TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

Mikko Huovinen

Large-scale Monitoring Applications in Process Industry

Julkaisu 867 • Publication 867

Tampere 2010

Tampereen teknillinen yliopisto. Julkaisu 867
Tampere University of Technology. Publication 867

Mikko Huovinen

Large-scale Monitoring Applications in Process Industry

Thesis for the degree of Doctor of Technology to be presented with due permission for public examination and criticism in Festia Building, Small Auditorium 1, at Tampere University of Technology, on the 12th of February 2010, at 12 noon.

Tampereen teknillinen yliopisto - Tampere University of Technology
Tampere 2010

ISBN 978-952-15-2304-5 (printed)
ISBN 978-952-15-2337-3 (PDF)
ISSN 1459-2045

ABSTRACT

Modern industrial processes are increasingly complex and managed with limited personnel resources. In order to ease the task of process management new networked communication and monitoring technologies have been developed and implemented in production plants. To a great extent this has been enabled by technological advances in electronics, communications and software techniques. The technological advances have resulted in an increasingly decentralized periphery and improved the possibilities in embedded data processing. On the other hand new integration techniques have resulted in improved integration of previously separate process assets. The integrated process assets and added intelligence form more manageable and compact units which improves the accessibility and usefulness of process data.

The theoretical part of the thesis discusses process monitoring methods and the trend of decentralization in the process industry. Several related methods and technologies are introduced and the evolution of process automation discussed. In the applied part two subprojects and their results are presented. The first subproject concentrates on the development of a large scale process monitoring application. The application was designed to be generic so that the chosen methods don't limit the implementation area to any specific processes and the deployment costs would be reasonable. The application was then deployed in a pulp drying process. The second subproject studies the possibilities of exploiting the features of intelligent field devices in process monitoring by incorporating field device level diagnostics into the developed monitoring concept.

The results of these projects were mixed. The results of the first subproject were very promising and all the participants were pleased with the developed application. The monitoring application learned the typical behaviour of the process and is a good tool for pointing out problem areas in the process. The results of the second subproject were not as good due to lack of interesting events in the process providing excitation for the monitoring system. The problem was a result of "too well" functioning machinery and thus nothing could be done about it. However it could be established that intelligent field devices can be used as a part of developed monitoring concept. The results of the subprojects suggest that such monitoring and performance measurement systems are useful tools but further work is needed, especially in the integration of separate information sources.

ACKNOWLEDGEMENTS

This work has been carried out at the department of Automation Science and Engineering at Tampere University of Technology in 2005-2009. I would like to acknowledge the financial support of Academy of Finland, Neles Oy 30th anniversary fund, Walter Ahlström fund and Wärtsilä fund without which this work would not have been possible.

My deepest gratitude goes to my supervisor, Prof. Pentti Lautala, and co-worker in the EMPRONET project, Yrjö Majanne, who offered me perspective and advice necessary for a starting researcher. I would like to thank the reviewers professor Kari Koskinen and professor Kauko Leiviskä for their valuable comments and patience. Also, I thank the M.Sc. thesis workers and members of the Metso Corporation staff (too many to mention) who participated in the projects. Finally, I would like to thank all my friends and family for their support and the staff at the department of Automation Science and Engineering for creating an enjoyable environment to work in.

Tampere, January 2010

Mikko Huovinen

TABLE OF CONTENTS

Notations

α	Valve position measurement
d	Distance to the nearest cluster center
d()	Distance function
e	Error between process output y and model output y_m
E	Quantization error
mv	Flow through a valve
op	Controller output signal
P1	Actuator pressure
P2	Actuator pressure
Ps	Supply pressure
pv	Measurement or controlled variable of the control loop
Q	Q distribution
R_c	Covariance matrix
T^2	Hotelling distribution
V	Matrix of code vectors
v	Code vector
v_c	Nearest code vector
x_i	Sample
x_i'	Normalized sample
y	Process output
y_m	Model output
Γ	Cluster
Γ_c	Nearest cluster
μC	Microcontroller
μ_j	Arithmetic mean
σ_j	Standard deviation
χ^2	Chi-squared distribution

Abbreviations

A&E	Alarms and Events
ACF	Autocovariance Functions
AEM	Abnormal Event Management
BAPI	Business Application Programming Interface
CD	Cross Direction
CLPA	Control Loop Performance Assessment
CLPM	Control Loop Performance Monitoring
COM	Component Object Model
CPA	Control Performance Assessment
CPU	Central Processing Unit
CRM	Customer Relationship Management

CSI	Control Speed Index
CTI	Control Travel Index
DA	Data Access
DAI	Distributed Artificial Intelligence
DCOM	Distributed Component Object Model
DCS	Distributed Control Systems
DD	Device Description
DDC	Direct Digital Control
DDE	Dynamic Data Exchange
DDL	Device Description Language
DDS	Device Description Services
DIM	Device Information Model
DOM	Device Operation Model
DTM	Device Type Manager
EDDL	Electronic Device Description Language
EMPRONET	Embedded control and monitoring systems in production machine networks
ERP	Enterprise Resource Planning
ESD	Emergency Shutdown
FCS	Field Control System
FDI	Fault Detection and Identification (in the context of process diagnostics)
FDI	Field Device Integration (in the context of integration techniques)
FDT	Field Device Tool
FF	Fieldbus Foundation
FIPA	Foundation for Intelligent Physical Agents
GUI	Graphical User Interface
HART	Highway Addressable Remote Transducer
HCF	HART Communication Foundation
HDA	Historical Data Access
HMI	Human-Machine-Interface
HTTP	Hypertext Transfer Protocol
I/O	Input/Output
IAE	Integral of Absolute Error
ICA	Independent Component Analysis
IEC	International Electrotechnical Commission
IT	Information Technology
JIG	Joint Interest Group
LOCA	Loss of Coolant Accident
MAS	Multi Agent System
MES	Manufacturing Execution System
MPC	Model Predictive Control
MRAS	Model Reference Adaptive Systems
MUST	Multistate Monitoring
NAMUR	An international user association of automation technology in process industries. Originally “Normenarbeitsgemeinschaft für Meß- und Regeltechnik in der chemischen Industrie“

NN	Neural Network
NNMF	Non-Negative Matrix Factorization
OCS	Open Control Systems
OEE	Overall Equipment Efficiency
OI	Oscillation Index
OLE	Object Linking and Embedding
OMG	Object Management Group
OPC	open connectivity via open standards
OPC UA	OPC Unified Architecture
PAS	Process Automation Systems
PC	Personal Computer
PCA	Principal Component Analysis
PCS	Process Control System
PI	Profibus International
PID	Proportional–Integral–Derivative controller
PIDEX	Petroleum Industry Data Exchange
PI&D	Piping and Instrumentation Diagram
PLC	Programmable Logic Controller
PLM	Product Life cycle Management
PLS	Partial least squares
PM	Preventive Maintenance
PNO	Profibus Nutzeorganisation
PR	Prestage coil
QCS	Quality Control System
QTA	Qualitative Trend Analysis
RCM	Reliability Centered Maintenance
RFC	Remote Function Call
RI	Robustness Index
SAC	Stochastic Adaptive Control
SCM	Supply Chain Management
SDG	Signed Digraphs
SDS	Smart Distributed System
SISO	Single-Input-Single-Output
SOA	Service Oriented Architecture
SOAS	Self-Oscillating Adaptive System
SPC	Statistical Process Control
SPS	Spool valve Position Sensor
STR	Self Tuning Regulator
TCI	Tool Calling Interface
TPM	Total Productive Maintenance
TQM	Total Quality Management
VI	Variability Index
VNC	Virtual Network Computing
XML	eXtensible Markup Language
ZVEI	Zentralverband Elektrotechnik- und Elektronikindustrie

1. INTRODUCTION

In recent years process industries have had to improve efficiency and product quality in response to tightening competition. In order to improve efficiency manufacturers have decreased the number of personnel operating and maintaining the processes while at the same time the requirements for higher production rates and quality have increased. Also modern industrial processes are increasingly challenging to manage, e.g. the interactions are typically strong and the amount of buffer capacity small due to reduced inventory and use of recycle streams and heat integration [Tho07]. Thus increasingly large and complex production lines are managed by fewer personnel. This may not be a problem in steady-state operation, but during abnormal situations the personnel often have to operate under considerable strain and perform quick corrective actions based on available process data. The burden of operating personnel is already quite high also in day-to-day operations as it is and besides the routine tasks there is not much time for other activities, such as optimization of processes or work routines. Also, often the information needed to manage the processes is based on years of experience and the know-how is only available in the minds of operators and process engineers which is problematic in many ways. For example training and breaking-in of new operators can be difficult.

In addition the basic set up of the whole delivery chain has changed so that increasingly products are no longer manufactured into stock but just in time to fulfil orders. This is more efficient in terms of bound capital but it also makes the whole delivery chain more sensitive to disturbances as there is limited buffer for disturbance attenuation. This emphasizes the importance of operational reliability. The problem with limited buffer capacity has always been present in power plant environment where changes in power demand have to be met almost immediately and failure to do so may cause a collapse of the power grid. In order to improve reliability in such critical processes monitoring systems that are able to detect incipient faults and equipment wear early on should be utilized. In case of a failure the system should point to the origin of the problem to aid the maintenance activities.

Tighter production and life cycle management demands have also had an effect on the delivery projects of equipment suppliers as new machinery is designed more accurately to minimize investment and operating costs. This means that new production machinery is designed to achieve a specified production level with minimal hardware investment and no designed reserve capacity. This has resulted in larger unit sizes and higher production rates and so has further emphasized the importance of performance and operational reliability. These factors set several demands for competitive machinery such as quick start-ups, high utilization rate, operational reliability, high end-product quality and energy efficiency. Deviations from these result in lower production levels and eventually lower profits.

Another effect of the changing marketplace is the automation and device suppliers' shift of focus from traditional delivery projects towards life cycle services which is an increasingly important part of modern business models. Life cycle management in process industry includes among other things design, start-ups, support services and maintenance but also continuous optimization and process development throughout the product life cycle. Whereas traditionally after delivery of a project the end-users have been responsible for operating, maintaining and optimizing the equipment mostly by themselves, nowadays various partners are increasingly involved in day-to-day operations of the plant [Mil04, LaB05]. An example of this is outsourcing of maintenance and optimization activities. Such partnership agreements require the use of various performance measurements to verify the results. Also tools for remote monitoring are often required as the partners usually have several clients in different locations and the personnel can't be on-site all the time.

These factors have made controlling and managing of whole delivery chains more challenging. Thus it is becoming obvious that new tools are needed for refining process data and providing decision support. On the other hand new technologies open up new possibilities to respond to the challenges. As a result of the new technologies process industries and process equipment suppliers are increasing the intelligence level of process equipment and management tools. Applications such as fault detection, fault diagnostic and performance monitoring systems are used to support process and life cycle management. These are useful for both the equipment suppliers and the end-users [Mil04, Roe07, Har03].

This chapter provides background on the topic. The first part discusses various process monitoring methods, the trend of decentralization and its implications. Then the projects from which the results were obtained from are shortly described. Finally the research problem and the applied part of this work are shortly presented after which the contributions of this thesis are listed.

1.1 Process monitoring and system integration

Recent developments in control systems, electronics, communications and software technologies have opened up new possibilities in the field of process monitoring and diagnostics [Pek04, Mil04, Lab05]. Collecting process data has become easier but still the used process data itself poses some challenges. Whereas previously a major problem with process monitoring used to be limited availability of process data caused by technical limitations, the problem nowadays is what to do with large sets of collected data. The size of the data sets is many times so large that it is easy to lose track of what is the essential and informative part of the mass. The exploitation of process databases is a critical component in successful operation of industrial processes over the long term but until fairly recently rarely had anything been done with them due to the nature of these data sets. The data sets are often enormous, the data is highly correlated and noncausal in nature, the information contained in any one variable may be small due to low signal-to-noise ratios and measurements are often missing on many variables. The data is often information-poor as the processes are typically operated at constant or nearly constant operating points and due to complexity and operating costs it is not usually possible to excite the system in a required mode to get the missing information. It is also typical that relatively few process disturbances or independent process changes routinely occur and the hundreds of measurements on the process variables are only different reflections of these few underlying events [Kou02, Efe09].

Enabled by technical developments the equipment suppliers started to increase the intelligence level of their products by adding embedded intelligence in field devices as early as the 1980's. I.e. they started to deliver intelligent field devices, for example intelligent valve positioners. However, until the fieldbus technologies started to emerge in the 1990's the exploitation of the new features was very limited due to limitations in field device level communications [Peko04]. One of the goals of intelligent field devices is to ease the effort of providing life cycle services, such as maintenance, in an increasingly demanding environment. Typically these "intelligent" functions include simple tasks like event counters and cumulative indicators to indicate the amount of wear but also more sophisticated functions such as self diagnostics and commissioning tools are used [Pek04, Har03, LaB05, Roe07, Mil04]. For example in the case of intelligent valve positioners faults like loss of supply pressure or electronic failure can be detected immediately. Also more incipient and harder-to-detect faults such as stick-slip behaviour can be detected and diagnosed more effectively.

Even though systematic exploitation of intelligent field devices benefits both the equipment suppliers and the end-users, in general the exploitation of this new information seems to be limited at the moment. The process industries can be conservative in their equipment acquisitions but in addition the features of already acquired and commissioned intelligent field devices often remain unexploited in practice. Human resistance against adapting new techniques and working routines are often blamed but also the limitations of current tools are acknowledged.

The process monitoring/diagnostic methods considered in this work can be classified into three general categories which are quantitative model based methods, qualitative model based methods, and process history based methods [Ven03c]. Out of the various process monitoring methods, the research and application of process history based methods such as Principal Component Analysis (PCA) [Pea01, Hot47] and Partial Least Squares [Wol82], seem to have been the most active in recent years [Ven03b]. Even though there are clear advantages in more holistic approach to process monitoring, in general the implemented industrial monitoring applications seem to be limited to smaller process entities (e.g. headbox of a paper machine). There is, however, some work done in the area of plant-wide monitoring applications, see for example [Tho07, Bau08, Tha09, Tho99, Tho01 Haa00, Jia07].

In general, there is not much literature on actual real life industrial applications of diagnostic systems and there seems to be a general lack of overall penetration of diagnostic systems in process industries. From industrial application viewpoint, the majority of monitoring/diagnostic applications in process industries are more concerned with fault detection than detailed analysis and are based on process history based (sometimes also called data based- or data-driven methods) approaches due to relatively easy deployment. Thus, even though quantitative model-based approaches such as observers have made an impact in mechanical and aeronautical engineering applications, they have had very little impact in process industries due to relatively minimal advantage over simple statistical approaches when considering the increase in required deployment effort. Furthermore, model-based approaches have been predominantly restricted to sensor and actuator failures. The impact of qualitative model-based approaches in terms of applications has also been minimal. Many of the academic demonstrations of these models have been on very simplistic systems and implementing them for industrial systems is plagued with problems related to computational complexity and generation of spurious solutions. Like with the analytical approach (quantitative models) most applications of qualitative models have been limited to systems with relatively small amount of inputs, outputs and states as constructing the fault models can require a large amount of effort. Recently graph-based approaches have been researched upon quite extensively. They have been applied in safety studies and tools are being developed for using these types of models in real-time decision making. [Ven03b, Chi01]

The introduction and increasing popularity of intelligent field devices has further promoted the ongoing decentralization in automation systems. As a result the evolution of automation systems is taking the next step from the nowadays popular Distributed Control System (DCS) structure towards the Field Control System (FCS) structure where the functionality is increasingly distributed to the lower levels of the functional hierarchy, i.e. functionality is distributed to individual field devices [Geo04]. Decentralization offers certain benefits but also presents some disadvantages. For example higher flexibility with the plant design and extensions of a plant is gained by the use of independent and reusable modules but the complete system may be more complex and difficult to handle [Par08]. As intelligent field devices possess features enabling more autonomous behaviour, this kind of decentralization can also be seen as transition towards industrial use of autonomous agent technologies even if the standardized agent terminology, communications etc. are not used in the implementation.

One of the problems introduced by high degree of decentralization is the lack of centralized control and global information repository requiring more communication to coordinate all the entities in a system [Par08]. Thus systematic and efficient use of distributed process assets requires integration of the automation system parts into a seamless entity. On the plant-floor level the proliferation of various fieldbuses and devices can make this tricky. Thus device integration techniques like FDT and EDDL have been developed and are gaining popularity [Zuc03, Ben08, Nie04]. Integration is also needed on the higher levels of the functional hierarchy where integration of information systems, such as Manufacturing Execution System (MES) and Enterprise Resource Planning (ERP), is essential for efficient management of various levels of an enterprise. This way the essential management information is available throughout the enterprise in more or less real-time enabling quick decisions based on the available data.

1.2 Research problem and project goals

There were two research problems in this thesis. The first research problem was to define requirements, desirable features, architecture and tools for a generic large scale industrial monitoring system and to test it in a pilot-plant application. The hypothesis was that the set requirements can be met and the main points of the desirable features achieved by using the defined architecture and tools. In the second part of the thesis the developed generic monitoring system was extended to the field device level. The second research problem was to study if and how intelligent field devices can be exploited as a part of the developed monitoring concept. The hypothesis was that the information produced by the intelligent field devices can be better utilized if it is combined with information from the other parts of the monitoring application. The research results were accrued in industrial research projects carried out in real life process environments.

The main project goal in the first part was to create a generic large scale monitoring application based on the requirement definitions and to deploy it in a pulp drying process in a way that satisfies the customer's needs. There are several factors that have to be addressed in development of such large scale applications and the greatest challenges seem to be usability, reliability and economical aspects. Especially the economical feasibility of the application must remain one of the main criteria which is a great challenge as the involved development, set up and maintenance activities are often very time consuming and expensive. In fact, most of the existing monitoring/diagnostic techniques would do poorly on the issue of ease of deployment [Ven03b] and therefore fail in the requirement of economical feasibility. Previously attempts have been made to solve the problems related to economical feasibility for example by using the minimum-variance controller for the evaluation of controller performance [Haa00]. Also the use of default parameters for different generic types of control loops (e.g. liquid flow and temperature control) has been suggested [Tho99]. Recently methods using process data to derive a qualitative model of the propagation path in the form of a causal map [Bau08] and methods combining readily available information about the connectivity of the process with the results from causal measurement-based analysis [Tha09] have been studied.

In order to gain better understanding of the problems related to large scale process monitoring the challenges and desirable features of monitoring systems were studied using literature (e.g. [Das00, Ger98, Jel06, Tho99, Haa00]) and own experiences [Huo04, Hie05, Huo06, Leh06]. As a result a set of requirements and desirable features for a large scale process monitoring application was created and updated during the projects. As the developed application was to be constructed using existing tools where applicable and developing new solutions where necessary, the created requirement and desirable feature set was studied in respect to the available monitoring tools and the needs of the target process. The most appropriate tools were then chosen and applied in the first project.

The generic monitoring application developed in the first phase consists of monitoring modules on several process levels. The highest of these is the plant level where performance indicators representing the performance of the whole machine are used. The next level consists of subprocess level indicators and the lowest level is the control loop level. In this phase device level monitoring was not applied due to unsuitable field devices at the plant. Considerable effort was put into the usability aspect by building a web based reporting portal for remote access.

The chosen tools were process history based as they were the most suitable to accommodate the set of requirements. As the developed application is a commercial product intended for large scale process monitoring, economical feasibility was a crucial criterion for the selection. A modified K-means clustering algorithm [Hie03] was chosen for the tasks of fault detection and operating point identification. Although scaling of inputs and selection of some algorithm parameters is needed, the clustering algorithm meets the economical specifications. The algorithm, once properly set up, is good for detecting abnormal situations but is somewhat lacking in diagnostic properties. Although the algorithm informs about the contribution of each variable to the clustering residual, much know-how and engineering work is often needed to determine the root cause.

Another chosen method was the conditional histogram algorithm [Fri03] which is a versatile tool for process monitoring and is applicable with relatively little work. The method compresses process data sorted according to identified operating points and is therefore also a very useful tool for process development. The critical part for the algorithm is the configuration work as there are several parameters to consider and the refined information provided by a poorly configured application offers little value. In this respect the configuration must often be done on trial-and-error basis but there are tools available to assist the configuration work provided that there is sufficient amount of process data available in the configuration phase. The previously mentioned clustering algorithm can be used for identifying the operating points and by so used as an input generator for the conditional histogram algorithm.

A commercially available control loop level performance monitoring application was also used. The LoopBrowser algorithm [Mar99b] is used for detecting decreased performance in individual control loops, for example oscillations and increase in control error and controller activity. As the algorithm is designed to monitor individual control loops and ignores process interactions, its use as a part of a large scale monitoring application requires manual work or additional automation of diagnostics. However as a proven technique with easy deployment features it was a suitable component for the monitoring application.

In the second part of the thesis the project goal was to find out how various events are observable in the extended monitoring system and to study how this information could be utilized in process monitoring. The applied part of this phase was carried out in a different factory where the control valves were fitted with ND9000-series valve positioners. These measure several variables indicating the performance, usage, wear and overall condition of the valve-positioner package [Met]. Even though the devices produce valuable information and the FDT technology provides access to the field device level data, simultaneous analysis of this data and other process data still requires manual work and therefore more development work is needed on this part. Also the limitations of fieldbus capacity and power supply limitations on the field device level at the moment limit the usefulness of this data. However, the positioners produce valuable information about their performance and overall condition even as a separate information source.

1.3 Project history and the author’s contributions

The presented subprojects (chapters 4 and 5) were a part of the EMPRONET project whose goal was to study the use of embedded control and monitoring systems in production machine networks. The applied part of the thesis started with a subproject where the focus was on developing generic monitoring methods for large scale processes. The development and industrial implementation of the monitoring application had already started in a separate pilot project during 2004 when the first prototype was developed and implemented to a pulp drying process as a master thesis work [Huo04]. At the start of the pilot project there were only limited tools available that would refine raw data from the pulp drying process to a more suitable form. Therefore much of the deeper know-how of how to operate the process was only available in the minds of the plant personnel and new methods for measuring and monitoring the pulp drying process were needed. The goal was to develop a generic performance monitoring and diagnostic application for large scale processes and to implement it to a pulp drying process. The project served also as pilot project for some prototypes of process monitoring modules. Therefore the use of these prototypes and already existing tools in general were favoured where applicable. One of the more general goals was to accumulate know-how and experiences of how to utilize the process monitoring methods. The project of 2004 created a framework for a generic large scale process monitoring application that was used in the first subproject.

The results of the prototype were so promising that the pulp producing company ordered the application to another plant and the first commercial application was set up during 2006. The application was based on the framework developed in the pilot project and was further developed by Metso Automation, Metso paper and Tampere University of Technology in co-operation with the plant personnel. Metso Paper and the plant personnel brought the process knowledge while the other participants were responsible for the technical side. The author and a thesis worker at the Tampere University of Technology were responsible for the implementation of the application using remote connections to the process database server at the plant. Metso Automation was consulted for technical assistance when needed and feedback was given to the research & development team for improving the existing tools. Metso Paper and plant personnel were consulted for technical and process details while developing the system, e.g. what and how to measure, what process measurements were directly available etc. The empirical results were gained mainly through analysing the information produced by the monitoring application and then consulting the plant personnel and experts at Metso Paper concerning the most interesting observations. The analysis was done utilizing remote connections, using both online (real-time) and offline features of the application. Utilizing the online features was mostly monitoring warnings and alarms given by the application. The online interface also gives information regarding the major contributors to the noticed abnormal situations but offline analysis was also often needed. This was done by examining the process database and also by transferring databases of the monitoring tools (conditional histogram and clustering tools) for further analysis. The analysis of conditional histogram database required tools that were deemed too valuable for use outside the Metso Corporation. Therefore these analyses were done by forwarding the database to the system developers and waiting for results.

The author’s contribution to the project involved participating in the design of the monitoring system, compiling a list of requirements and desirable features for a large scale monitoring application, implementing parts of the application, participating in the application set-up and configuration, coordinating the project at the Tampere University of Technology, guiding the thesis worker, outlining possibilities for future development and in some parts analysing the results. Also a master thesis work [Leh06] was completed during the project. The project and related issues are also described in [Huo06, Leh07, Huo07].

The second subproject was a continuation to the previous research projects and further development of the monitoring concept. The motivation was to study exploitation of field device level diagnostics information in the developed monitoring concept and to deepen the understanding of how and when various changes in the monitored process (stock preparation of a paper machine) can be observed in the developed monitoring system. The research data was received from a paper mill where the machine and the hardware are fairly new and modern. The machine in question is one of the first paper machines to utilize intelligent valve positioners extensively throughout the plant so in this respect it is a good place to carry out this research. Altogether there were data available from some 350 intelligent positioners. Due to limited online access to the process database the study was conducted by transferring large sets of process data from the process database and examining this data in offline mode. Interesting findings were then presented to process experts and possible scenarios explaining the observations were analyzed.

The author’s contribution to the second project involved participating in the design of the monitoring system and process data analysis, implementing parts of the monitoring system, participation in setting up and configuring the system, planning data collecting, implementing data analysis routines, participating in the result analysis, coordinating the project in the Tampere University of Technology, guiding the thesis worker and outlining possibilities for future development. A master thesis work [Leh08] was completed also during this project. The project is also described in [Huo09a, Huo09b].

The results from the developed monitoring applications were good although there were also some problems. The application developed in the first part has then become a commercial monitoring system and has received positive feedback from the end-users. Especially the feedback concerning the usability of the system has been good. From the system suppliers point of view the developed application forms a framework which enables relatively fast and cheap delivery of a similar monitoring system for various processes. This is enabled by the generic nature of the chosen methods and developed configuration tools that speed up the deployment phase. The results of the second phase were more limited due to various difficulties. However the project verified the usefulness of field device level measurements and gave some ideas of how to further exploit this information. Further development work under more controlled testing environment would be needed to improve the effectiveness of monitoring.

1.4 Contributions

The contributions of this thesis are listed below:

1) A new application for large scale process monitoring and its implementation on a pulp drying process are presented.
2) Present a set of requirements and desirable features for large scale process monitoring applications.
3) A study concerning the use of intelligent field devices as a part of larger process monitoring system is presented.
4) Review the advanced process management tools used in process industry.
5) Discuss decentralization, its implications and system integration in process industry.

The thesis is organized as follows: Chapter 1 introduces the topic while chapter 2 presents various advanced monitoring methods along with embedded intelligence in modern field devices and the basic concepts of life cycle management (contribution 4). The trends automation systems and their implications on process management are discussed in Chapter 3. As an essential part of the currently ongoing trend of decentralization, fieldbus technologies and state-of-the-art integration techniques are also briefly described (contribution 5). Some of the described techniques are then utilized in chapters 4 and 5 which describe the applied parts of the thesis. Chapter 4 focuses on the first research problem and presents a commercial large scale monitoring application implemented on a pulp drying process (contributions 1 and 2) while chapter 5 describes a project seeking an answer to the second research problem (contribution 3). The results are then analyzed and compared to the research problems and project goals in chapter 6. Chapter 7 provides a summary of this thesis.

2. PROCESS MONITORING

The review of process monitoring methods presented in this chapter forms the fourth contribution of the thesis. The focus in this work was on “more advanced” monitoring methods utilizing inherently available process data and often requiring complex calculations. Therefore more traditional methods using specific health assessment measurements, such as vibration and acoustic emission analysis, are omitted from the scope of this work. These are often associated with prognostics (ability to computationally predict future condition, i.e. the time between very early detection of incipient faults and progression to actual system or component failure states [Hes05]). Detailed review of this area was also left out of this thesis. The reason was that even though some prognostic applications have been reported especially in the field of aviation, the related technology and theory seem to not be mature enough for large scale deployment in process industry where the challenges are quite different. Although component/parts prognostic techniques have been developed and investigated extensively, the techniques such as neural networks, time/stress measurement devices, vibration monitoring, etc. are by their very nature equipment-specific and very few applications have been implemented. Moreover, they are primarily point solutions and are often too expensive and risky for system application [Su99].

Before continuing some definitions used throughout this thesis should be made. In the context of this work a common definition of process management is used: process management is the ensemble of activities of planning and monitoring the performance of a process. Process management is the application of knowledge, skills, tools, techniques and systems to define, visualize, measure, control, report and improve processes with the goal to meet customer requirements profitably [Wik09]. This definition is used especially in the context of business processes [Bec03], but it is also applicable to the management of industrial processes. Life cycle management is similar but focuses more on longer-term planning and includes cost elements throughout a product’s life cycle. There are a number of approaches to the concept of product life cycle depending on the perspective and temporal context. The meaning and importance of the concept varies according to the point of view but both the manufacturer and the buyer benefit from properly implemented life cycle management [Bli03]. The definitions of ‘monitoring systems’ and ‘diagnostic systems’ (alternatively ‘process monitoring’ and ‘process diagnostics’) typically vary depending on the source. In some sources ‘monitoring systems’ include only fault detection, in some sources it includes fault diagnostics also. In this text the terms ‘monitoring system’ and ‘process monitoring’ include all monitoring methods possessing fault detection features regardless of presence of diagnostic properties. The terms ‘diagnostic system’ and ‘process diagnostics’ imply the presence of both fault detection and fault diagnostic features. In the context of monitoring and diagnostics applications ‘system’ and ‘application’ mean essentially the same thing. Also, in the context of field devices the terms ‘intelligent’ and ‘smart’ mean the same thing. The terms ‘preventive’ and ‘predictive’ are also often used ambiguously. In this context ‘predictive maintenance’ is considered to be a subgroup of ‘preventive maintenance'. The use of the term 'predictive' in the context of maintenance implies that more advanced methodologies than in, say for example age-based maintenance, are used for predicting the need for maintenance. Therefore age-based maintenance would fall in the preventive maintenance but not in predictive maintenance category.

2.1 Reliability, product life cycle and maintenance

The increasing development and use of new management tools in process industry is a response to the challenges of modern production where the role of dependability is increasingly important. Dependability is comprised of many factors throughout a product's life cycle. The term dependability is a collective term used to describe availability performance and its constituent factors: reliability performance, maintainability performance and maintenance support performance [Ake06]. These are defined in [Sun88] as follows:

- Availability performance is defined as the ability of an item to perform its required function at a stated instant of time or over stated period of time.
- Reliability performance is defined as the ability of an item to perform a required function under stated conditions for a stated period of time. Reliability depicts the failure sensitivity of an item.
- Maintainability performance is defined as the ability of an item to be retained in or restored to a stated condition within a stated period of time, when the maintenance is performed in accordance with prescribed procedures and resources.
- The concept of maintenance support performance is most often interpreted to only stand for the ability or skills of the maintenance organization, which is only a limited definition. Maintenance support performance is also characterized by administrative and logistic delay, fault detection, spare parts shortage etc.

These are affected by several factors throughout the products life cycle, starting from the design and manufacturing phases. E.g. maintainability is in practice a feature fostered into the item during design/development and cannot be transplanted into an item later [Sun88]. Thus, it is important to effectively manage the whole life cycle of equipment from investment planning and design phases to decommission.

There are a number of approaches to the concept of product life cycle depending on the perspective and temporal context. The concept is quite different in meaning and importance from the point of view of the manufacturer and that of the buyer [Bli03]. As seen by the user or the buyer of the product, there are five phases in a product's life cycle: imagination, definition, realisation, use (or operation), disposal (or recycling). The first three phases are the same from the perspective of the manufacturer and the user, but the last two are different. When the user is using the product, the manufacturer will probably need to provide some kind of support. Later, for various reasons, the manufacturer will stop producing the product. Later still, it will stop supporting the product. The company "retires" the product - gradually reducing support levels, and eventually no longer providing any service. Not only are the last two phases different for the manufacturer and the user, but there is no simple chronological relation between them [Sta04]. Regardless of the perspective life cycle costing provides important information for long-term planning for both parties. Such an analysis includes not only the initial cost of the item but also operation and maintenance costs, replacement costs and many other cost elements. For the manufacturer these would include design, research and development, production, warranty costs, and all other costs incurred in producing and selling the item, amortized over units produced and time. For the buyer costs include initial cost, cost of all replacements, both under warranty and after expiration of warranty, and all other costs associated with use of the item over time. Furthermore, there are again many approaches from the manufacturer's point of view. For example, rather than look at a single product, one can view the life cycle from a broader perspective, with the product life cycle embedded in that of the product line and this, in turn, embedded in the technology life cycle [Bli03].

The word 'product' has many meanings and implications within Product Life cycle Management (PLM). The individual product used by the customer may be just one of an identical batch of thousands, or it may be a unique product. The product may be a successor or derivative of another product. It may be one of a product range or product line, and these may be parts of a product family. Once the product is known, its life cycle can be defined and the structure of the "extended enterprise" that participates in the life cycle management can be defined. This may include contractors, subcontractors, suppliers, partners, supplier's suppliers, customer's suppliers, customer's customers, etc. All of these organizations may be at different locations at any one moment in time, and they may all have different business architectures and owners. PLM manages the product across its life cycle in this extended and often complex environment. Once the organizational structure has been defined, the next component to address includes the activities, tasks and processes - the things that have to be done. After the tasks have been clarified, it is possible to define how to do them and people can be assigned to carry them out as a function of their skills, knowledge and competence. After these issues the information requirements for the product can be defined. This is then used to identify information systems to support the activities of developing and supporting the product. [Sta04]

PLM first appeared in a book in 'Product Lifecycle Management: Paradigm for 21st century Product Realisation' [Sta04]. PLM aims to maximise the value of current and future products for both customers and shareholders. Most important, PLM enables a company to maximize product value over the life cycle and to maximise the value of the product portfolio. PLM is also important because it gives transparency about what is happening over the product life cycle and therefore enables a company to be in control of its products across their life cycle. The benefits of PLM throughout the product life cycle are many, for example getting products to market faster in the beginning-of-life, providing better support for their use during the middle-of-life and managing their end-of-life better. Use of the term PLM implies that the activity of managing products across the life cycle is clearly-defined, well-documented, proactive, and carried out according to a particular design to meet specific objectives of increasing product revenues, reducing product-related costs, and maximising the value of the product portfolio. PLM has a holistic approach to the management of a product which distinguishes it from atomistic activities such as Product Data Management (PDM). In contrast, in the pre-PLM paradigm organizations didn't manage products in a joined-up way across the product life cycle which has several drawbacks. For example, product development and product support were often carried out in different parts of the organization even though they addressed the same products. With PLM, the organization manages the product in a coherent, joined-up way across the life cycle. PLM brings together what was previously separate. [Sta04, Sta07]

PLM is just one of the major business activities of a company and must interface seamlessly with other business activities and systems such as Supply Chain Management (SCM), Customer Relationship Management (CRM) and Enterprise Resource Planning. The interfaces that are required and the shape of the PLM activity in a particular organization, influence, and are influenced, by enterprise and inter-enterprise organisational structures and cultures, human resource issues, project management approaches, information structures and flows, work methodologies and techniques, standards, and information and communications systems. Many informal and formal standards may need to be respected during the activities in the life cycle. Some may be company-specific, some may be industry specific, some may be globally applicable. For example Java 2 Enterprise Edition and Microsoft .NET Framework may be used. SOAP (Simple Object Access Protocol) and XML (eXtensible Markup Language) may be required. There are also industry-specific standards such as the Petroleum Industry Data Exchange (PIDEX) standards. [Sta04]

Out of the life cycle stages of monitored products the operational life is the most relevant one for this research. The bathtub curve is widely used in reliability engineering to depict the reliability of a device or a system in three stages of the operational life [Smi01]. The first part is the decreasing failure rate, known as early failure or infant mortality phase. This can also be considered as commissioning phase. The "dead on arrival" products are caused by defects designed into or built into a product but faulty mountings also increase failure rates. The second part is the constant failure rate, known as random failures. This period is often called "normal life" or "useful life" and is characterized by a low, relatively constant failure rate with failures that are considered to be random cases of "stress exceeding strength". The third part is the increasing failure rate, known as wear-out failures. The wear-out period is characterized by an increasing failure rate with failures that are caused by the "wear and tear" on the product over time. Maintenance moves the product to the left on the failure rate function. The need to identify the optimal time for maintenance is evident from picture 1; one should carry out maintenance actions before the failure rate increases too much, but one should not move to a higher failure rate by early maintenance unless there is a specific reason [Geo].

