts that support SE and quality management techniques [4]. For example, software process knowledge can be externalized (converted from tacit to explicit) [5] by having a set of standard templates for different phases of development and support activities, such as a software requirements specification template.

This paper discusses the reengineering of a web-based software and product engineering knowledge base developed at ABB, Inc.[1] The initial version of the knowledge base, the Experience Database (henceforth referred to as ABB-1), contained experiences and training packages. However, the system lacked critical elements necessary for effective update and use by development units. The reengineered knowledge base, called the ASPI Knowledge Base (henceforth referred to as ABB-2), contains process templates, procedures, and documented valuable experiences of the employees in development units. More importantly, the supporting technology has been enhanced and processes have been developed to improve the utility of the system to ABB employees.

Our experiences reengineering System 1 have heightened our awareness that a comprehensive and successful knowledge management support system is multifaceted. Our findings support the framework established by the Knowledge Management Reference Model suggested by Abou-Zeid [6]. The rest of the paper is organized as follows. Section 2 provides background information on knowledge management support systems, particularly those in the SE domain. Section 3 describes the Abou-Zeid KM Reference Model. Sections 4 and 5 describe the initial and reengineered ABB knowledge bases. Section 6 portrays the implementation of this reengineered knowledge base within the context of the Abou-Zeid reference model. Finally, Section 7 provides the conclusion and plans for future work.

[1] http://www.abb.com

## 2. Related Work

This section discusses various industry efforts to support KM within their organizations and discusses different KM models typically used in the software engineering domain.

### 2.1 KM in Software Engineering

Several companies have initiated software engineering KM support systems (KMSS), as we now discuss. An objective of the KM initiative at VTT Electronics [4] was that the KM system needed to have minimal impact on the software development and processes. Therefore, the solution could not require new technologies. VTT utilized software process improvement (SPI) experts as knowledge-capturing agents, rather than soliciting knowledge from software developers. The SPI experts gleaned information from project final reports, error databases, discussion forums, and through speaking with people in the organization. This information was provided to customer projects on demand by packaging this knowledge for easy access. The required knowledge for the project was provided just-in-time. This needs-based KM approach adopted by VTT worked well for one of the customer projects inside the company. The customer project could retrieve the needed knowledge from the delivered knowledge package.

Infosys Technologies Ltd. is a software services company that manages organization-wide knowledge using its three centrally-operated knowledge repositories [7]: the Knowledge Shop (K-Shop), Process Asset Database, and People Knowledge Map. The K-shop provides employees access to its resources related to technology, domain, trends, culture, project experiences, internal and external literature through a web interface. The Process Asset Database captures the "as-is" project deliverables, such as project plans, design documents, and test plans. The People Knowledge Map is an intranet-based system for employees to search and locate experts in different fields. The K-shop has helped Infosys to increase its productivity by three percent due to effective reuse and to reduce defect levels by as much as 40 percent.

Similarly, International Semiconductor Technology is an Integrated Circuit assembly and testing company where knowledge intensity, specificity, and volatility are comparable to those in SE and other knowledge-intense fields. This organization utilized a KMSS to assist with knowledge creation, update, sharing, and reuse [8]. Their KMSS consists of a lessons-learned repository, a case repository, and an organizational directory. Soliciting experts for answers to problems that arise supports knowledge creation. The experts give their analyses, comments, and/or recommendations. The experts can then discuss the problems using groupware until a solution is reached. At this time, the exchange becomes "validated" knowledge, which is stored in the lessons-learned repository. By supporting the sharing of critical, task-specific knowledge previously held by individuals, the company aims at developing better systems at a faster pace.

These examples suggest that the structure of software engineering-related KM initiatives depends upon the context of the organization. However, in each case the primary goal was to have easy access to tacit and externalized knowledge of the company. Each KMSS usually involves a centralized repository of knowledge resources with an accessible interface to the users. The other goal was to facilitate socialization [14] by creating an environment conducive to interaction between persons in the company. For example, whenever a person has a problem that can be solved by some other expert in the same organization, he or she should be able to find and contact that expert and/or some documented knowledge provided by that expert.

### 2.2 Knowledge Management Frameworks

Many organizations have built their own KM frameworks to oversee KM processes, methodologies, tools and techniques within the company. Most of the frameworks are prescriptive in nature [9], implying that these frameworks are not holistic. Instead, the frameworks are task-based because they just consider the tasks which facilitate movement of knowledge within the organization. These frameworks do not consider factors such as linking knowledge management to business objectives, feedback loops within the organization, and cultural factors. Examples of organizations who have authored such frameworks include Ernst and Young, Knowledge Associates, and The Knowledge Research Institute, Inc. [9]

Even the descriptive frameworks that characterize knowledge management have been unable to view KM holistically. For example, "the National Technical Institute of Greece, Andersen Consulting, and The Delphi Group include cultural factors in their framework but not learning or linkages with strategic business objectives." [9]

Binney [10] developed a KM spectrum which acts as framework that helps organizations in understanding the range of KM options, applications and technologies available to them. This framework groups KM applications according to six groupings: transactional, analytical, asset management based, process based, developmental and innovation/creation-based. Although this framework transcends most of the KM applications available and their enabling technologies, it is again task-based rather than considering multiple facets of the organization.

The Knowledge Management Maturity Model (KMMM®) [11], developed at Siemens AG, does provide

a holistic approach to knowledge management. It defines five maturity levels of Knowledge Management, similar to those of the SEI's Capability Maturity Model [12] and defines the various key areas associated with the model. The phases involved in a typical KMMM® project for assessing an organization's maturity level are orientation and planning; motivation and data collection; consolidation and preparation; feedback and consensus; ideas for solutions and action proposals; and report and presentation. We considered guiding our reengineering work with the KMMM®. However, ABB places a higher priority on business value than on achieving a maturity level.

Abou-Zeid [6] composed a knowledge management reference model that is holistic and focuses on business value. We discuss this model more in detail in the following section because we present our research in relation to this model.

## 3. The Abou-Zeid Knowledge Management Reference Model

Abou-Zeid [6] has devised a Knowledge Management Reference Model (KMRM) that aims to answer the question *"How can a KMSS be developed, operated, and assessed?"* The model identifies the processes to be supported by any KMSS. These processes change the state of current knowledge within the organization to a state where knowledge exists in an updated and desirable form for everyone to access. For example, one such process could be identifying the experts in an organization and documenting their views on specific problems to form new knowledge for future reference. The Abou-Zeid KMRM also models the dynamics of these processes, thereby developing a framework of a KMSS that is specific to a particular business and that is "aware of" the issues of that business.

Abou-Zeid's model identifies the processes that should be supported by KMSS through its three-layer reference model. The three layers, cognitive domain layer, functional layers, and resources layer, are each discussed below.

*Cognitive Domain Layer:* The model's first layer, called the cognitive domain layer, addresses entities, called business things (or *B-things*), that relate to business issues or organizational goals. A B-thing may be a concrete or abstract entity. An example of a **B-thing** is **efficient and cost-effective software production**.

Each B-thing is associated with certain knowledge that enables or supports it. These are called knowledge things (or *K-things*, akin to knowledge nuggets). The "efficient and cost-effective software production" B-thing is supported through the reuse of already-available resources within the organization. For example, product development templates already developed within the organization should be available to any business unit. The **product development templates** are **K-things**.

*Functional Layer:* In an organization, the B-things are relatively stable; however, the associated K-things are in a state of continual change as knowledge evolves. The second layer, known as the functional layer, includes the processes required to change the state of the K-things. "Since knowledge requirements are difficult to determine fully a priori, associated processes have to be dynamic, evolving, and flexible."[6] These are called the *K-manipulating* processes. The processes supporting the K-manipulating processes [13] are called *K-enabling* processes. The K-enabling processes produce cultural and organizational enabling conditions for the K-manipulating processes. A **K-manipulating process** provides the ability for an employee or a business unit to **contribute a template** to the KMSS, keeping these K-things as current as possible. A **general KM awareness** within the organization can be termed as a **K-enabling process**.

*Resources Layer:* The third layer is the resources layer that consists of the enabling information/communication technologies (ICT) to support the K-manipulating and K-enabling processes. For example, the computing environment of an organization is a part of the resources layer.

In summary, the Abou-Zeid KMRM guides in the implementation of an integrative KMSS.

## 4. The Initial Experience Database

In our research, we studied a knowledge base reengineering effort at ABB. ABB-1 consisted of a Lotus Notes database with a custom-built HTML interface. The database was a collection of documented software development experiences and good practices from ABB companies. Experiences and best practices were collected through the submittal of structured experience "packages." These packages underwent a multi-phased review and approval process before being accepted as entries to the database. The database provided a limited search capability of the experience packages and did not offer a mechanism for identifying pertinent external resources and web links.

This experience database was a fairly ineffective K-thing. The problems with this experience database are listed below:

1. The review process for the experiences was too cumbersome. To submit an experience, an employee needed to first submit an abstract to a review board and, if approved, send the actual experience formatted according to a detailed packaging template. The experience packages then underwent a second review before final acceptance and publish. Turnaround times of a month or more

were not uncommon for publishing an experience. (unsuccessful K-manipulating process)

2. The database was not frequently updated because no one was formally responsible for it, further increasing the turnaround time. (unsuccessful K-manipulating process)
3. There was no incentive to submit experiences. (lack of K-enabling process)
4. The experiences were approved only if they were a corporate-level best practice, limiting the number of experiences in the knowledge base. (restrictive K-manipulating policy)
5. The experience base was not well publicized. (unsuccessful K-enabling process)
6. It was not easy to find relevant articles in the database because the structure was not intuitive. (ineffective technology)
7. The database was limited to ABB experiences; there were no external references to supporting material. (restrictive K-manipulating policy)

In summary, the function and resources layers of the initial system were lacking, reducing the utility of the system.

## 5. The Reengineering Effort

Due to the inefficacy of ABB-1, a KMSS reengineering effort was launched. The main author implemented a new knowledge base in the Lotus Notes/Domino environment on a Windows platform. The knowledge base is a Lotus Notes database with a template-driven Web interface for easy access to the knowledge artifacts within the database. The knowledge base was developed using the Common ABB Web Platform (CAWP), an application owned by the ABB Corporate Communication division; CAWP was also used to develop the ABB intranet. The use of CAWP provided visual consistency for employees. The contents of the knowledge base are available to all the employees via a well-known link on the company intranet.

The knowledge base consists of product development information for various ABB business units. This includes the various experiences submitted by the employees of ABB and the procedures, guidelines, templates and training materials that have been successfully used by ABB product development organizations. Change agents like developers and project managers can retrieve examples of processes, templates, and guidelines from the knowledge base. Contributions to the knowledge base can be submitted by anyone within the organization. The structure of the knowledge base ties most entries to two important reference points for the company: the ABB Gate Model [14, 15] and the SEI Capability Maturity Model Integration (CMMI) process areas [16].

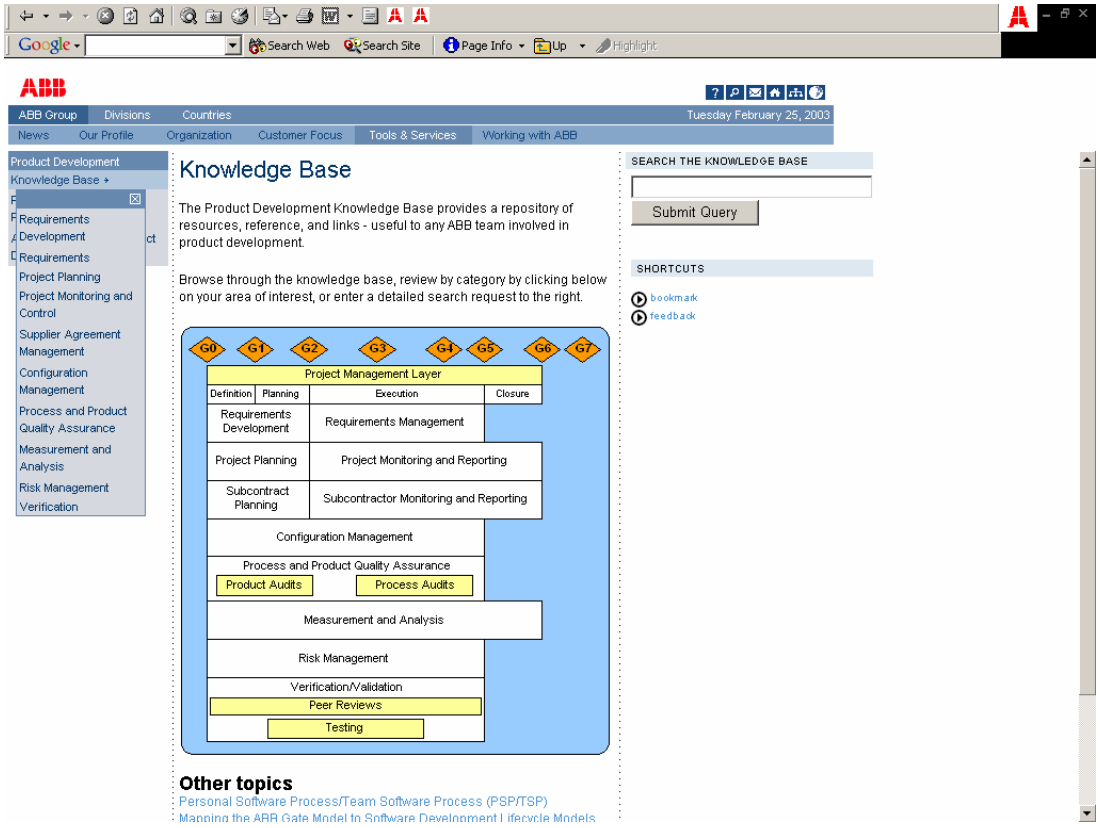


**Figure 1: Main Page of ASPI Knowledge Base (ABB-2)**

Figure 1 provides a screenshot of the main page of ABB-2. The left menu contains links to various process areas within the knowledge base, such as requirements management. The main page displays all the gates of the ABB Gate Model as diamonds that link to resources about each gate. Below each gate is the set of process areas that are related to that gate in the development cycle. Each of these process areas leads to a page of resources; each resource page contains links to ABB experiences, procedures, guidelines, templates and training materials, and other external (non-ABB) references. The knowledge base provides an effective search capability to query for particular documents, templates or keywords. A mechanism is provided for collecting feedback from the users so their questions and concerns can be addressed.

User-friendliness was a main priority in the knowledge base design. Based upon user feedback, the user interface design was modified several times to ensure that the knowledge base was easy to access and intuitive in the way it was organized. Another key objective was the provision of an employee incentive program to encourage contributions to this knowledge base; a list of contributors is posted on the front page of the knowledge base.

Knowledge base contents are regularly updated and new experiences submitted by the employees are incorporated after an efficient yet careful review. A review board, comprised of experts from the ASPI group, is responsible for this review process. The turnaround time is much shorter than that for the prior experience database, because the review board is formally responsible for the knowledge base. The approved submissions are posted on the knowledge base within a week. All ABB business units were asked to submit development process/practice artifacts that they have developed and found to be useful. Specifically, software development process templates were solicited and incorporated into the knowledge base. Subsequently, the business units can readily obtain these artifacts from one central repository.

## 6. Mapping to the Abou-Zeid Model

ABB started its KM initiative in 2000, gaining more attention recently. In 2002-2003, a product development knowledge base was created and implemented for more effective utilization of the experience database K-thing. The reengineered ABB-2 system is now discussed in the context of the three layers of the Abou-Zeid Model.

### 6.1 Cognitive Domain Layer

The model of ABB's cognitive domains is shown in a UML diagram in Figure 2. The Product Development Organizations within ABB form the enterprise of the KM initiative; they are the primary targets for ABB-2. The external B-things (generalizations of the generic B-thing) that this enterprise interacts with are its customers, suppliers, competitors and the market. For example, customers are a primary source of requirements to design and develop products. It is essential to have knowledge on how to elicit and manage requirements from the customer.

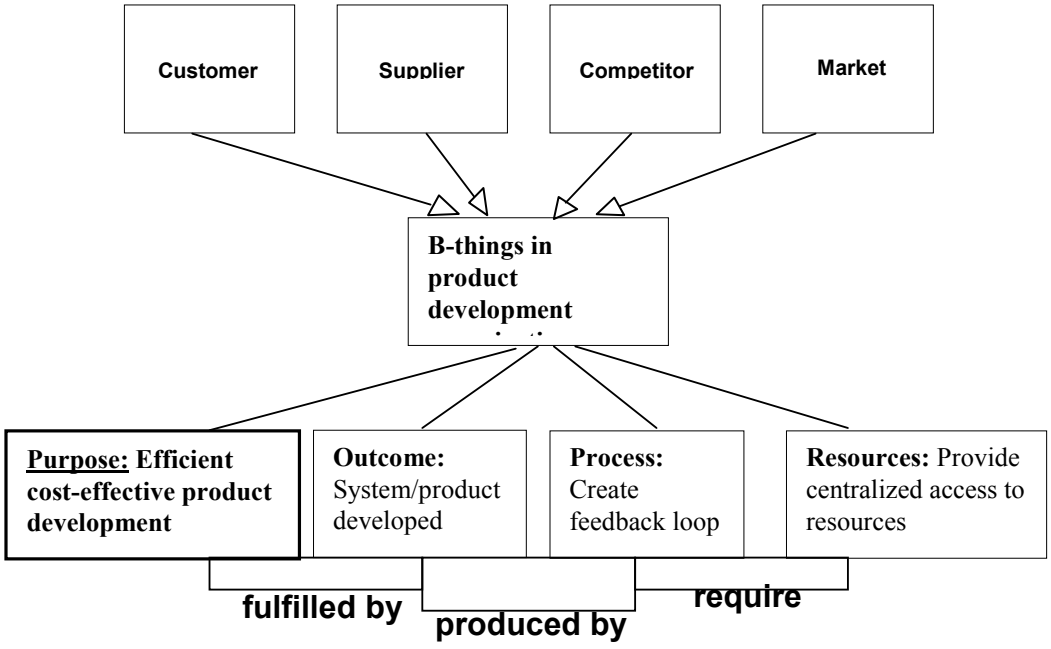


**Figure 2. Model of ABB's cognitive domains**

Often, sub-contractors are used to assist the development group and specific knowledge is used to manage subcontractors. ABB-2 provides knowledge about how to investigate competitors' intellectual property and products. Similarly, knowledge on how to incorporate market information into the development organization is presented in ABB-2.

An internal ABB B-thing is to develop new products or systems in an **efficient and cost effective manner**. The outcome of a product development organization is the developed product or system that is ready to be sold to a customer or market or that is ready for production. Knowledge is essential to the people in the development organizations to perform their work and a feedback loop is needed to incorporate experiences and knowledge into the knowledge repository that is accessed by the developers. As shown in the diagram, the B-thing consists of a purpose, outcome, process, and resources.

Figure 3 displays a simplified version of the ABB-2 in the context of Abou-Zeid KMRM. In this paper, we examine one internal B-thing, "efficient and cost effective product development." The set of actions required to address this B-thing are:

1. Providing centralized and easy access to all the experiences submitted by the employees within the organization. These experiences are integrated with the product-development related resources which were classified according to both the ABB Gate Model [14, 15] and SEI Capability Maturity Model Integration (CMMI) process areas [16].
2. Avoiding unnecessary development by facilitating reuse.
3. Making training materials developed by the ASPI group available to all business units.
4. Providing a mechanism for feedback and answering queries of the business units with respect to any of the three items listed above.

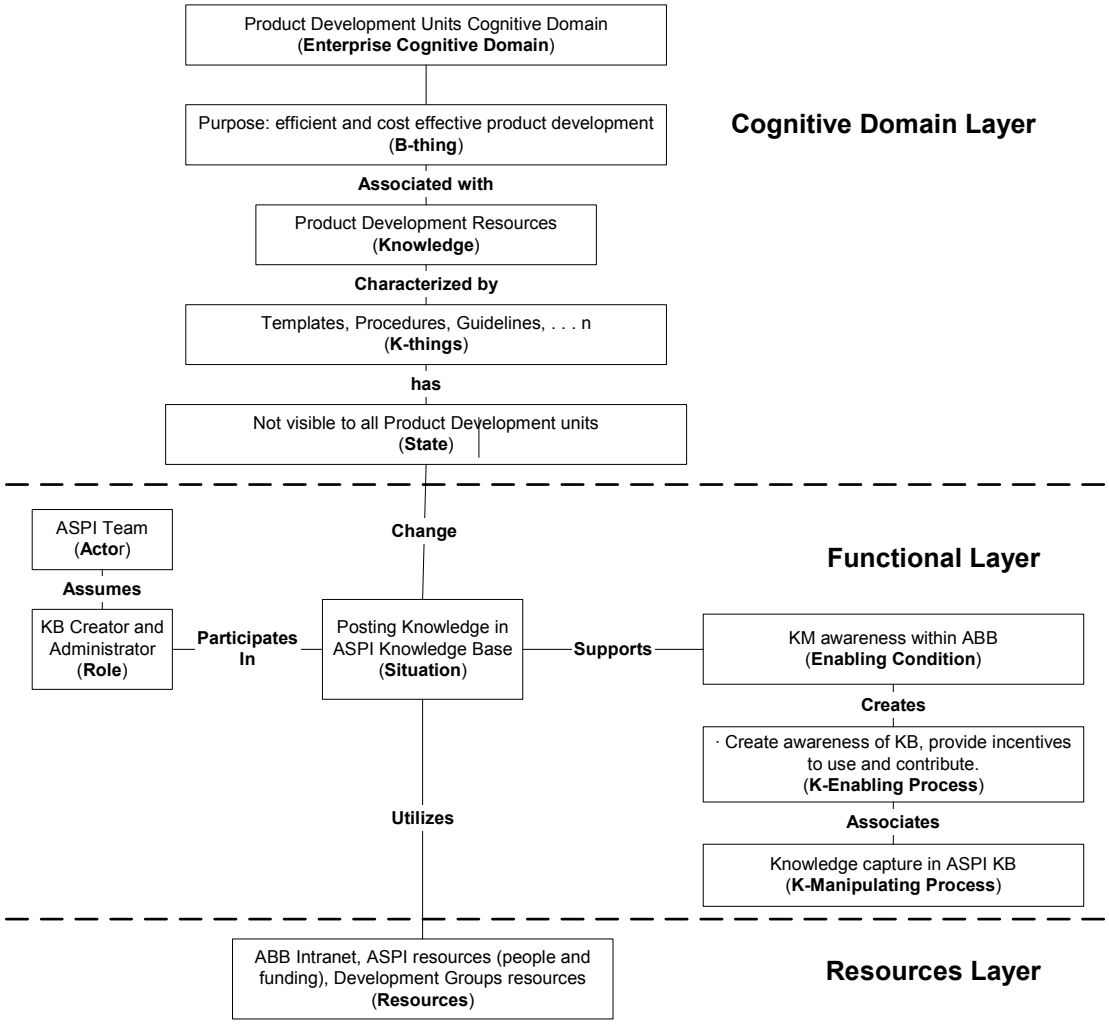


**Figure 3: ABB Knowledge Management Model**

The K-things that are associated with the B-thing are:

- Experiences in ABB-1;
- Templates for the different phases of product development;
- Training materials; and
- Tacit knowledge of the experts.