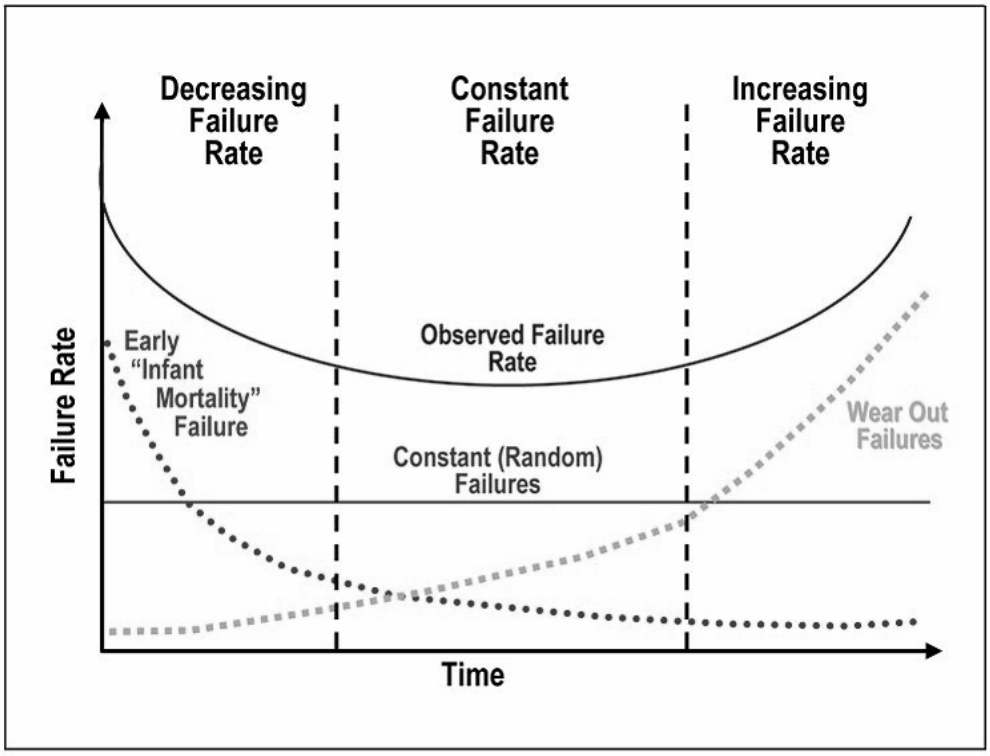

Figure 1 Bathtub curve [Wik08]

There has been some criticism towards the bathtub curve. For example in it's simplest form it does not recognize retirement, i.e. fewer operating hours per calendar time unit for a part in question. Retirement before wear out causes failure rate functions to decrease whereas with constant operating hours the rate would increase. Unreported failures could also cause failure rate functions to decrease, for example if product is replaced by a different product or if the product is repaired by a third party [Geo]. Also the use of bathtub curve on products consisting of several maintainable parts is more complicated. In such cases repair of only the failed part leads to randomized part ages and a pseudo-Poisson process for the system as a whole. More information on the application of bathtub curve can be found for example in [She99] and several books on the subject of reliability engineering.

Even though prognostic methods are not discussed in detail in this thesis, it is good to understand the relationship between predictive prognostics and diagnostics. One way to explore this is to envision an initial fault to failure progression timeline with three points; incipient fault, failure and catastrophic failure. Recent diagnostics technologies are enabling detections to be made at much earlier incipient fault stages than with more traditional diagnostic systems. In order to maximize the benefits of continued operational life of a system or subsystem, component maintenance will often be delayed until the early incipient fault progresses to a more severe state but before an actual failure event occurs. This area between very early detection of incipient faults and progression to actual system or component failure states is the realm of prognostics while diagnostic capabilities have traditionally been applied at or between the initial detection of a failure and complete system catastrophic failure [Hes05]. The approaches to prognostics include e.g. experience-based prognostics (e.g. fitting statistical distributions to data from legacy systems and using this information for future prediction), evolutionary prognostics (gauging the proximity and rate of current condition to known degradation or faults), feature progression and artificial intelligence based prognostics (using artificial intelligence, e.g. neural networks, to predict fault or feature progression), state estimator prognostics (using tracking filters, e.g. Kalman filters, to extract states and using state transition equations to predict their future values) and physics based prognostics (a technically comprehensive modelling approach) [Byi02].

Maintenance is a significant part of product life cycle management in process industry. In general there are two primary types of maintenance. Corrective maintenance comprises of actions (repair or replacement) taken to restore an already failed product or system to an operational state while preventive maintenance actions are intended to increase the length of lifetime and/or reliability. Preventive actions range from minor servicing, such as lubrication, to major overhauls generally requiring shutdown of an operational system. Preventive Maintenance (PM) is further divided in [Bli03] as follows:

- Clock-based maintenance: PM actions are carried at set times.
- Age-based maintenance: PM actions are based on the age of the component.
- Usage-based maintenance: PM actions are based on usage of the product.
- Condition-based maintenance: PM actions are based on the condition of the component. This requires monitoring of variables characterizing the wear process which in turn often have to be estimated. Sometimes also called predictive maintenance.
- Opportunity-based maintenance: A component is maintained as opportunity arises for example due to maintaining of other components in the system.
- Design-out maintenance: This involves carrying out modifications through re-designing the component resulting in better reliability characteristics. [Bli03]

Besides these basic maintenance strategies there are also higher level methodologies related to maintenance such as Reliability Centered Maintenance (RCM) [Aug99], Total Productive Maintenance (TPM) [Sei88] and Total Quality Management (TQM) [Oma04]. These are not discussed further here but they include a variety of factors outside the actual maintenance activities like personnel involvement, training and work practices.

From the point of view of this work and life cycle management the most relevant part of maintenance is condition based maintenance during the operational life but to some degree opportunity-based maintenance strategy is also inherently present in process industries. On one hand successful condition-based maintenance strategy is clearly an improvement over more traditional methods and therefore is an area of interest. On the other hand maintenance in large industrial plants often has to be opportunity-based as there are no equipment redundancies and the equipment must be out of service during the maintenance. This is why maintenance shutdowns, preferably scheduled, are often used to maintain a number of devices even if not all of them require immediate maintenance. Therefore one of the future interests in maintenance management improvement will probably involve improved prognostics to help determine whether the monitored item is likely to remain operable until the next planned shutdown.

As the process industries aspire to step up from traditional maintenance strategies towards more efficient predictive maintenance, new methods and applications are needed for monitoring purposes. For example 'Reliability Centered Maintenance Guide for Facilities and Collateral Equipment' [NAS00] mentions age exploration as an important part of a Reliability Centered Maintenance strategy. Adequate "open and inspect" interval of a process device can vary significantly depending on the process environment and usage of the device. Monitoring and determining when maintenance actions are actually needed is a challenging task, especially in large plants with plenty of different equipment in varying process environments often requiring extensive customizing of the monitoring tools. This is also one of the challenges of this thesis.

Due to the recent technological developments the traditional concepts and terminology of maintenance are changing. Advances in networking, sensor and integration technologies have resulted to a range of Internet-based services provided by the process device manufacturers. Forward-looking manufacturers have started to develop network capabilities and integrate Internet-based services with their products, transforming them into platforms for service delivery. As a result of these new technologies, closer customer relationships will be formed throughout the product's lifetime. "Service" usually means "repair" or "maintenance" today, but might mean something quite different with the opportunities presented by pervasive networking. With the rise of the life cycle management philosophies and new technologies the device suppliers need to consider how factories will be run in the future and the implications of this fundamental transformation to their business models. Machinery and equipment manufacturers need to change their existing business models and internal structure to deal with the new relationship with the customer. This growth in the importance and value of customer relationship is prompting a need to form partnerships with Information Technology (IT) or Internet services suppliers, application service providers and other industries. As a result, manufacturers need to change their concept of Customer Relationship Management from one that effectively ends with the sale to one that promotes an ongoing involvement with the customer and builds on continued service to develop valuable, revenue-generating relationships. The manufacturers will also need to re-evaluate their corporate governance, decision-making control and supply chain strategies in this new environment given the increased high-technology content in their products. By taking advantage of the developing technologies new services can be created and product differentiation achieved. [Mil04]

2.2 Intelligent field devices

The evolution of computers and the false presumptions related to it are illustrated in the article ‘The Future of Information in Intelligent Devices’ [Mil04] by using the movie ‘2001: A Space Odyssey’ as an example. The movie is placed in a spaceship equipped with a computer, HAL, whose intelligence seemed to embody the inevitable evolution of the computer. Instead, reality moved in an entirely different direction. Rather than ever more centralized large-scale intelligence represented by HAL, technology has evolved towards highly distributed, intelligent networked devices. A major reason is the cost benefit ratio of intelligence as reasonable computing power can now be achieved quite cheaply - but a machine like HAL would be extremely expensive. The result is that small computers and micro controllers are being embedded even in the simplest devices to add value and provide additional intelligence [Mil04].

In process automation this development is called distribution of functions or decentralization. Decentralization has had a significant effect on process automation and also has implications for this work as the applied part involves extensive use of intelligent field devices. Intelligent field devices are generally considered to be field devices embedded with a processor capable of data processing. The intelligence means that the device is not only a passive device but besides the usual measurement and control functions there exists additional functionality such as self diagnostics and signal processing. Intelligent device can be for example a measurement or a control device, a distributed I/O or a communication device such as a gateway link [Pek04].

Although first applications of intelligent field devices appeared already in the 1980’s, technically real possibilities for self diagnostics first appeared in the first half of the 1990’s. However back then the devices were not able to inform the users about possible problems through fieldbuses (a bidirectional and digital communications network used in the industry to carry information between field devices and control systems) but instead separate wiring was required. At the same time programmable devices with improved communication capabilities first started to appear on the markets. At this point, however, the device suppliers had their proprietary protocols for communication between higher level applications and the field devices. This resulted in closed architectures and inability to freely choose a device from other manufacturers to be deployed in an existing fieldbus. Also in the 1990’s an open protocol, HART (Highway Addressable Remote Transducer), gained some popularity in the process industry. In the early HART systems commissioning and maintenance of intelligent field devices was done on separate programming equipment with which the status and condition of the devices could be read. However the automation system rarely included a program capable of exploiting these features. Later ever increasing numbers of intelligent field devices has been appearing on the markets. However their use requires support for the various fieldbus standards and therefore the suppliers have separate devices for the most common fieldbuses. [Pek04]

2.2.1 Benefits of intelligent field devices

The benefits of intelligent field devices are many but their efficient utilization requires the use of fieldbuses. The fieldbus used by the devices can be any open digital communication technology. Intelligent devices connected to a fieldbus can collect, process and transmit data to other entities connected to the fieldbus. The use of embedded processor enables among other things pre-processing of measurements and improved control accuracy as the control can be executed locally in the device. Bidirectional communication of the fieldbus enables further functions such as centralized condition monitoring and executing device calibrations through the fieldbus. [Pek04]

Among the more popular recent adoptions in downtime reduction and improving the efficiency of maintenance has been the use of combination of intelligent field devices and asset management software. People in the power industry have for years appreciated the advantages of control systems that provide simple access to a wide variety of intelligent field devices for the purposes of configuration and commissioning, diagnostics and monitoring, calibration management and automated documentation. As the operations professionals are now recognizing the broad opportunities the number of plants implementing such technology is increasing. Asset management software enables quick and easy device configuration in commissioning phase and thereafter promotes the use of predictive maintenance by automating documentation and providing access to diagnostic information and calibration routines. Successful implementation will not only result in significant savings in start-up and maintenance costs but also in improvements in asset performance and process stability over time. [LaB05]

In the commissioning phase the new tools improve cost-efficiency and reduce the number of human errors. The sheer volume of device configuration and initial calibration tasks alone on a greenfield or a capital project often represents savings potential sufficient to justify the relatively modest incremental costs of intelligent devices and asset management software. Often the configuration of transmitters and valves is checked multiple times before startup, and calibration is verified at least once. These necessary activities will become quicker, easier and less error-prone by using asset management software from a workstation rather than by physically visiting and revisiting each device in the plant, using multitude of varying user interfaces. Additionally, all such activities made through asset management software can be automatically documented without any extra labour or risk of errors. [LaB05]

The benefits achievable in later stages of a plant life cycle are more allusive than the benefits of faster commissioning. Generally this requires an appreciation that intelligent devices need to be managed, and that current work practices in the heavy industries need to evolve from routine and corrective maintenance to a culture of predictive maintenance. Even though traditional maintenance strategies such as corrective (maintenance after failure) or clock-, age- or usage-based maintenance (maintenance after certain amount of use whether it is needed it or not) have advantages in simplicity and implementation costs they are not effective in modern manufacturing. In recent years the western countries have spent approximately five times more on maintaining existing assets than on expansions and new facilities. Even with major investments in maintenance, equipment failure is still the biggest contributor to refinery production losses. The benefits of predictive maintenance become even more evident when studying the work orders issued for instrumentation and valve maintenance in the oil and gas industry (in the USA) where half the maintenance work orders result in no action taken. This is a huge (and unfortunately typical) waste of the routine and reactive practices common in the heavy industries today. Evidence is available to support that predictive maintenance is a better approach and that proper management of intelligent field devices is essential for successful implementation [LaB05].

One of the most useful aspects of intelligent field devices is the use of condition and operating environment monitoring. The devices are designed to operate in certain operating conditions, such as limited temperature, and deviations from these conditions can lead to false operation or equipment failure. This can be avoided if the device can monitor its own surroundings and inform operators about critical operating conditions or approaching device failures through a fieldbus. Malfunctions can also occur as a result of an internal fault of a device. This can be detected by using self diagnostics which nowadays commonly appear as embedded functions in intelligent field devices. As the devices constantly provide this information, fault detection and diagnostics can be improved. E.g. device diagnostics with timestamps can help in locating faults as typically a fault will propagate with a small delay in the process complicating the analysis [Pek04]. Efficient condition monitoring utilizing self diagnostics of intelligent field devices promotes the use of predictive maintenance strategy and minimizes the risk of unexpected failures. As a result the useful life of the devices can be extended as they can remain in use beyond conservative clock-, age- and usage-based maintenance periods. Also unnecessary maintenance actions can be avoided as the maintenance can be done based on the actual condition of the device.

It is important to keep in mind that there are two separate groups of beneficiaries - users of the intelligent devices and their manufacturer. In some cases the benefits will coincide but will not always overlap. For example, new intelligence provided by the devices may allow the manufacturer to provide better customer service or improved maintenance, but in other cases, such as extracting usage information in order to make a better product that the user might never purchase, the benefits will accrue primarily to the manufacturer. The users benefit from faster commissioning and predictive metrics that assist maintenance. Data gets refined as deployed systems provide key process insights and allow users to better understand their processes and improve them. This generates a need for even greater access to similar data from other machines that support the process. The ability to “drill down” when additional information is needed will provide a significant resource for making process improvements. Access to local historical data enables users to optimize and tune the local process for better results through better understanding of process fundamentals. This information can help understand the effects of process variables such as differences in materials, environment, personnel and machine wear. Once such differences are analyzed, predictive metrics can be created to help prevent the problem recurring. If knowledge integration is at an adequate level, smooth flow of information in real time from the plant floor up to the business management systems will also make higher level managing easier. For example quarterly reports may be replaced by daily reports. [Mil04]

The manufacturer's benefits include e.g. easier process and device improvements by using data collected by the intelligent devices. In contrast previously the act of getting the process data has often been the most time consuming and costly part of the continuous improvement cycle. Having all the data readily available compresses the cycles in the improvement process enabling agility and responsiveness to market pressures and better implementations. Also useful, although a sensitive area, is the ability to collect information about customer's use patterns. This may provide manufacturers and their sales and support chain with useful marketing and product information via a live connection into the factory floor. This will also lessen the need to conduct customer surveys during future product planning and design. However, a discreet supplier will ensure that they have procedures for protecting the user's proprietary information before going ahead. It must be kept in mind that visionary adopters are the least likely to want to exchange information on their usage with outsiders since they will be seeking a competitive advantage in adopting a proprietary improvement in their manufacturing or processing methods. In some cases data stored in the backend of the device may also provide legal protection to the manufacturer. It will now be easier to prove when a malfunction took place and in many cases why, which will simplify processing warranty claims and even settling issues of legal liability. The Possibility to use remote diagnostics is also a major opportunity. The ability to monitor and service an intelligent device remotely and ahead off technical problems is a clear practical advantage. Perhaps more important than anything else is the boost that the network dimension gives to the device in terms of brand differentiation. In the future the public relations value brought by the manufacturer's association with cutting-edge technology concepts is likely to be a significant value. As regards to product differentiation, integration with enterprise-wide systems sets a product apart augmenting the core functions of the product. [Mil04]

Even though systematic use of intelligent field devices offers several potential benefits, the exploitation of their features still seems to be limited in practice. The process industries can be conservative in their equipment acquisitions and thus still somewhat reluctant to invest in fieldbus based automation and intelligent field devices. Also many of the existing features of already acquired and deployed intelligent field devices remain unused in practice. People in the process industry see human resistance (sticking to old work routines) as one of the main reasons. Furthermore often the organizations have not set anyone responsible for learning and systematically using the diagnostic features. Also the lack of new practices and easy-to-use tools are acknowledged. Thus, further work is needed both on the part of the suppliers and end-users. Better tools and further refined process information are needed for the management of field devices.

2.2.2 Intelligent positioners

Intelligent field devices come in variety of forms such as measurement sensors and operating devices. A common intelligent device in process industry is the valve positioner which is a device used for transforming the control signal of a controller into the corresponding desired position of a control valve. Typical control loop is consistency control where dilution flow is controlled by manipulating the valve opening to achieve a desired consistency in the main flow. Other common positioner applications in the process industry are emergency shutdown (ESD) valves and fan/damper actuators.

As predictive maintenance strategies that in the past appeared to be costly are now being seen as opportunities, many control valve users look for ways to determine which valves need service and more importantly which do not. As a result plant shutdowns involving removal of a full bank of control valves from the line for servicing are nowadays no longer necessary in many cases. Depending on the criticality of the process area, control valves need to be monitored in real-time or at periodic intervals to identify changes in performance early enough which helps avoid unplanned downtime and the associated expenses. Smart valves (valve, actuator and intelligent positioner) using diagnostic technology are helpful in the deployment of such predictive maintenance strategy. [Roe07]

Intelligent positioners use embedded sensors and on-board microprocessors to conduct a multitude of diagnostic tests and manage digital communications to/from host systems. The manufacturers apply the sensors in different ways in an attempt to differentiate their design and depending on the manufacturer there are variations of tests conducted and the information collected and communicated to the host system. Diagnostic information available from most manufacturers includes:

- Final control device tracking parameters for monitoring total (accumulated) travel and the number of reversals (cycles).
- Positioner health parameter tests to provide alerts of memory, microprocessor, and/or sensor problems and to permit operator alarms or, for more severe problems, unit or process shutdown actions.
- Final control device alerts to communicate travel deviations, travel to high/low limits, and/or when preset accumulated travel and cycle targets are exceeded.
- Dynamic error band tests to check hysteresis and deadband-plus-slewing in final control devices. Drive and output signal tests vary the transducer setpoint at a controlled rate and plot final control device operation to determine dynamic performance.
- Pre-defined final control device maintenance and signature tests.

Some digital fieldbus protocols also include pre-defined step tests for checking and/or establishing a final control device performance “signature” (e.g. in the case of FOUNDATION Fieldbus three pre-defined tests). Additionally manufacturers may include a dynamic step test designed to determine the current performance signature of the final control device. The present signature can be compared to the benchmark signature to determine when repairs and/or calibration will be required [Har03].

A typical problem in control valves is excessive static friction which often results into sticking valves having deadband and stick–slip behaviour. Deadband arises when a finite force is needed before the valve stem starts to move while stick–slip behaviour happens when the maximum static friction required to start the movement exceeds the dynamic friction once the movement starts [Tho07]. This kind of behaviour is typical also for fan/damper actuators in combustion processes, especially on waste fuel and coal-fired units where mechanical positioners become subject to sticking as components become soiled. When sticky positioner components are combined with loose or worn actuator linkages the actuator "hunts" for the proper position and causes the associated process variable to be always above or below setpoint. The result of hunting can be observed as poor control, additional wear and tear on moving parts, and/or other unnecessary cost and waste. Intelligent positioner technology can be applied for detecting and analysis of such abnormal events. Several intelligent positioner manufacturers also offer retrofit kits that allow replacement of mechanical fan/damper positioners with intelligent technology [Har03].

Some fairly sophisticated diagnostic and control methods utilizing intelligent positioners have been proposed for control valves. For example as early as mid 1990's diagnostic and control loop tuning methods utilizing smart valve positioners were proposed [Pyö95, Pyö97, Laa97]. These are model based methods requiring identification of model parameters of the positioner-actuator-valve package and thus resulting in some memory and processing power requirements for the device. This poses some problems for distributed on-line implementation especially since the typical way of using the same wiring for data transfer and power supply limits the available power in the field devices. Also the microprocessor technology of the era was quite limited compared to today's technology and the data transfer capabilities from the field devices were more limited than they are nowadays. Some of these limiting factors have been alleviated since then but the limitations in power supply and communications might still limit distributed on-line deployment of such methods. More recently Bartyś has tried using embedded Neural Network methods for modelling of positioners while acknowledging the limited computational power and memory resources of embedded microcontrollers [Bar09].

Better maintenance and safety practices are also required for emergency shutdown (ESD) valves. Partial-stroke testing is often used to ensure that ESD valves within safety systems are capable of operating on demand. Intelligent ESD valves can become part of an automated partial stroke testing solution with automated proof of performance designed to meet regulatory requirements [Har03]. Partial stroke testing means partial movement of the safety valve without interfering with the operation of the process. Successful movement is intended to prove that the valve can be operated. However, the validation of the partial stroke procedure cannot be generated by the software of the valve positioner. Normally the positioner software used in the field is not certified according to IEC 61508. The problem is in the restricted electric power available in field instruments, which limits the available computing power.

This limited computing power poses restrictions on the use of software tools, object-oriented programming, etc. Certified positioners on the market are available, but they are certified only for shutting down the valve in case the electric power at the input terminals of the positioner fails completely. This obstacle can be overcome by using additional instrumentation. For example a positioner with an inductive limit switch integrated into the positioner housing can be used as the signal of this limit switch does not depend on the software [Kar05].

Nuclear plant environments are unique as they have all the problems inherent in conventional plants, but also add in requirements for air operated valve testing, radiation, need for precise control and accident conditions. Traditional positioners have been used for many years but they have limitations on aiding the plant to track the performance of the valve and withstand the harsh operating environment. The combination of high temperatures and high radiation virtually rule out conventional smart positioners in many locations. E.g. the electronics needed for the digital portion of the device will be damaged by relatively low levels of radiation and will not withstand the effects of a LOCA (loss of coolant accident) event. Thus the electronics must be located in a mild environment and some other method adopted to get signals to and from the control valve. Special arrangements are needed to allow the electronics to be moved away from the valve without sacrificing accuracy or exposing the unit to environmental damage. [Mil06]

2.3 Advanced monitoring and diagnostic methods

The standard process controllers are designed to maintain satisfactory operating by compensating the effects of disturbances and changes occurring in processes. While many types of disturbances can be compensated by these controllers, faults occur in the processes which cannot be handled adequately. Such faults in a dynamical system are defined as a deviation of the system structure or the system parameters from the nominal situation. In structural deviations the set of interacting components of the plant or the interface between the plant and the controller are changed by the fault. Examples of such structural changes are blocking of an actuator and loss of a sensor. Parametrical changes, on the other hand, yield deviations of the dynamical input/output properties of the plant. Wear and leakage are some of the typical parametrical changes. In contrast to disturbances and modelling uncertainties, faults are more severe changes whose effect on the plant cannot usually be suppressed by a fixed controller. To be able to handle such abnormal events adequately, the faults need to be detected, diagnosed and removed [Chi01, Bla03]. This section focuses on methods for detection and diagnosis of process faults which enable the removal of faults, i.e. Abnormal Event Management (AEM).

Industrial statistics have shown that even though major catastrophes and disasters from chemical plant failures may be infrequent, minor accidents are very common, occurring on a day to day basis. For example in 1995 it was estimated that the petrochemical industry alone in the United States incurred approximately 20 billion dollars in annual losses due to poor AEM [Nim95]. Thus, people in the process industries view the automation of AEM using intelligent control systems, and thereby providing human operators assistance, as the next major milestone in control systems research and application. Early detection and diagnosis of process faults while the plant is still operating in a controllable region can help avoid abnormal event progression and reduce productivity loss. [Ven03c]

Efficient process monitoring will also promote the transition from traditional and ineffective maintenance strategies towards condition-based maintenance strategies. The goal of stepping up from the traditional maintenance strategies to condition-based maintenance strategies is to reduce maintenance costs with improvements, or at least with no reduction, in reliability. The implementation costs of condition-based maintenance strategy are higher but the rewards are potentially great. The savings come e.g. from avoiding equipment failures, unnecessary shutdowns, unnecessary maintenance work and large stocks of spare parts.

There are several methods developed for process monitoring but the procedure is often the same. The first phases are fault detection and fault identification. The last parts are fault diagnosis and process recovery. Fault detection means determining if a fault has occurred. Early detection of emerging problems may provide invaluable information with which serious process upsets can be avoided. Identification means identifying the most relevant variables for diagnostic purposes. Diagnostic part determines which fault has occurred, i.e. the reason for abnormal observations. In process recovery the fault effects are removed. Not all these have to be implemented or automated. For example diagnostic part can be left for the operators. [Chi01]

A case of a valve with increased stiction can be used as a simple example of diagnosis and recovery. Due to the presence of noise, varying setpoint and the difficulty of maintaining a data base of all possible patterns for a match, automated analysis of the process data can be problematic. Often the challenge in analysis of valve problems is to determine and quantify the type of fault present using only controller output and controlled variable signals. The major difficulty is that the process dynamics (e.g. integration in a level control loop) interfere with the analysis [Tho07]. Several of the analysis methods are reviewed in depth in [Hor07] and compared on a benchmark data set. Also, Hägglund describes a method for detection, estimation and compensation of dead band caused by backlash in control loops in [Häg07]. Gerry and Ruel [Ger01] have reviewed methods for combating stiction on-line including conditional integration in the PI algorithm and the use of a dead zone for the error signal in which no controller action is taken. For a short-term solution the oscillation should disappear in a non-integrating process if the controller is changed to P-only. While the oscillation may not disappear in an integrating process its amplitude may decrease. Also changing the controller gain changes the amplitude and period of the limit cycle oscillation. This observation can also be used for detection of a faulty control valve. More sophisticated control solutions for friction compensation can be found e.g. in [Pii06, Kay00, Häg02, Häg07, Sri05].

Although monitoring and diagnostics functionalities are not generally restricted by the reliability limitations of control operations, the applications suffer from other limitations. With respect to supporting process automation functionalities, systems have been suffering from decisions that were made in the past, e.g. alarm handling and processing is based on a rather simple approach developed when the first digital control systems were introduced. The strong emphasis on reliability and performance requirements in automation systems has resulted in systems that, especially in monitoring operations, are not sufficient any more with larger and more complex processes. Furthermore, devices and control systems as well as their monitoring have been built for continuous steady usage and their operation has been optimised for a steady state. Not enough attention has been focused on operations in change situations in terms of monitoring. These situations have been traditionally treated as special cases handled with special attention. However, this is not the case any more, as systems have become larger and more complex, and there are not enough personnel available on-site to handle the change situations. Fortunately with digital communication and almost fully software-based operation there are no big barriers any more to what functionalities may be developed and thus offered to users. This will result in a whole new set of functionalities and services for users working with process automation. New software technologies have been taken into use in the implementation of automation systems which are becoming increasingly distributed, networked and complex software systems. [Sei06, Pir08]

Monitoring/diagnostic algorithms are often classified into three general categories which are quantitative model based methods, qualitative model based methods, and process history based methods. In model based methods the model is usually developed based on some fundamental understanding of the physics of the process. In quantitative models this understanding is expressed in terms of mathematical functional relationships between the inputs and outputs of the system. In contrast, in qualitative model equations these relationships are expressed in terms of qualitative functions. Process history based methods on the other hand don’t assume a priori knowledge about the model, instead (theoretically) only the availability of large amount of historical process data is assumed. Various methods are then used to transform the data and present it as a priori knowledge to a diagnostic system. These methods can be used on several levels of the process hierarchy, e.g. for monitoring of individual devices or complex plants. [Chi01, Ven03c, Che99, Ger98, Gu04]

2.3.1 Quantitative Model based methods

Relying on an explicit model of the monitored plant, all model based fault detection and identification (FDI) methods (and many of the statistical diagnostic methods) require two steps. The first step generates inconsistencies between the actual and expected behaviour. Such inconsistencies are called residuals and they reflect the potential faults of the system. The second step chooses a decision rule for diagnosis. The check for inconsistency needs some form of redundancy. There are two types of redundancies, hardware redundancy and analytical redundancy. Hardware redundancy requires redundant sensors and the procedure checks for inconsistencies between these sensors measuring the same thing. This is simple but the extra hardware requires more space and results in additional hardware costs and weight. On the other hand, analytical redundancy (also termed functional, inherent or artificial redundancy) is achieved from the functional dependence among the process variables and is usually provided by a set of algebraic or temporal relationships among the states, inputs and the outputs of the system. Analytical redundancy can be divided into direct and temporal redundancy. [Ven03c, Ger98, Che99, Chi01]

A direct redundancy is accomplished from algebraic relationships among different sensor measurements. Such relationships are useful in computing the value of a sensor measurement from measurements of other sensors. This value is then compared with the measured value from that sensor and a discrepancy indicates a possible sensor fault. A temporal redundancy is obtained from differential or difference relationships among different sensor outputs and actuator inputs. With process input and output data, temporal redundancy is useful for sensor and actuator fault detection. The essence of analytical redundancy in fault diagnostics is to check the actual system behaviour against the system model for consistency. Inconsistencies expressed as residuals can be used for detection and isolation purposes. The generation of residuals requires an explicit mathematical model of the system. Either a model derived analytically using first principles or a black-box model obtained empirically may be used. Also, statistical methods are often required for the decision making. [Ven03c, Ger98, Che99, Chi01]

Process faults usually cause changes in the state variables and/or changes in the model parameters which can be seen as increased residuals. Based on the process model, one can estimate the unmeasurable state or model variables by the observed outputs and inputs using state estimation and parameter estimation methods. Kalman filters or observers have been widely used for state estimation [Fra89]. Least squares methods provide a powerful tool by monitoring the model parameter estimates online [Ise89]. Also techniques relying on parity equations for residual generation have been developed [Cho84]. Parity equations are obtained by rearranging or transforming the input-output models which are relatively easy to generate from online process data and easy to use [Ven03c].

There have been efforts to design residual generators capable of generating residuals which are conducive to fault isolation. Two such methods are the directional residual approach [Ger95] and the structural residual approach [Ger90]. The directional residual approach generates residual vectors which are confined to a fault specific direction that allows isolating faults from their locations in the multidimensional residual space. The structured residual approach produces residual vectors so that each residual element responds selectively to a subset of faults. It is required that only a fault-specific subset of the residual components is nonzero in response to a fault. Equivalently, the residuals corresponding to a specific fault will be confined to a subspace of the entire residual space. That allows forming of binary fault signatures, or so called residual structure for fault identification since each residual is completely unaffected by a different subset of faults. [Ven03c, Ger98, Che99, Chi01]

Many of the FDI methods use discrete black-box plant models and assume linearity of the plant. This assumption of linearity can complicate their implementation on nonlinear processes. The main difference between the first principles and the black-box models is that the parameters in the former bear certain physical meanings which can be very useful in diagnostic procedure or controller design while the latter's parameters can't usually be directly associated with known physical variables. The first principles models are obtained based on a physical understanding of the process. Factors like mass, energy and momentum balances as well as constitutive relationships (such as equations of state) may be used in the development of model equations [Ven03c, Chi01].

Model based methods offer good performance in the FDI procedure if the process and possible faults are known well enough. The main advantage of the approach is the ability to incorporate physical understanding of the process into the process monitoring scheme. In other words, when detailed analytical models are available, significant performance advantage over data-driven (process history based) measures can be achieved. However, often the modelled processes are complex and/or nonlinear requiring development of complex models. Sometimes the dynamics of the process are not known well enough to construct a first principle model. These factors will easily result in an inaccurate model or creating the model will become too expensive. This is why model based methods are often limited to only small and well known process entities [Chi01, Ven03c]. Though quantitative model-based approaches such as observers have made an impact in mechanical and aeronautical engineering applications, they have had very little impact in process industries. This might be due to inherent nonlinearity of chemical processes and thus relatively minimal advantage over simple statistical approaches when considering the effort required in model construction. Furthermore model-based approaches have been predominantly restricted to sensor and actuator failures [Ven03b]. The difficulties in developing the models make the feasibility of model based methods poor also in respect to his this work. However, readily available and easily obtainable models might be utilized. E.g. the model for valve-actuator-positioner used in [Pyö97] could possibly be deployed and duplicated for large scale use. Development of first principle models in this research was not seen reasonable.

2.3.2 Qualitative model based methods

The performance of model based methods depends much on the accuracy of the models. As mentioned before, creating the models is often problematic limiting the usefulness of the quantitative methods. Some of these problems can be avoided by using qualitative models as the method does not require development of detailed models. Instead the understanding of process relationships is expressed in terms of qualitative functions. [Ven03a, Chi01]

The development of knowledge-based expert systems was the first attempt to capture knowledge to draw conclusions in a formal methodology. An expert system is a computer program that mimics the cognitive behaviour of a human expert solving problems in a particular domain. It consists of a knowledge base, essentially a large set of if-then-else rules and an inference engine which searches through the knowledge base to derive conclusions from given facts. It was a popular research subject in the 1980's, see e.g. [Ven88, Hen84], but unresolved problems caused the activity to decrease later. One problem is that the size of the deduction tree grows rapidly with the behavioural complexity of the system. Other problem concerns the knowledge presentation. The knowledge in an expert system can be shallow (based on heuristics and expert testimony) or deep (based on structural, behavioural or mathematical models). The problem is that while shallow-knowledge systems are easier to develop, they don't have any understanding of the underlying physics of the system and therefore fail in cases where a new, undefined condition is encountered. On the other hand deep knowledge systems are significantly harder to develop. Ways to deepen the knowledge of an expert system include the use of governing equations based on physical laws to provide a set of constraints on the process variables and causal reasoning via signed digraphs [Ven03a, Chi01].

Signed digraphs (SDG) can be used to represent cause-effect relations or models, see e.g. [Iri79, Che07]. Digraph is a graph with directed arcs between nodes while SDG is a graph in which the directed arcs have a positive or negative sign attached to them. The directed arcs lead from the 'cause' nodes to the 'effect' nodes. Each node in the SDG corresponds to the deviation from the steady state of a variable. SDGs have nodes which represent events or variables and edges which represent the relationship between the nodes. They are much more compact than truth tables, decision tables, or finite state models. SDGs provide a very efficient way of representing qualitative models graphically and they have been the most widely used form of causal knowledge for process fault diagnosis. The drawbacks of the SDG implementations include lack of resolution, potentially long computing times and single fault assumption [Ven03a, Chi01].

Fault trees are used in analyzing the system reliability and safety. Fault tree is a logic tree that propagates primary events or faults to the top level event or a hazard, see e.g. [Sta02, Ves81]. The tree usually has layers of nodes. At each node different logic operations like AND and OR are performed for propagation. A general fault tree analysis consists of system definition, fault tree construction, qualitative evaluation and quantitative evaluation. Before the construction of the fault tree, the analyst should possess sufficient understanding of the system. In fact, a system description is also a part of the analysis documentation. A real-time version of a fault tree model is called symptom tree model. A variation of a symptom tree model called the weighted symptom tree model tries to solve the problem of more than one fault candidate by attaching trained weights to symptom-fault pairs [Ven03a, Chi01].

Qualitative physics or common sense reasoning about physical systems has been an area of major interest to the artificial intelligence community. Qualitative physics knowledge in fault diagnosis has been represented in mainly two ways. The first approach is to derive qualitative equations from the differential equations termed as confluence equations, see e.g. [Sim77]. Essential thing to note here is that qualitative behaviour can be derived even if an accurate mathematical model cannot be developed. Detailed information, such as exact expressions and numerical values, is not required. As an example, consider predicting level in a simple tank process after steady-state has been reached. As the inlet flow rate increases one can predict using qualitative reasoning that level would also increase (at least initially) without knowing the numerical values of cross sectional area of the tank, the outlet and so on. Also a method called precedence ordering [Soy73] has been used to order the variables from the view point of information flow among them. The central idea is that the information flow among these equations is not simultaneous, i.e. a recognition of the presence of asymmetry (partial or complete precedence order among the variables) in the equations. This asymmetry shows the channels of information flow and thus represents causality. The other approach in qualitative physics is the derivation of qualitative behaviour from the ordinary differential equations. These qualitative behaviours for different failures can be used as a knowledge source, see e.g. [Sac88] where piece-wise linear approximations of nonlinear differential equations are examined through the use of a qualitative mathematical reasoning to deduce the qualitative properties of the system. [Ven03a]

Like with the analytical approach (quantitative models) most applications based on qualitative models have been limited to systems with relatively small amount of inputs, outputs and states as constructing the fault models can require a large amount of effort [Chi01]. Also, many of the academic demonstrations of these models have been on very simplistic systems and implementing them for industrial systems is plagued with problems related to computational complexity and generation of spurious solutions. Therefore the impact of qualitative model-based approaches in terms of applications has been minimal. Still, graph-based approaches have been researched upon quite extensively and they have been applied in safety studies and tools are being developed for using these types of models in real-time decision making [Ven03b]. The feasibility of qualitative models is limited to small process entities with well known fault characteristics also in the context of this research. For example further development of fault diagnostics for the valve positioner might be feasible in the future. In more general sense development of easy-to-duplicate qualitative model structures for common processes and devices seems feasible as the models don’t need to be very exact or precise. For example, power plants often have similar process constructions and common features that might be exploited by developing generic models for the most common cases.

2.3.3 Process history based methods

In contrast to the model-based approaches where a priori knowledge about the process is needed, in process history based methods available historical process data is transformed and presented as a priori knowledge to a diagnostic system. This is known as feature extraction. The extraction process can be either qualitative or quantitative in nature. Two of the major methods that extract qualitative history information are the expert systems and trend modelling methods (e.g. Qualitative Trend Analysis, QTA). Methods that extract quantitative information can be broadly classified as non-statistical or statistical methods. Neural networks are an important class of non-statistical classifiers while principal component analysis, partial least squares and statistical pattern classifiers form a major component of statistical feature extraction methods [Ven03b].

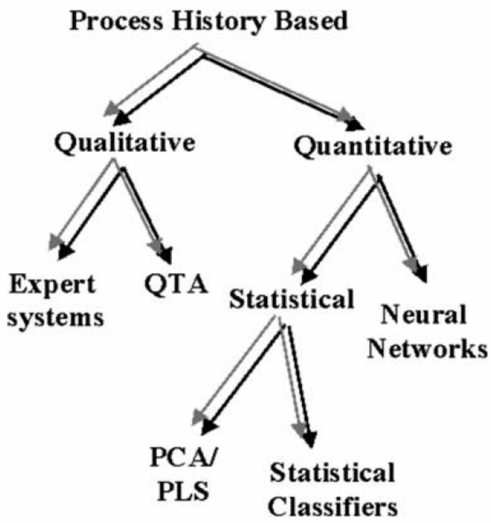

Figure 2 Classification of process history based methods [Ven03b]

Qualitative feature extraction

Rule-based feature extraction has been widely used in expert systems for many applications, e.g. [Ven88, Hen84]. An expert system is generally a very specialized system that solves problems in a narrow domain of expertise [Ven03b]. The main difficulty of using shallow knowledge expert system is in knowledge acquisition. Expert systems are usually better at collecting and archiving cases than in expressing the experience and cases explicitly into production rules. One way to solve this problem is to use machine learning techniques in which knowledge is automatically extracted from data [Chi01]. Therefore expert systems with machine learning capabilities can be considered also in the process history based monitoring category. The main components in an expert system development include knowledge acquisition, choice of knowledge representation, coding of knowledge in a knowledge base, development of inference procedures for diagnostic reasoning and development of input-output interfaces. However, there are limitations in the expert system approach as knowledge-based systems developed from expert rules are very system specific and difficult to update [Ric87]. The advantage though is the ease of development and transparent reasoning [Ven03b].

Alternative approach in qualitative feature extraction is the abstraction of trend information, QTA, see e.g. [Che90]. Qualitative abstraction allows a compact representation of a trend by representing only the significant events. Qualitative trend representation often provides valuable information that facilitates reasoning about the process behaviour for tasks such as diagnosis. In a majority of cases process malfunctions leave a distinct trend in the monitored variables. These distinct trends can be suitably utilized in identifying the underlying abnormality in the process and also be used for efficient data compression [Ven03b]. The basic idea of QTA is to identify predefined primitives from a given time window of time series data and to use the series of identified primitives for analysis. Although QTA provides quick and accurate diagnosis, the increase in computational complexity with the increase in the number of sensors used for diagnostics may prohibit its real-time application for very large-scale plants [Mau05]. QTA approach specializes more on diagnosis than detection and hence might be a useful tool where diagnosis is of importance. Also, QTA approach seems to be robust to routine variations in process operations. However, the time taken to customize and implement might be more than with approaches such as PCA [Ven03b]. This limits the feasibility of QTA applications especially for large scale monitoring.

Quantitative feature extraction

The quantitative approaches essentially formulate the diagnostic problem-solving as a pattern recognition problem. Statistical methods, such as Bayes classifier which uses the density functions of the respective classes, use knowledge of a priori class distributions to perform classification. Approaches like PCA, on the other hand, extract information about major trends in the data using a small number of relevant factors. Neural networks assume a functional form for the decision rule thus parameterizing the classifier. [Ven03b]

The basic concepts behind traditional univariate SPC methods (see e.g. [Mon91]) are still valid today but if applied to multivariate processes, the use of univariate control charts may be misleading. Therefore there is a demand for multivariate statistical techniques, such as PCA and PLS (partial least squares), which are powerful tools capable of compressing data and reducing the dimensionality while retaining essential information and making it easier to analyze than the original data set. They are also able to handle noise and correlation to extract information effectively. The main function of these multivariate statistical techniques is to transform a number of related process variables to a reduced set of uncorrelated variables to ease the task of process monitoring. To establish an “in-control” model for monitoring purposes, χ^2 distribution and/or T^2 statistics are used. With an “in-control” model established based on historical data during normal operation, process monitoring is achieved by comparing the factors against this nominal model. Also so called Q index (squared prediction error) is often used to indicate the error of the reduced model [Chi01, Tep99, Mar02]. Several variations of these methods have been proposed, such as variants able to deal with batch processes and non-linearities and recursive versions for time-varying processes [Ven03b, Lan02, Mar99a].

Fault diagnosis can be cast in a classical statistical pattern recognition framework as it is essentially a classification problem. The Bayes classifier [Fuk72] is an optimal classifier when the classes are Gaussian distributions and the distribution information is available. Distance classifiers, such as the Bayes classifier, use distance metrics to calculate the distance of a given pattern from the means (centers) of various classes and classify the pattern to the closest class. The most basic distance based classifier simply uses the Euclidean distance. This classification is exact only when the classes are spherically symmetrical with the same covariance. More appropriate distance measure for a Gaussian distribution is a scaled form of the Euclidean distance which transforms the problem (ellipsoidal distribution) to a spherically symmetric form before classification is be done. The development of such density functions is a challenging problem and has received considerable attention from researchers. One of the most popular non-parametric techniques for density estimation is Parzen windows [Par62]. Among the process history based approaches, statistical approach seems to have been well studied and applied. The reason for this might be that with the current state-of-art in applications, detection seems to be a bigger concern than detailed diagnosis. Hence, statistical approaches that are easy to build and which do very well on fast detection of abnormal situations have been successful in industrial applications [Ven03b].

It has been shown that a neural network (NN) is a universal approximator capable of approximating any continuous function [Hor89]. It is this universal approximation capability that is the most readily exploited property of the neural networks and has made multilayer perceptron a choice for modelling nonlinear systems and for implementing general-purpose nonlinear controllers [Hun92]. Also the application of neural networks for the problem of fault diagnosis has received interest, see e.g. [Cyb89, Car88]. The methods using NN for the fault detection can be classified based on how the NNs are employed. This division can be made into the following three groups: NN models of the system employed for residues generating (work as process model), NN employed for residues evaluation (work as classifier) and NNs exploiting these two approaches simultaneously [Bah09]. Neural networks used for fault diagnosis can also be classified according to the architecture of the network (such as sigmoidal and radial basis) and the learning strategy (such as supervised and unsupervised learning) [Ven03b, Chi01]. In supervised learning strategies, after choosing a specific topology for the neural network, the network is parameterized in the sense that the problem at hand is reduced to the estimation of the connection weights between network nodes. The connection weights are learned by minimizing the mismatch between the desired output and the actual output of the network. The most popular supervised learning strategy in neural networks has been the back-propagation algorithm and there are a number of papers that address the problem of fault detection and diagnosis using back-propagation neural networks. Most of the work on improving the performance of standard back-propagation neural network for fault diagnosis is based on the idea of explicit feature presentation to the NN, e.g. [Fan93]. Data processing and filtering is shown to lead to significant performance improvement and reduced training time [Ven03b].

Even though supervised NN approaches have several benefits, there are certain problems in applying them. One of the key practical problems is the generalization issue - the ability of a network to perform well in new situations. For accurate modelling of the plant the network needs to be trained with data which covers the entire range of possible network inputs. However, it may be difficult to obtain this data. In such situations it will be important for the network to be able to detect situations in which the inputs fall outside the regions where the network received training data [Hag99, Kar04]. For example in [El-98] a rule based approach was used for dealing with situations where the control signal exceeds their permissible limit. While there is little that can be done to improve the NN performance outside the range of the training data, the ability to interpolate between data points (generalization properties) can be improved by appropriate network training [Hag99]. Another common problem with using neural networks is that they act essentially as a black-box. The information stored in a neural network is a set of numerical weights and connections that provides no direct clues as to how the task is performed or what the relationship is between inputs and outputs. This limits the acceptance of such NNs since in many applications in science and engineering it is demanded to use techniques based on analytical functions that can be understood and validated [Fan05]. The described limitations of supervised NNs explain the reluctance of the industry to adopt neural networks into widespread use.

The NNs using unsupervised estimation techniques are popularly known as self-organizing neural networks as the structure is adaptively determined based on the input to the network. The objective of the unsupervised methods, such as self organizing maps [Koh84], is to give credit for patterns that are similar. The similarity measure is usually a distance measure. Whenever a pattern is seen that is not similar to any of the previously formed classes, a new class is formed and the pattern is retained as the defining pattern for the new class. In these kinds of architectures the crucial elements are the chosen distance metric and the threshold for similarity [Ven03b, Chi01]. Clustering on the other hand is a technique to group samples so as to maximize separability between these groups. Clustering algorithms specify the number of groups (no new classes are formed) and maximize an objective function that is a measure of separability of these groups. This way clustering becomes a well-defined optimization problem. Clustering procedures need a measure for estimating similarity between different data points so that credit can be assigned for similar patterns. Also representative patterns are needed against which the similarity of other patterns can be checked [Ven03b]. The most popular clustering algorithm proposed in the literature is the K-means [Dud73] clustering algorithm. The essential difference between these unsupervised estimation techniques and statistical pattern recognition is that the latter uses a priori class information to design a classifier while the former allocates samples to groups without this information. There are a number of papers on neural network applications in process fault diagnostics and it is not possible to provide an exhaustive review of all the approaches here.

Naturally there are several practical aspects concerning the process history based methods. Modern process plants produce large amounts of process data and often have large volumes of historical data stored in databases. The exploitation of this data offers a huge potential for improvements in the operation of any industrial process over long term but due to the nature of this data the possibilities have not been fully exploited. The success of applied process history based methods is not only strongly dependent on the amount of the process data, but also the data quality is important. The data sets are often enormous but the data is highly correlated and noncausal in nature. The information contained in any one variable is often small due to low signal-to-noise ratios and measurements may be missing on many variables. Many research efforts have recently focused on developing data based models by using latent variable methods such as principal component analysis (PCA) and partial least squares (PLS). These methods address the above problems in a straightforward manner and provide analysis tools that are relatively easy to present and interpret [Kou02].

From industrial application viewpoint, the majority of monitoring/diagnostic applications in process industries are based on process history based approaches and are more concerned with fault detection than detailed analysis. This is due to the fact that process history based approaches are easy to implement, requiring little modelling effort and a priori knowledge. Further, even for processes for which models are available, the models are usually steady-state models as it would require considerable effort to develop dynamic models specialized towards fault diagnostic applications [Ven03b]. These are also the reasons why the chosen methods for the applied part of this thesis were process history based. Especially the features of quantitative feature extraction methods make them attractive alternatives for the applied part of this thesis. The favourable features in the ease of deployment make them suitable for large scale use. For example the valve-actuator-positioner package could be modelled with relative ease using a suitable NN or the NN could be trained to identify the most common faults. However, such black-box approach would not be as useful for diagnostic purposes as the first principle model used in [Pyö97].

Finally, the authors of [Ven03a], [Ven03b] and [Ven03c] want to emphasize that no single method is necessarily adequate to handle all the requirements for a monitoring/diagnostic system. Though all the methods are only as good as the quality of information provided, some methods might better suit the knowledge available than others. Some of the methods can complement one another resulting in better diagnostic systems. Integrating these complementary features is one way to develop hybrid methods that could overcome the limitations of individual solution strategies. Hence, hybrid approaches where different methods work in conjunction to solve parts of the problem are seen as attractive. For example, fault explanation through a causal chain might be best done through the use of digraphs, whereas, fault isolation might be very difficult using digraphs due to the qualitative ambiguity. Analytical model-based methods might be superior in this respect and hence, hybrid methods might provide a general, powerful problem-solving platform [Ven03b]. The benefits of hybrid systems were acknowledged during the research. Thus the monitoring application developed during this thesis is a hybrid system consisting of multiple process history based methods complementing each other. For example the operating point of a process can be identified by clustering and this refined information can then be used by other methods for more accurate analysis.

2.3.4 Plant-wide disturbance detection

Besides the division based on the used monitoring methods, monitoring applications can also be examined from the perspective of the complexity and size of the monitored process. The challenges of monitoring are clearly different when comparing monitoring of e.g. a single control loop and larger process entities. The division between small and large scale processes is of course vague but the differences in the challenge are significant. Monitoring of simple single-input-single-output (SISO) control loops is well established in the process industries but the SISO approach has shortcomings as the control loops are typically interconnected. The basic idea of process control is to divert unwanted variability from key process variables into places capable of accommodating the variability such as buffer tanks. Unfortunately this is not often accomplished and it just appears somewhere else. This is increasingly the case in modern industrial processes which have strong process interactions due to characteristics such as low buffer capacity and increased use of recycle streams and heat integration [Tho07].

A plant-wide approach means that the distribution of a disturbance is mapped out and the location and nature of the cause of the disturbance are determined with a high probability of being right first time. The alternative is a time consuming procedure of testing each control loop in turn until the root cause is found. Some key requirements are detection of the presence of one or more periodic oscillations, detection of non-periodic disturbances and plant upsets and also determination of the locations of the various oscillations/disturbances in the plant and their most likely root causes. [Pau03, Qin98]

Detection

In timescale plant-wide disturbances can be classified into slowly developing (e.g. catalyst degradation or fouling of a heat exchanger), persistent and dynamic (e.g. plant-wide oscillations) and abrupt (e.g. a compressor trip). In the article 'Advances and new directions in plant-wide disturbance detection and diagnosis' [Tho07] the main subdivision of plant-wide disturbance detection is between oscillating and non-oscillating behaviour. An oscillation significant enough to cause a process disturbance can be seen in both in the time domain and as a peak in the frequency domain suggesting that either might be exploited for detection. By contrast the time trends affected by a non-oscillating disturbance often look somehow similar but in a way that is hard to characterize because of the multiple frequencies which are present. The frequency domain readily reveals the similarities in the spectral content, however, and therefore spectra are useful for detection of non-oscillating disturbances. The third class in the division is dynamic disturbances with non-stationary behaviour. For instance, an oscillation may come and go or may change in magnitude. This localization in time suggests that wavelet methods would be best for such cases. [Tho07]

Oscillation detection methods fall into three main classes namely those which use the time domain, those using autocovariance functions (ACF), and spectral peak detection. Most of the methods are off-line and exploit off-line advantages, such as the use of the whole data history to determine a spectrum or autocovariance function. One online method is the oscillation monitor of [Häg95] which has been implemented in the ECA400 PID controller (Alfa Laval Automation) and in the LoopBrowser [Mar99b] control loop performance monitoring application (Metso Automation). The latter is also used in the applied part of this research and is presented in chapter 4. Several methods, e.g. [Häg95, Häg05, Tho97, For99, Mia99, Sal05], are capable of detecting an oscillation one measurement at a time, but due to control loop interactions more is needed for plant-wide detection than the detection of oscillations in individual control loops. Recognition that an oscillation in one measurement is the same as the oscillation in another measurement is needed, even though the shape of the waveform may differ and when interferences such as other oscillations are present [Tho07].

Persistent non-oscillatory disturbances are generally characterized by their spectra which may have broad-band features or multiple spectral peaks. The plant-wide detection problem not only requires a suitable distance measure by which to detect similarity, but also determination and visualization of clusters of measurements with similar spectra. Spectral decomposition methods have been used to distinguish significant spectral features from broad-band noise that spreads all across the spectrum [Tho07]. Decomposition methods include principal component analysis (PCA) [Tho02], independent component analysis (ICA) [Xia05] and non-negative matrix factorization (NNMF) [Tan07].

Diagnosis

After detection the disturbance must be diagnosed. The main distinction of data-driven root cause diagnosis is between linear and non-linear root causes. Alternative approaches use process insights and knowledge. The diagnosis problem consists of two parts. First the root cause of each plant-wide disturbance should be distinguished from the secondary propagated disturbances which will be solved without any further work when the root cause is addressed. The second stage is testing of the candidate root cause loop to confirm the diagnosis. [Tho07]

Typical non-linear root causes of plant-wide disturbances include control valves with excessive static friction, on-off and split-range control, sensor faults, process non-linearities and hydrodynamic instabilities such as slugging flows. Sustained limit cycles are common in control loops having non-linearities, for example the stop–start nature of flow from a funnel feeding molten steel into a rolling mill. Valve diagnostic methods can be seen as a separate subdivision of root cause diagnosis for non-linear causes. [Tho07]

A non-linear time series means a time series that was generated as the output of a non-linear system. Also a distinctive characteristic is the presence of phase coupling between different frequency bands. Non-linear time series analysis concepts are covered e.g. in the book 'Nonlinear time series analysis' [Kan97]. For example surrogate data testing is used to detect non-linearity. In surrogate data testing a new data set is created having the same power spectrum as the time series under test but with the phase coupling removed by randomization of the phase. A key property of the test time series is then compared to that of its surrogates and non-linearity is diagnosed if the property is significantly different in the test time series. To make better use of phase information in signals also the bispectrum and the related bicoherence have been used to detect the presence of non-linearity in process data, see e.g. [Cho04a, Cho04b].

Root cause diagnosis based on non-linearity has been reported on the assumption that the measurement with the highest non-linearity is closest to the root cause [Tho05, Tho03, Zan05]. The assumption that the non-linearity is strongest in the time trends of measurements nearest to the source of a disturbance is based on disturbance propagation. The plant acts as a mechanical filter and as the limit cycle propagates to other variables the waveforms generally become more sinusoidal and more linear because plant dynamics destroys the phase coupling and removes the spectral harmonics which characterize a limit cycle oscillation. Empirically non-linearity measures do very well in isolation of non-linear root causes. However, a full theoretical analysis is missing at present. The waveform in a limit cycle is periodic but non-sinusoidal and therefore has harmonics which can be used in detecting non-linearity. However it is not always the case that the time trend with the largest harmonic content is the root cause because the action of a control loop may split the harmonic content of an incoming disturbance between the manipulated variable and the controlled variable [Tho07].

After non-linearity the most common root causes are poor controller tuning, controller interaction and structural problems involving recycles, i.e. linear root causes. The detection of poorly tuned SISO loops is routine using commercial Control Loop Performance Assessment (CLPA) tools, but determining whether an oscillation is generated within the control loop or originates from an external source has not yet been solved satisfactorily using only signal analysis of routine operating data [Tho07]. Promising approaches to date require some knowledge of the transfer function, e.g. [Xia03]. There has been little academic work to address the diagnosis of controller interactions and structural problems using only data from routine process operations. Some progress is being made by cause and effect analysis of the process signals using a quantity called transfer entropy which is sensitive to directionality to find the origin of a disturbance, see e.g. [Bau05, Bau07, Bau04, Sch00]. The outcome of the analysis is a qualitative process model showing the causal relationships between variables [Tho07].

The methods using process insights and knowledge can also be seen as complementary, rather than just alternative, methods to data-driven methods. After all qualitative process information is implicitly used in diagnosis when an engineer considers the results from a data-driven analysis. An interesting possibility is to capture and make automated use of such information resulting in a qualitative model. Qualitative models include signed digraphs (SDG) [Mau04, Sri06, Ven03a], Multilevel Flow Modelling [Pet00] and Bayesian belief networks [Wei05]. Enhanced diagnostics using signal-based analysis if a qualitative model is available has been shown in [Chi03] and [Lee03]. Thornhill and Horch [Tho07] believe that qualitative models of processes will in the future become almost as readily available as the historical data. The technology for generating such models is already in place in Computer Aided Engineering tools such as ComosPT (Innotec) and Intools (Intergraph). XML has recently started to be used in such intelligent PI&D (Piping and Instrumentation Diagram) tools for export and exchange of process drawings and PI&Ds. This development offers new possibilities for qualitative modelling and large scale monitoring as the plant topology in a process diagram can now be exported into a vendor independent XML-based data format, giving a portable text file that describes all relevant items, their properties, the connections between them and the directions of those connections [Fed05, Tha09]. Especially recently published methods using process data to derive a qualitative model of the propagation path in the form of a causal map [Bau08] and methods combining readily available information about the connectivity of the process with the results from causal measurement-based analysis to reduce spurious solutions [Tha09] seem interesting.

3. SYSTEM INTEGRATION

The introduction of decentralized periphery has changed factory automation and stamps today's plant structures. Besides the reduction of the wiring expenditure, it is the subdivision of the whole system into small manageable applications that has lead to reduction in expenses compared to a construction with central periphery. Big and difficult to maintain control programs in the central units give way to smaller, decentralized applications in intelligent field devices [Sch03]. Such decentralized periphery requires bidirectional communications and integration of process assets which have a central role in gaining full potential from decentralized automation systems.

The communications in control systems have evolved from analog signals to proprietary fieldbus systems with proprietary field devices, and more recently towards open systems utilizing open fieldbuses which enable direct communication between devices from different suppliers [Ron99]. The ability to connect to intelligent devices has become a necessity but this poses a dilemma for the equipment manufacturers. Providing an “open platform” system for user access exposes them to potential support problems, increased support costs and possibly even unhappy customers. If the device took extensive customer support to use, the company would never make any money and the customer would never recommend the company and their products to others. On the other hand providing a secure method of analyzing machine metrics can lead to higher user satisfaction and more sales [Mil04].

There has been a lot of work done in the area of communications and system integration by several communities in the form of interface standards. However there are still lots of different fieldbus standards and communication protocols for connecting field devices to control systems requiring additional layers of integration. E.g. Microsoft has done much work to standardize the way systems communicate by providing layers of network infrastructures that let systems share data in a secure way. This will foster the ability of systems to share information with each other while limiting the access via standard security measures [Mil04]. These issues are discussed in this chapter and form the fifth contribution of the thesis.

3.1 Communications and hierarchy in automation systems

The automation technology has evolved from mechanical and pneumatic systems to analog electronic systems and later to various generations of digital computer based systems. Although still is use today, the limitations of analog technologies, specifically the one-way communication, make it unsuitable for modern process control. A significant milestone in process data transfer was reached when open HART technology was introduced. In HART technology a digital message is sent “on top” of traditional analog signal so that the analog information is not disturbed. It was thought that this was only an interim period in a move towards fully digital fieldbuses but the technology still has a significant position in the process industry [Ron99]. Besides the HART fieldbus, Profibus and FOUNDATION Fieldbus are now the most common fieldbus techniques in process automation. Wireless communications, nowadays popular in personal computer use, are up-and-coming techniques also in the process industry. As a result the ISA100 committee (part of the Instrumentation, Systems, and Automation Society, ISA) was formed in 2005 to establish standards and related information that will define procedures for implementing wireless systems in the automation and control environment with a focus on the field level.

Fieldbus is a general concept used to describe a bidirectional and digital communications network that carries information between field devices and control systems. It is a data-communication channel which has a transmission path, a topology and a data transfer protocol. A modern fieldbus is fully digital system in which a large part of the functionality is transferred to the field devices themselves, i.e. the functionality is distributed. At the same time the amount of data transfer has grown on a new level. Compared to the traditional wiring and analog data transfer the users of a fieldbus have to have a protocol defining how the resources of the shared transmission path are shared. There are several proprietary protocols and the industry has long wanted a unified standard. Use of open standard fieldbuses guarantees interoperability of devices from different manufacturers and so enables their use in the same system. Fieldbuses can be divided into sensor buses (e.g. CAN, ASI), device buses (e.g. DeviceNet, Profibus DP) and process buses (FOUNDATION Fieldbus, Profibus PA). [Ron99]

The interconnected information systems and process assets in process control form a functional hierarchy which can be seen as an increase of comprehensiveness of information when climbing the hierarchy levels. The higher hierarchy level possesses the essential information from the lower level and gives instructions to the lower level. Going down the hierarchy the information and instructions become ever more detailed and the timeframes shorter. There are several hierarchy models used for describing data transfer and information content in process control, although these are somewhat vague and not well established [Ron99]. The joint committee of the American National Standards Institute (ANSI) and the ISA, provides standard models and terminology for inter- and intramanufacturing enterprise integration in the ANSI/ISA-95 standard. The model of ISA95 standard defines hierarchical levels at which decisions are made and addresses the interfaces between the levels. It is based on the Purdue Reference Model for Computer Integrated Manufacturing. In the standard the functional hierarchy levels numbering starts from level 0, which is the production process itself and ends in level 4, Business logistics [Whi06, ISA99].

Business Planning and Logistics level (4)

The highest level of the hierarchy is Business Planning and logistics level. This level is sometimes also referred as Global Management Level or Enterprise Level and is mainly used by administrative information- and production management systems utilizing the Internet. Level 4 activities include business and logistics functions, such as collecting and maintaining large amounts of data, e.g available inventories, inventory usage, schedules etc. Activities also include optimizations of scheduling, preventive/predictive maintenance and inventory. [Whi06, ISA99, Ron99]

There are several software packages available for managing information flows. ERP software systems are designed to improve and unify the functions of an enterprise, the goal is to integrate modules of an organization into a unified system using a unified database, see e.g. [Kla00]. ERP purports to support all business functions of an enterprise, especially procurement, material management, production, logistics, maintenance, sales, distribution, financial accounting, asset management, cash management, controlling, strategic planning, and quality management. ERP also often supports industry specific functions like patient management in hospitals or student administration at universities. The concept of ERP is somewhat vague, and some functionalities are sometimes excluded.

Manufacturing Operations Management level (3)

The next level is the Manufacturing Operations Management level, sometimes also called Management Level or Plant Level. The activities include for example reporting on area production including variable manufacturing costs, collecting and maintaining area data (e.g. production, inventory, manpower, raw materials, spare parts, energy consumption), performing off-line analysis, personnel functions such as work period statistics, establishing detailed production schedules, optimizing individual production areas while carrying out the production schedules of the level 4 functions and modifying production schedules to compensate for plant production interruptions. [Whi06, ISA99]

Manufacturing execution systems (MES) are located on this level between the enterprise level (ERP) and the Process Control System (PCS). MES is a production floor control system used for manufacturing operations. It is the bridge connecting ERP and PCS, and the key of integrated automation system based on ERP/MES/PCS concept, see e.g. [Qiu04, Cha07, Wan07]. MES optimizes the whole production process, collects large amounts of real-time data from the production process and can handle unexpected events in real-time. With the bidirectional communication of the ERP and PCS levels, MES gains corresponding data from the two layers and feeds back processing results and production instructions in order to realize productive plan dispatch and management. MES includes functions like resource allocation, operations scheduling, production dispatching, document control, data collection, labour management, quality management, performance analysis, process management, product tracking, intelligent maintenance etc. The scale of the implemented MES depends on the software supplier.

Process Control and Field levels (2 and 1)

Levels 1 and 2 are responsible for the actual control and operating of the plant. Level 2 consists of monitoring, supervisory control and automated control of the production process [Whi06, ISA99]. The process Control Level connects control rooms, process stations, data management- and engineering stations. Human-machine-interfaces (HMI) are deployed using graphical presentations of process parts and measurements. The operator can observe the course of a process through trends and alarm displays. PLCs (Programmable Logic Controller) or process computers control and optimize processes in the control level. Feedback control, the definition of optimal setpoints, logical sequences and interlockings are executed on this level.

Level 1 is responsible for sensing the production process and manipulating the production process. Sensors are used for measuring the process variables and actuators perform the commands from the control level. Field level is also the part of an automation system that is responsible for transferring data between the automation system and field devices through fieldbuses. [Whi06, ISA99, Ron99]

3.2 Decentralization

The purpose of automation, among other things, is to reduce the number of routine tasks done by humans in processes where deficiencies of humans may cause mistakes. This can be done by replacing human operators where the tasks can be modelled or described by algorithms, i.e. the interface between human operators and machines is moved higher up in the hierarchy. On the other hand there is a serious effort in trying to lower the hierarchy, i.e. to move tasks done by automation to the field devices near the process. This has been enabled by the developments in communications but also in the size and cost of embedded electronics [Ron99, Pyy07]. As a result nowadays various functions are embedded in individual field devices and more and more of the automation systems’ functions can be decomposed and distributed to the lower levels of the functional hierarchy.

The theoretical aspects of decomposing complex systems into smaller hierarchical entities have been studied in the literature. The theory of hierarchical systems was established in Mesarovic, Macko, and Takahara's work called ‘Theory of Hierarchical, Multilevel Systems’ [Mes70] which provided a set of theoretical principles. This work was continued in ‘Control and Coordination in Hierarchical Systems’ [Fin80] which focused on more practical aspects. There are at least five basic ways to decompose modelling and control problems, namely decomposition into physical components, decomposition into goals, decomposition into operating regimes, phenomena-based decomposition and decomposition in terms of mathematical series expansions. These approaches have distinct but also overlapping characteristics and must therefore be combined in a number of natural ways. E.g. both component-based and operating regime-based decompositions must be combined with either series expansion-based or phenomena-based models or controllers since the local models or controllers, or component sub-models must at some point be represented in terms of equations based on series expansions or phenomena. In other words the first three approaches are mainly used to break down the problem by introducing high level representations that bring the problem to a higher and more user friendly level but at some point the problem can not be decomposed further. At this point one must deal with the sub-model or sub-controller in terms of low level representations. The task of optimal decomposition is difficult and the optimal structure rarely obvious. Despite the developed theories, so far heuristic approaches have been prevalent in process automation even though they cannot directly guarantee that the overall solution meets the specifications [Mur97, Jut79].

The decentralization in process automation can be said to have started with the use of minicomputers in the 1960's but the first real *Distributed Control Systems* were introduced in the 1970's. In a DCS programmable digital control functions are distributed to PLCs or PC-stations, often called process stations. Analog and digital field devices connect to these through I/O systems. In order to reduce cabling, connections from sensors in the field are collected into separate remote I/O systems which are then connected to the process stations through data lines. Such networks relaying simple sensor information are called sensor networks. The control signals are computed in an automation system consisting of process stations, each of which has its own process area to control. These process stations can also have a data transfer connection between them. These internal data connections of automation systems are called control networks. Operating of process stations is done on a monitoring station which can be process station specific or may be shared by several process stations [Pyy07, Ron99]. A simple DCS architecture is presented in figure 3 [Geo04]. Alternative names for the hierarchy levels are used in the figure.

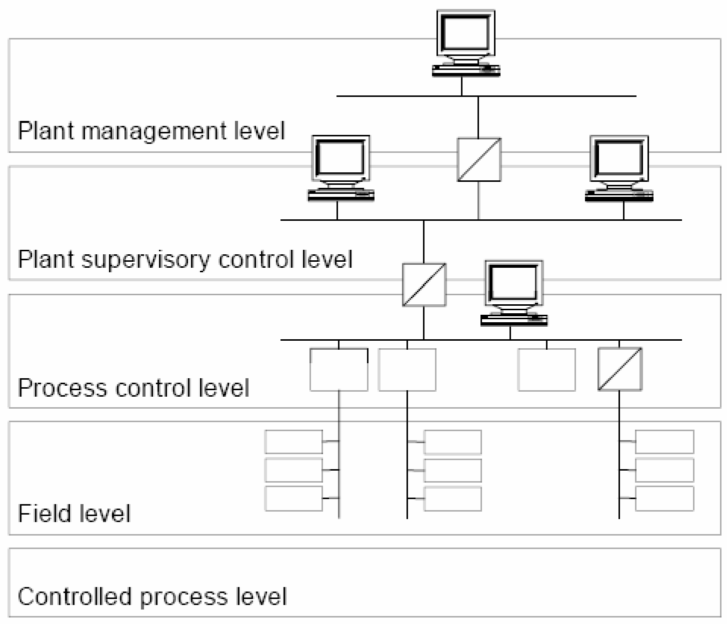

Figure 3 Distributed Control System structure [Geo04]

After DCS the next step in the evolution of control systems is the *Field Control System* (FCS). In some sources it is also called *Smart Distributed System* (SDS), although the name has not become established. In FCS programmable functions have been extensively distributed from the process stations to intelligent field devices. The field devices no longer are only information producers or executors of commands but can execute their own software applications. A software application can even be distributed to several devices. This principle and the relating description methods are standardized IEC in the IEC61499 document. Figure 4 depicts the evolution of control systems. The evolution is characterized by three basic generation changes. DDC: Direct Digital Control, DCS: Distributed Control System and FCS: Field Control System. *Open Control Systems* (OCS) has been mentioned as the step after FCS. The main feature of OCS is that the openness permits separating the design of software part from selection of automation parts. [Geo04, Ber01, Pyy07]

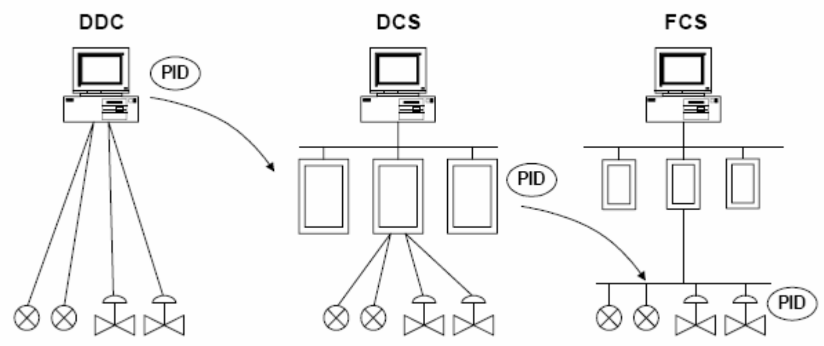

Figure 4 Evolution of process control systems [Geo04]

For implementation of an FCS environment a solution with bidirectional communications and deterministic features is needed. In some sources these are called fieldbuses and in some device networks. In an FCS environment there are two separate information flows; data- and event flow. The limitations of some buses in this respect have been noted in ISO and a shared responsive link consisting of double data link has been suggested as a solution in the ISO/IEC FCD 24740 -standard. The other data link would transfer the data flow and the other event flow. When designing a control system environment the solution is influenced by the possibilities to have parallel data links, possibilities to send prioritized and independent data- and event messages and also the possibility to supply power to the devices through the data link. In any case decentralization should not be an end in itself but the intended goals should be well defined and justified. [Pyy07]

Even though fieldbuses offer several advantages such as savings in investments and design work, from the point of view of this research the main advantage is the bidirectional communication enabling the distribution of functions. From the end-users' point of view one significant advantage is the use of embedded device diagnostics which extends the scope of the user to the field device level and enables easier monitoring of the devices [Ron99]. Other benefits include easier management of the automation entity (division into smaller parts) and more accurate control (measurements can be manipulated and control signals computed locally) [Ron99]. Some fieldbuses already provide and support standardized function blocks for embedded control in positioners. Distributed intelligence also offers higher flexibility with the plant design and extensions of a plant by the use of independent and reusable modules. The independence of the modules enables component testing before commissioning and thereby accelerates the integration process [Sch03]. In a broader sense decentralization offers several other potential benefits. For example some applications can be better solved using a distributed solution approach - especially those tasks that are inherently distributed in space, time, or functionality. Also, if a system is solving subproblems in parallel, the overall task completion time is potentially reduced. Any system consisting of multiple, sometimes redundant entities, offers possibilities of increasing the robustness and reliability of the solution due to the ability of one entity to take over from another failing entity. For many applications, creating a monolithic entity that can address all aspects of a problem can be very expensive and complex; in contrast, creating multiple, more specialized entities capable of sharing the workload offers the possibility of reducing the complexity of the individual entities [Par08].

The advantages of distributed intelligence are, to some extent, offset by some disadvantages. For example, even though the individual entity cost and complexity may be less, managing the complete system may be more difficult and complex due to the lack of centralized control or of a centralized repository of global information. Further, distributed intelligent systems require more communication to coordinate all the entities in the system. Increasing the number of entities can lead to increased interference between entities, as they must act without complete knowledge of the other entities' intents. Systems of multiple entities will typically experience increased uncertainty about the state of the system as a whole [Par08]. Finally, despite the possibilities, the real life applications of distributed systems have unfortunately shown that they have been often designed and applied to please the designers and manufacturers instead of meeting the needs of the end-users. As a result the end-users often have to adjust to the acquired technology instead of the technologies adjusting to the needs of the end-users [Pyy07].

Relation to agent technologies

In a sense high degree of decentralization and use of intelligent field devices can also be seen as a move towards industrial use of agent technologies even if the standardized terminology, communications etc. are not used in the implementation. Agent technologies in process automation have been studied quite extensively lately, e.g. [Sei06, Pir08]. Although agent technology in automation doesn't have an exact definition, it is often associated with the following properties [Jen98, Woo99a]:

- Intelligence: ability to make decisions based on observations and internal knowledge.
- Autonomy: ability to solve problems without direct guidance. Also the ability to refuse a service.
- Reactivity: ability to react to various stimuli.
- Proactivity: ability to take initiative and to be goal-oriented.
- Sociality: ability to communicate with other agents.

Agents are decentralized, individual either physical or purely software entities with a powerful autonomous character, which enables them to make independent decisions and deliberated planning concerning their actions based on the nature of the environment they are responsible of. They are also characterized as intelligent due to their flexibility and fast adaptability. Their actual power, however, comes with their social behaviour and their ability to organize and coordinate themselves in a hierarchical society. To achieve this agents are well equipped with communication skills. They cooperate through negotiations, which facilitate them to achieve optimal solutions for their society and environment. Such an agent society, also called a Multi Agent System (MAS), had originally purely software based background but is nowadays being used to extend current Process Automation Systems (PAS) [Cha03]. Clearly the current intelligent field devices don't possess all the listed features of agents. However, given that the technical problems (e.g. limited processing power and fieldbus bandwidth) are solved, extensive use of intelligent field devices offers a potential platform for industrial implementation of agent technologies in the future.