### 6.2 Functional Layer

The ASPI group, which is responsible for the KMSS reengineering effort, owns the knowledge associated with the various B-things for the product development organizations. The group assumes the role of the administrator for ABB-2 and is responsible for eliciting knowledge from various sources, including databases, employees, and reports and posting the knowledge to ABB-2 after a careful review. This layer is responsible for converting the K-things to a visible state so that any employee on the intranet of ABB can easily access them.
The K-manipulating processes are:

- Identifying product-development related resources to be made available through ABB-2;
- Making training materials available through ABB-2;
- Setting up a review board, which controls new additions to ABB-2; and
- Publicizing ABB-2 once within the organization.

The K-enabling processes can be identified as:

- Spreading awareness about the KM initiative taken up by the ASPI group among other business units; and
- Providing incentives to the employees to contribute to the knowledge base.

### 6.3 Resources Layer

The resources layer consists of the infrastructure of ABB, including the intranet of ABB, the Lotus Notes/Domino development environment and CAWP, the web development tool. The ASPI resources, including people employed for the project and funding the KM initiative also falls into this layer.

## 7. Conclusion and Future Work

By sharing our experiences, we aim to make organizations aware of the importance of all three of the layers of the Abou-Zeid model. A retrospective analysis of ABB-1 based upon the Abou-Zeid model revealed significant deficiencies in the functional and resources layers. Because of these deficiencies, ABB did not effectively capitalize on the resources invested in this initial KM system. These deficiencies were addressed in the reengineering effort, resulting in the more effective ABB-2. A comparison of the processes and things at each level can be found in Table 1. Note the stability of the B-thing, the addition of K-things and K-processes, and enhancement of the ICT.

Initial feedback on ABB-2 has been positive. ASPI team members and change agents at ABB organizations planning process improvements have found the knowledge base to be easy to find on the intranet, easy to use, and full of information and resources supportive of their tasks.

**Table 1: Comparison Chart**

| **Things and Processes (Level)** | **Experience Database (ABB-1)** | **ASPI Knowledge Base (ABB-2)** |
|---|---|---|
| B-things (1) | Efficient and cost-effective product development | Efficient and cost-effective product development |
| K-things (1) | • Experiences<br>• Training materials | • Experiences<br>• Templates<br>• Training materials<br>• Tacit knowledge |
| K-manipulating processes (2) | • Cumbersome review process<br>• Not well publicized<br>• Difficult to navigate<br>• No frequent updates | • Identify resources<br>• Centralized access to resources<br>• Efficient review<br>• Proper publicity<br>• Frequently updated |
| K-enabling processes (2) | No K-enabling processes identified | Awareness of the KMSS and recognition for contribution |
| ICT (3) | Lotus Notes | Lotus Notes, CAWP |

The structure of the knowledge base, relating development activities to CMMI process areas and ABB Gate Model gates, has simplified the knowledge search process as these two are well known within the organization. New contributions to the knowledge base are reviewed and incorporated within one week and initial web statistics show a broad range of users across multiple country sites. We are continuing the evaluation of ABB's knowledge management initiative through a detailed quantitative assessment. The three aspects being monitored are usage, effectiveness and return on investment.

*Usage:* Usage metrics include the number of people using the knowledge base and the frequency of use

gathered via server logs. We will also analyze the profiles of the people accessing the knowledge base, which includes factors such as position in the organization and group to which they belong. This will help us to better tailor the knowledge base to the employees who access it.

*Effectiveness:* Effectiveness is measured through feedback from the users of the knowledge base. This feedback will help us analyze the average percentage of times the employees find what they want from the knowledge base. Effectiveness metrics include ease of use, accessibility of the knowledge base, and average amount of time required to find the required resource.

*Return on investment (ROI):* To examine ROI, we have solicited all business units to provide all the templates they presently use for their development purposes. When we feel this collection is complete, we can find out how many different copies/versions of the same template have been created in the organizations. Feedback from the product development organizations about the time taken to develop their templates will help us find an estimate of the average amount of development time required for each template. The time saved henceforth by avoiding redevelopment of templates can be factored in calculating ROI. The other factor is time saved in the future due to presence of a central repository of resources. This involves feedback from the product development organizations as to what was the time required to access resources compared to now, with the presence of ABB-2. We are investigating additional ROI metrics, which can also account for the presence of feedback loops within the organization due to ABB-2.

**Acknowledgements**

We would like to thank Dr. Abou-Zeid for helping us by further explaining his model in our context and for reviewing and improving this paper.

## References:

[1] M. Lindvall, I. Rus, R. Jammalamadaka, and R. Thakker, "Software tools for Knowledge Management," DoD Data Analysis Center for Software, Rome, N.Y. 2001.

[2] I. Rus and M. Lindvall, "Knowledge Management in Software Engineering," *IEEE Software*, vol. 19, no. 3, pp. 26-38, May/June 2002.

[3] T. K. A. Hamid, "The Slippery path to productivity improvement," *IEEE Software*, vol. 13, no. 4, pp. 43-52, 1996.

[4] S. Komi-Sirvio, A. Mantyniemi, and V. Seppanen, "Toward a Practical Solution for Capturing Knowledge for Software Projects," *IEEE Software*, vol. 19, no. 3, pp. 60-62, May/June 2002.

[5] H. Takeuchi and I. Nonaka, *The Knowledge Creating Company:How Japanese Companies Create the Dynamics of Innovation*: New York: Oxford University Press, 1995.

[6] E.-S. Abou-Zeid, "A Knowledge Management Reference Model," *Journal of Knowledge Management*, vol. 6, no. 5, pp. 486-499, 2002.

[7] S. Ramasubramaniam and G. Jagadeesan, "Knowledge Management at Infosys," *IEEE Software*, vol. 19, no. 3, pp. 53-55, May/June 2002.

[8] C.-P. Wei, P. J.-H. Hu, and H.-H. Chen, "Design and Evaluation of a Knowledge Management System," *IEEE Software*, vol. 19, no. 3, pp. 56-59, May/June 2002.

[9] B. Rubenstein-Montano, J. Liebowitz, J. Buchwalter, D. McCaw, and T. K. M. M. Team, "A Systems Thinking Framework for Knowledge Management," *Decision Support Systems*, vol. 31, no. 1, pp. 5-16, May 2001.

[10] D. Binney, "The knowledge management spectrum - understanding the KM landscape," *Journal of Knowledge Management*, vol. 5, no. 1, pp. 33-42, 2001.

[11] K. Ehms and D. M. Langen, "Holistic Development of Knowledge Management with KMMM®." http://www.kmadvantage.com/docs/km_articles/Holistic_Development_of_KM_with_KMMM.pdf, 2002.

[12] M. C. Paulk, B. Curtis, M. B. Chrissis, and C. V. Weber, "Capability Maturity Model for Software, Version 1.1," *IEEE Software*, vol. 10, no. 4, pp. 18-27, July 1993.

[13] V. Krogh, I. Nonaka, and K. Ichijo, *Enabling Knowledge Creation*: Oxford University Press, Oxford, 2000.

[14] R. A. Carter, A. I. Antón, A. Dagnino, and L. Williams, "Evolving Beyond Requirements Creep: A Risk-Based Evolutionary Prototyping Model," presented at IEEE 5th International Symposium on Requirements Engineering (RE'01), Toronto, Canada, 2001.

[15] C. Wallin, F. Ekdahl, and S. Larsson, "Integrating Business and Software Development Models," vol. 19, no. 6, pp. 28-33, November/December 2002.

[16] M. B. Chrissis, M. Konrad, and S. Shrum, *CMMI®: Guidelines for Process Integration and Product Improvement*: Addison Wesley Professional, 2003.

# The Kinetics of Nanocrystallization and Microstructural Observations in FINEMET, NANOPERM and HITPERM Nanocomposite Magnetic Materials.

M.E. McHenry; F. Johnson; H. Okumura; T. Ohkubo; V.R.V. Ramanan*; D.E. Laughlin
Dept. of Material Science and Eng.; Carnegie Mellon Univ.; Pittsburgh, PA 15213
* ABB Power T&C Inc. /ETA/USTRA/ABB, Raleigh, NC 27606

## Abstract

This paper presents experimental observations of the nanocrystallization process in materials having Fe-Si (FINEMET), α-Fe (NANOPERM), and α'-FeCo (HITPERM) nanocrystals coupled through an amorphous phase. Crystallization kinetics and chemical partitioning during crystallization are described. Isothermal nanocrystallization is discussed in the framework of the Johnson-Mehl-Avrami-Kolmogorov (JMAK) model and constant heating rate experiments analyzed in the context of the Kissinger model.

## Introduction

Nanocrystalline magnets are being investigated for transformers, inductors, and high moment underlayers for pependicular recording media, etc., where low coercivities, $H_c$, large saturation inductions, $B_s$, large resistivities and good thermal stability are needed. Superior soft magnetic properties stem from chemical and structural variations on a nanoscale. Important magnets are derived by crystallizing amorphous precursors resulting in nanocrystalline grains of a (BCC, $DO_3$ or CsCl) ((Fe(Co),X) phase consuming 20-90% of the total volume, homogeneously dispersed in an amorphous matrix. These 2-phase materials are designated ***metal/amorphous nanocomposites***. Applications have been identified utilizing the patented Fe-Si-B-Nb-Cu alloys (tradename FINEMET@TM [1]) and FeMBCu, alloys (tradename NANOPERM@TM [2]). CMU efforts under AFOSR(ONR) MURI funding, resulted a new nanocrystalline $(Fe_{1-x}Co_x)_{88}M_7B_4Cu$ (M = Nb, Zr, Hf) soft magnetic material called HITPERM [3]. HITPERM has high frequency response and high temperature induction. Magnetic metal/amorphous nanocomposites [4] have excellent soft magnetic properties as measured by the figures of merit of combined induction and magnetic permeability. Other figures of merit include high frequency magnetic response and retention of magnetic softness at elevated temperatures. Frequency response is ultimately limited by Snoek's law and uniaxial anisotropy must be developed in a controlled manner to attain the high frequency limit. High temperature operation is advantageous either to allow integration with other high temperature electronic components or to allow greater efficiency in going to higher frequency.

Averaging magnetocrystalline anisotropy over grains coupled within an ***exchange length*** is the root of magnetic softness in these materials [5,6]. In HITPERM nanocomposites [4] nanocrystalline α and α'(B2)-FeCo are formed with improved high temperature magnetic properties. Here we the following fundamental issues: (1) models of the kinetics of nanocrystallization; (2) observations of the role of chemical partitioning occurring during crystallization in determining the temperature dependence of the intergranular coupling of the nanocrystals. HITPERM has been developed for applications with high permeability requirements and with needs for large inductions at high temperatures.

## Crystallization and Nanocrystallization Kinetics

Crystallization is a solid state phase transformation often controlled by nucleation and growth kinetics. The progress of an isothermal phase transformation is represented by plotting the volume fraction (i. e. of the primary crystallites) transformed, X(t,T), as a function of temperature, T, and time, t, in a *TTT curve*. The reaction kinetic arguments necessary to derive a

TTT curve are based on Johnson-Mehl [7]-Avrami [8]-Kolmogorov [9] (JMAK kinetics). The JMAK model for isothermal (primary) crystallization kinetics considers modifications to simple first-order reaction rate kinetics, where the rate of transformation depends linearly on the volume fraction of parent phase remaining, (1- X(t)), with a rate constant k which is thermally activated:

$$k = k_0 \exp\left(\frac{-Q^*}{k_B T}\right) \qquad [1]$$

and Q* is the activation energy barrier to crystallization. Modifications within JMAK consider (1) the dimensionality of the growing particle; (2) the time dependence of the nucleation rate, (3) whether growth is linear or parabolic and (4) the eventual impingement of growing particles at long times. The JMAK equation, for X(t), is written in the general form:

$$X = 1 - \exp\left(-(k(t - t_i))^n\right) \qquad [2]$$

where $t_i$ is an incubation time, the exponent n is observed to vary between 1 and 4 and is used to corroborate a nucleation and growth mechanism and dimensionality. From determination of X(t) at various temperatures k(T) can be determined and $Q^*_{JMA}$ inferred from the Arrhenius [Eq. 1]. Kissinger plots [10] are also often used to determine activation energies for crystallization, $Q^*_K$.

The JMAK model gives a representation with predictive capability of the nanocrystallization kinetics for NANOPERM alloys as originally discussed by Suzuki, et al [11]. It should be noted that the assumptions of the JMAK model do not completely represent the micromechanisms for nanocrystallization [12]. In particular, the slowing down of the rate of transformation at longer times in JMAK is attributed to growth from *phantom nuclei* in the already transformed volume and extended volume to account for the eventual impingement of the growing nanocrystals. It is now clear that the slowing down of the nanocrystallization rate is due to the redistribution of early transition metal (and metalloid) species to the intergranular amorphous phase where it plays the role of a diffusion barrier to further growth of the nanocrystalline phase. Future kinetic models need to be re-examined in light of early transition metal redistribution and so-called *soft impingement.* The decreasing rate of transformation at long times is attributed to impingement of diffusion fields (i.e. early transition metals) surrounding growing nanoparticles rather than the particles themselves.

Microscopic mechanisms for primary nanocrystallization are at the heart of microstructural control for optimizing magnetic properties in metal/amorphous nanocomposites. In general, alloys are chosen to be of hypoeutectic compositions (to maximize the amount of the high moment transition metal). Primary crystallization gives rise to a 2-phase microstructure consisting of the magnetic nanocrystals and an intergranular amorphous phase enriched in early transition metal and metalloid species. A mechanistic model for nucleation and growth has been established in most detail for FINEMET alloys [13, 14].

**Crystallization Observations in Fe-based Metallic Glasses**

Typical Metglas alloys are hypoeutectic (Fe-rich), so as to have large inductions, and observed to crystallize in a 2-step process involving primary crystallization of α-Fe. The primary crystallization reaction; **Am --> Am' + α-Fe;** is followed by secondary crystallization of a metalloid enriched amorphous phase; Am' [15]. Luborsky and Lieberman [16] studied crystallization kinetics of $Fe_xB_{1-x}$ alloys using differential scanning calorimetry (DSC). Activation energies of ~2.5 eV were determined for x = 72-82 independent of x, for x = 82-88 $Q^*_K$ was observed to decrease with x. This indicates the importance of near eutectic compositions for

amorphous phase stability. Ramanan and Fish [17] observed that replacement of B by Si increases activation energy barriers in Metglas alloys, with JMAK morphology indices n ranging from 2.5 - 3.5 consistent with a 3-d nucleation and growth mechanism. Donald and Davies [18] studied the primary crystallization temperature, $T_{x1}$, as a function of transition metal substitution, X, in $M_{78-x}X_xSi_{10}B_{12}$. They explained variations in $T_{x1}$, in the framework of Hume-Rothery rules correlating $T_{x1}$ with the ordering of the cohesive energies of the pure species, X, and atomic size. Two important observations pertinent to the discussion below are: (1) Cu additions which promote nucleation of the primary nanocrystals by clustering resulted in significant reductions in $T_{x1}$, with additions as small as 0.5-1.0 at %; (2) Early transition metals, TE (e.g. Zr, Hf, Mo), additions impede growth (increasing activation energies for diffusion) and result in the largest primary crystallization temperatures.

**Nanocrystallization of FINEMET Alloys**

$Fe_{73.5}Si_{13.5}B_9Nb_3Cu_1$ alloy compositions have been studied extensively for over a decade. Atom probe analysis and high resolution TEM showed [14] that after short annealing (10 min at 550°C) Cu clusters in nanometer scale regions. In the optimally heat treated material these Cu-rich clusters have increased in size to about 5 nm. Nanocrystals with bcc structure were reported to nucleate around these clusters due to the composition fluctuation. Another model proposed that the Cu clusters catalyze the nucleation of the $DO_3$ phase (instead of BCC) through providing a low energy interface for heterogeneous nucleation [13]. Nb, which refines the microstructure, is proposed to stabilize the amorphous state slowing the growth of the Fe-Si $DO_3$ crystallites and to promote Cu clustering by altering the solubility of Cu [13]. The kinetics of Cu clustering and crystallization in FINEMET grades with higher Fe content have also recently been studied [19].

Here we report on FINEMET ribbons with a composition of $Fe_{73.5}Si_{16.1}B_{6.4}Nb_{2.9}Cu_{1.1}$ (at%), which have 2.6 at% higher Si content (replacing B) than in the original FINEMET composition. Fig. 1, shows the existence of the $DO_3$ structureas evident from X-ray diffractometry. Data taken with a slow step scan (inset) exhibits superlattice reflections from (111) and (200) planes of the Fe-Si $DO_3$ structure on the annealed (550°C for 90 min) ribbons. This confirms the existence of the $DO_3$ phase although the solid solution of BCC Fe phase with Si may still coexist (the major Bragg peaks exactly coincide with each other). HREM has revealed large quantities of amorphous matrix surrounding ferromagnetic nanocrystals different in composition due to chemical partitioning during nanocrystallization. A HREM image of one

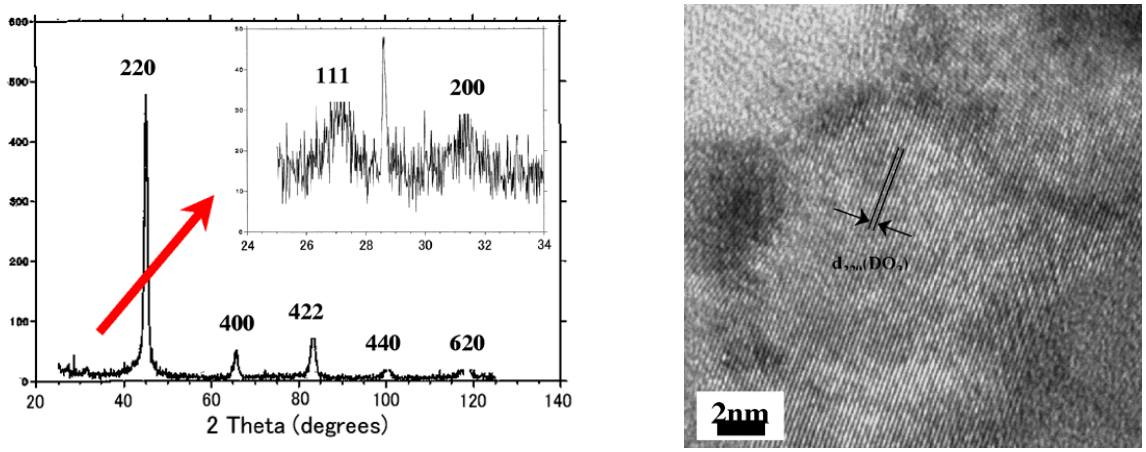


**Fig. 1:** (a) XRD evidence of $DO_3$ nanocrystals in annealed $Fe_{73.5}Si_{16.1}B_{6.4}Nb_{2.9}Cu_{1.1}$ (b) HREM BF image of one of an ~ 15 nm particle showing lattice fringes of (220) $DO_3$ planes.

of the largest particles (the diameter of ~15 nm) with the lattice fringe of (220) $DO_3$ planes is shown in Fig. 1(b). The d-value of 0.200-0.201 nm is consistent with x-ray results. The same (111) and (200) plane superlattice reflections were observed in selected area electron diffraction pattern of the sample. Nano-probe EDS (Energy Dispersive Spectroscopy) analyses indicate that Cu clusters are often enveloped in nanocrystals consistent with the model of Ayers et al. [13].

TEM bright field (BF) images exhibit a uniform distribution of grains, and an average nanocrystal grain size of ~ 7 nm as determined from the dark field (DF) image (Fig. 2(a)). The size distribution of 710 grains is shown in Fig. 2(b). The grain morphology appears less spheroidal and less uniform with higher Fe content. For amorphous ribbons slowly heated to 625°C (just below $T_{x2}$), a homogeneous 10~25 nm distribution of $DO_3$ grains is obtained. The composition of the $DO_3$ phase estimated from its Curie temperature suggests that it regains the original Fe/Si ratio in the amorphous ribbon. From the estimated volume fraction and the magnetization, the intrinsic Fe moment in the $DO_3$ phase was calculated (considering the moment reduction due to Si neighbors), and shown to be consistent with the data of [20]. The microstructure is dramatically changed when the sample is further heated to 825°C (> $T_{x2}$).

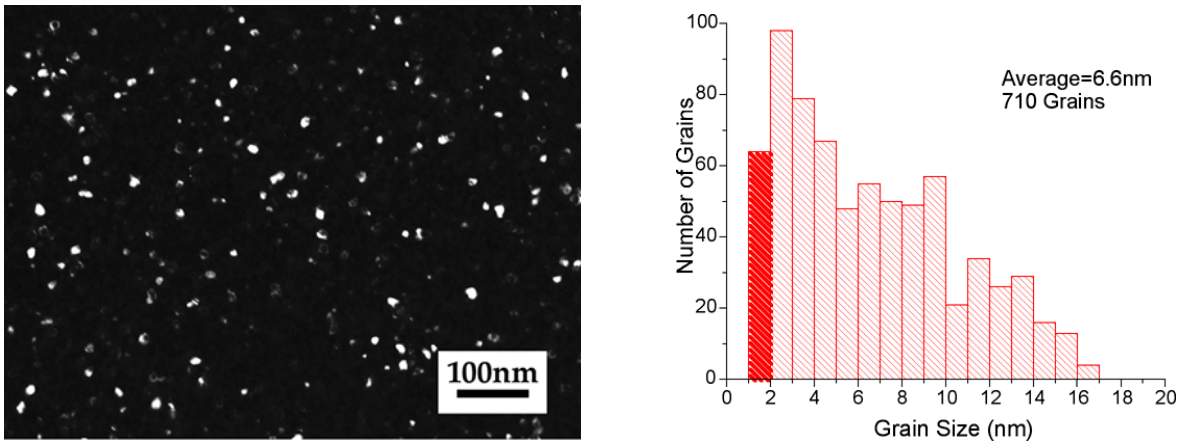


**Fig. 2:** (a) Dark field image in sample of Fig. 1 and (b) size distribution of 710 grains.

**Nanorystallization of NANOPERM Alloys**

Vibrating sample magnetometry (VSM), synchrotron x-ray diffractometry (XRD), and differential scanning calorimetry (DSC) have been used to observe crystallization kinetics and results reviewed by Hsiao, et al. [21]. The volume fraction of nanocrystals transformed in the crystallization process was inferred magnetically, thermally and structurally. Constant heating rate, 3-D synchrotron diffraction (Kramer et al., [22], similar studies were performed by Koster, et al. [23] on FINEMET) of the crystallization of NANOPERM ribbon showed the amorphous to nanocrystalline transformation, primary and secondary crystallization phases, and coarsening phenomena. Features included: (1) the appearance of Fe (110) and Fe (200) peaks as the primary crystallization of α-Fe occurs at $T_{x1}$~510°C, (2) $T_{x2}$ occurs at ~ 710°C with the crystallization of $Fe_2Zr$ and $Fe_{23}Zr_6$, (3) the narrowing of the of Fe(110) and Fe(200) peaks due to coarsening that occurs after secondary crystallization. Both x-ray and TEM have been used to identify the primary and secondary crystallization reactions, respectively, to be: **Am --> Am' + α-Fe; Am' --> $Fe_{23}Zr_6$ + $Fe_2Zr$.** Isothermal VSM observations took advantage of the Curie temperature of the amorphous phase, $T_{c,am}$ of the NANOPERM alloy being less than its primary crystallization temperature, $T_{x1}$. The magnetization is directly proportional to the volume fraction, X, of the primary α-Fe crystalline phase. The JMAK model for isothermal