The most suitable application areas of agents in process industry consist of tasks requiring cooperative distributed problem solving. The limitations of current PASs lies in the inability to support complicated negotiations between distributed entities. Furthermore there is limited support for abnormal situation handling, e.g. device faults. Another problem is the lack of flexible connections between different automation systems and computational entities as they are mostly fixed, stable and defined at design-time. Especially the local decision making capability of autonomous agents could be used for particular automation functions, such as self-diagnostics of intelligent devices, which could lead to more accurate analysis than with external diagnosis [Cha03]. Even with the need for additional coordination, the use of agents might reduce the amount of additional fieldbus traffic as a whole if more detailed diagnostic information is only transmitted on request. In an example situation an agent representing a mixing tank notices an unexpected drop in the level of the tank. The agent has knowledge of other agents representing inflows and outflows of the tank and thus can inquire the status of the corresponding devices. If the fault can not be determined from the results of the inquiry, a simple leak test could be executed by stopping the inflow and outflow. To avoid the invasive leak test, the inquiring agent could send a description of the situation asking for fault candidates from the other agents and try to determine the most likely cause from the answers [Syr05b]. The potential benefits of exploiting intelligent field devices as a platform in this simple example are evident. For example intelligent positioners and embedded software applications could be the agents representing the inflows and outflows. At the moment the existing technology poses restrictions on the embedded software in intelligent field devices but with future developments and proper execution this kind of interaction and autonomy will not only promote improved fault diagnostics but also active fault tolerance.

Another important aspect relating to future requirements of a PAS is the increasing software orientation. Development of robust and scalable software systems for process automation is an ideal environment for agent application [Cha03]. E.g. agents can be seen as a way to combat software aging in applications that seldom accept downtimes. The software aging is defined as progressive performance degradation, mainly caused by exhaustion of system resources (e.g. resource deadlocks, memory leakage). The idea is that an agent population can be monitored, faulty agents isolated and reloaded in a healthy state, hence rejuvenated without interfering with the other parts of the application [Cap09]. It has also been claimed that if data and resources are distributed, or if a number of legacy systems must be made to work together, agents are an appropriate technology for building such a system [Woo99b]. Typically, agents use match-making and brokering techniques (much like the SOA approach, section 3.3.3) and utilize semantic information [Pir08]. In the current research of MAS applications in process automation agents have been proposed as a society of intelligent controllers that are expected to operate at higher, non real-time control levels. MASs have been proposed to select suitable control algorithms and their parameters for continuous controllers of lower level automation. According to this approach agents are expected to have knowledge about the applicability of control algorithms and the knowledge is represented with rules that map from a situation description to a control algorithm. MASs have also been proposed as means to coordinate several continuous controllers and their modifications [Sei06]. It has been assumed that with this role MAS could enhance e.g. the flexibility and responsiveness properties of control functions. The MAS organisations of the proposed applications have usually been hierarchies consisting of three levels representing the whole system, subprocesses and equipment [Tic02].

Thus, it can be argued that agent technology might be suitable to a number of process automation functions. Multi agent systems are nowadays a popular subject in the research of Distributed Artificial Intelligence (DAI) and standardization of agent systems is being carried out by several organizations such as Foundation for Intelligent Physical Agents (FIPA) and Object Management Group (OMG) [Syr05a].

3.3 Integration standards in process industry

After the popularity of the Windows platform several techniques have been developed to promote interface standardization but the first real milestone was the DDE (Dynamic Data Exchange) standard. It was lacking in some respects which in the mid '90s led to the development of OPC (originally *OLE for Process Control*, now *open connectivity via open standards*) which defines an interface for real time data transfer in Windows environment [Iwa06]. However, the scope of OPC is somewhat limited on the field device level which poses some challenges to management of field devices. Due to the limitations on the field device level, the configuration of field devices has traditionally been done with specialized tools often using their own communication protocols and sometimes even their own hardware. This has several disadvantages such as incompatibility of field devices, high number of user interfaces and version management and update problems [Zuc03, Kas04]. At the same time intelligent field devices and various management applications are becoming ever more integral part of automation systems and full advantage of these systems will be achieved only if the hardware, automation system and asset management applications are properly integrated. This is why new open integration concepts are welcome. Also the increased re-use of software components has raised the importance of standardized component interfaces [Rep06, Iwa06].

To effectively manage modern process assets the standardized interfaces have to properly reach the field device level and problems with numerous different fieldbus protocols have to be dealt with. One of the related hot topics in the development of automation systems is the electronic management of structures of automation systems, devices and their functional configuration. Favourable technical developments have enabled storing of large files in the field devices and by so have further promoted decentralization and the need for new integration techniques. These files can contain for example life cycle information of a device and they can be managed through fieldbuses as long as the applications are equipped with appropriate services and the transferred messages are suitable for the protocol used [Pyy07]. This was the motivation for developing EDDL and FDT techniques. In the future increased need for integration of various management systems is evident also on the higher hierarchy levels of the functional hierarchy where Service Oriented Architecture (SOA) is one of the used methods.

Although integration techniques were not the focus of this research, they were exploited in the applied part and have a central role in modern process automation. Also, future developments may result in integration of the developed monitoring systems and higher level management systems. These issues and currently used integration techniques are discussed in the following sections.

3.3.1 OPC

OPC is the technological basis for efficient link of automation components and it enables the integration of office products and information systems on company level. By using OPC process data can be presented e.g. in Excel sheet and stored in process history databases. The use of OPC technology means the manufacturers of OPC servers and clients are not limited as to the complexity, functionality and execution of their OPC component as long as the definitions for the interfaces, methods and implementation regulations are observed. The most diverse OPC components of different manufacturers can work together and no additional programming for adaptation of their interfaces is necessary. Thus OPC provides horizontal and vertical interoperability between software and hardware components in the automation industry and allows users to use products from different vendors [Iwa06, Zuc03].

OPC was developed in the 1990's and has achieved the acceptance of manufacturers and users of industrial automation application. The series of OPC definitions started from OPC Data Access (OPC DA) protocol which is focused on transferring real time data. Later the definitions expanded to alarm information (Alarms and Events, A&E), history data (Historical Data Access, HDA) and other individual protocols. These protocols were designed to be independent from each other which sped up the establishment of the separate parts and enabled flexible addition of new application areas. Altogether there are nine OPC specifications at the moment and they are written under the guidance of the OPC Foundation that consists of more than 400 companies. The foundation's goal is to resolve the incompatibility problems by providing easy information connectivity throughout the enterprise [Aro06, Mat05, Zuc03]. The existing OPC specifications are based on Microsoft's Distributed Component Object Model (DCOM) which poses some limitations to the client and server applications. E.g. these are required to be on the intranet (inside the corporate firewall) and there is also lack of interoperability to platforms not COM/DCOM based [Iwa06].

Mutually independent protocols have enabled easy expansion to new application areas without changes to the other existing parts but in practise the definitions have become too separate and handling of different types of information as completely separate parts is not practical. In the end it is basically same kind of information that is being transferred, only the perspective is different. Thus there is a desire to combine these but the traditional OPC definitions do not support this very well. This has limited the use of OPC protocol and it has mainly been used on control room levels above the field device level. To some extent OPC has also been used above the control room level but the MES and ERP levels are not reached yet. Also there are hardly any OPC applications in the field devices themselves, but instead these still communicate with several separate protocols. [Aro06, Mat05]

To solve the problems with separate definitions, the OPC UA (Unified Architecture) was defined which combines the separate definitions as a unified definition. The problems could not be solved by modifying the existing definitions due to the expiring DCOM technology. Therefore it was decided that a totally new definition was needed which also enables setting of more ambitious goals. The information models used in the previous definitions are combined in the new version so that properties of a single measurement can be handled as a part of a single information model. The basis of the new definition is a very abstract basic information model which enables flexible definition of information models for different application areas. OPC UA is not just another standard but entirely new and holistic architecture designed to replace the group of existing separate definitions. [Aro06, Mat05]

The UA definition includes a data transfer protocol which is not tied to any specific technology such as DCOM. A binary protocol capable of faster transfer speeds than DCOM was included in the definition to ensure fast data transfer. Another basic protocol in OPC UA is Web Services technology which is based on the SOAP protocol. This is widely supported technology which enables easy integration with other applications. To improve data security the definitions of OPC UA include encryption methods. Also the reliability of data transfer has been improved by using so called heartbeat signals to monitor connections actively. [Aro06, Mat05]

The transition towards OPC UA started with the development and release of the OPC XML DA specification which breaks the tradition of COM/DCOM communications on which the traditional OPC has relied. Web services and XML remove the limitation for data and information to be isolated behind a corporate firewall and therefore improve cross-platform connectivity over the Internet. In the short term the OPC Foundation still believes that for low level, high performance applications COM remains the clear choice. However for higher level communications over the Internet and for communications where performance is less important than flexibility XML and its related transports have a clear advantage. [Iwa06]

3.3.2 Field device integration

Even though control systems manufacturers and systems integrators have made inroads to making process control data available through OPC technology, it concerns itself more with the free exchange of live process data between Windows based applications, and to a much lesser extent the configuration of the specific field devices providing the data. The basic reason for this is that OPC hides the communication network behind the interface. This is fine for basic operating applications that are only interested in the data behind the interface but configuration applications need information about the devices and network topology behind the interface. The device information can be accessed through OPC if additional Device Description Services (DDS) application is integrated into the OPC server or the client application may be double ended (meaning it can present itself as an OPC client to the main server and as an OPC server to other clients). However, these solutions are not in accordance with the OPC standard and still can't publish information about the network topology. Therefore the configuration applications have in the past often been proprietary and connected straight to device networks instead of operating through OPC interface [Zuc03, Har02].

To tackle these problems there are currently two prevailing field device integration standards on the market, FDT/DTM and EDDL. In both of these the intelligent field devices have a device description file which typically contains information like device parameters and calibration information. Depending on the source a device description (DD) can mean a general device description without any links to the actual implementation technique (FDT or EDDL) or in some cases it refers strictly to a text based Device Description written in EDD language and used by the EDDL technology. In this text DD refers to the former. An EDDL file is a human readable text, defining device variables, structures and methods encapsulating device data. An EDDL file can be read by an interpreter which extracts the information needed to configure the described device [Pyy07, Kas04]. In FDT/DTM the field devices and related information are represented by software applications called Device Type Managers (DTM) which are more flexible than EDDL files [Pyy07]. DTM is a software component developed by the device manufacturer containing device specific application software [FDT07a]. The FDT specification defines a common interface to provide a separation of frame applications (FDT) and device drivers (DTMs). To gain access to devices an FDT frame application can load DTMs [Kas04].

FDT/DTM

The purpose of FDT/DTM is to complement OPC by providing device manufacturer and fieldbus protocol independent access to intelligent field equipment data and thereby providing a convenient way to manage the life cycle of field assets. Whereas OPC focuses on transferring critical process and business information horizontally across the enterprise and vertically to and from the process control system, FDT extends this vertical integration down to the field equipment domain by providing OPC clients with vertical access to the data and information in the multiplicity of field equipment. The OPC and FDT technologies work together to expand the availability of information. FDT enables control systems manufacturers and device manufacturers to develop plug-and-play data access plumbing that makes the field equipment information domain readily available to OPC clients. FDT-compliant field devices provide data and information to OPC servers, which in turn honor OPC clients' requests. The FDT specification, like OPC, is riding the open Microsoft technology road map. [Zuc03]

FDT technology helps automation users avoid a situation where proprietary interfaces for field devices are needed and users of some systems can not benefit from applications of some other device vendors. FDT defines the interfaces between device applications, the control system platform and the physical fieldbus devices. It allows device vendors to create applications that have a common interface to the system and any system that supports these interfaces can integrate the applications. The applications have the same behaviour, look and feel in every system [Yok06]. Broadly speaking, the FDT system can be compared to a printer driver system in normal office applications. The printer is delivered with an according driver which implements standardized interfaces so that any office application can make use of it. In FDT, the hardware (the field device) is delivered with a driver called Device Type Manager (DTM), which has the standardized FDT interface. This enables any FDT-enabled application (FDT Frame Application) to use it. In other words, the FDT specification simply standardizes the way data is exchanged between software components for field devices and control systems or asset management tools. Currently the supported protocols include FOUNDATION Fieldbus, HART, Profibus DP, Profibus PA, Profinet IO, DeviceNet, Ethernet IP, Interbus, AS-interface and ControlNet [FDT07a, FDT07b].

The FDT technology consists of a frame application and three kinds of DTMs.

- DTMs for a device class with direct access to a communication component are named Communication DTMs.
- DTMs used for routing between different protocols (i.e. from PROFIBUS to HART) are named Gateway DTMs. These ensure that the details of the PC, network, interface cards, and pass through protocols of the host system are transparent to the Device DTM.
- A DTM that represents a field device is called Device DTM. A Device DTM interacts with a Communication DTM or Gateway DTM to access its field device.

These DTMs run in an FDT Frame Application (e.g. device configuration tools, control system engineering tools, operator consoles or asset management tools) which manage all device instances, store the instances' data and take care of data versioning. The Device Type Manager (DTM) is a software component developed by the device manufacturer containing device specific application software and is typically supplied with the device. [FDT07a]

As Microsoft Windows has established itself as the de facto industry standard, COM (Component Object Model) has been chosen to be basis of FDT. COM has proven client-server architecture and manages the integration of software components into the FDT Frame Application. It also enables inter-process communication and dynamic object creation. However, the dependence on Windows technology can be considered the biggest drawback of FDT. As such, it is subject to the inevitable upgrades and operation system version changes which means that FDT platforms and device applications will also need to be updated occasionally and software upgrades require certain amount of regression testing of DTMs to ensure that a new functionality doesn’t break functionality of already existing functions. Despite these disadvantages, Windows technology’s advantages have increased the popularity of FDT. A major advantage is its versatile possibilities to implement complex functions whereas text based EDDL has limited possibilities for complex functionality such as complex algorithms and graphical presentations. [FDT07a, Nie04, Yok06, Pyy07]

EDDL

The Electronic Device Description Language was created to describe devices and to avoid the need for creating separate device drivers for every device type. This is achieved by providing rules of conversation for systems that interact with these fieldbus devices. EDDL is a device description language with which the suppliers can create a device description for the devices needing integration to the automation system. [Yok06, Zie06, Mäk06, Hak07]

The DD files of EDDL contain a description of device parameters such as measurement limits and calibration information. Typically a DD describes the device, device variables, structures and how to access the information. A host application is needed for reading the DD file in order to learn how to retrieve and interpret information from the device. All software tools equipped with an interpreter accepting EDDL grammar can extract the device data out of the EDDL file. Besides the data storage there are several differences in the working mode of FDT/DTM and EDDL concepts. E.g. EDDL instances do not communicate with their devices while DTMs do. The way of storing data in EDDL means that an interpreter is needed and several layers of software must exist between the EDDL file and actual communication with a device. On the other hand EDDL programming needs less effort to learn and handle than the development of DTMs does. [Kas04, Zie06, Yok06]

EDDL is well suited for simpler applications and its benefits include operating system independence and therefore the chance to avoid update problems associated with Windows. The EDDL enhancements are also backward compatible, ensuring that existing DDs will continue to work and allow users to have a continuous forward migration path with no investment loss [Hel07]. However, as EDDL is a fairly simple language it offers somewhat limited features compared to the FDT/DTM. As in all engineering decisions, the toolset should be chosen based on the product requirements. If the application is simple and can be accomplished with EDDL, this is a reasonable choice and has some benefits, including a degree of platform independence [Yok06].

The data transfer in both FDT and EDDL techniques uses fieldbuses also used for operative use of automation. Even though the collection and use of device management data is periodic, due to the bursty nature of the communication it may interfere with the main task of the fieldbus. In some cases it may be necessary to use a second line of communication alongside with the operational fieldbus to ensure the deterministic properties of the system. However there seems to be no such applications available that support this [Pyy07]. Both the EDDL and FDT/DTM organizations are currently working with the OPC Foundation to ensure compatibility with the OPC UA applications [FDT06, Aro06].

Other techniques

The two established device integration technologies, FDT and EDDL, exhibit different focuses, although they also possess many overlapping features. This has led to a situation where both technologies have competed on the market instead of complementing each other. To overcome the limitations a new concept, FDI (Field Device Integration), is being developed that combines the advantages of EDDL – platform independence, ease of use, and robustness – with the advantages of FDT – unlimited functionality, extensibility, and market differentiation – in a clearly structured client-server architecture with open communication. The basis for the technology was first presented to the public at NAMUR's 2006 Annual General Meeting under the name FDD UA. [Ben08, Ben07]

In addition to the presented field device integration techniques another concept called Tool Calling Interface (TCI) has been introduced recently. TCI defines a mechanism that allows an engineering tool to start a device specific standalone application for device operation. It is a very simple solution but as TCI only defines how to invoke the proprietary device tool automatically, it does not contribute to device integration since the device data and functionality are only available within the device tool. Asset Management application, for example, can not access device data via TCI even though it is precisely this type of open access to device data and functionality that represents one of the main goals for field device integration. [Ben08]

3.3.3 SOA

As the importance of information availability increases in process automation, system integration is increasingly important also on the higher hierarchy levels. Integration on the higher levels has traditionally been very expensive due to proprietary applications and interfaces. However, work has been done also in this area and SOA is one of the used techniques in the industry. It offers some features (loosely coupled, modularised, and service-based operation) similar to agent technology [Nik07, Pir08].

Traditionally the intelligence has been embedded only in the software applications and the task of the networks has just been to transfer information between applications. Therefore integration of new applications to existing ones has meant negotiations with the owner of the existing applications and extensive integration work. The SOA approach avoids this by using techniques which “loosely” connect services and service users. SOA applications are independent from communication protocols, operating systems and software techniques. The service is disconnected from its interface so that it is possible to update the service without any modifications to the client applications. At the moment the most established solutions are WebService and XML based. [Fly05, Nik07]

SOA offers a way to systematically define data structures and interfaces for software applications and therefore promotes creation of libraries of standardized solutions. Based on such a service-oriented architecture, a service consumer does not need to know about the characteristics of a service provider it is communicating with because the underlying infrastructure will locate the appropriate service provider on behalf of the consumer. In addition, the infrastructure hides as many technicalities as possible from a service consumer. [Fly05, Nik07]

Key Service-Oriented terminology includes services, service provider, service consumer, service locator and service broker. Services are logical entities, the contracts defined by published interfaces while service provider is the software entity that implements a service specification. Service consumer is the software entity that calls a service provider; traditionally this is called a “client”. A service consumer can be an end-user application or another service. Service locator is a specific kind of service provider that acts as a registry and allows for the lookup of service provider interfaces and service locations. Service broker is a specific kind of service provider that can pass on service requests to additional service providers. Collaboration in SOA follows a “find, bind and invoke” model, where a service consumer performs dynamic service location by querying the service registry for a service that matches its criteria. If the service exists, the registry provides the consumer with the interface contract and the endpoint address for the service. [End04, Fly05]

In an industrial environment SOA enables easier integration of existing applications, such as process automation systems and quality control systems, into production management systems. An example of a SOA application in process automation is the metsoDNA automation system where previously separate information- and automation systems are combined into a network of applications consisting of various modules. Integration with older systems in a genuine SOA application requires a WebService interface to be built after which the application can be connected. As a one time solution this is acceptable but greater benefits are achieved only if the interface is reusable. For this reason Metso Automation is employing a combination of traditional hub architecture and SOA, called SOA-hub architecture. This brings several advantages and the approaches complement each other. The service providers can offer WebService interface either integrated into the software or as a partner (or an adapter). When the interface is an adapter in a hub, significant savings can be achieved through better reusability. When comparing the amount work needed for creating a single interface and an adapter there is not much difference but if the adapter is designed properly, good reusability can be achieved by simply changing the adapter parameters. There are also advantages in security and maintenance issues as all communications go through a single point. This enables for example creation of centralized filtering of messages going through the hub [Nik07].

The use of SOA-hub approach offers benefits for example in SAP environment integration where a well designed SAP adapter can be re-used in several SAP systems just by changing a few parameters. A SAP adapter uses the RFC (Remote Function Call) interface of SAP which offers the possibility to use the internal BAPI (Business Application Programming Interface) functions of the SAP. By using RFC interface one does not have to write any internal BAPI functions of the SAP system but all the functionality is based on the already existing functions. The use of small adapters enables also the use of more complicated operations as reading can be done using one adapter, the operations are done in the client application and the results can be fed back to the SAP by using another adapter. [Nik07]

4. LARGE SCALE MONITORING APPLICATION

As the focus in the applied part of the thesis was on developing generic monitoring methods for large scale processes, the requirements and desirable features for such an application were studied. A literature survey was conducted and the results of this survey were then combined with own experiences to form a set of requirements and corresponding features. These are described in the following section which forms the second contribution of the thesis. Related issues are also briefly discussed in [Huo07]. The application and some use cases are described in sections 4.2 and 4.3 which form the first contribution of the thesis. Chapter 4 along with the analysis in section 6.1 also answers the first research problem.

4.1 Requirements and desirable features

In general, there is not much literature on industrial applications of monitoring/diagnostic systems. This could be due to the proprietary nature of the development of in-house systems, but there also seems to be a general lack of overall penetration of monitoring/diagnostic systems in process industries [Ven03b]. Many of the articles and literature on process monitoring applications seem to focus on performance of individual control loops (Control Performance Assessment, CPA). However, whereas the theoretical aspects of CPA algorithms have been discussed quite effectively in many publications the descriptions of industrial CPA applications and software packages are relatively scarce [Jel06]. Also, in general the scope of process history based (the most common approach for monitoring/diagnostics in process industry) systems as applied in the industry is mainly restricted to sensor failures and there are very few industrial applications in published literature that deal with parametric failures [Ven03b]. Often the cases in literature focus on monitoring of individual processes or equipment instead of larger entities. There are plenty of articles where a monitoring application and its specific target process are described but there are not many descriptions of wider scale applications. This is the case even though the use of more holistic approach holds greater potential for improvements.

The overall lack of implemented diagnostic systems in process industries might be due to a gap between academic research and industrial practice. Adaptability of the systems and ease of deployment can be seen as two of the most important considerations from an industrial viewpoint but these are seldom addressed in academic research. As a result most of the techniques would do poorly on the issue of ease of deployment [Ven03b]. It is often the case that much of the attention is given for the performance requirements (e.g. fast fault detection) of the application, see e.g. [Das00, Ger98, Jel06]. This is naturally a very important aspect but more attention should be paid to economical and usability aspects as these have a significant role in whether the technology will be commercially successful or not. One should be able to show that the customer gets a reasonable return for the investment before anyone is going to invest in such projects. This is especially important for large scale process monitoring where the configuration work can be extensive. Earlier attempts in this aspect include the use of minimum-variance controller for the evaluation of controller performance and using default monitoring parameters for different types of control loops to address the issue of amount of configuration work [Haa00, Tho99].

The importance of these issues was acknowledged and they received significant attention. It was determined that one of the primary goals was to develop a monitoring application with such features that allow it to be delivered to different kinds of processes with reasonable costs (i.e. with minimal customizing work) even as a part of normal automation delivery project. It was also concluded that probably the most problematic prerequisite for successful commercial monitoring applications is achieving economical feasibility while retaining acceptable performance level. The requirements are very demanding and in some cases contradictory. For example, in general the performance of monitoring applications depends on the quality of configuration and customizing work but on the other hand the economical aspects demand that the amount of setup work is reasonable. The main divisions in the set of requirements and desirable features are economical, performance and usability aspects. The most relevant parts for this thesis are summed up at the end of section 4.1.4.

4.1.1 Economical aspects

To achieve economical feasibility the used methods should be easy to set up, configure and duplicate (reusability). The ease of deployment is also noted in [Ven03b]. In many cases it can be difficult to convince a customer even if the long term benefits would outweigh high initial costs. Therefore low initial costs can be seen as a door to wider acceptance and use of monitoring applications in process industry. The importance of easy set up and configuration increases if the application is meant for large scale monitoring with several different types of processes whose monitoring modules need separate configuration and customizing. Easy duplication is useful if there are several instances of identical equipment or the application is meant to be used in large scale, i.e. to implement it to several similar processes or devices in other facilities. These requirements can be achieved by appropriate choice of methods, e.g. number of monitoring parameters should not be excessive. Additional setup tools can be used to aid setup.

The application maintainability issues are also very important as nowadays it is not uncommon to expand, rebuild or make changes in production lines for increased capacity. The design of the application should enable easy modifications and scalability. This can be achieved with modular structure so that changes in a module have minimal effect on other system modules. From the point of view of the monitoring algorithms and application maintainability industrial environment is very challenging even in normal every day use. Modern day processes are usually in a state of constant change and the monitoring application must be able to adapt to the changing environment as independently as possible. The need for adaptability and desire for modular structures are also mentioned in several sources, e.g. [Pir08]. After initial setup monitoring applications without adaptability are usually effective for only a short period after which the performance starts to decrease. A good example of natural changes that the algorithms should be able to adapt to is seasonal changes. The monitoring methods should be able to adapt to naturally changing situations without false alarms but also without increase of missed alarms. This means that in many cases static alarm limits are usually not sufficient.

Maintenance issues also include software updates and support for the chosen technologies. Software updates are usually unavoidable and must be easily manageable. As there are potentially several remote users, possibly in separate organizations, the software maintenance of the client machines can become a significant issue. IT-security issues must also be considered as the monitoring application often must be connected to other information systems and client machines [Mil04]. Therefore the software updates must be managed so that the client machines, where ever they may be, can not cause serious security risks. Achieving this is easier if major software updates are centralized and only concern the server of the application. As the expected lifespan of a successful monitoring application is long, sufficient consideration must also be taken when selecting software techniques so that the application can be effectively supported even years later.

4.1.2 Performance criteria

Commonly used performance criteria state that the number of missed and false alarms should be minimal. On one hand the application should in an optimal case always detect serious problem situations. On the other hand the alarms created by the application should always represent real problem situations because high number of false alarms decreases the efficiency of operating personnel and generally creates mistrust towards the application. These goals lead to a compromise determined by the chosen methods and their parameters. Use of properly verified methods is naturally preferred. Traditional process monitoring methods are based on different statistical approaches and include the use of distributions and different control charts such as Shewhart, cumulative-sum and exponentially weighted moving-average control charts [Mon91]. However, these traditional methods have a drawback in that they do not take into consideration the possibility of changing operating points which can cause the specification of good or normal performance to change. This is a typical situation in a modern industrial process where different grades and products are produced using the same equipment. Therefore the traditional univariate SPC is usually not sufficient for modern process industry [Hie03].

As mentioned before the requirement of easy deployment is also related to performance. In general the performance of a monitoring application improves with the amount and quality of configuration work but excessive deployment costs can not be justified. For example model based methods often produce good results but their use is limited because creating the models is often too expensive, or in some cases impossible. On the other hand less detailed process modelling leads to problems with model inaccuracies [Chi01, Che99]. The requirement of adaptability is also related to the performance criteria as the definition of good or normal operation can depend on operating point (e.g. grade changes, seasonal changes) and without consideration this can decrease the application performance significantly.

The application should produce reliable, sufficiently precise and repeatable results. For the algorithms repeatability means that the same results should be achieved using the same data (given that with adaptive methods sufficient adaptation has already taken place). Besides the used algorithms these requirements also include the whole measurement chain and instrumentation in general. In practise this means that only reliable measurements should be used. Use of unreliable measurements, for example measurements having a tendency to drift, should be avoided where possible as their reliability is questionable in long run even if the performance is good during the application deployment. However alternative measurements are not always present in which case compensation of known measurement errors may be considered. This can be useful especially for slowly developing measurement errors if there are more reliable but otherwise not so suitable measurements available. An example of this is bias estimation and compensation in pulp consistency measurement by exploiting scanner measurements from a quality control system. When the needed measurements are not directly available from the monitored process the use of soft sensor approach may be considered as in many cases the interactions of process variables can be used to calculate the missing measurement indirectly. An example of soft sensor application is calculation of steam enthalpy from temperature and pressure measurements.

The ability to handle nonlinearities is also an important performance aspect as real life processes are typically nonlinear. Nonlinearities will limit the usefulness of linear models to a specific operating point but on the other hand nonlinear models are often complicated to develop and use. One possible solution is to use several adaptive linear models corresponding to identified operating points. More discussion on performance can be found in [Das00] where the following desired features for monitoring applications are listed:

- Early detection and diagnosis.
- Isolability: discriminating between different faults.
- Robustness: performance degrades gracefully instead of abrupt and total failure.
- Novelty identifiability: ability to discriminate between normal and abnormal states and if abnormal state is caused by known or unknown malfunction.
- Multiple fault identifiability: ability to identify multiple simultaneous faults
- Explanation facility: ability to provide explanations how the fault originated and propagated to the current situation.
- Adaptability: ability to adapt to changing environment.
- Reasonable storage and computational requirement: trade-off between the computational complexity and system performance.

Also a set of desired features for CPA algorithms collected from several sources is available in [Jel06] and these can also be considered as desirable features for any large scale monitoring applications. These are somewhat similar to the list described in [Das00]. Notable differences in the requirements are:

- No need for plant tests.
- Non-invasiveness: the CPA procedures should run without disturbing the normal operation of the control loops.
- Use of raw data: the use of archived (usually modified) data is not advisable.
- Low error rate: low number of false and missed alarms.
- Appropriate presentation of the results to the user (human–machine interface).

It is worth noting that in this source the importance of usability and human-machine interface is mentioned.

4.1.3 Usability

Despite its importance usability seems to be a many times overlooked aspect in academic research. It is understandable that in academic research projects the usability aspect is not always seen as very important as the users in the development phase only include people who are familiar with the case and have direct access to the application without much consideration to security issues. However when developing commercial applications for industrial use the significance of these factors increases. As in this case the application is designed to be used by people from different organizations and different backgrounds the importance of usability is critical. This means that the application should be easy to use without any specific IT know-how besides basic Windows use. The usability aspects are important for monitoring tools in general because auxiliary applications with poor usability tend to have low utilization rate. The interface is the key to the user acceptance, and therefore must be intuitive and easy to use. A successful interface provides a summary of problem areas that may exist in the plant, as well as a detailed presentation of the data collected and the analysis done. Thus the level of information should be controllable; in a steady situation abstract reports are often enough, unlike in an abnormal situation, where specific details may be needed. Along with intuitive operation user configurability is also a desired feature [Jel06, Pir08].

Typical data sets collected from modern industrial processes are huge but monitoring and analyzing a process from raw data alone is difficult and time consuming. Due to the increased amount of available process data the collected information should be refined in a way that makes it easier to handle. In problem situations the interactions of process variables can create an overflow of alarms if the effects of interactions are not considered. The resulting bombardment of alarms can actually make the job of operators even more difficult as it is hard to analyze the importance and meaning of numerous alarms in a hurry. Thus, especially the exploitation of online (meaning that the observations are analyzed in real time) features of the application requires that the disturbance origin can be identified with reasonable accuracy and work. The ability to distinguish between fault/disturbance origin and propagated disturbances is a very desirable feature particularly for larger scale monitoring applications. Besides abnormal and transient situations where the online features of the application are important the refined information can also be helpful in offline process analysis, optimization and development. Integrated offline analysis tools are needed to fully exploit these possibilities, especially in multivariable cases.

To get a comprehensive overview of a process situation the application should integrate components of performance measurement, fault detection and diagnostic systems. This has several advantages for usability. In case the fault detection part detects deviations in the process it is useful to have tools for assessing the seriousness of the situation as this will help in planning the responses to the situations. The information can be used to determine if the problem needs to be dealt with immediately or if the corrective actions can be done during the next planned shutdown. On the other hand if there is a decrease in performance the diagnostic modules can be used for identifying the cause even if the situation has caused no actual alarms. A distributed setup of monitoring modules corresponding to the process hierarchy is intuitive and helpful in getting the overall picture of the process states. From the module hierarchy it is also easier to observe the effects of changes (process faults, re-tuning of control loops etc.) on different levels of the process hierarchy. This feature can also be useful for process development. Thus it can be argued that the modularity and module setup corresponding to the process hierarchy help fulfilling the requirement of intuitive system use.

Transparency of the methods and algorithms is also a requirement as in science and engineering it is often demanded to use techniques based on functions that can be understood and validated. It also increases confidence for the results and the application in general. This can be a critical feature when trying to convince customers to invest in a monitoring application. For example the problem with lack of transparency typical for artificial neural networks hinders their usability as monitoring tools [Fan05].

It can be argued that the accessibility of the produced data is as important as the data itself, especially for the online features of the application. Even if optimal process monitoring methods are used, the application can not be very successful if the information is not available in the right place at the right time. Furthermore, as the users of the application can be located physically in different places, it is clear that it is not enough to just store the information in a local database and an interface for remote users is also required. Also connections to other systems are useful. If the online features play an important role in the application, information concerning critical faults should be made available to the operators as soon as possible to enable a quick response. This requires a functioning interface between the monitoring application and the automation system. Due to long life cycles of control systems in process automation, support connections to pre-existing information and control-level systems are often required, meaning the system should provide integrated information access [Pir08]. Access to logged process events and process database is a desired feature especially for remote monitoring purposes.

While monitoring systems utilizing online features have certain real-time issues, monitoring systems mainly provide supportive information and are not directly involved in the control of processes. Therefore real-time issues are typically not as important and reasonable delays are not seen as problematic. The operators are watching the process in real-time while maintenance personnel are looking at things from a longer-term perspective trying to estimate the correct timing for maintenance. Furthermore, management personnel often focus on higher-level data that concern elements of the physical process less directly. Therefore, typically for monitoring systems that focus on supporting maintenance and life cycle management, processing can be performed in the background while utilization of online features places tougher requirements for the system [Pir08].

4.1.4 Other issues

As the developed application is intended for large scale process monitoring (in a sense that it is suitable for large scale processes) but also for large scale deployment (in a sense that the application should be easy to deploy on several production sites), it would be useful if the measurements and calculations comply with some standards or common definitions to enable benchmarking. Benchmarking between different production lines can be a useful tool for comparing between different process or control solutions and in this way assist process development. Therefore standards for different processes should be defined and documented. Everything having a significant effect on a measurement results should be standardized and documented starting from the measurement instrumentation and hardware to the presentation of results. Required measurements should be listed but also related issues like sampling frequencies and filtering in the hardware should be documented as these can have a significant effect on the calculations. Also the calculation parameters like calculation frequencies and calculation windows should be documented too. If the process database uses data compression the compression parameters should be documented. Finally, there are typically several exceptional situations in the process usage which must be considered separately and the related practices need to be documented. For example the use of scanner measurements of a paper machine after a web break usually requires a certain amount of time before the scanner warms up and gives reliable results.

In more general terms there are many potential pitfalls in the development of monitoring applications. There have been projects where good solutions for monitoring very specific processes or events have been achieved but the application has quickly made itself useless. This can happen after deeper understanding of the process has been achieved through process analysis and designing the monitoring application, e.g. creating a process model. Better understanding of the process can then result in changing the process setup and consequently in disqualification of the developed monitoring application as the specific problem no longer exists or the developed application no longer represents the actual process. Therefore from the point of view of a monitoring system supplier generic methods should be preferred and detailed and very specific models avoided. It should be remembered that improved data accessibility also poses additional risks as remote connections to production sites potentially form security risks. Therefore the use of remote connections requires proper attention to the management of the connections. The involving intellectual property rights are also important issues. The end-users usually have a specific product and a certain way of producing it, a recipe. Even though this information is not usually particularly interesting for the partners, the end-users may be concerned that their know-how will leak through the monitoring applications. Furthermore, outsourcing of IT-services typically means that using remote connections further increases the amount of organizations involved in a project. For avoiding further problems responsibility issues between involved organizations should be clearly stated before starting a project. Unresolved responsibility issues can be difficult for example if the end-user (plant etc.) and a partner (automation supplier etc.) have separate subcontractors for maintaining remote connections and there is a difficult problem with the connections. In such cases it may occur that neither party wants to claim responsibility for a problem. The most relevant parts of the discussed requirements and desired features are summed up in table 1.

REQUIREMENT	**FEATURE**
ECONOMICAL ASPECTS	
Easy deployment	modular structure; number of parameters; setup tools
Scalability and ease of modifications	modular structure; number of parameters; setup tools
Application maintenance	adaptability; modular structure; software update management; server/client structure; supported software technologies
PERFORMANCE ASPECTS	
Number of missed and false alarms	use of verified methods; parameter compromise; ability to distinguish original and propagated disturbances
Reliable and repeatable results	use of verified methods; use of reliable measurements
Ability to handle nonlinearities	nonlinear and adaptive methods; identification of operating points
USABILITY ASPECTS	
Ease of use	intuitive GUI; module setup corresponding to process hierarchy; refined information; integrated components of performance measurement, fault detection and diagnostics; controlled level of information
Data accessibility	realtime GUI; reporting interface with remote connections
Transparency	no black-box methods
OTHER	
Standards for benchmarking	documentation of methods and practises
Generic application	no detailed and specific 1st principle models

Table 1 Requirements and corresponding features

4.2 Description of the process and the monitoring system

The main function of pulp drying is to modify the fiber properties and to reduce transport costs. The process itself is a lot like a basic paper machine. After the pulp manufacturing process the pulp is stored in mass towers from where it is taken to the headbox via two smaller tanks. The main function of the two tanks is to act as a buffer to filter out disturbances in pulp stock consistency. The pulp is diluted in several stages before the headbox so that the consistency drops from around 10-12% to around 1%. The headbox then dispenses the pulp stock as evenly as possible to the wire where the web formation takes place with the help of suction and pressing rolls. After formation the web goes through the dryer where it is dried with hot air. Usually there is a cooler after the dryer before the cutter and baling. The process was divided in to the following subprocesses:

- Stock preparation.
- Wet end (headbox, wire, rolls).
- Dryer.
- Cutter.
- Baling, wrapping, binding.
- Broke line.

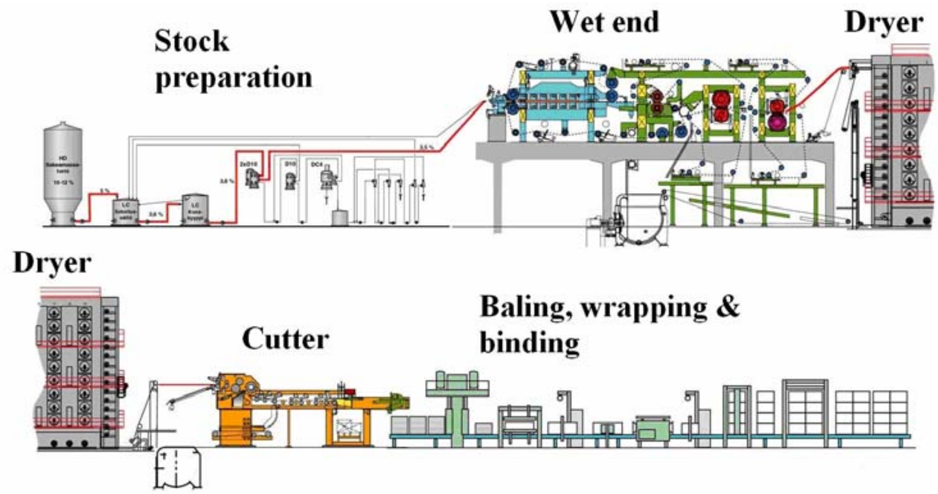

Figure 5 Pulp drying process

The subprocesses are sequential, i.e. the pulp advances from one subprocess to the next and the performance of a subprocess can have an effect on the performance of a following subprocess. The exception to this is the broke line which is used in the event of serious quality or runnability problems to recycle pulp from the wet end or after the dryer (in which case the pulp has already been dried) back to the mass towers through the pulpers [Gul00]. The use of recycled pulp may cause further disturbances because the pulp properties change every time it is dried. Also the addition of broke pulp into the main pulp stock flow may disturb the main flow. These properties cause a disturbance feedback which can make isolating the original disturbance or fault difficult.

4.2.1 Monitoring modules

The monitoring modules can be divided into performance measurements producing simple performance data and more complex fault detection and analysis modules. The performance measurements consist of several modules representing plant and subprocess level performance. The application hierarchy is described in more detail in section 4.2.2. Most of the performance measurements are relatively simple and static calculations merely producing information for the end users and in some cases for the fault detection and analysis modules. Thus most parts in the requirement list are more focused on fault detection and analysis modules where higher level functions, such as interpretation of the performance measurements, are carried out. However, the economical and usability requirement aspects are still relevant for each component. In this respect the ease of deployment for the performance measurements is assured by the use of a tag alias list and transparency is achieved with the use of the developed calculation environment (section 4.2.2).

The economical aspects were the main driving forces in selecting the fault detection and analysis modules, i.e. before a method was considered it had to meet the economical requirements. There were naturally other requirements, such as adequate performance and ability to handle nonlinearities, to be met as well. For example traditional Statistical Process Control (SPC, e.g. Shewhart charts) was ruled out because even though it is easy to apply, it does not consider dependencies between process variables and therefore the number of false alarms would be unacceptable in a dynamic environment. Statistical projection methods were ruled out primarily due to technical reasons and difficulties with nonlinear processes. Model-based methods were not suitable because they require a process specific model and therefore are not within specifications. NN methods were not chosen because of their lack of transparency and unpredictable results with data from never before seen operating points and situations. The chosen methods are process history based and include conditional histograms [Fri03, Fri07] (also a basis for a tool called Multistate monitoring, MUST), cluster analysis [Hie03] and control loop monitoring tools [Mar99b].

Performance measurements

In order to define meaningful performance indicators one must consider what are the purpose and goals of the measured processes. E.g., the main function of stock preparation is to deliver pulp stock flow to the headbox as evenly as possible. Disturbances in the pulp stock flow can cause runnability problems and even cause web breaks. A good measure for stock preparation performance is therefore relative standard deviations (standard deviation divided by measurement mean) of the pulp stock flow. The performance measurements were divided into plant level measurements and subprocess specific measurements.

The plant level measurements are further sorted into usability, efficiency and quality measurements. The quality measurements don’t, however, measure the quality of the end product but are instead more associated with the process runnability. After all, the produced pulp sheets are broken in a pulper before their eventual use and the quality of the end product is a property of individual fibers. This is created mainly before the drying process and can only be ruined in the drying machine by extreme conditions. Some of the measurements were directly available from the Quality Control System.

Usability:

- Time use: process up time (determined from the status of headbox mixing pump), process idle time (process running but no production). Definitions of calendar time and available time (see the OEE-definitions below this list) are used. The actual measurements are process up time in relation to available time and process idle time in relation to process up time.
- Amount and duration of web breaks.
- Broke line usage is measured as amount and duration of activations. These are also sorted according to production speeds resulting in a histogram presentation.

Efficiency:

- Overall Equipment Efficiency, OEE, indicator represents the performance of the overall process compared to its ideal performance. More detailed description below the list.
- Production is measured as produced air dry tonnes of pulp per day [ADt/d].
- Energy consumption and efficiency are measured as energy consumption in the dryer [MW], relative energy consumption [GJ/ADt] and relative consumption of steam [t/ADt].

Quality:

- Quality is measured as CD and MD profile variations of moisture in the end product. These can also be used for measuring performance of wet end and dryer although weak performance in the stock preparation can affect these measurements.
- Other quality measurements are the profile variations of total weight and dryweight. These can also be used to measure the performance of stock preparation and wet end although the performance of the dryer can affect the total weight measurements.

Overall Equipment Efficiency, OEE, indicator represents the performance of the overall process compared to its ideal performance. The indicator consists of three parts: availability, performance and quality rate. The overall efficiency is calculated as a product of these terms. For the calculation of availability the definition of intended operating time of the process is needed. For simplicity this can be replaced by calendar time but for more informative results factors like planned shutdowns, strikes and lack of orders should be considered as these factors affect the result but are not caused by the monitored process. The availability indicator itself is defined as the ratio of realized operating time and intended operating time. For the performance part the maximum production rate needs to be defined. This can be replaced by the design specifications of the machine but in such case the performance indicator can be over 1 (or over 100%) if the machine can be operated at higher speeds than the specifications. The performance indicator is defined as the ratio of achieved pulp flow through the machine and the maximum production rate. In the case of pulp drying machine the quality factor is defined by the ratio of actual achieved production (pulp flow through the machine minus broke pulp) and potential production (pulp flow through the machine). All the time definitions such as available production time are based on Zellcheming 2005 standard [Zel05].

On the lower level the following subprocess performance measurements were used. The first two represent the stock preparation process.

- Mass towers: relative standard deviation of pulp consistency was used. See details of using relative standard deviation below this list.
- Feed from the machine chest: relative standard deviations of pulp flow and pulp consistency were used.
- Wet end: CD profile variations in dry weight were used.
- Dryer: the energy consumption measurements described in the plant level measurements were also used for the dryer performance measurement.
- Cutter: the number of broke line activations and duration of them caused by the cutter were used. Similarly to the plant level measurements, these can be sorted according to productions speeds and more specific causes.
- Baling: number and durations of disturbances caused by baling. These can be additionally sorted separate causes and individual devices.

Using relative standard deviation instead of regular standard deviation removes the measurement units and enables easier comparisons. Also this way the undesired deviations are proportioned to their effect. For example in the case of consistency measurement the effects of standard deviation with magnitude of 0.5 are much more significant if the corresponding average is 2 when compared to a situation with an average of 4. Relative standard deviation is a good measure for processes that have a constant setpoint for most of the time. Calculating the standard deviation part from setpoint and measurement difference is more suitable for processes that have constantly changing setpoint.

Conditional histograms

The main purpose of the MUST algorithm [Fri03, Fri07], which is based on conditional histograms, is to monitor processes with changing definitions of acceptable performance. This definition can change with various operating points, e.g. production speed and grade changes. A simple example of changing operating points and their effects on a monitored variable is the outside temperature whose definition of normal clearly depends on the time of year. In this simple case where there is only one major affecting variable that can explain most of the normal variation to a certain degree, the definition of normal can be set manually for each operating point. E.g. the limits of normal variation can be set for every month or week of the year. However, if there are more major affecting variables or if the precision of the system simply needs improvement by adding more affecting variables, setting the limits manually becomes impractical very quickly.

The basic idea of conditional histograms is that the method learns the typical behaviour of a process by collecting normal-performance (reference) histograms representing normal situation for each operating point. In the figure 6 the month is the operating point. Essential part here is to set meaningful parameters (number, width and location of histogram bins) for the performance histogram and to identify the operating point of the process. Once these issues have been dealt with the algorithm learns the typical behaviour automatically and can detect changes in the monitored variable. In the figure the temperature measurements are classified into 11 bins that have 5 °C intervals from -25 °C to 30 °C and the operating points are classified by the month.

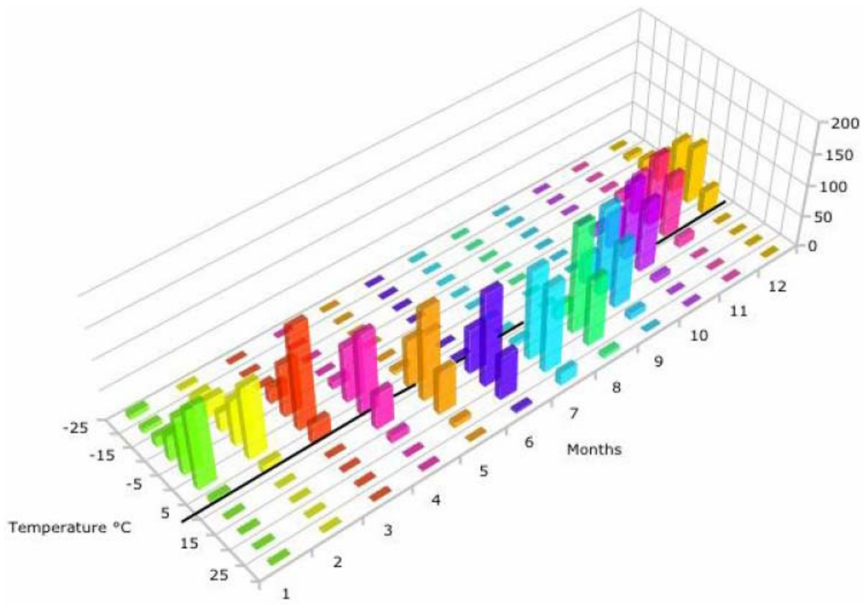

Figure 6 A conditional histogram presentation of temperatures in Helsinki [Fri07]

The operating points are determined from the selected affecting variables by means of clustering or by using fixed crisp or fuzzy limits. The selection of affecting variables is important and deep process knowledge is needed here. Adequate configuration of operating point limits or clustering parameters (number of clusters, forgetting factor etc.) is equally important as this will determine how effective the algorithm is in separating different operating points. The configuration of performance histogram parameters is also very important because the algorithm should be able to construct a meaningful histogram based on the hits to the bins. For example, the information value of a histogram is minimal if all the data samples happen to fall into the same histogram bin. Due to the importance of successful configuration additional tools were developed to assist the deployment phase. After the configuration is finished the reference histograms representing the normal operating conditions are automatically constructed for every operating point from the process history data. A forgetting factor can be applied to increase adaptation. Thus, with these features the method meets the economical requirements.

When the application is in online use the current performance histogram of the monitored variable is first created for a given time period (e.g. the last 24h period). Then the corresponding operating points and their active times during the period are determined from the affecting variables. This information is then used to build a combined reference histogram from all the reference histograms by time weighing them according to the active operating points of the reporting period. Now there are two histograms, one representing the measured performance from the given time period and one representing the "normal performance". These histograms are then compared to tell if the performance has changed and the user doesn't have to be aware of changing operating points during the reporting period. This also makes creating and comparing reports from longer time periods easy. The comparison of the histograms produces a single number which tells how the process is performing compared to its normal behaviour. The number tells if the measured histogram contains smaller or larger values compared to the reference distribution. For periodically executed modules this number is stored into the database which allows the user to monitor the development of the situation. Thus it can be argued that the method meets the usability requirements. The performance aspects have been verified in earlier projects.

Besides online monitoring the conditional histogram tool is also very useful as an analysis tool as it can store performance data from very long periods in a compact form and divided into different operating points. This tool will enable storing and using of large data sets in a convenient way, for example the effects of different variables on the performance variable can be analyzed. This of course requires that the configuration work has been successful. An example of this is given later.

Cluster analysis

One of the chosen methods was cluster analysis with a modified K-means algorithm described in [Hie03]. Generally speaking clustering is used to find similar sample groups, called clusters or classes, from a large data set. The learned cluster centers can also be used to represent the original data set resulting in significant data compression. In this application the cluster analysis modules are used for learning the typical process behaviour which can be used for identifying the process operating points but also for detection of deviations from normal behaviour. The information produced by the clustering modules can also be used for diagnostic purposes.

The methods for clustering are iterative algorithms that try to place code vectors into sample space so that they model the data set as accurately as possible. There are several different algorithms in the literature. The algorithm used here has a prefixed number of code vectors and tries to minimize the quantization error E

$$E(\boldsymbol{V}) = \frac{1}{2}\sum_{i=1}^{n} d(\boldsymbol{x}(i), \boldsymbol{v}^c)^2 , \tag{1}$$

where v^c is the nearest code vector for sample x(i) and V is the matrice formed by the code vectors. The distance d between sample x(i) and code vector v^c is usually determined by using Euclidean distance (2) or Mahalanobis distance (3)

$$d(x(i),v^c)^2=(x(i)-v^c)(x(i)-v^c)^T \tag{2}$$

$$d(x(i),v^c)^2=(x(i)-v^c)R_c^{-1}(x(i)-v^c)^T, \tag{3}$$

where R_c is the covariance matrice of the samples [Koh01, Jai88]. This distance can also be used to tell how different a new sample is from a previously learned situation. Together with appropriate threshold the distance, also known as residual, can be used to detect abnormal situations and generate alarms. A contribution plot tells how each input has contributed to the distance and improves usability. Thus the method satisfies the usability requirements.

Scaling of input signals is needed so that variables moving on numerically wide area don't dominate the clustering process. Although scaling of the inputs is required, the configuration is simple due to configuration tools that assist the deployment phase. To tackle the problem of non-spherical clusters with a simple solution it can be assumed that sufficient amount of clusters is enough to represent non-spherically-symmetrical clusters. The user simply specifies a sufficiently large number of clusters that will represent "in-control" situation after sufficient amount of learning. Other assumption is that the monitored process is in normal state most of the time and therefore the routine learns only the "in-control" model as the clusters with only few samples are omitted. The algorithm learns the normal positions of the clusters independently and as it can be assumed that the process is in normal operation most of the time, deviations from these normal positions can be interpreted as abnormal or new situations. The clustering modules have a forgetting factor to increase adaptivity and to lessen the effects of starting values. The number of clusters and the forgetting factor are user specified configuration parameters but choosing them does not present significant problems [Hie03]. Thus the economical requirements are met. The performance has been verified in earlier projects.

Following is a simple example of the idea how to utilize clustering technique to detect a fault in a fictional process. For clarity the process has only two measurements and four operating points which can be found automatically by the data clustering method. The process data and the learned clusters (operating points) from normal operating conditions can be observed in figure 7. The clusters represent normal and acceptable behaviour of the process. The process can be monitored by the residual signal. In the case of a fault or abnormal situation the process measurements start to deviate from the clusters and the residual will increase. This is illustrated in the figure 8 where the process starts oscillating around operating point 1. Typical reason for this kind of behaviour is sticking control valve.

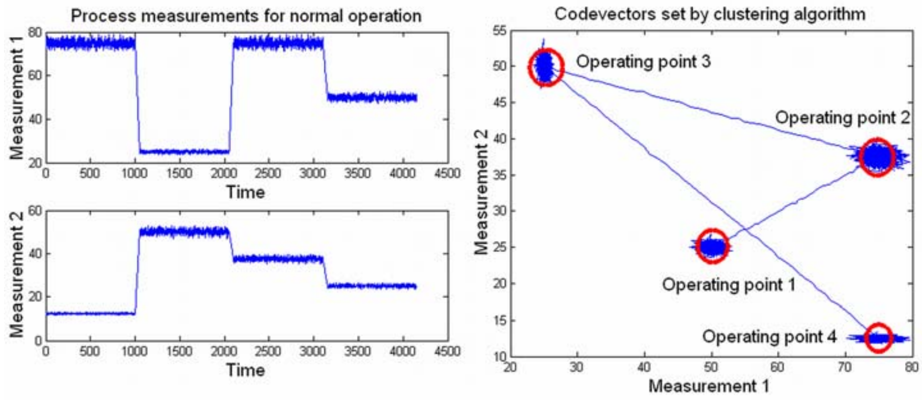

Figure 7 Data from normal operation and the learned cluster centres

In the residual plot of figure 8 abrupt but temporary spikes can be observed that cross over the alarm limit before the actual fault. These are caused by the changing operating points and are considered normal behaviour. Because the duration of such changes in normal situations is limited, alarms caused by changing operating points and random measurement noise can be filtered out.

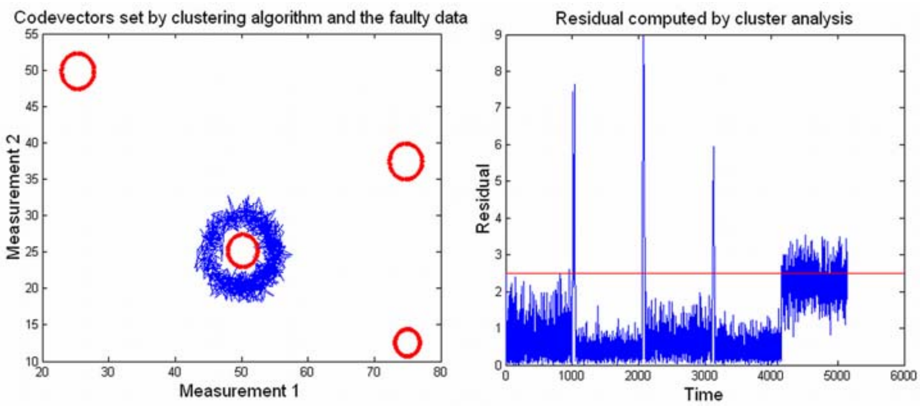

Figure 8 Data from abnormal operation, the learned cluster centres and the residual plot with corresponding alarm limit

The algorithm learns typical areas in the measurement space and is able to adapt to slow changes in the process. The speed of adaptation has to be set so that the algorithm doesn’t adapt to actual process faults too quickly but can adapt to slower changes, e.g. seasonal changes. Even with the use of adaptation slowly developing faults can also be detected by monitoring movements of the clusters instead of only using the residual.

Control loop performance monitoring tool

A control loop monitoring application, called the LoopBrowser [Mar99b], was used as a part of the monitoring application as it was an already validated commercial application that fit well in the overall system architecture. The application produces data about the performance of a control loop and represents them by six indicators:

- Control Speed Index (CSI) is a measure of the speed of the control loop. It is based on a process model, target speed and maximum theoretical speed of the control loop.
- Robustness Index (RI) is a measure of robustness of the control loop and is based on the same information as CSI.
- Control Travel Index (CTI) is a measure of how much the controller is doing work.
- Integral of Absolute Error (IAE) index is a dimensionless index scaled by a nominal value from the well-known IAE measure.
- Variability Index (VI) is measured from the minimum and maximum error value during a period of time. The value is scaled by nominal variability determined from typical behaviour process data. The index reveals errors with large amplitude and short duration better than the IAE index.
- Oscillation Index (OI) is a measure of how much the loop oscillates based on the Control Loop Performance Monitoring (CLPM) method [Häg94] for automatic detection of control loop oscillations.

The CSI and RI indexes require a process model so their use was not considered in this context.

The LoopBrowser application has installation tools to assist the configuration work and practically requires no maintenance after setup. Therefore the tool meets the set economical requirements. As it was previously designed as an independent tool, the LoopBrowser also has a separate reporting interface which can be used for creation of detailed control loop level reports. Besides the normal index calculation cycles, scheduled and manually executed summary reports can be created. In a typical report a summary of the control loop in question is shown representing the chosen reporting period. One part of the report presents the overall performance of the control loop and the rest is more detailed information about the causes of non-optimal performance. Also an online interface is available in which a graphical control measure called The Control Diamond is presented. The system also creates alarms from poorly performing control loops which can then be sent to responsible personnel by e-mail [Mar99b]. Although the application offers no real transparency, the produced information is intuitive and easy to interpret. Thus the application meets also the usability requirements sufficiently.

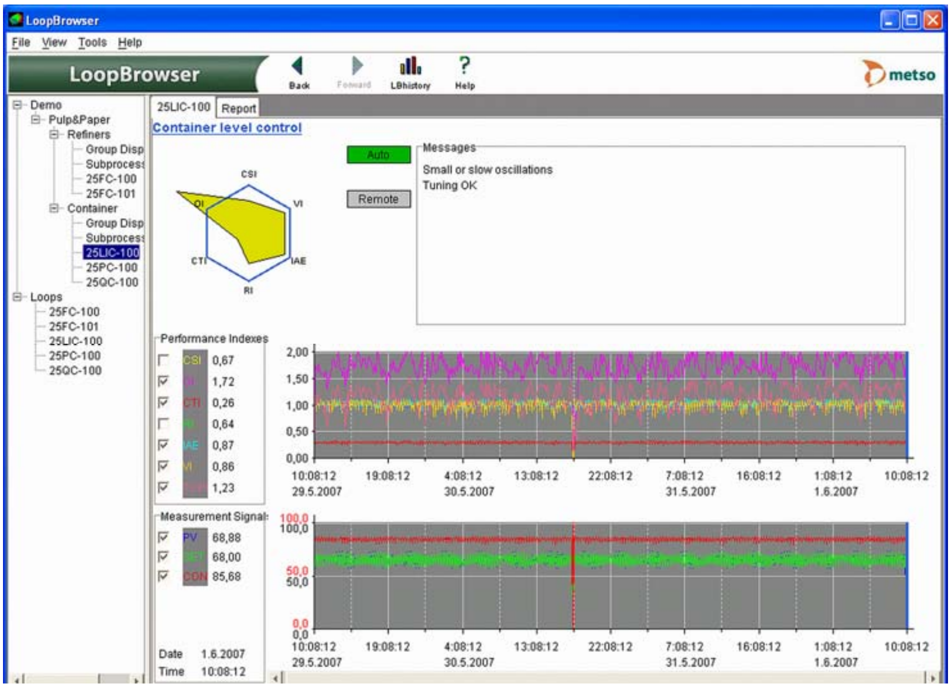

Figure 9 The LoopBrowser online user interface

4.2.2 Application hierarchy and the server features

The application is modular and the modules are functionally distributed, i.e. the setup of the modules corresponds to the physical process hierarchy. However, the current application is not hierarchical in functional sense. E.g. the time frames are not consistently shorter going down the module hierarchy levels and the levels are not interdependent (in most cases the higher level modules don't use information produced by the lower level modules and the modules don't control or intervene other modules). Thus there is no specific need for additional coordination between the modules. The decomposition into modules is purely heuristic and mainly based on the process hierarchy and available monitoring tools. E.g. some of the functions of intelligent field devices have inherent autonomy, are inherently distributed and are therefore seen as separate entities. However the characteristics of physical component and goal decomposition [Fin80] can be observed from the module hierarchy.

Modularity has advantages which enable adding, removing and modifying the modules with causing minimal interference to the other modules. This increases adaptability and scalability of the application and enables implementation of the modules with any kind of method, for example model based methods if there is need for one. The module setup corresponding to the process hierarchy increases usability as the overall state of the process can be intuitively observed from the user interface, e.g. which subprocesses are affected by an abnormal situation. The position of an alarming module in the process hierarchy will help to identify the source of possible problems and also assess the severity of the situation as the propagation and effects of disturbances can be roughly observed on the process hierarchy levels. These benefits and especially the need to utilize previously implemented modules where possible were the main reasons for choosing the system architecture. The modules and their position in the module hierarchy are presented in table 2. Also a simple example of a process hierarchy is presented in figure 10.

The highest level of the application is the plant level which is represented by the plant level performance indicators, conditional histograms and a clustering module used to identify the operating point. The operating point identification is executed in 1 minute cycles and the result is used in interpreting the performance indicators. Momentary production and energy consumption measurements are also calculated and stored in the relatively short 1 minute cycle because they also act as auxiliary variables to ease the load in generating long term reports. However interpretation of these indicators usually has much longer time span. Also many other auxiliary variables indicating things such as web break, idle running, broke percentage and the number of used mass tower are automatically calculated in varying but short cycles. The OEE-indicator uses many of the auxiliary variables and is automatically calculated in 24 hour cycles and 24 hour time window to create a trend. The usability measurements also use many of the auxiliary variables but have no specified execution cycles or calculation windows as they are calculated on request. Conditional histogram modules are periodically executed to interpret some of the plant level performance measurements and use 1-24 hour cycles and time windows. E.g. the quality measurements are received directly from the Quality Control System and are interpreted by conditional histogram modules in 1 hour cycles and 24 hour window. Some of the energy consumption and quality measurements are also used for subprocess performance measurements. Besides the time based executions the user can execute any calculation and create a report manually for any desired time period. Even though some of the calculations have much shorter execution cycles it can be said that the general time frame on the plant level is 1-24 hours.

The next level is the subprocess level where clustering and conditional histogram modules are used besides the subprocess performance measurements. The relative standard deviations are calculated in 1 minute cycles and 10 minute windows. The cutter and baling measurements are only calculated on request when creating a report and thus have no specified cycle or calculation window. Again, the user can execute any calculations and report creation for any desired period. The conditional histogram modules are used for interpreting the performance measurements and are executed in 10 minute to 1 hour cycles and 1-24 hour windows depending on the monitored measurement. At this level the clustering modules are used for detecting abnormal situations and are executed in 1 minute cycles. The purpose of using clustering modules instead of only using conditional histograms is to enable faster fault detection. This is due to the fact that certain amount of process history data is needed to build a histogram which results, in a way, to averaging of the measurements. This will lead to slower fault detection even if the execution frequency is high. The general time frame on this level can be said to be from 1 minute to 24 hours.

The next level in the concept is the control loop level where the LoopBrowser modules are applied. The modules are executed in 1 minute cycles and the typical time window needed for the calculations varies from 1-15 minutes [Met02] and thus the timeframe on this level is 1-15 minutes.

The lowest level in the monitoring concept is the field device level where intelligent field devices with device diagnostics are used. However, this level was not implemented during the first project. The time frame of this level would be 4-24 hours. It is obvious that the time frames in the module hierarchy don't correspond to the assumption associated with hierarchical systems stating that higher levels have longer time spans than the lower ones. This was the case especially for the field device level data collection. The technical reasons behind this and the use of field device level monitoring is described in chapter 5.

MODULE	PURPOSE
PLANT LEVEL	
Plant level performance measurement modules	Overall plant performance measurements
Clustering module	Operating point identification
Conditional histogram modules	Interpreting plant level performance measurements
SUBPROCESS LEVEL	
Subprocess performance measurement modules	Subprocess performance measurements
Conditional histogram modules	Interpreting subprocess performance measurements
Clustering modules	Identifying abnormal situations
CONTROL LOOP LEVEL	
Control loop performance modules (LoopBrowser)	Control loop performance measurements
FIELD DEVICE LEVEL	
Field device level performance modules	Field device level performance and ambient conditions

Table 2. Monitoring and performance measurement modules sorted according to process hierarchy position

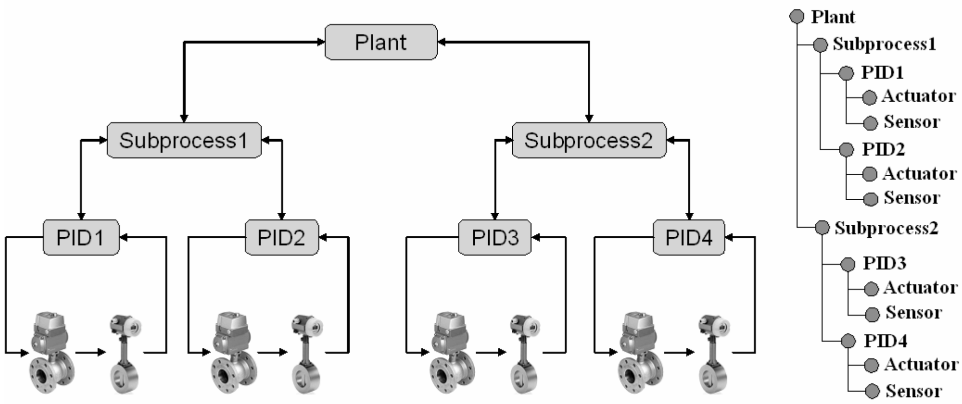

Figure 10 A simple example of a process hierarchy and corresponding module hierarchy

Separate user interfaces for online monitoring and reporting/analysis purposes are provided on the server. The real-time interface (figure 11) gives an insight how the process is performing at the moment and gives alarms if there are significant enough deviations in the process. The tree structure of the module setup can be seen on the left as a set of “traffic lights” which change their colour according to the results of the module. In the figure a conditional histogram module is chosen for closer inspection. From the top part of the figure it is easy to see that the current measured distribution clearly has greater values than the reference distribution. This can also be observed from the performance index below which indicates that the change has started about three days ago. If this is a new acceptable and repeating operating mode the algorithm will adapt to it given enough time but for now an alarm is given.

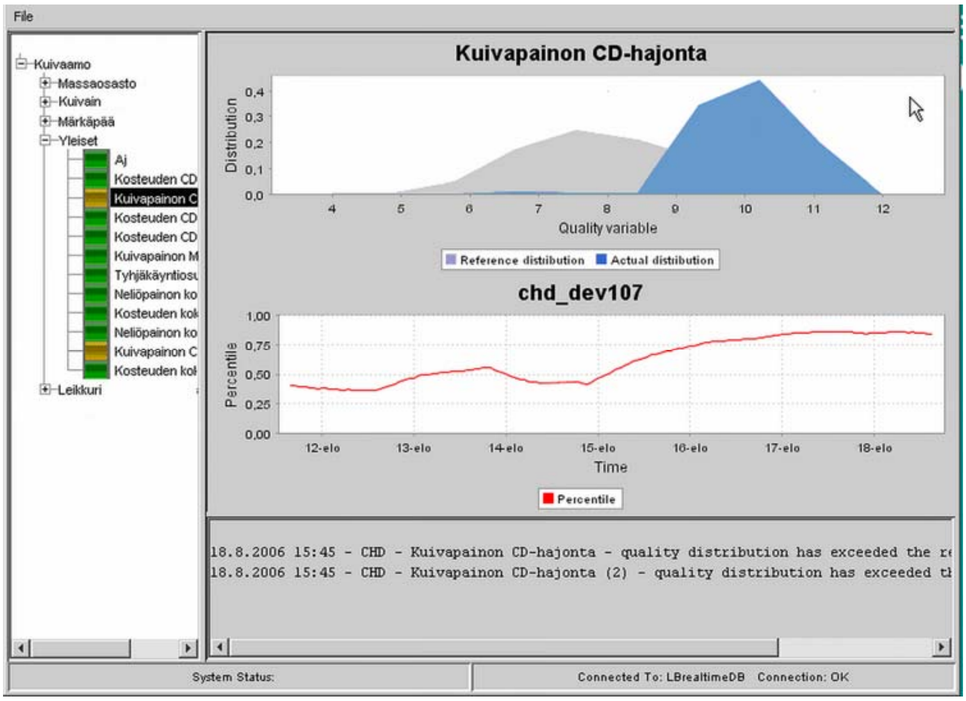

Figure 11 The online user interface with conditional histogram module selected. Below: the performance index, above: the reference distribution and the measured distribution

The reporting interface is web based and easy to access. Ready made report forms are provided in the interface and a user only has to select a report and define the desired reporting period. The reports were designed in co-operation with the plant personnel and they contain refined information from the process hierarchy levels. If necessary, adding or modifying the reports can be done with little work. Once a connection to the server has been established a web report can be opened in client machine’s regular browser where it is presented in html or pdf format. Major software updates only concern the server and there is no need for any specific client programs. The client machines only have to make sure that their browsers are compatible with the server. A regular, fairly recent version of any common browser is sufficient making software update management easy. In the figure 12 an OEE report for one month period is created. The module setup corresponding to the process hierarchy can be seen on the left side of the screen.

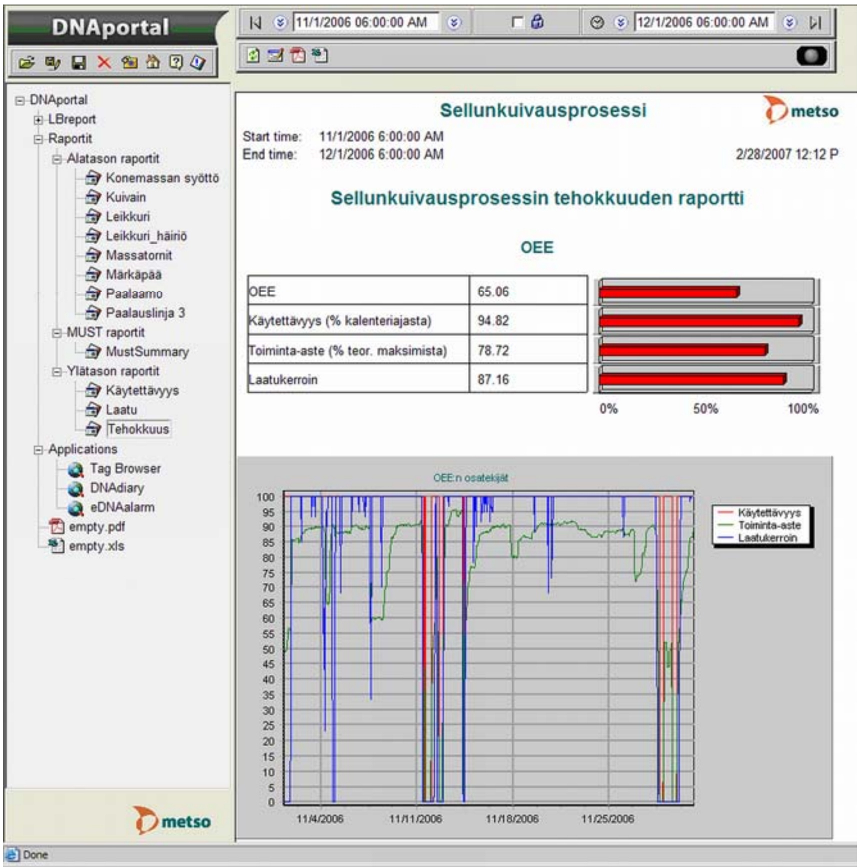

Figure 12 A web based report of the OOE indicator

The major modules, their relations and information flows are presented in figure 13. All of the monitoring modules use a common process database to read data and store results. Variables with fast dynamics that are used for the calculations are first stored in the process database's temporary buffer to ensure precise results. To make sure the database does not fill too quickly the variables in the buffer are compressed after the calculations. Besides the common process database the clustering modules and conditional histogram (MUST) modules use internal databases to store their internal states. These can be used in offline process analysis with additional tools which are normally in restricted use (dashed lines on the left). These relatively small separate databases are typically analyzed by transferring them to client machines of process experts and analyzing them with the additional tools. Besides being independent tools the clustering, LoopBrowser and performance measurement modules are also used as input generators for other modules. The dashed lines representing this should go through the process database but are drawn directly for the sake of clarity. The separate LoopBrowser interface is shown in the figure alongside with the real-time graphical user interface (GUI) and the generic reporting module. The generic reporting module can create reports on any variables in the process database once the report is designed and executed in the reporting module. Data access module provides remote access to the application. A VNC connection was used during the research to gain remote access to the server.

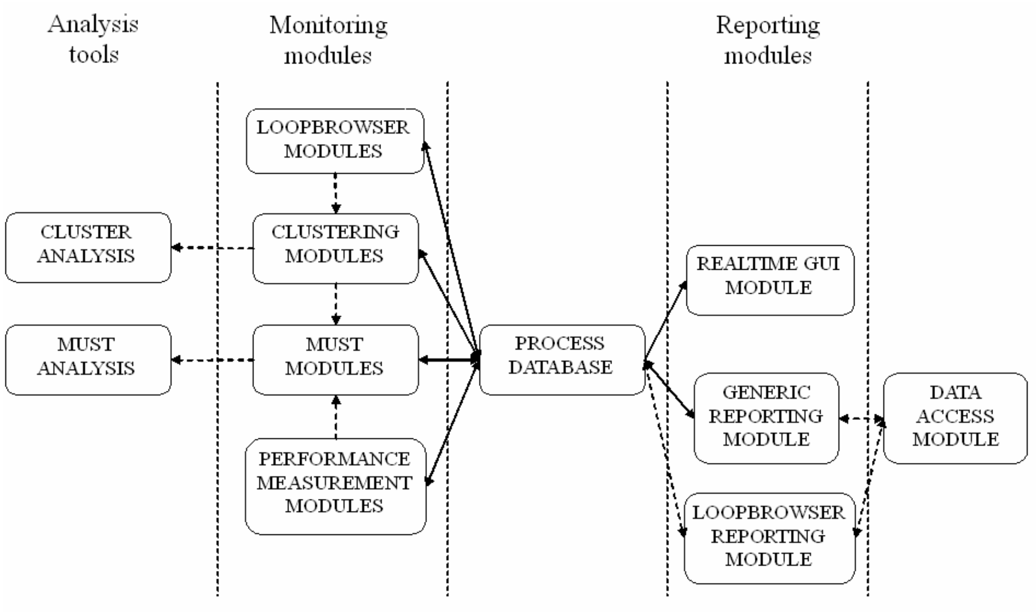

Figure 13 Module information flow

Most of the performance calculations and all of reports in the reporting interface are designed in a calculation environment which reads data from the process database, executes the calculations, stores the results to the database and/or presents them to the user. In figure 13 the calculation environment is included in the process database module. The calculation environment provides a reliable and transparent service for various calculation needs and the use of process history database allows storage and viewing of long time trends. Good transparency is achieved by a graphical interface, with a library of readymade function blocks and ability to program new function blocks. Adding of new calculations and possible modifications to the existing ones are also easy to carry out in the calculation environment. Visually and functionally the calculation environment resembles Matlab/Simulink environment. Simple calculations with faster execution cycles were implemented as SQL-scripts in the process database to avoid unnecessary computational load.