crystallization kinetics was compared with the Kissinger [10] model for non-isothermal crystallization kinetics using data from the three characterization methods. The activation energy for nanocrystallization of an $Fe_{88}Zr_7B_4Cu$ alloy was determined to be in the range of Q = 2.8 - 3.4 eV, with the crystallization kinetics determined as proceeding by immediate nucleation and 3-D diffusive growth, with the morphology index n = 1.5.

**Nanocrystallization of HITPERM Alloys**

HITPERM alloys $(Fe_{1-x}Co_x)_{88}M_7B_4(Cu)$ (M = Nb, Zr, Hf, Ta) are currently being investigated by our group [24, 25]. Fe and Co are the ferromagnetic species. The composition x, is important in determining the magnetic induction. Fe and Co don't partition uniformly between the nanocrystals and residual matrix, Co shows a preference for the matrix, this has important implications for increasing the Curie temperature of the amorphous phase). Most early data considers x = 0.5 where it has been definitively shown that the FeCo nanocrystals have an ordered B2 structure [27]. APFIM and TEM studies on HITPERM [26] show FeCo nanocrystalline grains to nucleate in the amorphous precursor without the need for Cu as a nucleation agent. Glass-forming elements segregate to the intergranular phase in a fashion similar to other nanocrystalline alloys and act to impede nanocrystal growth. Crystallography and microstructure have been studied by x-ray diffraction, synchrotron x-ray diffraction, x-ray absorption fine structure [276] atom probe field ion microscopy (APFIM) [25], transmission electron microscopy (TEM) and high resolution transmission electron microscopy (HRTEM).

Modifications have been made to the HITPERM composition to explore the effects on crystallization kinetics. Additions of Mo and Ta to the base HITPERM composition have been reported previously [25]. In new alloys Nb was substituted for Zr in some alloys, as well as Si and Al being substituted for B. Differential calorimetry was used to examine crystallization kinetics. Heating rates ranged between 5 and 40 °C/min. The Kissinger analysis was employed to calculate activation energies for primary crystallization. Activation energies of 2.7 – 3.8 eV were determined The base HITPERM composition has an activation energy of 3.8 eV/atom. Increasing electron concentration by substituting Nb or Mo for Zr was seen to decrease the activation energy. Removing Co was observed to decrease the activation energy.

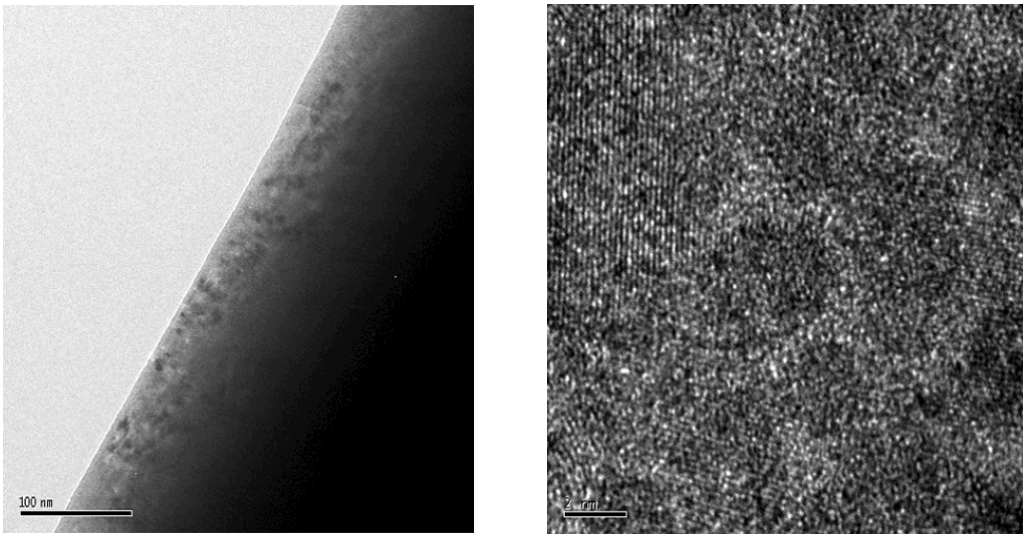


**Fig. 3:** (a) Cross sectional bright-field TEM of as-cast HITPERM near free side. (b) HRTEM image of roll-side of as-cast sample.

TEM and HRTEM studies were performed on as-cast and partially crystallized HITPERM ($Fe_{44.5}Co_{44.5}Zr_7B_4$) samples. The samples were prepared as cross-sectional specimens and examined with a TECNAI F20 200 keV microscope. The as-cast sample was observed to be partially crystalline near the side of the ribbon that solidified farthest from the wheel [Fig. 3(a)]. Nanocrystalline grains were observed to extend 100 to 500 nm into the ribbon. This is in agreement with other observed surface crystallization in this type of nanocrystalline alloy [28]. The side of the ribbon that solidified closest to the wheel was observed to be mostly amorphous, with some regions containing isolated nanocrystalline grains that appeared in HRTEM [Fig. 3(b)]. The as-cast ribbon interior was observed to be amorphous.

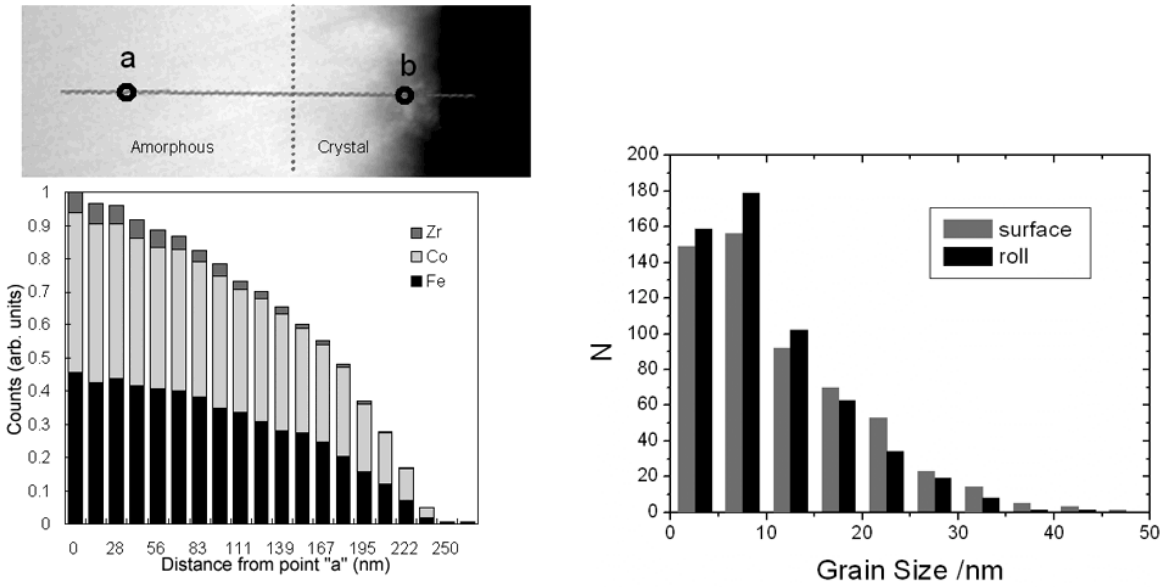


**Fig. 4:** (a) Cross sectional bright-field TEM of as-cast HITPERM near free side. (b) HRTEM image of roll-side of as-cast sample.

To investigate surface crystallization, EDX spectra for Fe, Co and Zr were collected in a line leading from the interior to the free surface side of the ribbon. It was found that the Fe and Co composition remained uniform but the Zr concentration decreased near the surface side of the ribbon [Fig. 4(a)]. It follows that the glass-forming ability of the alloy was reduced near the surface of the alloy, leading to nucleation of grains. Sporadic crystallization near the roll side of the ribbon is attributed to the presence of air gaps between the solidifying ribbon and copper chill roll locally lowering the quench rate. A partially crystallized ribbon sample was prepared by heating an as-cast specimen in a DSC for 1000 s at 500 °C (~ 20% crystallization was inferred from thermomagnetic measurements of similar NANOPERM alloys). Conventional TEM shows a distribution of nanocrystalline grains embedded in an amorphous matrix. Importantly, the grain diameter distribution [Fig. 4(b)] near the roll and free sides of the ribbon are similar. This indicates that nucleation proceeds uniformly across the ribbon.

A typical HITPERM microstructure, after (20 % volume fraction, nanocrystals) nanocrystallization, is shown in Fig. 5(a) with a DF image and electron diffraction pattern shown in Fig. 5(b). Fig. 5(c) shows an EELS spectra for Zr overlaid on a BF TEM image. Zr is seen to be prevalent at the interface between the nanocrystalline grains and the amorphous matrix. A similar observation for NANOPERM was by made by Zhang et al [29] using an APFIM. These observations support a model of Zr collecting in front of the growth interface.

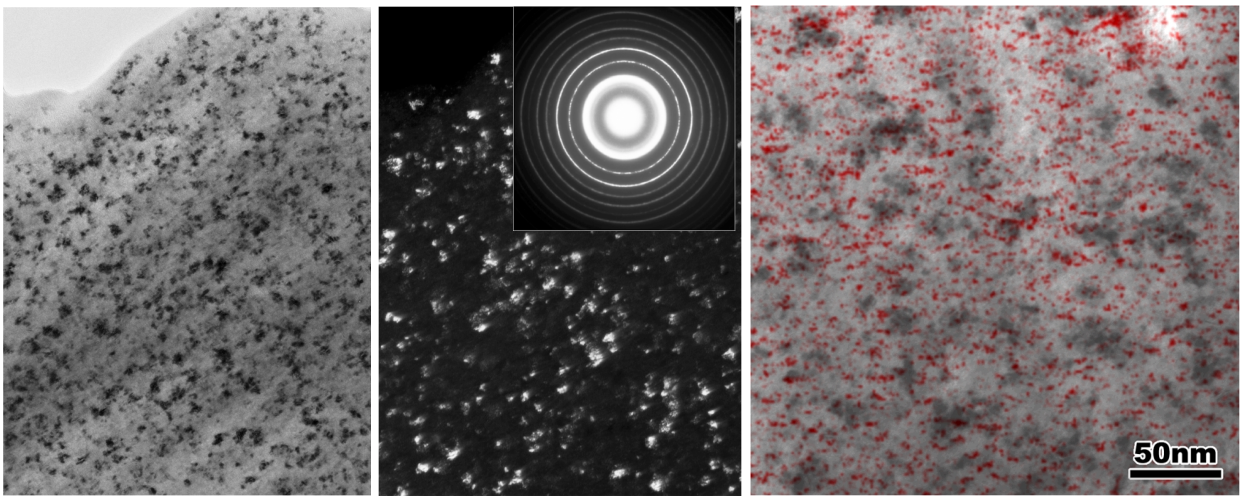


**Fig. 5:** (a) typical HITPERM microstructure, after (20% volume fraction nanocrystals) nanocrystallization, (b) dark field image and electron diffraction pattern of the same and (c) 3-window Zr EELS spectra showing Zr distribution (dark spots) in amorphous matrix.

**References**


1. Y. Yoshizawa, S. Oguma, K. Yamauchi, *J. Appl. Phys*., vol. 64, pp. 6044-6046, (1988).
2. K. Suzuki, A. Makino, N, Kataika, A. Inoue, and T. Masumoto, *Mat'l. Trans. JIM*, **32**, 93, (1991).
3. M. A. Willard, M. -Q. Huang, D. E. Laughlin, M. E. McHenry, J. O. Cross, V. G. Harris and C. Franchetti, J. Appl. Phys. **85**, 4421, (1999).
4. M. E. McHenry, M. A. Willard, and D. E. Laughlin, Prog. Mat. Sci. **44**, 291, (2001).
5. G. Herzer, IEEE Trans. Magn**., 26**,1397, (1990).
6. G. Herzer, Physica Scripta **T49**, 307, (1993).
7. W. A. Johnson, and R. F. Mehl, R.F. Trans. Am. Inst. Miner. (Metall.) Eng. **135**, 227 [1939].
8. M. J. Avrami, M. J. Phys. Chem., **7**, 1103, [1939]; **8**, 212, [1940]; **9**, 177, [1941].
9. A. N. Kolmogorov; Bull Acad. Sci. U.S.S.R., Phys. Ser. **3**, 355, (1937).
10. H. Kissinger, *Anal. Chem*. **29**, 1702, [1957].
11. K. Suzuki, A. Makino, A. –P. Tsai, A. Inoue, and T. Masumoto; Mat. Sci. and Eng. **A179**, 501, (1994).
12. N. Lecaude and J.C. Perron; Mat. Sci. and Eng. **A226-228**, 581-585 (1997); N. Lecaude and J.C. Perron, Proc. of the International Symposium on Metastable, Mechanically Alloyed and Nanocrystalline Materials), Enfield, N.H., USA: Trans Tech Publ., 1998. (Materials Science Forum ; V. 269-272).
13. J. D. Ayers, V. G. Harris, J. A. Sprague, W. T. Elam and H. N. Jones; Acta Mater. **46**,1861 (1998).
14. K. Hono, K. Hiraga, Q. Wang, A. Inoue, and T. Sakurai; Acta Met. Mat. **40**, 2137, [1992].
15. U. Koster and U. Herold, **Glassy Metals I; Topics in Physics 46**; Springer-Verlag, Berlin (1981).
16. F. Luborsky and H. Lieberman, Appl. Phys. Lett. 33, 233, (1978).
17. V. R. V. Ramanan and G. Fish, JAP 53, 2273, (1982).
18. I. W. Donald and H. A. Davies, Phil. Mag. **A42**, 277, (1980).
19. M. Ohnuma, K. Hono, S. Linderoth, J. S. Pedersen, Y. Yoshizawa and H. Onodera, Acta Materialia **48**, 4783, (2000).

20. A. Paoletti and L. Passari, Nuovo Cimento, **32**, 25, (1964).
21. A. Hsiao, M. E. McHenry, D. E. Laughlin, M. J. Kramer, C. Ashe, and T. Okubo, 2002 Intermag (Amsterdam), Invited, to appear in IEEE Trans. Mag. [2002].
22. M. J. Kramer, L. Margulies, R. W. McCallum, *Review of Scientific Instruments*, **70**, 3554, (1999).
23. U. Koster, U. Schunemann, M. Blank-Bewersdorff, S. Brauer, M. Sutton, and G. B. Stephenson, Mat. Sci. and Eng. **A133,** 611, (1991).
24. F. Johnson, P. Hughes, R. Gallagher, D. E. Laughlin, M. E. McHenry, M. A. Willard, V. G. Harris; IEEE Trans. Magn. **37**, 2261, (2001a).
25. F. Johnson, P. Hughes, R. Gallagher, D. E. Laughlin, M. E. McHenry, M. A. Willard, V. G. Harris; IEEE Trans. Magn. **37**, 2261, (2001b).
26. D. H. Ping, Y. Q. Wu, K. Hono K, M. A. Willard, M. E. McHenry, and D. E. Laughlin, Scripta Mater **45**, 781, (2001).
27. M. A. Willard, M. Gingras, M. J. Lee, V. G. Harris, D. E. Laughlin and M. E. McHenry. MRS Res. Symp. Proc. **577**, 469, (1999).
28. J.D. Ayers, J.H. Konnert, P. Dantonio, A. Pattnaik, C.L. Vold, H.N. Jones, J. Mat'l Sci. **30**, 4492, (1995).
29. Y. Zhang, K. Hono, A. Inoue, T. Sakurai, Scripta Mater. **34**, 1705, (1996).

Proceedings of the 2005 IEEE/ASME
International Conference on Advanced Intelligent Mechatronics
Monterey, California, USA, 24-28 July, 2005



# Machining with Flexible Manipulator: Toward Improving Robotic Machining Performance

Hui Zhang, Jianjun Wang, George Zhang, Zhongxue Gan
ABB Inc. Corporate Research
Robotics and Automation Group
2 Waterside Crossing, Windsor, CT 06095

Zengxi Pan, Hongliang Cui, Zhenqi Zhu
Stevens Institute of Technology
Castle Point on Hudson,
Hoboken, NJ 07030

***Abstract*—This paper presents the critical issues and methodologies to improve robotic machining performance with flexile industrial robots. Compared with CNC machines, the stiffness of industrial robots is significantly lower, resulting in unacceptable quality and lower productivity. The problem is treated with a novel methodology that consists of stiffness modeling, real-time deformation compensation for quality and controlled material removal rate for efficiency. Experimental results show that higher productivity as well as better surface accuracy can be achieved, indicating a promising and practical use of industrial robots for machining applications that is not possible at present.**

## I. Introduction

The automotive industry represents the fastest-growing market segment of the aluminum industry, due to the increasing usage of aluminum in cars. The drive behind this is not only to reduce the vehicle weight in order to achieve lower fuel consumption and improved vehicle performance, but also the desire for more sustainable transport and the support from new legislation. Cars produced in 1998, for example, contained on average about 85 Kg of aluminum. By 2005, the automotive industry will be using more than 125 Kg of aluminum per vehicle. It is estimated that aluminum for automotive industry alone will be a 50B$/year market.

Figure 1: Manual cleaning operations in a foundry plant

Most of the automotive aluminum parts start from a casting in a foundry plant. The downstream processes usually include cleaning and pre-machining of the gating system and riser, etc., machining for high tolerance surfaces, painting and assembly. Today, most of the cleaning operations are done manually in an extremely noisy, dusty and unhealthy environment (Figure 1). Therefore, automation for these operations is highly desirable. However, due to the variations and highly irregular shape of the automotive casting parts, solutions based on CNC machining center usually presented a high cost, difficult-to-change capital investment.

To this end, robotics based flexible automation is considered as an ideal solution for its programmability, adaptivity, flexibility and relatively low cost, especially for the fact that industrial robot is already applied to tend foundry machines and transport parts in the process. Nevertheless, the foundry industry has not seen many success stories for such applications and installations. Currently, more than 80% of the application of industrial robots is still limited to the fields of material handling and welding. (Figure 2)

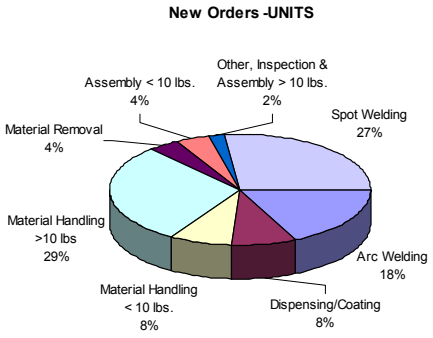


Figure 2: 2003 Robot Applications in North America. A total of 12,367 robots valued at $876.5 million were ordered. When sales to companies outside North America are added in, the total for North American robotics suppliers is 12,881 robots valued at $913 million. NOTE: These numbers include results from North America and Outside North America. Source: Robotic Industries Association

The major hurdle preventing the adoption of robots for material removal processes is the fact that the stiffness of today's industrial robot is much lower than that of a standard CNC machine. The stiffness for a typical articulated robot is usually less than 1 N/μm, while a standard CNC machine center very often has stiffness greater than 50 N/μm.

Most of the existing literature on machining process, such as process force modeling [1, 2], accuracy

improvement [3] and vibration suppression [4] are based on the CNC machine. Research in the field of robotic machining is still focused on accurate off-line programming and calibration [5, 6]. Akbari *etc al* [7] describe a tool angle adjustment method in a grinding application with a small robot. In that case the process force is very small. Matsuoka *etc al* [8] study the characters of an articulated robot in a milling process avoiding large process force by using an end mill with small diameter and high spindle speed. Without the capability of realtime force control, the method to eliminate the force effect on the robotic machining process has not been fully addressed in the research community or in industry.

Machining processes, such as grinding, deburring, polishing, and milling are essential force tasks whose nature requires the end effector to establish physical contact with the environment and exert a process-specific force. The inherent lower stiffness of the robot has presented many challenges to execute material removal applications successfully. The first one is the structure deformation and loss of accuracy due to the required machining force. The predominant cutting action in machining involves shear deformation of the work material to form a chip. The contact between the cutting tool and the workpiece generates significant forces. As a result, a perfect robot program without considering contact and deformation will immediately become flawed as the robot starts to execute the machining task. Unlike multi-axis CNC machine centers, such deformation is coupled and varies even subjected to the same force in different workspace locations. Such coupling results in deformation not only in the direction of reaction force and can generate some counter-intuitive results.

Secondly, the lower stiffness also presents a unique disadvantage for machining of casting parts with complex geometry, which means non-uniform cutting depth and width. As a result, the machining force will vary dramatically, which induces uneven robot deformation. What this means in one example is that the flatness of the machined plane is so inferior that it renders the robotic process unable to meet the minimum requirement.

In general practice, machine tools maximize the material removal rate (MRR) during roughing cycles by applying all of the available spindle power to the machining process. When machines use carbide tools for roughing operations, the available spindle power is usually the limiting factor on MRR. In conventional robot programming and process planning practice, the cutting feed rate is constant even with significant variation of cutting force from part to part, which dictates a conservative cutting feed rate without violating the operational limits. Therefore, it is desirable to maximize MRR and minimize cycle time by optimizing the cutting feed rate based on a programmed spindle load. By optimizing the feed rate in real time, one could compensate for conservative assumptions and process variations to help reduce cycle time. Every part, including the first, is optimized automatically, eliminating the need for manual part program optimization.