Because the economical aspects were critical and the performance of the application is strongly dependent on successful configuration work, additional configuration tools were developed although these are not presented in figure 13. The clustering modules need the number of clusters and nominal values (mean, variance) of inputs for scaling purposes. LoopBrowser modules also need nominal values to distinct real oscillation from process noise. For conditional histogram modules the user need to set the number and location of histogram bins for constructing the histograms. Tools were developed that assist by visualizing the process data and by enabling fast trial-and-error procedure. This of course requires that there is sufficient amount of process data already available in the database. To further assist the set up process for other pulp drying processes in the future, an alias-list of necessary measurements (tags) was created. Therefore in an optimal case deploying the application to another pulp drying process and making the necessary connections only requires updating of this alias-list by changing the position tag names. In practice, of course, at least the nominal values need updating also and the user needs to verify that the instrumentation requirements for the new process are met.

4.3 Use cases

Some use cases of the application are presented in this section. The results are discussed in more detail in chapter 6. The first use case illustrates decreased performance detection in a multivariable environment using conditional histograms. In the figure 14 a conditional histogram module detected that the cross direction (CD) variations of the dry weight have increased compared to the historical results obtained from same operating conditions. The higher plot indicates the trend of momentary performance measurement and lower one presents the corresponding measurement histogram and the reference histogram. The difficult thing is not to detect the change in the momentary performance measurement but to determine if the change is acceptable due to changing operating point. In this case it was found out that the change was indeed a new situation caused by certain test runs made by the operators. Thus the decision of the algorithm to notify of this change was correct because this was a never before seen situation.

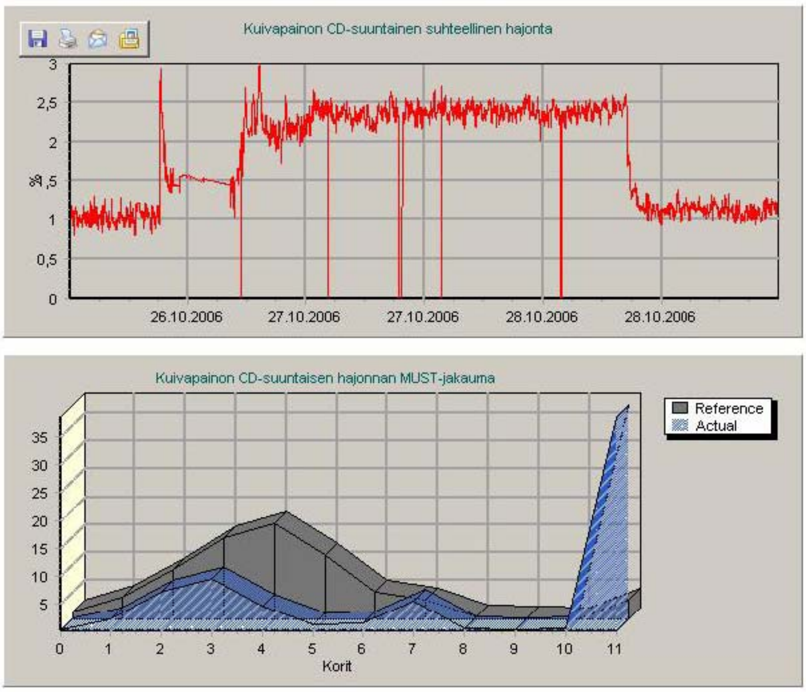

Figure 14 Above: performance measurement data, below: reference and measured distribution

The second use case illustrates how conditional histograms are useful for analysing affecting variables in a multivariable environment. The conditional histogram method can be helpful in analysing the significance of affecting variables for the monitored variable. In the figure 15 is an example of analysis tools included in the conditional histogram toolkit applied on a performance indicator and four affecting variables. The plots can be used to visualize the effects of affecting variables on the monitored indicator but deeper process understanding is of course needed for further analysis. The vertical lines describe the values of each affecting variable during current reporting period and the dashed lines describe where a certain percentage of the measured indicator values have been historically in respect to the affecting variables.

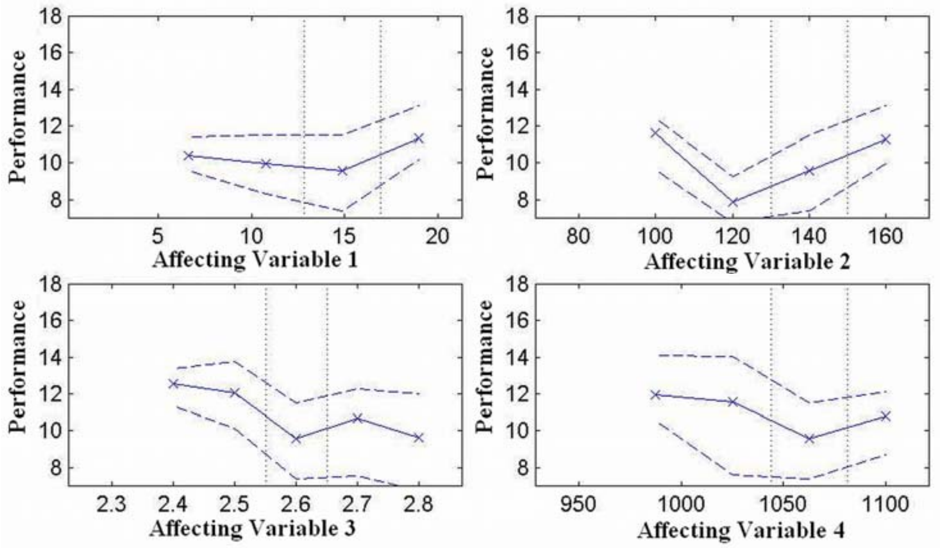

Figure 15 A conditional histogram analysis of a performance indicator with four affecting variables

The next example shows the comparative results of the conditional histogram algorithm when the operating point is identified using different methods. The algorithm calculates a number (scaled 0 to 1) informing if the current histogram contains smaller or larger values compared to the reference distribution. In the example of figure 16 the lines are the trends of this comparison using the clustering method to identify the operating point (the blue line) and using crisp limits for the affecting variables to identify the operating point (the red line). The results are similar suggesting that both methods are usable for the identification of operating points. Some of the operating points identified by using the clustering algorithm are presented in table 3. Altogether 30 identified adaptive operating points were considered to be sufficient to represent typical operating of the machine. The *tower #* refers to the used mass tower, *broke ratio* refers to the amount of broke pulp going to the headbox and *wrapping material* refers to producing wrapping material for the pulp bales. Others are self explanatory.

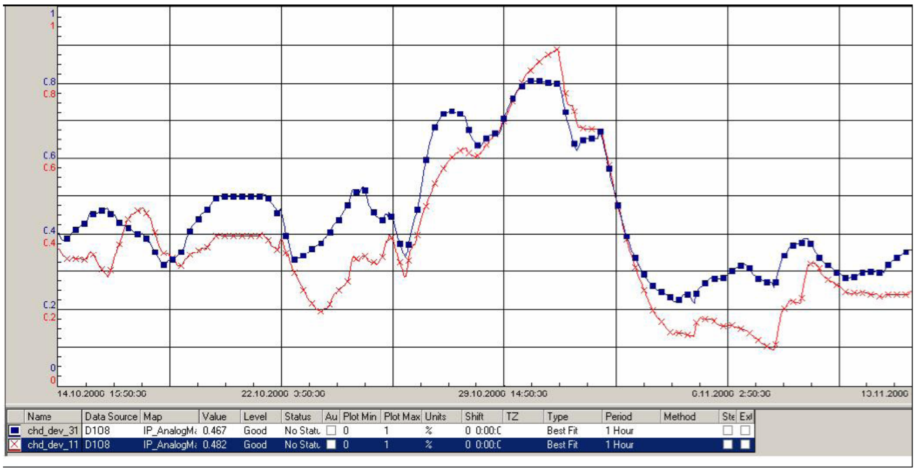

Figure 16 Comparison of conditional histogram results

Operating point name	Tower #	Dry weight [g/m^2]	Broke ratio [%]	Wire speed [m/s]	pH	Wrapping material	Web break
MT2, slow	2	1042,31	0,5	160,7	5,02	0	0
MT2, fast, high dry weight	2	1094,12	1,0	192,8	5,08	0	0
MT1, Wrapping material	1	1051,36	9,5	180,3	5,14	1	0
Web break	3	832,95	97,3	136,4	5,50	0	1
MT2, slow, lots of broke	2	1017,56	28,8	150,8	5,46	0	0

Table 3 Some identified operating points

In the next case a clustering module monitoring the headbox notifies of an abnormal situation where the pressure level in the headbox drops unexpectedly, figure 17. The change is controlled and initiated by the operators but nevertheless the monitoring module notifies of the new situation, as it is supposed to, before it learns the new operating point. When analysing the situation also quick periodical changes were observed in the pressure measurement. These however were so quick and seldom that the clustering algorithm had not been able to detect them. This points out that the method is not so suitable for detecting very quick and temporary disturbances. The reason for the disturbance was found out to be process interactions and actions were later taken to counter them.

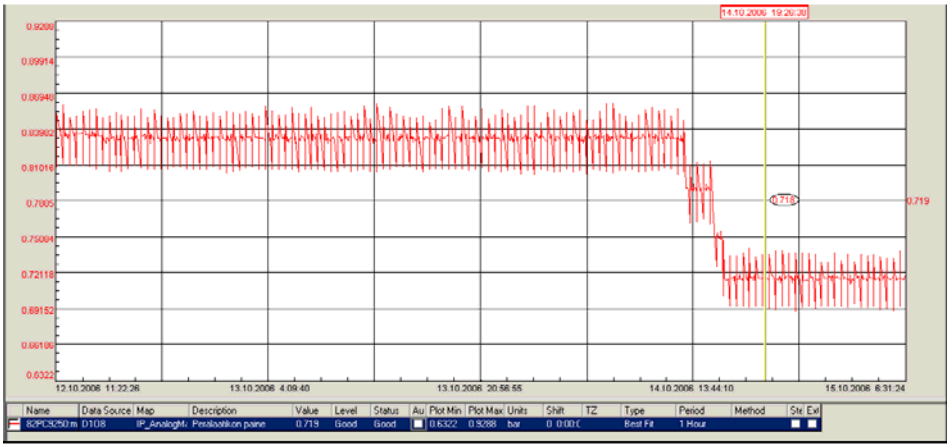

Figure 17 Pressure drop and disturbances in the headbox

The next case represents a problem noticed by a control loop level module. A valve shows signs of weak performance by causing oscillations, figure 18. The symptoms are a classic example of increased stiction in a valve. Due to the control error the controller’s integrator changes the control signal until the valve moves. Once the valve moves the controlled variable overshoots and the cycle starts again in the other direction. Later it was found out that the valve was also somewhat oversized. The scales of the signals are not the same in the figure.

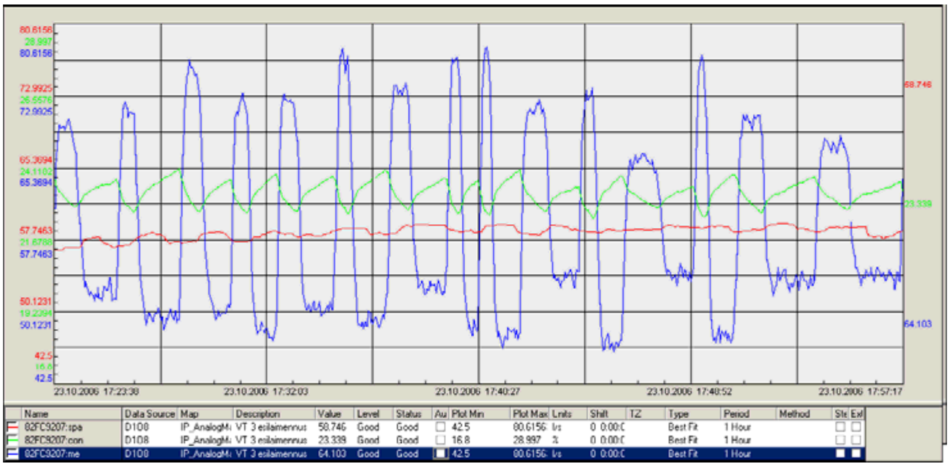

Figure 18 Setpoint (red), measurement (blue) and control (green) signals of an oscillating control loop.

5. INTEGRATION OF FIELD DEVICE LEVEL DIAGNOSTICS

The next chapter focuses on the second research problem by studying the possibilities of utilizing intelligent field devices as a part of the monitoring concept. The research problem is then answered in the analysis section 6.2 which along with sections 5.2, 5.3 also contains the third thesis contribution. The intelligent field devices extend the concept's scope by forming the lowest module level where self diagnostic features and performance measurements of Neles ND9000H series valve positioners were used. The higher levels of the concept remained the same although the highest plant level was not implemented and was replaced by feedback from the process operators and other plant personnel. Also some parts of the higher levels were not as extensive as described in chapter 4. A master thesis work [Leh08] was completed during this research. Related issues are also discussed in [Huo09a, Huo09b].

5.1 Neles ND9000 valve positioner

Neles ND9000 valve positioners and many of their features were utilized in this part of the work. The main features are presented in this section and the main components in figure 19. The principle of operation is as follows: the micro controller of the positioner (μC) reads the input (controller) signal, valve position sensor (α), pressure sensors (Ps, P1, P2) and spool valve position sensor (SPS). The difference between the input signal and the valve position sensor (α) measurement is detected by the control algorithm and a new value for prestage (PR) coil current is calculated based on the information. Changed current to the PR changes the pilot pressure to the spool valve which in turn moves the spool and the actuator pressures (P1, P2) change accordingly. The spool opens the flow to the driving side of the double diaphragm actuator and opens the flow out from the other side of the actuator. The increasing pressure will move the diaphragm piston. By using a control algorithm the μC modulates the PR-current from the steady state value until a new position of the actuator according to the input signal is reached. [Met]

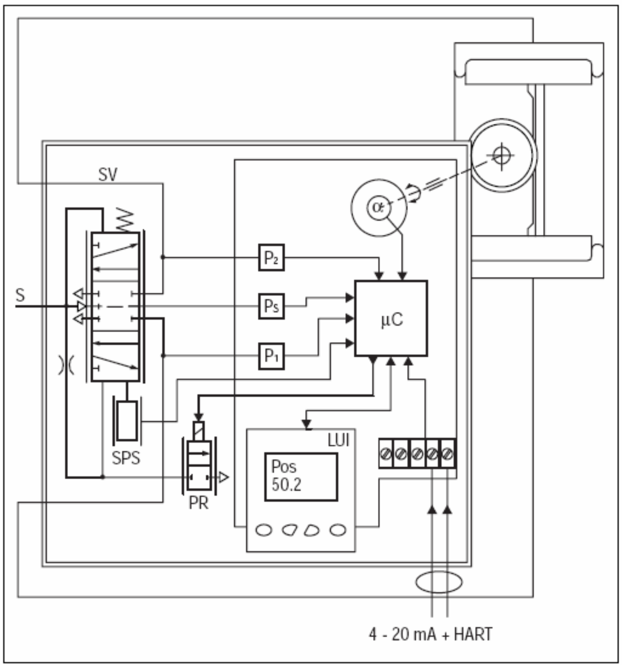

Figure 19 The principle of operation [Met]

The positioner has several embedded performance-, process condition- and "mileage" measurements. The performance indicators measure variables such as control deviation of the valve (controller output vs. actual valve position) in steady- and dynamic states separately, figure 20. The time spent in these states during the last 24 hours is also measured which is essential for detailed analysis. The mileage measurements indicate the usage and wear of the device by calculating the number of reversals and travel of the valve and spoolvalve. Besides the information offered about usage and wear, these measurements also contain information about the behaviour of the device. For example changes in reversals of a valve in respect to reversals of the setpoint can inform about changed behaviour. Histograms representing the historical operating area of the valve are also available. These are especially useful for determining if the sizing of the valve is correct.

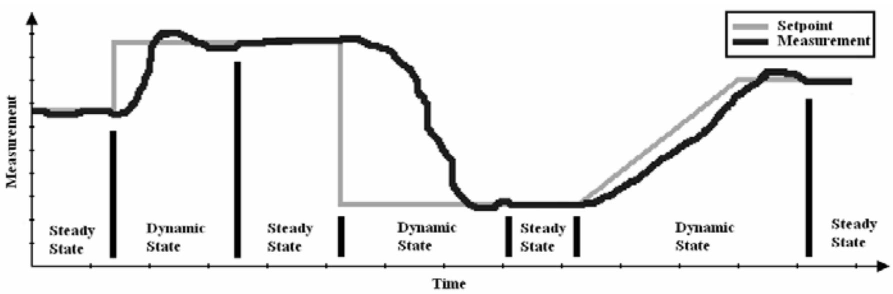

Figure 20 Steady State and Dynamic State deviations

The process condition measurements consist of variables like temperature, stiction, supply pressure and loads during steady state, opening and closing. E.g. steady state load is significant for single-acting actuators as it can be used to estimate the actuator spring force and spring state. The load and stiction measurements are pneumatic load measurements useful for estimating internal frictions in the valve package. Spool valve (used for controlling the actuator pressure on the both sides of the double diaphragm actuator) position is measured as a percentage of the movement range and is typically in the middle of the range. Changes in the trend may suggest pneumatic actuator leak, especially with single-acting actuators.

Some of the collected information is stored in the actual device while longer term data is stored in a central Fieldcare database using fieldbus connections. Thus the use of intelligent field devices adds to the fieldbus traffic as there is additional diagnostics information transmitted through the fieldbus. The database holds information about a number of positioners in a factory or process area. Besides the aforementioned embedded performance measurements advanced statistical diagnostic tools that operate in the database can be acquired by a customer but were not utilized in this research. Deviations in the measurements are detected by comparing the measurements to their reference values, which are typically determined automatically during first three months of operation. The tuning of the valve position controller is set by selecting one of the strategies which are aggressive, fast, optimum, stable and maximum stability. There are also automatic routines for calibration, tuning and testing the positioner. The software and communications are based on the FDT standard. The FDT frame application enables easy configuration and access to field device level information. The positioner is available for Foundation Fieldbus, HART and Profibus PA fieldbuses [Met06].

5.2 Developed monitoring system and analysis tools

The functionally distributed monitoring system consists of four levels; field device level, control loop level, subprocess level and human know-how level. A simplified block diagram of a control loop and the position of the three lowest module levels are presented in figure 21. Due to the feedback connections and interactions it is not necessarily easy to determine the reason for detected abnormal situations even in a single control loop. For example the tunings of both control loops in the figure have a significant effect on the quality of control.

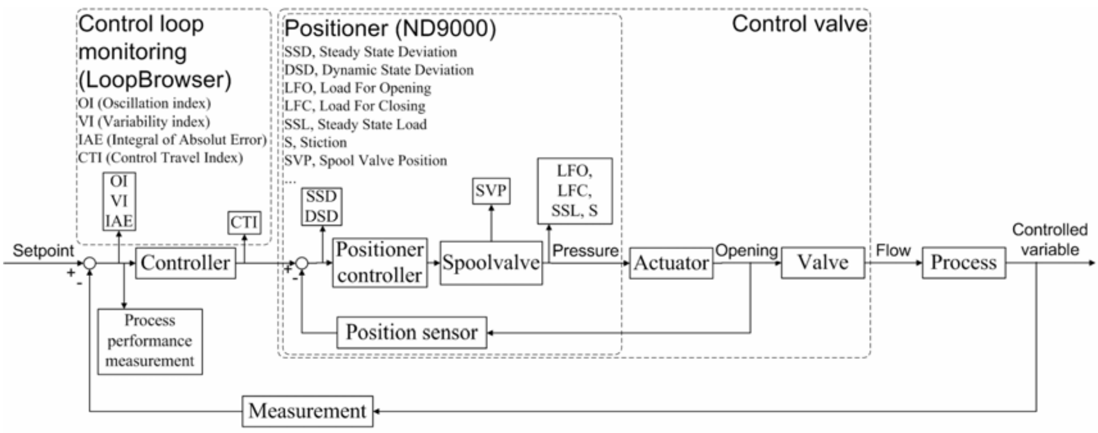

Figure 21 Simplified block diagram of control loops and measurement modules

The field device level consists of ND9000H valve positioners and their self diagnostic measurements presented in the previous section. The sampling frequency in the Fieldcare database storing the field device data was as low as six samples per day due to limited fieldbus capacity. The raw data in the positioners has higher sampling frequency but is stored only temporarily in the memory of the positioner where it can be accessed through a separate FDT interface.

The control loop level consists of control loop performance measurements calculated by the LoopBrowser application. The control loop performance monitoring application measures how well the control loop performs with measurements that represent oscillation, accuracy and work done by the controller. The LoopBrowser application was described in chapter 4.

The subprocess performance measurement level consists of modules calculating relative standard deviation of the controlled measurements. These are calculated by dividing the standard deviation by the average from a calculation period. The use of relative standard deviation was explained chapter 4. In this case relative standard deviation measurements of pulp consistency were used to describe non-ideal functioning of the stock preparation process.

The highest level is the human know-how level which seems to always be the hardest part to incorporate into any monitoring system and is often the weakest link in remote monitoring applications. There are several reasons for this. One of them is that usually there is only limited access to logged process events, e.g. the log can be only a hand written notebook in the control room. Another reason is the lack of time to participate on the part of the operators who really know the process. Therefore it can be difficult to get information about noticed abnormalities if the situation is not critical for the process.

As mentioned before, the use of distributed intelligent systems in general requires additional communication to coordinate all the entities in the system. However, as in this case the used monitoring modules are merely information producers and not directly concerned with control, there is no particular need for specific coordination of the modules. Instead this project can be seen as a preliminary study of developing the monitoring concept towards hierarchical functionality. Development of higher level functions requiring such hierarchical coordination is one of the possible future development directions.

The actual setup of the monitoring system was such that each control valve was monitored with the positioner diagnostics and corresponding control loop level measurements. These two lowest levels were available in every monitored control valve and were analyzed “simultaneously”, meaning simultaneous changes were searched for and analyzed. Some of these valves were located near a subprocess level measurement and in these cases the results were obtained by analyzing two or three of the lowest levels simultaneously. The setup process of the monitoring system was simple and fast. The embedded diagnostics are already present in the actuators as they are part of the basic configuration and no additional work was required. The rest of the applications is located in the process database hardware and their setup was fast due to partly automated setup process and readily available modules.

The first interest was to study if there were similarities between measurements on different levels to verify the validity of the project plan and the used methods. Similarities were found although there were some problems with the lack of abnormal situations where the similarities are easiest to find. In other words the plant and the equipment are so new and well performing that there is a lack of excitation (e.g. faults) for the monitoring system. This is a typical problem in development of monitoring systems as the development and research phases are usually limited in time and the monitored processes cannot be disturbed by experiments. Even though the problem was present also in this case, there were similarities found between the module levels first by manually investigating changes in the data and later by calculating correlations between the measurements. Also a Matlab routine for detecting simultaneous changes was created.

After some analysis it was concluded that in some cases there are clear correlations between device level and control loop level measurements and that sometimes the changes can be seen on the lower levels before they have an effect on the higher levels. Again, this kind of setup corresponding to the process hierarchy can be useful in determining the seriousness of a detected abnormality. The subprocess level measurements on the other hand were so stable that there were very few cases where significant similarities could be studied. However a few cases were found where the control loop and subprocess level measurements have interesting simultaneous changes. It was also concluded that there are certain measurement pairs on different levels that seem to correlate more than others. E.g. control loop level oscillations can be observed on the field device level as increase in spoolvalve movement and dynamic state indicators, which is a fairly trivial observation but further validates the used methods.

As there were numerous positioners to monitor, tools for detecting trends and deviations in positioner measurements were needed. Therefore a Matlab routine was developed that goes through all the positioner measurements, determines the corresponding nominal values during normal operation and checks if the measurements differ significantly from their nominal values. For example standard deviation and average values from a few month “normal period” can be used to create a definition of acceptable measurement values. Fixed nominal values were used for the measurements that are more independent of the operating environment. Also a routine that detects smaller but simultaneous changes in various measurements was created. Filtering was used so that outliers and single samples with deviation in a measurement don’t trigger an alarm immediately. Polynomial fitting routines were also used for trend detection and this seems to be a promising tool for detecting incipient problems. These routines were also useful in checking for similarities between the module levels during abnormal situations.

5.3 Use cases

The results were not as good as in the initial phases of the monitoring system development mainly due to the lack of interesting abnormal situations and consequently information-poor process data. There were however a number of interesting observations, some of which are presented here. After detecting and analyzing the observed abnormalities feedback from the plant personnel was acquired. In some cases the reasons for the observations were found to be related to degrading performance and maintenance actions. If there were no clear reasons found in the actions of the plant and maintenance personnel, process and device experts of Metso Automation were consulted and possible scenarios that could explain the situation were considered. Especially situations that could be explained by a malfunction or degrading equipment were of interest. Based on talks with the experts some recommended procedures for certain known events were also considered. For example if there is a clear change in the position of a positioner spoolvalve and the static state deviation of the valve simultaneously increases, it was determined that a pneumatic leak was a likely cause and the maintenance actions should start from inspecting if this is actually the case.

Most of the observations were from the field device level and results from incipient fault situations or maintenance actions. The effects of these observations were usually not visible on the higher module levels and becoming observable on the higher levels would have required a long time of fault progression. Devices with such observations would therefore be often classified as “normally functioning”. Also it was found out that in many cases the situation could be corrected by a simple device calibration and often the observed step-shaped changes were in fact the result of a device calibration. The results are further discussed in chapter 6.

5.3.1 Jamming valve

Only one actual fault was discovered in a positioner whose steady state deviation had slightly increased. In closer inspection slight stick-slip behaviour was also detected (target position and valve position signals, black and blue trends in figure 22). Check up revealed that the space between valve and actuator was full of pulp causing some jamming in the actuator mechanism. Even though the non-optimal functioning of the positioner posed no significant control problems, it was decided to remove the pulp to see the effects. The effects of the clean up can be observed from the device level before-after figures (22 and 23) where the stick-slip behaviour has been removed. The plots are generated directly from the positioner memory and thus the data has significantly shorter sampling periods than the Fieldcare database. The performance has also improved on the control loop level where the correlation between control and measurement signals has improved, figure 24.

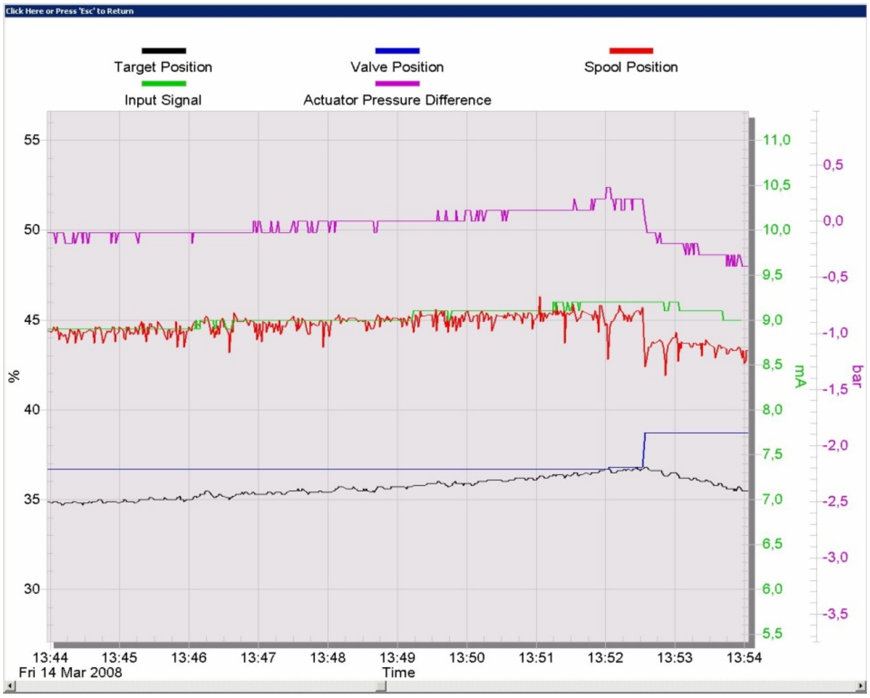

Figure 22 Slight stick-slip behaviour observable from the positioner measurements (black and blue trends)

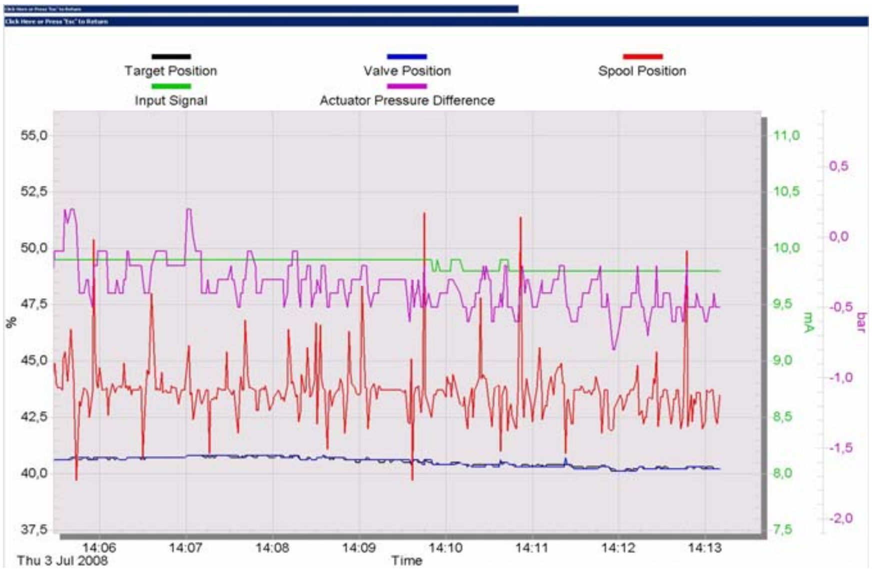

Figure 23 Stick-slip behaviour removed

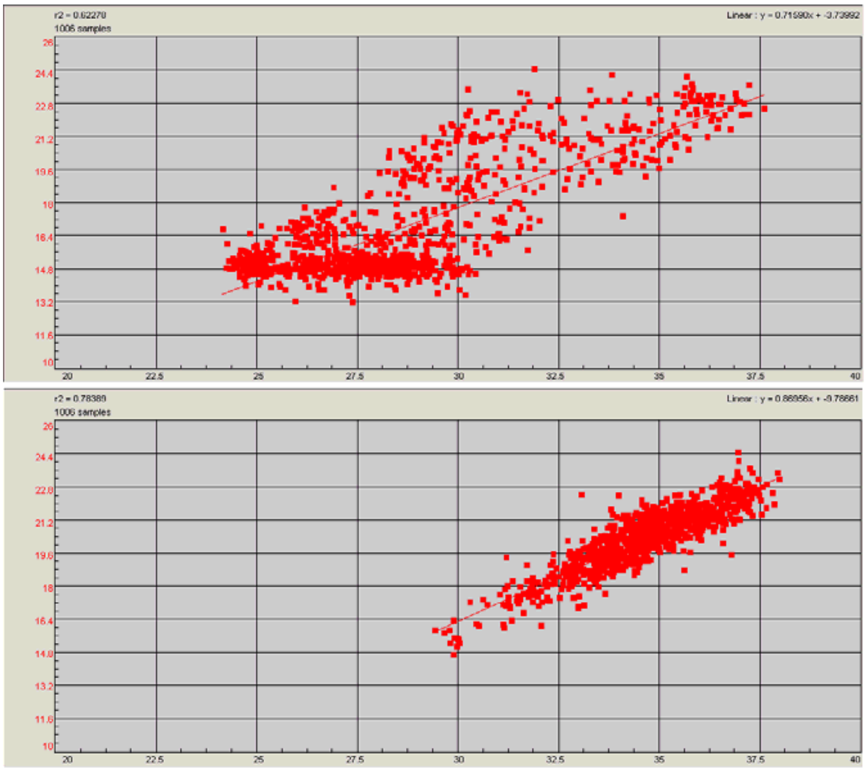

Figure 24 Improved (below) correlation between control and measurement signals on control loop level

5.3.2 Calibration of a positioner

In the next case a positioner showed early signs of weakening performance in the valve position deviations and drifting in the spoolvalve position measurement. The left side plots in figure 25 are the average position of the spoolvalve and standard deviation of the spoolvalve position. The right side plots are the average steady state deviation of the valve position and the average dynamic state deviation of the valve position. Although the effects were not yet visible on the control loop or subprocess level, it was decided to recalibrate the positioner and see if the problem would be solved. The calibration was done in beginning of October and the effects were clearly observable in the deviation and spoolvalve measurements. Even though there still seemed to be some minor drifting in the spoolvalve position the calibration seemed to correct the situation and the detected situation was dealt with before its effects could be clearly seen on higher levels. Although this was a good thing for the end-users and validates the used monitoring methods, the case also illustrates one of the problems in this study and in development of monitoring applications in general as the progression of faults can not be observed freely.

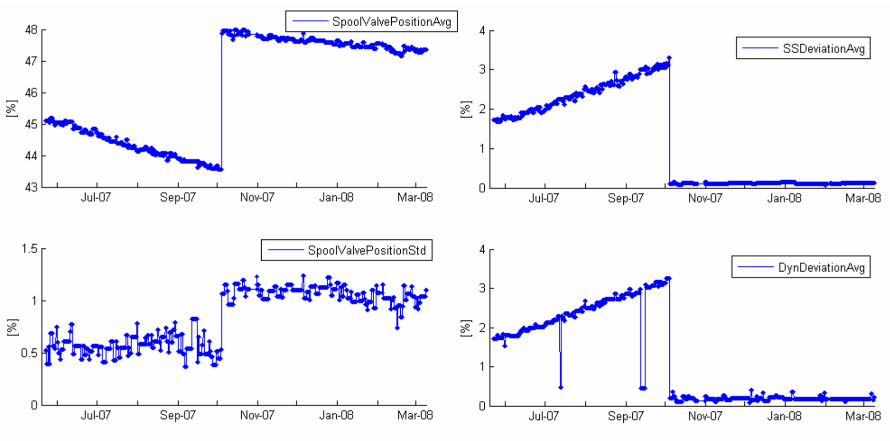

Figure 25 The effects of the calibration

5.3.3 Change of positioner tuning

During the research some of the positioners had their tuning parameters changed which had observable effects on some measurements. The positioner in figure 26 had its adaptability parameters changed in the beginning of July. The plots are the average spoolvalve position, standard deviation of the spoolvalve position and the average steady state deviation of the valve position. After the parameter change the spoolvalve and steady-state deviation measurements had a step-shaped change. After this some of the measurements had slight drifting but these are insignificant in scale and the values start to settle after awhile. The increase in the steady state deviation was insignificant but the behaviour of the valve stabilized otherwise which can be observed e.g. in the amount of spoolvalve movement.

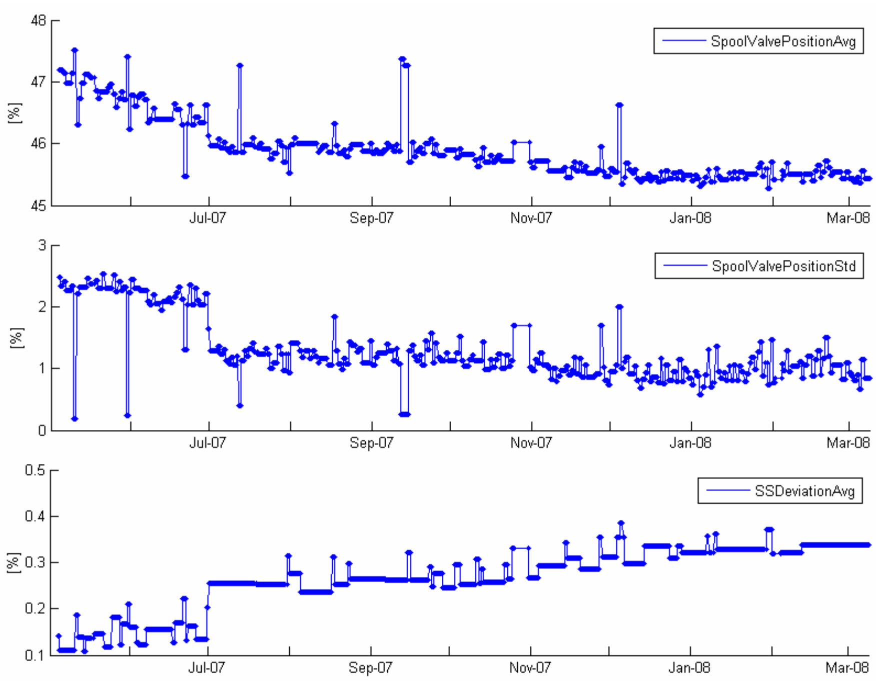

Figure 26 Effects of positioner parameter change

5.3.4 Changes observable on the lowest module levels

There were a few cases where the effects of changes could be observed on more than one module level. The first example of this is a dilution control loop which has observable changes on the field device level and control loop level. In the figure 27 there is a change in the average stroke lengths (travel/reversals) of setpoint and valve after the end of October. The changes could also be observed on all of the control loop level measurements (figure 28) which indicate that the control quality has improved. The cause for the change could be modifications of the positioner or control loop parameters but there was no documentation to support this. On the other hand the operating area of the valve had moved to a more favourable area.