To summarize, in this paper, our focus is to address two issues: 1) to improve the robotic machining quality with the low stiffness, low accuracy robot; 2) to improve the robotic machining efficiency by providing real time optimization to maximize material removal rate. If industrial robots could be made to provide end-effector position accuracies under contact situations equal to the end-effector position repeatabilities they already provide, then robotic machining would result in significant cost savings for a lot of potential applications. Moreover, if such applications can be proven to be economically viable and practically reliable, one would expect to witness many success stories in the years to come.

This paper is organized in six sections. Following this introduction section, section II provides the in-depth analysis for the modeling and measurement techniques of the robot stiffness matrix, which would determine the structure deformation when a machining process is taking place. Building on that result, Section III addresses the technique for real-time deformation compensation so that the quality and accuracy of the robotic machining operation can be improved. Section IV presents the real-time optimization methodology for a controlled material removal rate so that the efficiency of such operations can be improved. Experimental results are presented in Sections V followed by a summary in Section VI.

## II. Robot Stiffness Model

As was stated before, one of the focuses for this paper is to improve the robotic machining accuracy by reducing machining force induced deformation. While thermal induced error is the largest error component for CNC machining, motion error due to machining force contributes most of the total machining errors in robots. For example, a 500N cutting force during a milling process will cause a 1 mm position error for a robot instead of a less than 0.01mm error for a CNC machine. In order to achieve higher dimensional accuracy, the deformation due to the interactive force must be compensated.

Since force measurement and subsequent compensation is carried out in 3-D Cartesian space, a stiffness model, which relates the force applied at the robot tool tip to the deformation of the tool tip in Cartesian space, is crucial to realize deformation compensation. The model should be accurate enough for the prediction of robot structure deformation under arbitrary load conditions. At the same time, it needs to be simple enough for real time implementation. Detailed modeling of all the mechanical components and connections will render a model too complicated for real-time control, and difficult for accurate parameter identification.

Industrial robotic systems are designed to achieve high

positioning accuracy and high strength. Elastic properties of the arms are insignificant. Therefore, the dominant contribution factor for a large deflection of the manipulator tip position is the joint compliance, e.g., due to gear transmission elasticity. Modeling of robot stiffness could be reduced to six rotational stiffness coefficients in the joint space. From the control point of view, this model is also easy to implement, since all industrial robot controllers are decoupled to SISO joint control at the servo level. As a result, the joint deformation could be directly compensated on the joint angle references passed to the servo controller. Note here that the axis of the force sensor coincides with the axis of joint 6, the stiffness of the force sensor and its connection flange could be modeled into joint 6. Figure 3 shows the structure of a 6-DOF ABB IRB 6400 robot with black arrows representing the location of compliant joints.

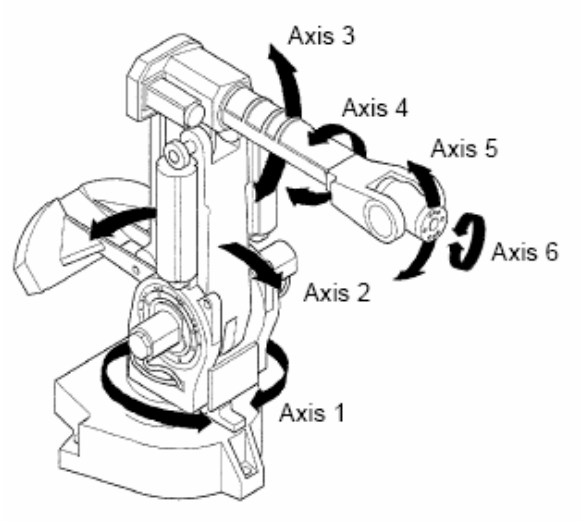


Figure 3: Structure of 6-DOF ABB IRB 6400 manipulator

Next, we will derive the stiffness model in Cartesian space based on joint compliance parameters.

In joint space, the model could be represented as:

$$\tau = K_q \cdot \Delta Q \quad (1)$$

Where: $\tau$ is the torque load on the each joint; $K_q$ is a 6×6 diagonal matrix; $\Delta Q$ is the 6×1 deformation vector.

While in Cartesian space:

$$F = K_x \cdot \Delta X \quad (2)$$

Where $F$ is the 6 D.O.F. force vector, $\Delta X$ is the 6 D.O.F. deformation of the robot in Cartesian space, and $K_x$ is a 6×6 stiffness matrix.

From the definition of the Jacobian matrix, we have:

$$\Delta X = J(Q) \cdot \Delta Q \quad (3)$$

Where $J(Q)$ is the Jacobian matrix of the robot.

At the steady state, after compensating the tool gravity force, the robot joint torques will exactly balance external forces applied on the tool tip. The principle of virtual work gives us:

$$F^T \cdot \Delta X = \tau^T \cdot \Delta Q \quad (4)$$

From (1), (3), (4), we have:

$$K_x = J(Q)^{-T} K_q J(Q)^{-1} \quad (5)$$

For an articulated robot, $K_x$ is not a diagonal matrix and it is configuration dependent. This means that: first, the force and deformation in Cartesian space is coupled, in other words, the force applied in one direction will cause the deformation in all possible directions; second, the stiffness is also a function of robot kinematics $J(Q)$, it changes significantly in the entire workspace. However, even though at different locations, the stiffness matrix will take different values (see Table 1 for one example), these changes can be sufficiently modeled by Eq. (5), with the assumption that $K_q(i,i)$, representing the stiffness of joint $i$, is a constant value. Thus, if $K_q$ can be measured accurately, the deformation of robot TCP under external force at any location in the workspace could be estimated as,

$$\Delta X = J(Q) K_q^{-1} J(Q)^T \cdot F \quad (6)$$

| | | | | | |
|---|---|---|---|---|---|
| 2.86E+02 | -8.78E+02 | 8.12E+02 | 1.49E+06 | 3.01E+05 | 3.97E+05 |
| -8.78E+02 | 2.39E+02 | 5.48E+03 | -1.85E+05 | -5.46E+05 | -4.09E+05 |
| 8.12E+02 | 5.48E+03 | 4.91E+02 | -4.56E+06 | 4.31E+05 | 8.62E+05 |
| 1.49E+06 | -1.85E+05 | -4.56E+06 | 7.53E+07 | 5.85E+08 | -7.62E+08 |
| 3.01E+05 | -5.46E+05 | 4.31E+05 | 5.85E+08 | 1.68E+08 | 3.20E+08 |
| 3.97E+05 | -4.09E+05 | 8.62E+05 | -7.62E+08 | 3.20E+08 | 1.30E+08 |

Table 1 One example of the Cartesian Stiffness Matrix (Units are N/mm, N/rad, N·mm/mm, and N·mm/rad respectively.)

Experimental determination of joint stiffness parameters is critical in fulfilling real-time position compensation. In this model, the joint stiffness is an overall effect contributed by motor, joint link, and gear reduction units. It is not realistic to identify the stiffness parameter of each joint directly by dissembling the robot; the practical method is to measure it in Cartesian space.

Figure 4: Experiment Setup of robot stiffness measurement

To be able to measure small deformations in 3-D space, the end-effector is equipped with a sphere-tip tool shown in Figure 4. The tool tip is set to a fixed point in the workspace, and the manipulator joint values are recorded. A given load in the range of 100N~400N is applied to the tool, causing the sphere-tip to move away from the original point. The original and deformed positions are measured with ROMER, a portable CMM 3-D digitizer, and the 3-DOF translational deformations are calculated. From, Eq. (6), $K_q$ could be solved by least square method. The same experiment was repeated at several different locations in the robot workspace.

The same procedure is taken at different locations in the robot workspace, the deviation of the results is small, which means a set of constant model parameters could model the robot deformation with small error. (Figure 5)

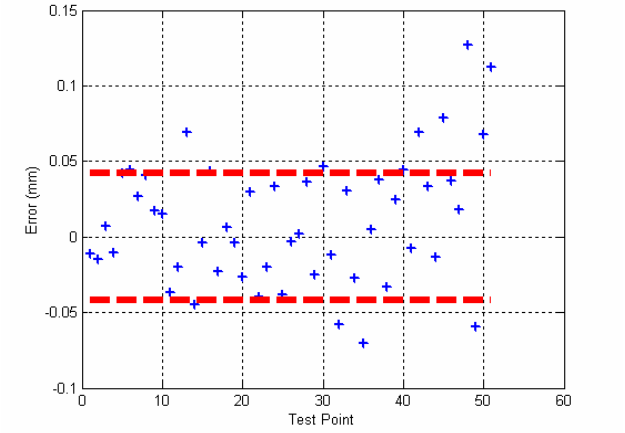


Figure 5: Error of stiffness modeling

## III. Robot Deformation Compensation

The major position error sources in robotic machining process can be classified into two categories, (1) Machining force induced error, and (2) motion error (kinematic and dynamic errors, etc.). The motion error, typically in the range of 0.1 mm, is inherent from the robot position controller and would appear even in non-contact cases. While the machining force in the milling process is typically over several hundreds of Newton, the force-induced error, which could easily go up to 1 mm, is the dominant factor of surface error. Our objective is to estimate and compensate the deformation in real time to improve the overall machining accuracy.

The existing research of robot deformation compensation is focused on gravity compensation, deflection compensation of large flexible manipulators, etc. Not much attention has been paid to the compensation of process force induced robot deformation due to the lack of understanding and model of robot structure stiffness, the lack of real time force information and limited access to the controller of industrial robot.

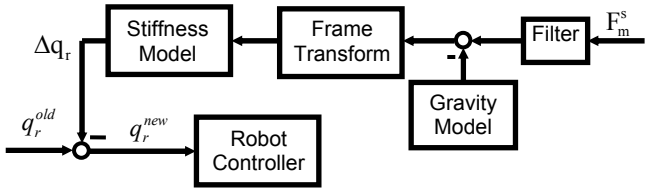


Figure 6: Principle of real-time deformation compensation

The block diagram of real time deformation compensation is shown in Figure 6. After filtering the force sensor noise and compensating the gravity of the spindle and the cutter, the force signal was translated into the robot tool frame. Based on the stiffness model identified before, the deformation due to machining force is calculated in real time and the joint reference for the robot controller is updated.

## IV. Controlled Material Removal Rate

In pre-machining processes, maximum material removal rates are even more important than precision and surface finish for process efficiency. MRR is a measurement of how fast material is removed from a workpiece; it can be calculated by multiplying the cross-sectional area (width of cut times depth of cut) by the linear feed speed of the tool:

$$MRR = w \cdot d \cdot f \qquad (7)$$

Where $w$ is width of cut (mm), $d$ is depth of cut (mm), $f$ is feed speed (mm/s).

Conventionally, feed speed is kept constant in spite of the variation of depth of cut and width of cut during foundry part pre-machining process. Since most foundry parts have irregular shapes and uneven depth of cut, this will introduce a dramatic change of MRR, which would result in a very conservative selection of machining parameters to avoid tool breakage and spindle stall. The concept of MRR control is to dynamically adjust the feed speed to keep MRR constant during the whole machining process. As a result, a much faster feed speed, instead of a conservative feed speed based on maximal depth of cut and width of cut position, could be adopted. (Figure 7)

Since the value of MRR is difficult to measure, the MRR is controlled by regulating the cutting force, which is readily available in real-time from a 6-DOF strain gage force sensor fixed on the robot wrist. Placing the analysis of the material removal process on a quantitative basis, the characterization of cutting force is important for research and development into the modeling, optimization monitoring and control of metal cutting.

The challenges for designing a robust controller for MRR is the fact that cutting process model varies to a large degree depending on the cutting conditions. Efforts for designing an adaptive controller will be presented in a separate paper.

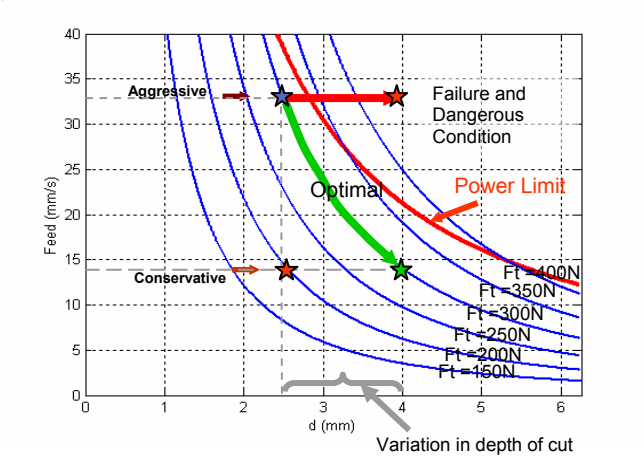


Figure 7: Controlled material removal rate

As the feed speed $f$ is adjusted to regulate the machining force, MRR could be controlled under a specific spindle power limit avoiding tool damage and spindle stall. Also, controlled MRR means predictable tool life, which is very important in manufacturing automation. Figure 8 shows the block scheme of machining force control with controlled material removal rate (CMRR).

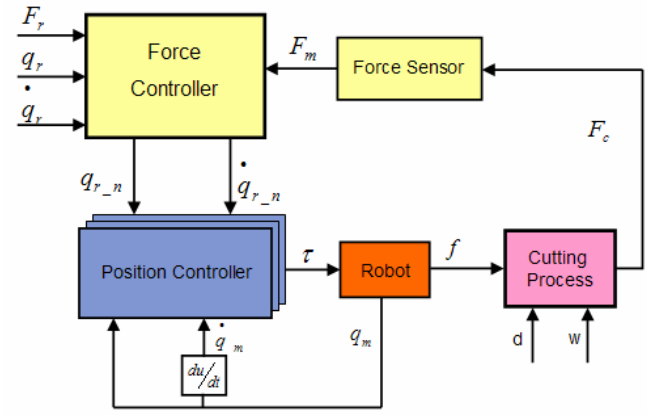


Figure 8: Force control for robotic machining

The structure of cutting force in a milling operation is represented as linear first-order model:

$$F_c = K \cdot w \cdot d \cdot f \frac{1}{\tau_m s + 1} \qquad (8)$$

where $\tau_m$ is the machining process time constant. Since one spindle revolution is required to develop a full chip load, $\tau_m$ is 63% of the time required for a spindle revolution [9]. Since the force control is implicitly implemented, the control loop bandwidth is limited by a position servo control, which is around 10 Hz for an industrial robot. The force process gain may be seen as $\theta = K \cdot w \cdot d$ , which is sensitive to the process inputs. With the proper selection of reference feed speed $f_r$ and reference force $F_r$, a PI controller is adopted to regulate the cutting force $F_c$, while force process gain $\theta$ changes.

## V. Experimental Results

In the previous sections, the robot deformation subject to an arbitrary process force loading is modeled and the model parameter is experimentally measured. With this model, the online deformation scheme is implemented on the robot controller. Secondly, the concept for controlled material removal rate is presented and implemented. In this section, the experimental results are presented to validate the aforementioned schemes.

Figure 9 shows the setup of a milling test. A spindle is fixed on the robot arm and the workpiece is fixed on a steel table. For illustration, a 6063 aluminum block is used for testing purpose.


Figure 9: Experimental setup for robotic milling

Tests on an aluminum block with the depth of cut changed from 2 mm to 3 mm shows, when force control is activated, the cutting force is regulated in spite of the variance of depth of cut. (Figure 10) The milling test of aluminum with variation of width of cut shows similar results.

As a result, the feed speed could always be setup as fast as the limit of spindle power. In a foundry parts milling or deburring process, the robot won't have to move at a very conservative speed to avoid tool breakage or spindle stall. The cycle time decreased by CMMR is typically around 30% to 50% for different workpieces.

In the deformation compensation test of milling an aluminum block, a laser displacement sensor is used to measure the finished surface. The surface error without deformation compensation demonstrates counter-intuitive results; an extra 0.5mm was removed in the middle of the milling path. Conventional wisdom says that a flexible machine would also cut less material due to deformation, since the normal force during cutting will always push the cutter away from the surface and cause negative surface error. However, in the articulated robot structure, the deformation is also determined by the structure Jacobian, in a lot of cases, a less stiff robot could end up cutting more

material than programmed. The coupling of the robot stiffness model explains this phenomenon, the force in feed direction and cutting direction will result in positive surface error in that robot configuration. Since the feed force and the cutting force are the major components in this setup, the overall effect will cut the surface 0.5 mm more than the commanded depth. In our definition, negative surface error means less material was removed than the commanded position. The result after deformation compensation shows a less than 0.1 mm surface error, which is in the range of robot path accuracy. (Figure 11) Further test conducted on the foundry cylinder head workpiece shows that the surface accuracy improved from 0.9mm to 0.3mm, which is below the 0.5mm target accuracy for pre-machining application.

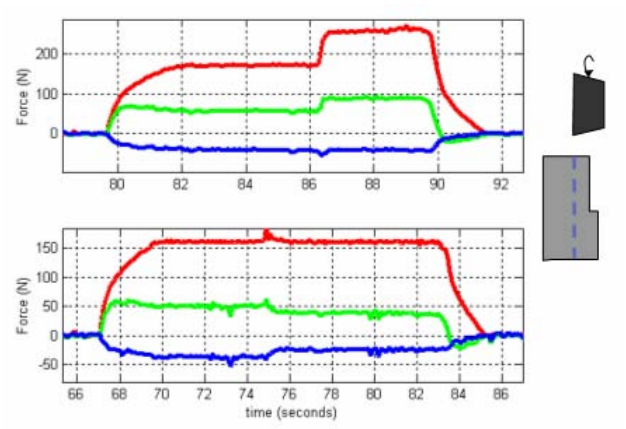


Figure 10: Force control result of variant depth of cut

## VI. SUMMARY

In this paper, critical issues and methodologies to improve robotic machining performance with flexile industrial robots are presented, where previous force control research is expanded into machining application. Practical machining experiments were conducted in the lab to validate the concept and design methodology; both deformation compensation and controlled MRR algorithm are demonstrated in experiments. The experiment results showed great reduction of cycle time, as well as better surface accuracy. The results outline a promising and practical use of industrial robots for machining applications that is not possible at present.

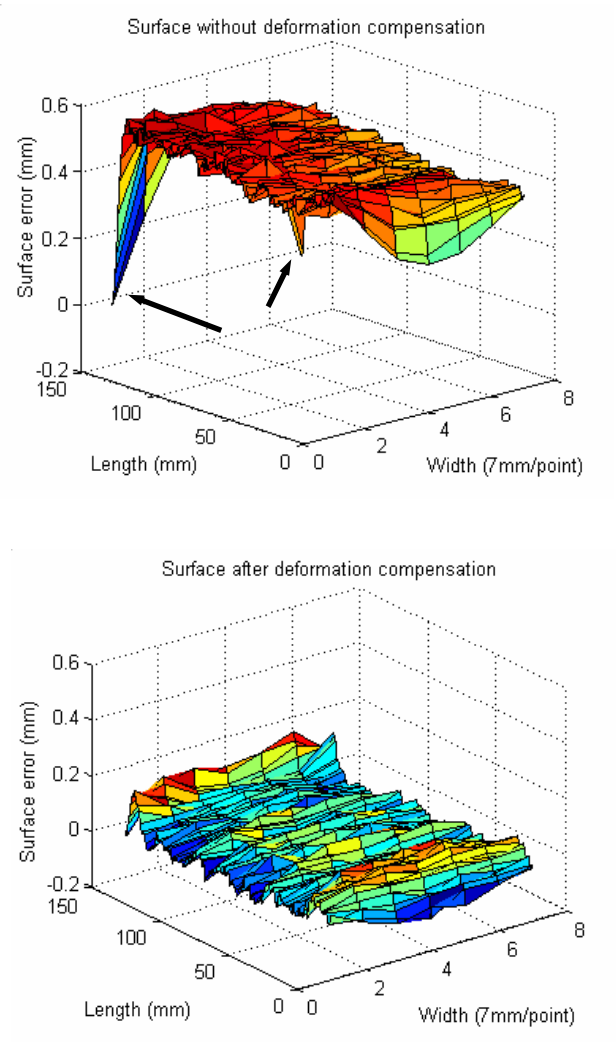


Figure 11: Deformation compensation results

## REFERENCES


[1] Sung I. Kim, Robert G. Landers, A. Galip Ulsoy, 2003, "Robust Machining Force Control with Process Compensation," *Journal of Manufacturing science and engineering,* Vol 125, pp. 423-430

[2] Jeffrey L. Stein, Kunsoo Huh, 2002, "Monitoring Cutting Forces In Turning: A Model-Based Approach," *Journal of Manufacturing science and engineering*, Vol 124, pp. 26-31

[3] Seung-Han Yang, 1996, "Real-time compensation for geometric, thermal, and cutting force induced errors in machine tools," Ph.D. dissertation, *The University of Michigan*

[4] E. Budak, Y. Altintas, 1998, "Analytical Prediction of Chatter Stability Conditions for Multi-Degree of Systems in Milling. Part I: Modeling, Part II: Applications," *Transactions of ASME, Journal of Dynamic Systems, Measurement and Control,* vol.120, pp.22-36

[5] Y.H. Chen, Y.N. Hu, 1999. "Implementation of a Robot System for Sculptured Surface Cutting. Part 1. Rough Machining". *Int. Journal of Advanced Manufacturing Technology*, Vol 15. Pp. 624-629

[6] M. Sallinen, T. Heikkilä, 2000, "Flexible Workobject Localisation for CAD -Based Robotics", Proceedings of SPIE Intelligent Robots and Computer Vision XIX: Algorithms, Techniques, and Active Vision. Boston, USA, 7 – 8 Nov. 2000. USA. Vol. 4197 (2000), pp. 130 - 139

[7] Ali A. Akbari, Shizuichi Higuchi, 2000,"Autonomous Tool Adjustment in Robotic Grinding," *The int. conf. on Precision Engineering(ICoPE)* ,121-126

[8] Shin-ichi Matsuoka, Kazunori Shimizu, Nobuyuki Yamazaki, Yosh inari Oki, 1999, "High-speed end milling of an articulated robot and its characteristics," *Journal of Materials Processing Technology* Volume: 95, Issue: 1-3 pp. 83-89

[9] L.K. Daneshmend, H.A. Pak, 1986, "Model Reference Adaptive Control of Feed Force in Turning," *ASME Journal of Dynamic Systems, Measurement, and Control*, Vol. 108, No. 3, pp. 215-222.

[10] D. E. Whitney, "Force feedback control of manipulator fine motions," *ASME J. Dynamic Syst. Measure. Control*, vol. 99, no. 2, pp. 91–97, 1977