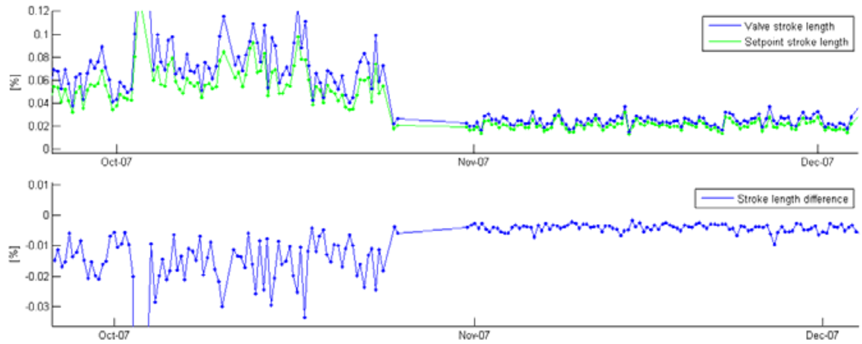

Figure 27 Effects on the field device level

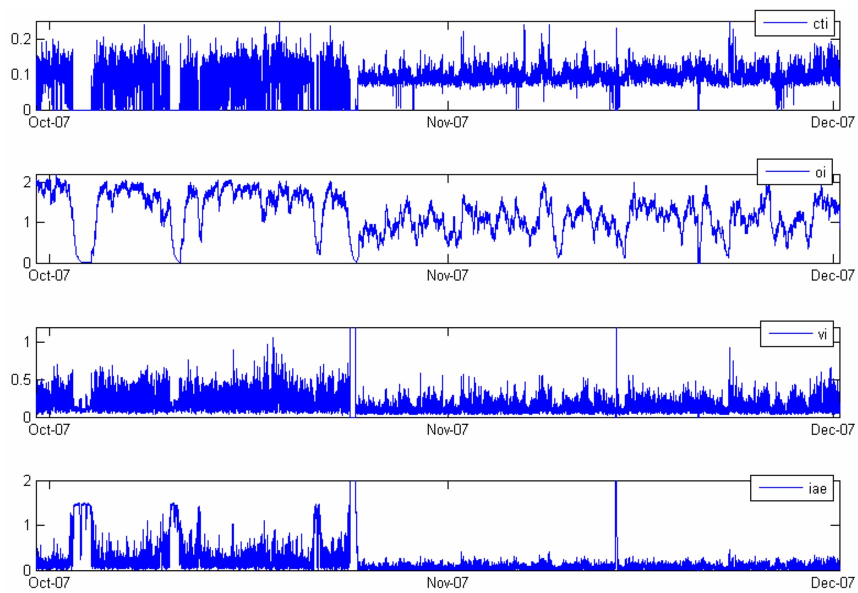

Figure 28 Control loop level performance

5.3.5 Changes observable on the higher module levels

The next case features one of the few cases where changes could simultaneously be observed on the higher levels. In the figures 29 and 30 the timescale is changed to measure days from the latest measurement due to shorter time period. On the control loop level there were clear changes between days -45 and -44, similar changes were also observable between days -40 and -39. The same events could also be observed on the subprocess module level (relative standard deviation of pulp consistency), figure 30. The cause is probably some event related to the operational issues but the exact reasons remained unsolved.

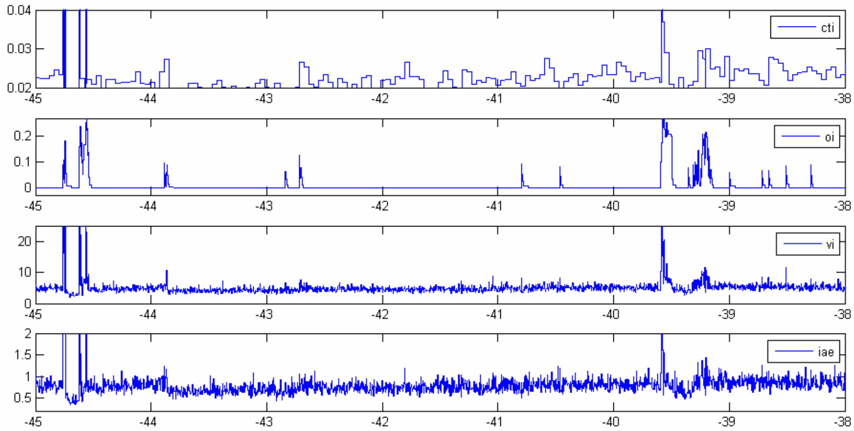

Figure 29 Control loop level performance

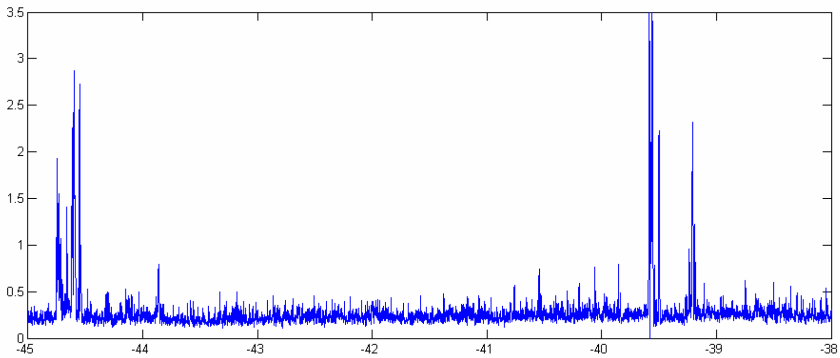

Figure 30 Subprocess level performance, relative standard deviation of pulp consistency

5.3.6 Oversized control valve

Next is a case where the effects of an oversized control valve could be observed. The valve was oversized due to limited availability of smaller control valves during the construction. The changes were visible in the field device and control loop level measurements (figures 31 and 32) after December when the performance clearly weakens. In figure 31 the highest plot is the time spent in dynamic state during the last 24 hours, the middle one the average spoolvalve position and the lowest one the average steady state deviation. The changes in the latter two are insignificant from operational point of view but are still observable.

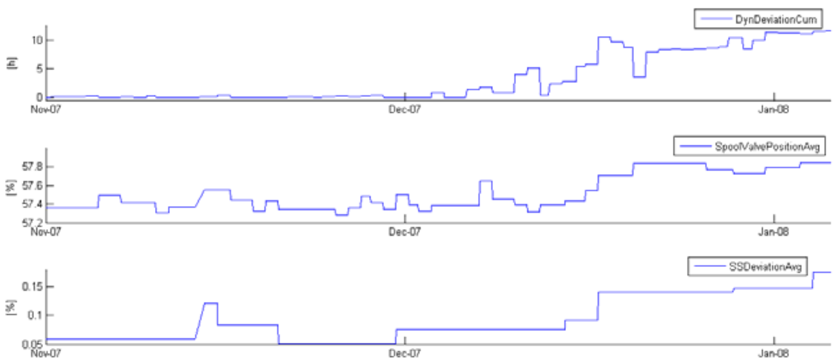

Figure 31 Effects on the field device level

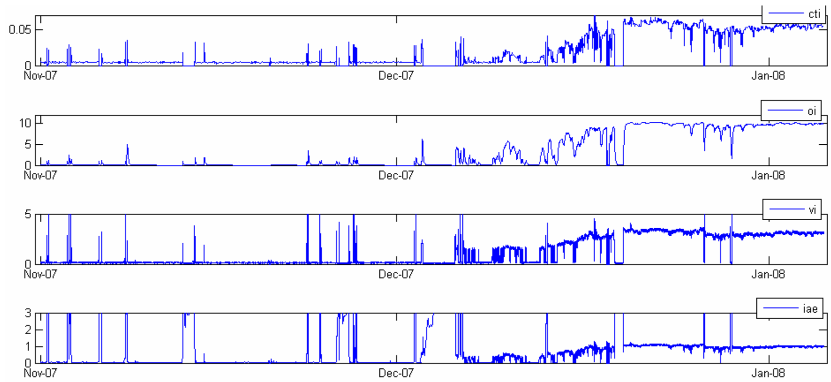

Figure 32 Effects on the control loop level

Weakened performance could be easily observed also from raw control loop process data, figure 33. The increased oscillation around the setpoint is clearly visible. It is also clear from the control signal that the valve is operating on the edge of its operating area where the control is not very stable and any small change in the process could cause the oscillation. This particular valve is used for temperature control and the change was caused by seasonal changes (time of year). The process returned to the previous state in the spring confirming the diagnosis. In this case the diagnosis was simple and easily achievable using only process data but during the next summer a decrease in the performance was noted again. As there were no obvious reasons found in the process data the plant personnel had to be consulted for the reasons. It was found out that due to maintenance work the flow to the cooler was temporarily cut off by a hand operated valve which caused the abnormal situation. There was no collected information available from the status of neither the hand valve nor the reasons for the change. This simple case demonstrates the importance and essential role of human participation in many diagnosis cases.

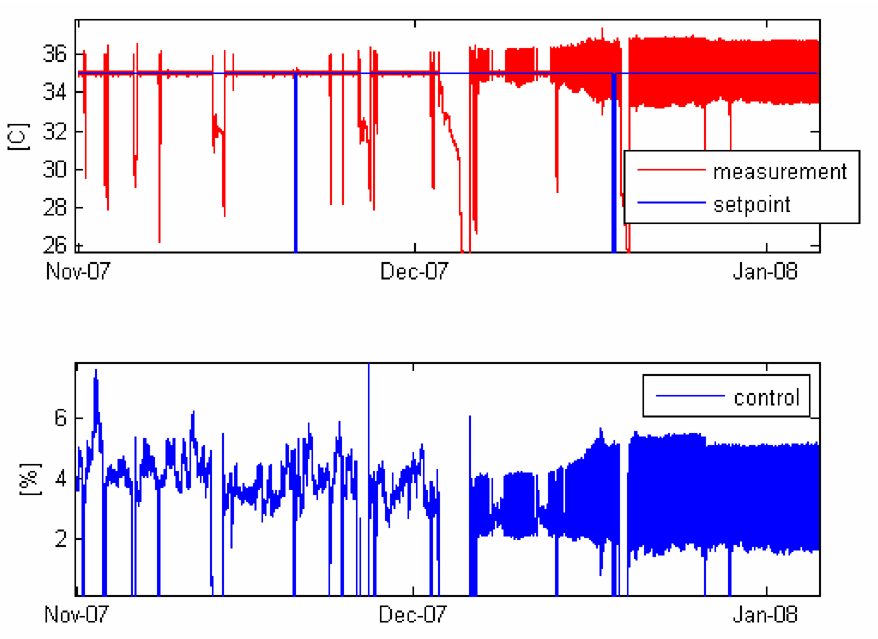

Figure 33 Raw process data

6. DISCUSSION OF RESULTS

The results are assessed against the research problem, research hypothesis and project goals in this chapter. The reasons affecting the results and possible future development directions are also discussed.

6.1 Results of the large scale monitoring system development

Overall the results were good. The requirements, system architecture, desirable features and tools were defined for a generic large scale monitoring system. The system was then implemented and deployed in a pulp drying process confirming the first research hypothesis. The project goals were also met. This section provides a more detailed analysis of the results.

The setup and configuration of the application took some time, but was the situation had improved since the 2004 pilot project due to accumulated experience and improved tools. The setup tools were also further developed during the setup process. The benefits of these tools will become significant when new installations of the application are made and are essential in achieving economical feasibility. Besides the quicker setup times the tools will improve the performance of the algorithms as also the quality of configuration work is likely to improve. The calculation environment offers transparency and the possibility to use offline data files which was very useful in speeding up the development process when there were problems with the remote connections. The used methods are able to adapt to changing environments and are applicable to any other processes with adequate measurements while the modular structure enables scalability and easy modifications to the system. Also the server solution and the chosen software techniques make software maintenance easily manageable. Therefore it can be argued that the economical requirements are achievable with appropriate tools and experience.

All the parties were pleased with the application and it is currently in active use. The performance measurement modules produce reliable and informative data about the process and together with the conditional histogram modules give warnings about abnormal situations in a challenging environment. The conditional histogram algorithm is a versatile tool for process monitoring and is applicable with little work using the setup tools. The method compresses process data sorted according to identified operating points and is therefore very useful tool for process development. The clustering algorithm is used for quicker problem detection and also seems to be a promising tool although more work is needed on this part. The algorithm is also used for identifying the operating point of the process and therefore used as an input generator for the conditional histogram algorithm. Although scaling of inputs and selection of some algorithm parameters is needed, the setup tools assist the configuration process. The control loop performance tool was already in use as an independent tool and its performance has been verified earlier. The information produced by the tool is also used as an input for the clustering modules. However, as it is designed to monitor individual control loops ignoring process interactions, its use as a part of a large scale monitoring application requires a lot of human participation or additional automation of diagnostics. The used methods make no assumptions about linearity of the monitored process and automatically learn the typical behaviour of the processes resulting in more meaningful alarms compared to static alarm limits. The application is currently used mainly by process experts in offline process analysis use. More extensive exploitation of the online features in quick decision making would require improvements in the performance by fine tuning of the monitoring parameters and embedding of additional intelligence. Also the system performance would require more thorough validation and system integration. Thus, even though the performance requirements and desirable features were not all entirely met, the performance is at an acceptable level and the requirements should be achievable with further development.

The end-user feedback was also very positive concerning the usability of the application. The users of the application were satisfied especially with the web based reporting tools. For example the plant personnel and the partners typically hold a monthly meeting where they discuss how the agreed goals were met, how the future goals should be set and what measures should be taken to achieve these goals. It is important that all the parties have access to the data beforehand so the participants can present problems they consider important and have the information ready to verify this. The whole procedure is more effective in a lot of ways when the participants can seek the information they are interested in and present it in a clear form. Transparency and clarity of the application are also features that have received positive feedback. For example the effects of corrective actions or changing operating procedures can be easily verified from the monitoring system.

The used methods can also learn previously unknown and unexpected dependencies between process variables. Especially the conditional histogram tools are useful also for process development as they enable examination of large amounts of process data in a convenient way. For example the interactions of variables can be examined by a statistical analysis tool with visualization features. A simple way to exploit such information is to optimize processes by avoiding non-optimal operating points revealed by the tools. This is especially useful in an environment where different work shifts with different operators tend to operate the equipment slightly differently even while producing the same end product. Also the less sophisticated indicators, such as device specific alarm classifications and counters, have provided surprisingly valuable information as no proper statistics of them were available before due to the distributed nature of the information. This is very useful in identifying problematic equipment and bottlenecks of the process. The application can tell for example which equipment causes the most alarms, how long the problems persist and how much production is lost. Once deviations in the fault detection and/or diagnostic modules are detected the severity of the situation can be roughly assessed by examining how high in the process hierarchy the deviation can be detected. On the other hand alarms in the lower level modules direct the user towards reasons for the deviations. The module setup corresponding to the process hierarchy therefore gives a good general view of the process. Thus it was concluded that the usability requirements are sufficiently met.

With all of the main points in the requirement list sufficiently met, the first research hypothesis was confirmed. It was also concluded that the set of requirements and desirable features (section 4.1) for the application is applicable to large scale monitoring applications in general although not all of the specified features are necessarily mandatory. Furthermore, not all of the points made in section 4.1 are directly related to the application, some are instead only observations regarding the application development and maintenance on a more general level.

6.1.1 Development ideas

There are several possible development directions. While the developed application forms a framework enabling relatively fast and cheap delivery to other pulp drying processes it would be interesting to deploy the application to other process environments, e.g. a power plant environment. Implementing the existing application to a new process environment, even without further developments to the actual monitoring concept, would produce a framework for the new environment but also valuable information about the versatility of the application and experience of using the configuration tools. This would aid in further development of the application. The possibilities of defining standards for benchmarking purposes should also be an interesting research subject.

The possibility of incorporating field device diagnostics to the application was later studied and is an attractive topic for further research. Properly implemented integration of device level diagnostics with higher level diagnostic modules should increase the intelligence level of the system through better utilization of information regarding the operating environment of the devices. This would enable more accurate monitoring and analysis of individual process devices. The role of proper integration will be essential in utilizing field device level diagnostics in larger monitoring systems.

Possibilities of utilizing the benefits of integration between information systems should also be considered. In this respect SOA provides organizations a flexible and cost effective way to build info systems and applications. In an industrial environment SOA enables integration of existing applications such as DCS and Quality Control System (QCS) to ERP systems. This kind of higher level integration opens up possibilities such as communications between the monitoring application and MES/ERP systems. This is useful to the monitoring application also; for example the OEE indicator can be wrongly interpreted if planned maintenance shutdowns are not considered or there are no orders to fulfill. Integration with DCS and other systems is also possible but one must be careful not to build overly complicated systems that could compromise the functioning of other systems critical for the plant operation.

The developed large scale monitoring application consists of various monitoring modules but to a large degree their main focus is on detection of abnormal situations in individual process parts and devices. For example the control loops are monitored by numerous LoopBrowser and clustering modules but more detailed analysis is still in many ways manual work. Thus the methods and practices still need some work if more automatic root-cause analysis is desired. Such developments are very desirable for plant-wide monitoring applications designed for online use, for example operator support systems. If such developments are to be made in the future, the information presented in section 2.3.4 will gain more practical significance in respect to the applied part of this work.

6.2 Results of intelligent field device utilization

The possibilities of exploiting intelligent field devices as a part of the developed monitoring system were also studied but due to some problems the research hypothesis could only be partially confirmed. However, some positive observations supporting the hypothesis were obtained. It was concluded that intelligent field devices are suitable for the developed monitoring concept but additional work and system development are needed for industrial deployment. Similarly, the project goals were only partially reached as the faced problems limited the results of the event observability study and the consequent study concerning the usability of the observations. However, links between certain observations and events could be established and with further research this information should be usable for process monitoring purposes. Also it could be confirmed that certainly the maintenance of individual field devices is easier with the embedded monitoring features if they are utilized properly.

Some of the faced problems were irresolvable in the available timeframe, e.g. the lack of abnormal situations was one such problem and the biggest one for the project. One of the other problems was that numerical analysis of such control loops and positioners that are often in manual or force mode requires some additional work. Proper interpretation of performance measurements in such control loops requires acknowledgement of controller modes, e.g. whether the controller is in force, manual or auto mode. Automatic analysis has to neglect such periods or take them into consideration in some other way. Also valves that are used as a shutoff valves most of the time and are infrequently in control use have issues in sampling. For example it is impossible to get stiction readings if the valve is not moving. Similar situation occurs with loops that are only active periodically. Some data related problems were also encountered in the correlation calculations. Irrelevant abnormal situations, such as shutdowns, often cause relatively huge changes in the measurements and by so can dominate the calculations. In general the discontinuous operating of the machine resulted in some additional work as data from these situations had to be filtered out or otherwise considered before analysis took place.

Dissimilar sampling periods on different module levels also posed some problems. The sampling period on the lowest module level was measured in hours while some of the sampling periods on the higher levels are measured in minutes and seconds. The low sample rate is caused by somewhat limited fieldbus capacity reserved for diagnostic purposes. Also the fact that the field device database was a separate entity and not integrated into the process database complicated matters.

The highest level of monitoring, human know-how, presented some problems as expected. This was emphasized by the use of remote diagnostics (no one from the research group was actually on-site during the research). The data was collected periodically and analyzed mainly in off-line mode resulting in a delay from an abnormal event to the detection. This is problematic as it can be impossible to find out what was the cause for a minor abnormality that happened weeks ago.

Even though the research was not as successful as the first one, it was successful as a preliminary study of how to integrate and utilize diagnostic and performance information from different levels of a functionally distributed process monitoring system. General knowledge was gained on what is needed for future developments and what are the current weaknesses of the concept. As a result of the study it can be argued that intelligent field devices fit the developed monitoring concept. However, with the limited results the second research hypothesis is only partially confirmed and further work is needed on the subject.

6.2.1 Development ideas

The research showed that there is unexploited potential in intelligent field devices for process monitoring purposes. Further integration and exploitation of device level diagnostics opens up new possibilities for process monitoring. E.g. higher level diagnostics could combine device and control loop level information to enable better distinction between control loop level problems (poor tuning etc.) and device failures. However at the same time increasing use of distributed systems complicates the utilization of process data as the information can be spread on multiple devices making the data less cohesive.

Due to the more complicated nature of distributed data further research and work is required to improve the utilization rate of the produced information. Better integration of information sources is the first requirement. At the moment simultaneous analysis of the separate databases still requires manual work and thus better usability through better integration is needed. Also better incorporation of the human know-how requires better tools and practices. This is a common problem and needs to be properly resolved before the full potential of remote diagnostics can be exploited. The integration of lower level databases is a more technical question but remains an unresolved problem. Also routines and work practices of how to systematically exploit the information should be developed. E.g. it should be defined how the data is utilized and people assigned to carry it out. It is often the case that information from monitoring/diagnostic systems is only utilized after problem situations which is not very efficient usage of the information. However it can be argued that even with their potential not fully exploited, intelligent field devices offer new and effective ways of improving process monitoring.

Better use of the integrated information sources will be achieved if the sampling periods are more similar requiring an increase in the sampling frequency on the field device level. However, this could result in unacceptable congestion in the fieldbuses if the transfer capacity is not sufficient. This invokes the question of increasing fieldbus capacity and/or embedding additional functionalities to the devices in the future as they have access to much more detailed measurement and diagnostic information than it is currently possible to transmit through fieldbuses to databases. If additional measurement information, such as flow and/or pressure, is delivered to the positioners, new diagnostics features could be added to them. E.g. the diagnostics could monitor inconsistencies in the valve's pneumatic pressure-opening-flow signature. This would not necessarily add to the fieldbus traffic if the measurements and the positioner are connected to the same fieldbus segment transmitting the measurement information to the control system. Some limited amount of diagnostics data could then be transmitted and stored in a database while detected abnormalities would be reported immediately. Such processing of larger amounts of data in the devices would require more embedded memory and processing capacity. However, the requirement for more processing capacity can be impossible to realize using current technology if the power is fed to the field device through fieldbus wiring. At the moment the accuracy and resolution of the device diagnostics, especially the analysis of momentary events, are somewhat limited by the low sampling frequency. In short term more detailed information can be accessed in the memory of the devices but much of this information is lost in when the data is compressed and stored in the database.

It would be very useful if further research could be done in a more controlled environment enabling experiments, e.g. inducing faults to the process. Further research effort is needed to define what variables are needed for efficient monitoring on different module levels and these variables should be calculated and monitored automatically. The effects of various process events should also be studied in more detail. If the effects of most common faults or abnormal events on the module levels are known, a further developed monitoring system would be helpful in identifying various situations. This would require construction of an expert or an agent system possessing knowledge of the events and corresponding observations on the module levels. The system could then point out the most likely fault candidates to the operators. Potentially even very specific reasons (e.g. a pneumatic leak) for a detected problem could be automatically determined and recommendations for corrective actions (e.g. replacing a gasket) given. Creating such a support system that is also credible would require a significant amount of further research. If some significant situations could be reliably identified, the system could propose some measures to be taken to achieve fault tolerance. At this point such a system could be feasible especially for limited systems, such as control loops with intelligent positioners, which enable reusability and large scale deployment without extensive customizing work. For example, a sticky valve causing oscillations could be temporarily be dealt with by changing the positioner's control parameters in the inner loop or by changing the control scheme from PI- to P-control as described earlier. Such automated actions would also require better coordination and communication between the participating entities and would develop the concept towards a hierarchical system where the functionality of the modules would be hierarchically determined. Provided that in the future the intelligent field devices meet the needed technical requirements sufficiently, such systems could also be implemented in a truly distributed fashion on the principles of agent technology, i.e.

distributed autonomous software agents negotiating about possible faults and counter measures. This could improve the fault identification part through better flexibility and adaptability. It also might reduce the amount of additional diagnostics related fieldbus traffic as a whole if more detailed diagnostic information is only transmitted on request. The communication features of agents could be exploited also in system integration. On the other hand, such use of autonomous agents requires more communication and coordination between the entities. It can also increase the amount of uncertainty about the overall state of the system making it more difficult for the industry to accept the concept.

In general, setting of nominal values representing normal operation can be problematic in monitoring applications as the operating environment often affects the definition of normal operation. This is often neglected and default values are used to avoid excessive configuration work. This practice, however, impairs the performance of monitoring applications as the alarm values are set conservatively to avoid high number of false alarms. The ND9000 series positioners determine some of their nominal values by using collected data from the first three months of operation. This practice is not always reasonable since especially in green field sites the production doesn’t necessarily reach the normal state during the first months. Also it is possible that the new device itself doesn’t function properly from the start. These situations may result in inappropriate nominal values. One possible way to alleviate this problem could be to form a database consisting of reasonable nominal values in different operating environments. The operating environments should be classified according to most significant variables, such as valve size, valve type, average flow and pressure through the valve and so on. This would also require systematic use of process design data and measurement information from a large set of positioners classified according the operating environment. Also routines and tools for automatically setting the nominal values to the positioners during commissioning would be needed. Some preliminary analysis of the device level measurements divided into groups based on valve types and sizes was done during the research. This revealed fairly trivial information, for example that the force required to move a valve depends on the type and size of the valve. However, it was quickly determined that the available time was not sufficient for meaningful analysis and development. Even though in the end this research direction was not particularly useful for this research, this kind of information could possibly be used in monitoring applications for determining nominal values more accurately.

7. SUMMARY AND CONCLUSIONS

This thesis was a part of the EMPRONET project whose goal was to study the use of embedded control and monitoring systems in production machine networks. The theoretical part of the work included surveys of intelligent field devices and various advanced monitoring methods. Also decentralization along with current communication and integration techniques were investigated. The applied part of the thesis continued the work of the 2004 master thesis work ‘Performance Monitoring of Pulp Drying Process’ [Huo04] and consisted of two subprojects. The first subproject consisted of developing a generic large scale process monitoring application and implementing it to a pulp drying process. The second subproject studied the possibilities of utilizing intelligent field devices as a part of the developed large scale monitoring concept.

7.1 Life cycle management and intelligent field devices

The role of PLM is increasing also in process automation. In order to effectively meet the requirements of modern production it is increasingly important to consider asset management during the whole production equipment life cycle. Thus, the basic concepts of reliability, maintenance and life cycle management were discussed. Although the definition of life cycle depends on the point-of-view, both the manufacturers and end-users of process equipment benefit from properly implemented life cycle management. This requires appropriate tools for supporting PLM in process industry environment. PLM tools can be used e.g. for commissioning and collecting feedback data to product development as well as for daily process management. However, effective PLM of modern production assets not only requires modern automation solutions, but also changes in management philosophies, work practices and business models.

With the rise of the life cycle philosophies the suppliers need to consider how factories will be run in the future and the implications of this fundamental transformation to not only their products but also to their business models. The concept of Customer Relationship Management needs to change from one that effectively ends with the sale to one that promotes an ongoing involvement with the customer and builds on continued service. The manufacturers need to change their existing business models and internal structure to deal with the new customer relationship. New services that support such continued customer relationship can be created and product differentiation achieved by taking advantage of the developing technologies [Mil04].

Acknowledging the possibilities enabled by the favourable technical developments the field device manufacturers have started to increase the amount of embedded intelligence in their products to gain competitive edge. The resulting intelligent field devices will have a significant and increasing role in implementation of PLM in the process industries. Intelligent field devices are process devices, such as valve positioners or measurement devices, that have embedded microprocessors, memory and software modules to enable data processing. This embedded intelligence means that the devices are not only passive digital devices but have additional functionality such as self diagnostics and signal processing. The technical developments have also resulted to a range of Internet-based services provided by the manufacturers. Thus, forward-looking manufacturers have started to develop network capability and integrate Internet-based services with their products, transforming them into platforms for service delivery. As a result closer customer relationships will be formed throughout the product's lifetime [Mil04].

Even though intelligent field devices seem to offer promising possibilities for efficient PLM of production assets their full potential is not exploited yet. Even the features of already acquired and commissioned intelligent field devices often remain unused. Human resistance, responsibility issues and lack of new practices and easy-to-use tools are seen as contributing factors. Further development work is needed both on the part of the suppliers and end-users in order to maximize the potential of intelligent field devices. As the technology improves the features of intelligent field devices will be utilized in increasing numbers along with advanced monitoring and control methods.

7.2 Process monitoring

The increasing complexity of modern industrial processes poses new challenges for process management. The reasons include e.g. reduced inventory, use of recycle streams and heat integration, faster production speeds, multi-product production lines etc. To meet the challenges and to enable transition from traditional and ineffective maintenance strategies towards condition-based maintenance strategies, various process monitoring methods are used. The monitoring methods are numerous and can be divided into model based and process history based methods [Ven03c]. Process monitoring methods can also be viewed from the point of view of the monitored process. For example plant-wide process monitoring introduces additional challenges and requirements when compared to monitoring of smaller process entities but at the same time more holistic monitoring approach also offers greater rewards once implemented properly. However, the theories, applications and practices of plant-wide monitoring are still somewhat lacking and require more research. Although the issue of large scale monitoring is not yet addressed sufficiently in the research community there is some progress made, see e.g. [Tho07, Bau08, Tha09]. Especially the methods capable of extracting qualitative propagation path models from process data and combining readily available information about the connectivity of the process seem promising. Also the increasing use of XML based PI&D tools offer new possibilities for qualitative modelling and large scale monitoring.

From industrial application viewpoint, the majority of monitoring and diagnostic applications in process industries are based on process history based approaches due to easy implementation requiring little modelling effort and a priori knowledge. Also, with the current state-of-art in applications, detection seems to be a bigger concern than detailed diagnosis which is why less complex methods are often seen as adequate. The literature on industrial applications of monitoring/diagnostic systems and their actual industrial implementations are not many in numbers. This is probably due to the proprietary nature of the development of in-house systems and a gap between academic research and industrial practice. Also, many of the academic demonstrations of the applications have been on very simplistic systems and implementing them for industrial systems is beset by several problems. In this respect overlooked areas in the academic research seem to be the issues affecting the economical feasibility such as ease of deployment and adaptability of the systems. Therefore many of the existing monitoring applications are often seen as expensive to deploy and maintain [Ven03b].

Also, there seems to be little articulation in the literature about the benefits that can be accumulated through deployment of monitoring/diagnostic systems. There are some general guidelines based on experience on the economic impact due to abnormal situations, but there are no case studies that analyze specific benefits attainable through implementation of diagnostic systems. More research is needed on this issue in line with the work that has been carried out analyzing the benefits of implementation of advanced control systems [Ven03b]. The lack of such references hinders the development work and makes it harder to get the industry to commit to research and development projects. Thus, it can be argued that there is unexploited potential in the field of process monitoring and the number of monitoring applications is likely to increase as some of the problems are solved and the related technology improves. Once more successful industrial cases are made public proving the benefits there should be large markets for such applications.

7.3 Distributed systems and system integration

The trends of increasing decentralization in process automation and the need to integrate process assets into a seamless entity were discussed. Decentralization has been ongoing for some time with the popularity of DCS structure and has been enabled by the introduction of programmable digital control systems. More recently especially the developments in fieldbus technologies have been instrumental in promoting this development. One of the latest steps in the evolution of fieldbus technology is the application of wireless transfer techniques in process industries. Nowadays the decentralization is going further with the development and deployment of intelligent field devices which enable the distribution of various functions to the field device level. Such adding of digital intelligence to final control devices adds capabilities that help reduce process variability, increase production uptime, reduce operating costs, and improve safety [Har03].

The evolving communication technologies offer new possibilities and the suppliers need to be aware of the importance of integrating pervasive networking and intelligent devices into factories [Mil04]. At the same time increasing decentralization means that the available information is more distributed and there is often a lack of centralized repository. Further, distributed intelligent systems in general require more communication to coordinate all the entities in the system and full potential of the distributed intelligence is only achieved if hardware, automation system and asset management applications form a seamless entity [Par08, Rep06]. Thus, various integration techniques have been developed. Nowadays the most significant integration techniques in process automation are the OPC, FDT and EDDL techniques which were discussed. These standards help integrating the traditionally separate process assets into an easier-to-manage entity. System integration on higher levels was also discussed and an integration technique, SOA, briefly presented. In the future the role of system integration will further increase on all hierarchy levels.

7.4 Applied part of the thesis

The main research problem was to define requirements, desirable features, architecture and tools for a generic large scale industrial monitoring system and to test it in a pulp drying process. A comprehensive performance measurement and process monitoring application was developed for the pulp drying process. The application provides the plant personnel and partners additional tools for monitoring and analysing the process. The application is a functionally distributed monitoring system based on process history based methods.

It was estimated in the beginning of the study that from the point of view of the application development the most challenging parts are economical, reliability (performance in a dynamic environment) and usability aspects. The usability aspects such as presentation of results (Human Machine Interface) would require attention as the user group consists of people with various backgrounds and people using remote connections. From the point of view of the remote users, the most challenging part seems to generally be the inclusion of un-coded human know-how such as knowledge of why something happened in the process. Some technical solutions, such as electronic process diaries, have been developed for this purpose but their use seems to be limited making it hard to find out why something happened in the process e.g. a week or a month ago. Therefore currently monitoring application development and detailed analysis of the process usually require some “babysitting” and active consultation with the process operators. The preconceptions of the challenges were confirmed during the development work and a corresponding set of requirements and desired features for large scale monitoring applications was compiled.

The goals of the project were achieved and the research hypothesis confirmed. The involved parties were pleased with the results and the application has been taken into active use. The application presents the overall status of the process and individual subprocesses based on measured facts giving new tools for process management. The application also provides more detailed information which helps identifying problems and focusing maintenance efforts on the most critical areas. This is a huge improvement over process management relying to a large degree on unmeasurable human know-how. The application is also useful as an optimization tool and the collected information can be exploited in future equipment and process development.

The second part of the study was continued on the base of the developed monitoring concept, although at a different plant and without certain parts of the monitoring application of the first phase. The second research problem was to study the feasibility of exploiting intelligent field devices as a part of the developed monitoring concept. At this point the functionally distributed monitoring system consisted of field device level, control loop level and subprocess level measurements. The project goal was to study how various events are observable in the monitoring system. Also the role of human know-how was studied as the highest level of the monitoring hierarchy. This proved, once again, to be the most difficult part of the study.

The results were not as good as in the first study and not all of the goals were met. A major problem was that the plant and its equipment were functioning "too well" and real faults or abnormal situations rarely occurred during the project period. This limited the results but there were some events that had credible explanations and were visible at various module levels. The results suggest that further research of the functionally distributed monitoring system under more controlled environment (enabling free experimenting) should result in improved monitoring and diagnostic features. For example better distinction between field device level and control loop level problems could be achieved and some of the diagnostic procedures could be automated. This would however require better integration of the information sources and more uniform requirements for the data. As a result of the study it can be argued that intelligent field devices fit the developed monitoring concept but further work is needed on the subject. All in all the results were encouraging but there is still room for improvements. Both studies showed that there are still some major issues to deal with in the development of monitoring systems.

REFERENCES

[Ake06] Akersten, P. "Maintenance-related IEC dependability standards", *Engineering Asset Management, Proceedings of the First World Congress on Engineering Asset Management (WCEAM)*, pp. 115-119, July 11-14 2006, London: Springer.

[Aro06] Aro, J. "OPC Unified architecture – uuden sukupolven tiedonsiirtoprotokolla automaatiojärjestelmille ja vähän muuhunkin", *Automaatioväylä*, Automaatioväylä Oy, (2006) 4, pp. 7-8 (in Finnish).

[Aug99] August, J. *Applied reliability-centered maintenance*, Tulsa (OK): PennWell, 1999.

[Bah09] Bahrampour, S., Salahshour, K. and Moshiri, B. "Distributed Fault Detection and Isolation Algorithm Using Modified Gath-Geva Clustering Technique", *Preprints of SAFEPROCESS 2009, 7th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*, IFAC, Barcelona, Jun.30-Jul.3 2009.

[Bar09] Bartys, M.Z. "Concept of Feedback Fault Tolerant Actuator", *Preprints of SAFEPROCESS 2009, 7th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*, IFAC, Barcelona, Jun.30-Jul.3 2009.

[Bau05] Bauer, M. "Data-driven methods for process analysis", Doctoral Thesis, University of London, 2005.

[Bau07] Bauer, M., Cox, J.W., Caveness, M.H., Downs, J.J. and Thornhill, N.F. "Finding the direction of disturbance propagation in a chemical process using transfer entropy", *IEEE Transactions on Control Systems Technology*, IEEE, 15 (2007) 1, pp. 12-21.

[Bau04] Bauer, M., Thornhill, N.F. and Meaburn, A. "Specifying the directionality of fault propagation paths using transfer entropy", *DYCOPS7, Seventh International Symposium of Dynamics and Control of Process Systems*, IFAC, Cambridge (Ma), July 5–7 2004.

[Bau08] Bauer, M. and Thornhill, N.F. "A practical method for identifying the propagation path of plant-wide disturbances", *Journal of Process Control*, Elsevier, 18 (2008) 7/8, pp. 707-719

[Bec03] Becker, J., Kugeler, M. and Roseman, M. (eds.), *Process Management*, Berlin: Springer, 2003.

[Ben08] Bender, K., Grossmann, D. and Danzer, B. "Field Device Integration: Combining the benefits of EDDL and FDT", *The industrial Ethernet Book*, IEB Media, (2008) 45.

[Ben07] Bender, K., Grossmann, D., and Danzer, B. *Future Device Integration V1.1*, http://www.itm.tum.de/forschung/future-device-integration/whitepaper/itm_Whitepaper_FutureDeviceIntegrationV1.1.pdf, February 2007.

[Ber01] Berge, J. *Fieldbuses for Process Control*, ISA, 2001.

[Bla03] Blanke, M., Kinnaert, M., Lunze, J. and Staroswiecki, M., *Diagnosis and fault-tolerant control*, Berlin: Springer, 2003.

[Bli03] Blischke, W. and Murthy, D.N. *Case Studies in Reliability and Maintenance*, Hoboken, NJ: John Wiley, 2003.

[Byi02] Byington, C.S., Roemer, M.J. and Galie, T. "Prognostic enhancements to diagnostic systems for improved condition-based maintenance", *Proceedings of the 2002 Aerospace Conference*, IEEE, Big Sky, MT, March 2002, pp. 2815-2824.

[Cap09] Caponetti, F., Bergantino, N. and Longhi, S. "Fault Tolerant Software: a Multi Agent System Solution", *Preprints of SAFEPROCESS 2009, 7th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*, IFAC, Barcelona, Jun.30-Jul.3 2009.

[Car88] Carpenter, C.A. and Grossberg, S. "The art of adaptive pattern recognition by a self-organizing neural network", *Computer*, IEEE Computer Society Press, 21 (1988) 3, pp. 77-88.

[Cha07] Chai, Y., Zhou, Y. and Wang, Y. "A Collaborative Strategy for Manufacturing Execution Systems", *CSCWD 2007, 11th International Conference on Computer Supported Cooperative Work in Design*, IEEE, Melbourne, Australia, April 26-28 2007, pp. 904-907.

[Cha03] Chakraborty, S. "Agent-based approach to fault recovery in a process automation system", Doctoral thesis, Helsinki University of Technology & Technische Universität Darmstadt, 2003.

[Che99] Chen, J. and Patton, R. J. *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Boston: Kluwer Academic Publishers, 1999.

[Che07] Cheng, H. Nikus, M. and Jämsä-Jounela, S. "Enhanced Causal digraph reasoning for fault diagnosis with application on the paper machine short circulation", *3rd International Workshop on Networked Control systems : tolerant to faults*, IFAC, Nancy, June 20-21,2007.

[Che90] Cheung, J.T. and Stephanopoulos, G. "Representation of process trends part I. A formal representation framework", *Computers and Chemical Engineering*, Pergamon Press, 14 (1990) 4/5, pp. 495-510.

[Chi03] Chiang, L.H. and Braatz, R.D. "Process monitoring using causal map and multivariate statistics: Fault detection and identification", *Chemometrics and Intelligent Laboratory Systems*, Elsevier, 65 (2003) 2, pp. 159–178.

[Chi01] Chiang, L.H., Russel, E.L. and Braatz, R.D. *Fault detection and diagnosis in industrial systems*. London: Springer, 2001.

[Cho04a] Choudhury, M.A.A.S. "Detection and diagnosis of control loop nonlinearities using higher order statistics", Doctoral thesis, University of Alberta, 2004.

[Cho04b] Choudhury, M.A.A.S., Shah, S.L. and Thornhill, N.F. "Diagnosis of poor control loop performance using higher order statistics", *Automatica*, Elsevier, 40 (2004) 10, pp. 1719–1728.

[Cho84] Chow, E.Y. and Willsky, A.S. "Analytical redundancy and the design of robust failure detection systems", *Transactions on Automatic Control*, IEEE, 29 (1984) 7, pp. 603-614.

[Cho98] Chowdhury, F.N. and Aravena, J.L. "A Modular technology for fast fault detection and classification in power systems", *Transactions on Control System Technology*, IEEE, 6 (1998) 5, pp. 623-634.

[Cyb89] Cybenko, G. "Approximation by superpositions of a sigmoidal function", *Mathematics of Control, Signals and Systems*, Springer UK, 2 (1989), pp. 303-314.

[Das00] Dash, S. and Venkatasubramanian, V. "Challenges in the Industrial Applications of Fault Diagnostic Systems", *Computers and Chemical Engineering*, Pergamon Press, 24 (2000) 2, pp. 785-791.

[Dud73] Duda, R.O. and Hart, P.E. *Pattern classification and scene classification*, New York: Wiley Interscience, 1973.

[Efe09] Efendic, H. "Data-Based Structural Analysis, Modeling and Fault Diagnosis in Complex Industrial Plants", *Preprints of SAFEPROCESS 2009, 7th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*, IFAC, Barcelona, Jun.30-Jul.3.2009.

[El-98] El-Sayed, M.A.H. "Artificial neural network for reactive power optimization", *Neurocomputing*, Elsevier, 23 (1998), pp. 255-263.

[End04] Endrei, M., Ang, J., Arsanjani, A, Chua, S., Comte, P., Krogdahl, P., Luo, M. and Newling, T. *Patterns: Service- Oriented Architecture and Web Services*, IBM, 2004, http://www.redbooks.ibm.com/redbooks/pdfs/sg246303.pdf

[Fan93] Fan, J.Y., Nikolaou, M. and White, R.E. "An approach to fault diagnosis of chemical processes via neural networks", *American Institute of Chemical Engineers Journal*, Wiley, 39 (1993) 1, pp. 82-88.

[Fan05] Fan-Yin T. and Kwan-Liu M. "Opening the Black Box - Data Driven Visualization of Neural Network", *IEEE Visualization 2005*, IEEE, October 23-28 2005, p. 49.

[FDT06] FDT group, *FDT Newsletter*, FDT group, September 2006, http://www.fdt-jig.org/newsletter_attach/FDT_Newsletter_September_2006.pdf

[FDT07a] FDT Group, *FDT technical description*, FDT group, 2007, http://www.fdtgroup.org/_PDF/download_information/FDT_Technical_Description.pdf

[FDT07b] FDT Group, *FDT User Brochure*, FDT group, 2007, http://www.fdtgroup.org/_PDF/download_information/FDT_User_Brochure.pdf

[Fed05] Fedai, M. and Drath, R. "CAEX—A neutral data exchange format for engineering data", *ATP International Automation Technology*, 3 (2005), pp. 43–51.

[Fin80] Findeisen, W., Bailey, F.N., Brdys M., Malinowski, K., Tatjewski, P. and Wozniak, A. *Control and Coordination in Hierarchical Systems*, Chichester: Wiley, 1980.

[Fil05] Filippetti, F., Franceschini, G. and Tassoni, C. "Recent developments of induction motor drives fault diagnosis using AI techniques", *Transaction of industrial Electronics*, IEEE, 47 (2000) 5, pp. 994-1004.

[Fly05] Flytström, T. "Palvelukeskeisen arkkitehtuurin (SOA) arviointi", Master thesis, Tampere University of technology, 2005 (in Finnish).

[For99] Forsman, K., Stattin, A. "A new criterion for detecting oscillations in control loops", *Proceedings of ECC'99, European control conference*, Springer, 1999.

[Fra89] Frank, P.M. and Wunnenberg, J. "Robust fault diagnosis using unknown input observer schemes", In: Patton, R.J., Frank, P.M. and Clark, R.N. (Eds.), *Fault diagnosis in dynamic systems: theory and applications*, New York: Prentice Hall, 1989.

[Fri03] Friman, M. "A New Method for Condition Monitoring of Unit Processes and Field Devices", *Automaatio '03 seminar*, Finnish Society of Automation, Helsinki, 2003 (in Finnish).

[Fri07] Friman, M. and Mustonen, H. "Kokemuksia prosessin kunnonvalvonnasta", *Automaatio '07 seminar*, Finnish Society of Automation, Helsinki, 2007 (in Finnish).

[Fuk72] Fukunaga, K. *Introduction to statistical pattern recognition*, New York: Academic press, 1972.

[Geo] George, L. *The Bathtub Curve Doesn't Always Hold Water*, American Society for Quality, http://www.asq.org/reliability/articles/bathtub.html

[Geo04] Georgiev, B. "Field Control Systems", *Proceeding of Process Control 2004 - The Sixth International Scientific – Technical Conference*, University of Pardubice, Pardubice, June 8-11 2004.

[Ger01] Gerry, J. and Ruel, M. "How to Measure And Combat Valve Stiction On Line", *Proceeding of ISA 2001*, ISA, Houston, TX, 2001.

[Ger98] Gertler, J. *Fault detection and diagnosis in engineering systems,* New York: Marcel Dekker, 1998.

[Ger95] Gertler, J. and Monajemy, R. "Generating directional residuals with dynamic parity relations", *Automatica*, Elsevier, 31 (1995) 4, pp. 627-635.

[Ger90] Gertler, J. and Singer, D. "A new structural framework for parity equation-based failure detection and isolation", *Automatica*, Elsevier, 26 (1990) 2, pp. 381-388.

[Gu04] Gu, Y., Zhao, Y. and Wang, H. "Visual method for process monitoring and its application to Tennessee Eastman challenge problem", *Proceedings of 2004 International Conference on Machine Learning and Cybernetics, Volume 6*, IEEE, Shanghai, August 26-29 2004, pp. 3423 - 3428.

[Gul00] Gullichsen, J. and Fogelholm C. *Chemical Pulping*, Fapet, Helsinki, 2000.

[Haa00] Haarsma, G. and Nikolaou, M. "Multivariate controller performance monitoring: Lessons from an application to snack food process", Submitted to *Journal of Process Control*, 2000.

[Hag99] Hagan, M.T. and Demuth, H.B. "Neural Networks for Control", *Proceedings of the American Control Conference*, AACC, San Diego, CA, June 1999, pp. 1642-1656.

[Hak07] Hakonen, M. "Laitekuvaus - avoin ratkaisu kenttälaitteiden hallintaan", *Automaatioväylä*, Automaatioväylä Oy, (2007) 1, (in Finnish).

[Har03] Harrold, D. "Making Valve Controllers/Positioners Smarter is Smart Business", *Control Engineering*, Reed Business Information, January 2003.

[Har02] Hartikainen, K. "Kenttäväylälaitteiden hallintaan soveltuva avoin teknologiaratkaisu – FDT, osa II", *Kenttäväylä*, Metso Endress+Hauser Oy, December 2002, (in Finnish).

[Hel07] Helson, R. "Intelligent devices unleashed", *Industrial Automation Asia*, Eastern Trade Media, December 2006/January 2007, http://www.iaasiaonline.com/news/decjan07/pdf/50.pdf

[Hen84] Henley, E.J. "Application of expert systems to fault diagnosis", In: *AIChE annual meeting*, San Francisco, CA, 1984.

[Hes05] Hess, A., Calvello, G. and Frith, P. "Challenges, Issues, and Lessons Learned Chasing the “Big P”: Real Predictive Prognostics Part 1" *Proceedings of the IEEE Aerospace Conference*, IEEE, New York, March 5-12 2005.

[Hie03] Hietanen, V. "Klusterointimenetelmien hyödyntäminen yksikköprosessien suorituskyvyn seurannassa", Master thesis, Helsinki University of Technology, 2003 (in Finnish).

[Hie05] Hietanen, V., Mutikainen, T., Happonen, H. and Huovinen, M. "Kokonaisvaltainen järjestelmä sellunkuivausprosessin diagnostiikkaan ja suorituskykyseurantaan", *Automaatio '05 seminar*, Finnish Society of Automation, Helsinki, 2005 (in Finnish).

[Hor07] Horch, A. "Benchmarking control loops with oscillations and stiction", In: Ordys, A., Uduehi, D. and Johnson, M.A. (ed.), *Process control performance assessment*, London: Springer, 2007.

[Hor89] Hornik, K., Stinchcombe, M. and White, H., "Multi-layer feedforward networks are universal approximators", *Neural Networks*, Elsevier, 2 (1989) 5, pp. 359–366.

[Hot47] Hotelling, H. "Multivariate quality control illustrated by the testing of sample bombsights" In: Eisenhart, O. (Ed.), *Selected techniques of statistical analysis*, New York: NY, McGraw-Hill, 1947.

[Hun92] Hunt, K.J., Sbarbaro, D., Zbikowski, R. and Gawthrop, P.J. "Neural Networks for Control System - A Survey", *Automatica*, Elsevier, 28 (1992) 6, pp. 1083-1112.

[Huo06] Huovinen, M. and Hietanen, V. "Monitoring and diagnosis of large scale industrial systems", *Preprints of SAFEPROCESS 2006, 6th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*, IFAC, Beijing, Aug.30-Sep.1.2006.

[Huo04] Huovinen, M., "Performance Monitoring of Pulp Drying Process", Master thesis, Tampere University of Technology, 2004 (in Finnish).

[Huo07] Huovinen, M. and Hietanen, V. "Issues of Large Scale Monitoring", *IFAC 3rd International Workshop on Networked Control Systems : Tolerant to Faults*, IFAC, Nancy, June 20-21 2007.

[Huo09a] Huovinen, M., Cedeflöf, H. and Hietanen, V. "Älykkäämpää diagnostiikkaa kenttälaitteiden avulla", *Automaatio '09 seminar*, Finnish Society of Automation, Helsinki, 2009 (in Finnish).

[Huo09b] Huovinen, M., Hietanen, V. and Harju, T. "Intelligent Field Devices in Process Monitoring", *Preprints of SAFEPROCESS 2009, 7th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*, IFAC, Barcelona, Jun.30-Jul.3 2009.

[Häg95] Hägglund, T. "A control-loop performance monitor", *Control Engineering Practice*, Elsevier, 3 (1995) 11, pp. 1543–1551.

[Häg02] Hägglund, T. "A friction compensator for pneumatic control valves", *Journal of Process Control*, Elsevier, 12 (2002) 8, pp. 897–904.

[Häg94] Hägglund, T. "Automatic Monitoring of Control Loop Performance", *Control Systems 94-Conference on Control Systems in the Pulp and Paper Industry*, pp. 190-196, 1994.

[Häg07] Hägglund, T. "Automatic on-line estimation of backlash in control loops", *Journal of Process Control*, Elsevier, 17 (2007) 6, pp. 489-499.

[Häg05] Hägglund, T. "Industrial implementation of on-line performance monitoring tools", *Control Engineering Practice*, Elsevier, 13 (2005) 11, pp. 1383–1390.

[Iri79] Iri, M., Aoki, K., O'Shima, E. and Matsuyama, H. "An algorithm for diagnosis of system failures in the chemical process", *Computers and Chemical Engineering*, Pergamon Press, 3 (1979), pp. 489-493.

[ISA99] ISA, *Enterprise - Control System Integration Part 1: Models and Terminology*, ISA, 1999, http://www.mel.nist.gov/sc5wg1/isa95part1-d14.pdf

[Ise89] Isermann, R. "Process fault diagnosis based on dynamic models and parameter estimation methods", In: Patton, R.J., Frank, P.M. and Clark, R.N. (Eds.), *Fault diagnosis in dynamic systems: theory and applications*, New York: Prentice Hall, 1989.

[Iwa06] Iwanitz, F. and Lange, J. *OPC : fundamentals, implementation and application*, Heidelberg: Hüthig, 2006.

[Jai88] Jain, A.K. and Dubes, R.C. *Algorithms for Clustering Data,* Englewood Cliffs, NJ: Prentice Hall, 1988.

[Jel06] Jelali, M. "An overview of control performance assessment technology and industrial applications", *Control Engineering Practice*, Elsevier, 14 (2006) 5, pp. 441-466.

[Jen98] Jennings N. and Wooldridge M. "Applications of Intelligent Agents", In: Jennings N. and Wooldridge M. (ed.), *Agent Technology - Foundations, Applications, and Markets,* pp. 3-28, Berlin: Springer, 1998.

[Jia07] Jiang, H., Choudhury, M.A.A.S. and Shah, S.L., "Detection and diagnosis of plant-wide oscillations from industrial data using the spectral envelope method", *Journal of Process Control*, Elsevier, 17 (2007) 2, pp. 143-155.

[Jut79] Jutila E., Aurasmaa, H., Kortela, U., Uronen, P. and Silvan, H. *Hierarkiset ohjaus- ja säätöjärjestelmät*, Helsinki: Insko, 1979 (in Finnish).

[Kan97] Kantz, H. and Schreiber, T. *Nonlinear time series analysis*, Cambridge: Cambridge University Press, 1997.

[Kar04] Karray, F.O. and de Silva, C. *Soft computing and intelligent systems design: theory, tools and applications*, Harlow, New York: Pearson/Addison Wesley, 2004.

[Kar05] Karte, T. and Kiesbauer, J. "Partial stroke testing for final elements", *Special print of conference proceedings of Petroleum and Chemical Industry Conference (PCIC) Europe 2005*, IEEE, October 26 - 28 2005.

[Kas04] Kastner, W. "EDDL inside FDT/DTM", *Proceedings of 2004 IEEE International Workshop on Factory Communication Systems*, IEEE, September 22-24 2004, pp. 365-368.

[Kaw90] Kawato, M. "Computational Schemes and Neural Network Models for Formation and Control of Multijoint Arm Trajectory", In: Miller, W.T., Sutton, R.S. and Werbos, P.J. (ed.) *Neural Networks for Control,* pp. 197-228, Boston: MIT Press, 1990.

[Kay00] Kayihan, A. and Doyle, F.J. "Friction compensation for a process control valve", *Control Engineering Practice*, Elsevier, 8 (2000) 7, pp. 799–812.

[Kla00] Klaus, H. , Rosemann, M. and Gable, G.G. "What is ERP?", *Information Systems Frontiers*, Springer, 2 (2000) 2, 141-162.

[Koh84] Kohonen, T. *Self-organization and associative memory*, New York: Springer, 1984.

[Koh01] Kohonen, T. *Self-Organizing Maps, 3rd ed.*, Berlin: Springer-Verlag, 2001.

[Kou02] Kourti, T. "Process Analysis and Abnormal Situation Detection", *Control Systems Magazine*, IEEE, 22 (2002) 5, pp. 10-25.

[Kra90] Kraft, L.G. and Campagna, D.P. "A Comparison between CMAC Neural Network Control and Two Traditional Control Systems", *Control Systems Magazine*, IEEE, 10 (1990) 2, pp. 36-43.

[Laa97] Laaksonen, J. and Pyötsiä, J. "Control loop tuning with smart valve positioner", *ISA TECH*, 1997.

[LaB05] LaBelle, B.L. *Intelligent Field Devices and Asset Management Software: Discover the Benefits of Utilizing the Combination*, ISA, 2005, http://www.isa.org/filestore/Division_TechPapers/GlassCeramics/BLISAPOWID2005.doc

[Lan02] Lane, S., Martin, E.B. and Morris, A.J. "Adaptive performance monitoring of a dynamic manufacturing process", *Pulp & Paper Canada*, Business Information Group, 103 (2002) 5, pp. 34-37

[Lee03] Lee, G.B., Song, S.O. and Yoon, E.S. "Multiple-fault diagnosis based on system decomposition and dynamic PLS", *Industrial and Engineering Chemistry Research*, ACS, 42 (2003) 24, pp. 6145–6154.

[Leh08] Lehtonen, K., "Älykkäät kenttälaitteet osana prosessin monitorointia", Master thesis, Tampere University of Technology, 2008 (in Finnish).

[Leh06] Lehtonen, S., "Sellunkuivauskoneen monitorointijärjestelmä", Master thesis, Tampere University of Technology, 2006 (in Finnish).

[Leh07] Lehtonen, S., Hietanen, V., Happonen, H., Huovinen, M. and Ahvenlampi, J. "Kokonaisvaltainen järjestelmä teollisten prosessien diagnostiikkaan ja suorituskyvyn seurantaan", *Automaatio '07 seminar*, Finnish Society of Automation, 2007 (in Finnish).

[Lun91] Lundh, M. "Robust adaptive control", Doctoral thesis, Lund Institute of Technology, 1991.

[Mar99a] Martin, E.B., Morris, A.J. and Kiparissides, C. "Manufacturing performance enhancement through multivariate statistical process control", *Annual Reviews in Control*, Elsevier, 23 (1999) 1, pp. 35-44.

[Mar02] Martin, E.B., Morris, A.J. and Lane, S. "Monitoring Process Manufacturing Performance", *Control Systems Magazine*, IEEE, 22 (2002) 5, pp. 26-39.

[Mar99b] Marttinen, A. and Friman, M. "New Control Performance Measures for Industrial SubProcesses", *Proceedings of the Automation Days 1999 conference*, Finnish Society of Automation, Helsinki, September 14-16 1999, pp. 226-231.

[Mat05] Mattila, M. "OPC:n uudet tuulet – Mikä ihmeen OPC?" *Automaatioväylä*, Automaatioväylä Oy, (2005) 4, pp. 35-38 (in Finnish).

[Mau05] Maurya, M. R., Rengaswamy R. and Venkatasubramanian, V. "Fault Diagnosis by Qualitative Trend Analysis of the Principal Components", *Chemical Engineering Research and Design*, Elsevier, 83 (2005) 9, pp. 1122-1132.

[Mau04] Maurya, M.R., Rengaswamy, R. and Venkatasubramanian, V. "Application of signed digraphs-based analysis for fault diagnosis of chemical process flowsheets", *Engineering Applications of Artificial Intelligence*, Elsevier, 17 (2004) 5, pp. 501–518.

[Mes70] Mesarovic, M.D., Macko, D. and Takahara, Y. *Theory of hierarchical, multilevel, systems*, New York: Academic Press, 1970.

[Met02] Metso Automation, *LoopBrowser 2.0 User's Manual*, Metso Automation, 2002.

[Met06] Metso Automation, *DTM User Guide*, Metso Automation, 2006.

[Met] Metso Automation, *Intelligent valve controller ND9100H Rev. 1.1*, Metso Automation.

[Mia99] Miao, T. and Seborg, D.E. "Automatic detection of excessively oscillatory feedback control loops", *IEEE conference on control applications*, IEEE, Hawaii, August 22-27 1999, pp. 359–364.

[Mil04] Miller, M. "The Future of Information in Intelligent Devices", *automationtechies.com e-news,* 5 (2004) 12.

[Mil06] Miller, S. *Advanced Digital Valve Controller*, CCI, 2006.

[Mil87] Miller, W.T. "Sensor-Based Control of Robotic Manipulators Using a General Learning Algorithm", *IEEE Journal of Robotics and Automation*, IEEE, 3 (1987) 2, pp. 157-165.

[Mon91] Montgomery, D.C. *Introduction to Statistical Quality Control*, New York: John Wiley & Sons, 1991.

[Mur97] Murray-Smith, R. and Johansen, T.A. *Multiple Model Approaches for Modelling and Control*, Taylor and Francis, 1997.

[Mäk06] Mäkelä, M. ”Kenttäväylälaitteiden hallinta helpottuu”, *Prosessori*, Sanoma Magazines, (2006) 12, pp. 43-45 (in Finnish).

[NAS00] NASA, *Reliability Centered Maintenance Guide for Facilities and Collateral Equipment*, NASA, February 2000, http://www.hq.nasa.gov/office/codej/codejx/Assets/Docs/RCMGuideMar2000.pdf

[Nie04] Niemelä, I., "FDT/DTM ja DDL: ystäviä, ei vihollisia!", *Automaatioväylä*, Automaatioväylä Oy, (2004) 4, (in Finnish).

[Nik07] Nikunen J., Happonen H. and Rintamäki I. "Tehokkuutta häiriöhallintaan SOA -arkkitehtuurilla" *Automaatio '07 seminar*, Finnish Society of Automation, Helsinki, 2007 (in Finnish).

[Nim95] Nimmo, I. "Adequately address abnormal situation operations", *Chemical Engineering Progress*, AIChE, 91 (1995) 9, pp. 36-45.

[Oma04] Omachonu, V.K., Ross, J.E. *Principles of total quality*, Boca Raton (FL): CRC Press, 2004.

[Par08] Parker, L.E. "Distributed Intelligence: Overview of the Field and its Application in Multi-robot Systems", *Journal of Physical Agents*, Red de Agentes Físicos, 2 (2008) 1, pp. 5-14.

[Par62] Parzen, E. "On the estimation of probability density function and the mode", *Annals of Mathematical Statistics*, Institute of Mathematical Statistics, 33 (1962) 3, pp. 1065-1076.

[Pau03] Paulonis, M.A. and Cox, J.W. "A practical approach for large-scale controller performance assessment, diagnosis, and improvement", *Journal of Process Control*, Elsevier, 13 (2003) 2, pp. 155–168.

[Pea01] Pearson, K. "On lines and planes of closest fit to systems of points in space", *Philosophical Magazine*, Taylor&Francis, 2 (1901) 6, pp. 559-572.

[Pek04] Pekkola, K. "Kenttäväylien diagnostiikka", Master Thesis, Tampere University of Technology, 2004 (in Finnish).

[Per92] Persson, P. "Towards autonomous PID control", Doctoral thesis, Lund Institute of Technology, 1992.

[Pet00] Petersen, J. "Causal reasoning based on MFM", *Proceedings of the cognitive systems engineering in process control (CSEPC 2000)*, pp. 36–43, November 22-25 2000.

[Pii96] Piipponen, J. "Controlling processes with nonideal valves: Tuning of loops and selection of valves", *Proceedings of the control systems '96*, Nova Scotia, April 29-May 2 1996, pp. 179–186.

[Pir08] Pirttioja, T. "Applying Agent Technology to Constructing Flexible Monitoring Systems in Process Automation", Doctoral Thesis, Helsinki University of Technology, 2008.

[Pyy07] Pyyskänen S. *Teollisuuden laiteverkot : johdatus väylätekniikkaan*, Helsinki: Suomen automaatioseura, 2007 (in Finnish).

[Pyö95] Pyötsiä, J. and Cederlöf, H. "Fuzzy logic expert system and failure analysis of dynamic system", *Proceedings of the industrial computing conference*, ISA, October 1-6 1995, pp. 357-363.

[Pyö97] Pyötsiä, J. and Cederlöf, H. "Non-linear model based diagnostic concept for control valves", *ISA TECH*, 1997

[Qin98] Qin, S.J. "Control performance monitoring—A review and assessment", *Computers and Chemical Engineering*, Pergamon Press, 23 (1998) 2, pp. 173–186.

[Qiu04] Qiu, R.G. and Zhou, M. "Mighty MESs; state-of-the-art and future manufacturing execution systems", *Robotics & Automation Magazine*, IEEE, 11 (2004) 1, pp. 19-25.

[Rep06] Repo, O. "Älykkäät venttiilinohjaimet kunnossapidon kannalta", *Automaatioväylä*, Automaatioväylä Oy, (2006) 5, pp. 38-39 (in Finnish).

[Ric87] Rich, S. H. and Venkatasubramanian, V. "Model-based reasoning in diagnostic expert systems for chemical process plants", *Computers and Chemical Engineering*, Pergamon Press, 11 (1987) 2, pp. 111-122.

[Roe07] Roeller, T. "Smart valve diagnostics: users guide", *Plant Services magazine*, Putman Media, 2007, http://www.plantservices.com/articles/2007/017.html?page=1

[Ron99] Ronkainen, T. "Kenttäväylät automaation tiedonsiirrossa", Master thesis, Tampere University of Technology, 1999 (in Finnish).

[Sac88] Sacks, E. "Qualitative analysis of piecewise linear approximation", *Journal of Artificial Intelligence in Engineering*, Computational Mechanics Ltd, 3 (1988) 3, pp. 151-155.

[Sal05] Salsbury, T.I. and Singhal, A. "A new approach for ARMA pole estimation using higher-order crossings", *Proceedings of the American Control Conference, ACC 2005*, IFAC, Portland, OR, June 8-10 2005.

[Sch03] Schneider, K. "Intelligent field devices in factory automation - modular structures into manufacturing cells", *Proceedings of Emerging Technologies and Factory Automation Conference ETFA '03*, IEEE, Lisbon, September 16-19 2003, pp. 101- 103.

[Sch00] Schreiber, T. "Measuring information transfer", *Physical Review Letters*, APS, 85 (2000) 2, pp. 461–464.

[Sei88] Seiichi, N. *Introduction to TPM : total productive maintenance*, Cambridge (MA): Productivity Press, 1988.

[Sei06] Seilonen, I. "An Extended Process Automation System: An Approach based on a Multi-Agent System", Doctoral Thesis, Helsinki University of Technology, 2006.

[She99] Sherwin, D.J., "A Constructive Critique of Reliability-Centered Maintenance", *Proceedings of Annual Reliability and Maintainability Symposium*, IEEE, Washington, January 18-21 1999, pp. 238-244.

[Sim77] Simon, H.A. *Models of discovery*, Boston: Reidel Publishing Company, 1977.

[Smi01] Smith, D.J. *Reliability, Maintainability and Risk*, Oxford: Butterworth-Heinemann, 2001.

[Sor90] Sorsa, T. "Neuraaliverkot vikadiagnostiikassa", Master Thesis, Tampere University of technology, 1990 (in Finnish).

[Soy73] Soylemez, S. and Seider, W.D. "A new technique for precedenceordering chemical process equation sets", *American Institute of Chemical Engineers Journal*, Wiley, 19 (1973) 5, pp. 934-942.

[Sri05] Srinivasan, R. and Rengaswamy, R. "Stiction compensation in process control loops: A framework for integrating stiction measure and compensation", *Industrial and Engineering Chemistry Research*, ACS, 44 (2005) 24, pp. 9164–9174.

[Sri06] Srinivasan, R., Maurya, M. R. and Rengaswamy, R. "Root cause analysis of oscillating control loops", *Proceedings of the International Symposium on Advanced Control of Chemical Processes (ADChEM-2006)*, IFAC, Gramado, Brazil, April 3–5 2006.

[Sta02] Stamatelatos, M., Vesely, W., Dugan, J., Fragola, J., Minarick, J. and Railsback, J. *Fault Tree handbook with Aerospace applications*, NASA, 2002.

[Sta04] Stark, J. *Product Lifecycle Management: 21st Century Paradigm for Product Realisation*, New York: Springer, 2004.

[Sta07] Stark, J. *Global Product: Strategy, Product Lifecycle Management and the Billion Customer Question*, London: Springer, 2007.

[Su99] Su, L.P., deMare, G. and Carey, D.R. "Prognostics framework", *AUTOTESTCON '99 IEEE Systems Readiness Technology Conference*, IEEE, San Antonio, TX, August 30 - September 2 1999, pp. 661-672.

[Sun88] Suntio, T. and White, K. "Reliability and availability performance as a tool to optimize power systems", *10th International Telecommunications Energy Conference INTELEC '88*, IEEE, October 30-November 2 1988.

[Syr05a] Syrjälä, S. "Ohjelmistoagentit teollisuuslaitosten poikkeustilanteiden hallinnassa", Master thesis, Tampere University of Technology, 2005 (in Finnish).

[Syr05b] Syrjälä, S. and Kuikka, S. "Ohjelmistoagentit teollisuuden poikkeustilanteiden hallinnassa", *Proceedings of the Automaatio 05 Conference*, Finnish Society of Automation, Helsinki, September 6-8 2005 (in Finnish).

[Tan07] Tangirala, A.K., Kanodia, J. and Shah, S.L. "Non-negative matrix factorization for detection and diagnosis of plantwide oscillations", *Industrial & Engineering Chemistry Research*, ACS, 46 (2007) 3, pp. 801-817.

[Tep99] Teppola, P. "Multivariate Process Monitoring of Sequential Process Data-A Chemometric Approach", *Acta Polytechnica Scandinavica – Chemical Technology Series*, Finnish Academies of Technology, No. 270, 1999.

[Tha09] Thambirajah, J., Benabbas, L., Bauer, M. and Thornhill, N.F. "Cause-and-effect analysis in chemical processes utilizing XML, plant connectivity and quantitative process history", *Computers and Chemical Engineering*, Pergamon Press, 33 (2009) 2, pp. 503–512.

[Tho05] Thornhill, N.F. "Finding the source of nonlinearity in a process with plant-wide oscillation", *IEEE Transactions on Control System Technology*, IEEE, 13 (2005) 3, pp. 434–443.

[Tho07] Thornhill, N.F. and Horch A. "Advances and new directions in plant-wide disturbance detection and diagnosis", *Control Engineering Practice*, Elsevier, 15 (2007) 10, pp. 1196-1206.

[Tho97] Thornhill, N.F. and Hägglund, T. "Detection and diagnosis of oscillation in control loops", *Control Engineering Practice*, Elsevier, 5 (1997) 10, pp. 1343–1354.

[Tho03] Thornhill, N.F., Cox, J.W. and Paulonis, M. "Diagnosis of plant-wide oscillation through data-driven analysis and process understanding", *Control Engineering Practice*, Elsevier, 11 (2003) 12, pp. 1481–1490.

[Tho99] Thornhill, N.F., Oettinger, M. and Fedenczuk, M.S. "Refinery-wide control loop performance assessment", *Journal of Process Control*, Elsevier, 9 (1999) 2, pp. 109–124.

[Tho01] Thornhill, N.F., Shah, S.L., and Huang, B. "Detection of distributed oscillations and root-cause diagnosis", *4th IFAC workshop on on-line fault detection & supervision in the chemical process industries*, IFAC, Seoul, June 7-8 2001.

[Tho02] Thornhill, N.F., Shah, S.L., Huang, B. and Vishnubhotla, A. "Spectral principal component analysis of dynamic process data", *Control Engineering Practice*, Elsevier, 10 (2002) 8, pp. 833–846.

[Tic02] Tichy, P., Slechta, P., Maturana, F. and Balasubramanian, S. "Industrial MAS for planning and control" In: Marik, V., Stepankova, O., Krautwurmova, H. and Luck, M. (Eds.), *Multi-Agent Systems and Applications II*, Springer, Germany, 2002.

[Ven03a] Venkatasubramanian, V., Rengaswamy, R. and Kavuri, S.N. "A review of process fault detection and diagnosis Part II: Qualitative model and search strategies", *Computers and Chemical Engineering*, Pergamon Press, 27 (2003) 3, pp. 313–326.

[Ven03b] Venkatasubramanian, V., Rengaswamy, R., Kavuri, S.N., Yin, K. "A review of process fault detection and diagnosis, Part III: Process history based methods", *Computers and Chemical Engineering*, Pergamon Press, 27 (2003) 3, pp. 327-346.

[Ven03c] Venkatasubramanian, V., Rengaswamy, R., Yin, K. and Kavuri, S.N. "A review of process fault detection and diagnosis, Part I: Quantitative modelbased methods", *Computers and Chemical Engineering*, Pergamon Press, 27 (2003) 3, 293-311.

[Ven88] Venkatasubramanian, V., and Rich, S.H. "An object-oriented two-tier architecture for integrating compiled and deep-level knowledge for process diagnosis", *Computers and Chemical Engineering*, Pergamon Press, 12 (1988) 9/10, pp. 903-921.

[Ves81] Vesely, W.E., Goldberg, F.F., Roberts, N.H. and Haasl, D.F. *Fault Tree Handbook*, U.S. Nuclear Regulatory Commission, 1981.

[Wan07] Wang, X., Sun, L., Qingjin, M. and Wenqian, C. "Study of MES for Cement Industry", *IEEE International Conference on Automation and Logistics*, IEEE, Jinan, Shandong , China, August 18-21 2007, pp. 1278-1282.

[Wat82] Watanabe, K. and Himmelblau, D.M. "Instrument Fault Detection in Systems with Uncertainties", *International Journal of System Science*, Taylor&Francis, 13 (1982) 2, pp. 137-158.

[Wei05] Weidl, G., Madsen, A.L. and Israelson, S. "Applications of object-oriented Bayesian networks for condition monitoring, root cause analysis and decision support on operation of complex continuous processes", *Computers and Chemical Engineering*, Pergamon Press, 29 (2005) 9, pp. 1996–2009.

[Whi06] Whitfield, E.J., Sahu, S.N. *ANSI/ISA-95 SUPPORT IN mySAP ERP IN SAP MANUFACTURING*, 2006, http://www.sap.com/usa/solutions/business-suite/erp/operations/pdf/ANSI_ISA_95.pdf

[Wik08] Wikipedia, 2008, http://en.wikipedia.org/wiki/Image:Bathtub_curve.jpg

[Wik09] Wikipedia, 2009, http://en.wikipedia.org/wiki/Process_management

[Wol82] Wold, H. "Soft modeling, the basic design and some extensions", In: Joreskog, K. and Wold, H. (Eds.), *System under indirect observations*, Amsterdam: North Holland, 1982.

[Woo99a] Wooldridge, M. "Intelligent Agents", In: Weiss, G. (ed.), *Multiagent Systems - A Modern Approach to Distributed Artificial Intelligence,* pp. 27-77, Cambridge: The MIT Press, 1999.

[Woo99b] Wooldridge, M.J. and Jennings, N.R. "Software engineering with agents: pitfalls and pratfalls", *IEEE Internet Computing*, IEEE, 3 (1999) 3, pp. 20–27.

[Xia05] Xia, C. and Howell, J. "Isolating multiple sources of plant-wide oscillations via spectral independent component analysis", *Control Engineering Practice*, Elsevier, 13 (2005) 8, pp. 1027–1035.

[Xia03] Xia, C. and Howell, J. "Loop status monitoring and fault localisation", *Journal of Process Control*, Elsevier, 13 (2003) 7, pp. 679–691.

[Yok06] Yokogawa, *Fieldbus moves into the next generation*, SA Instrumentation & Control The Official Journal of the SAIMC, January 2006, http://instrumentation.co.za/article.aspx?pklArticleId=3778&pklCategoryId=58

[Zan07] Zang, X. and Howell, J. "Isolating the source of whole-plant oscillations through bi-amplitude ratio analysis", *Control Engineering Practice*, Elsevier, 15 (2007) 1, pp. 69–76.

[Zel05] Zellcheming, PI. *Production Indices for Paper Production*, The German Pulp and Paper Chemists and Engineers Association & The Finnish Paper Engineers' Association, 2005, www.zellcheming.com

[Zha04] Zhao, Q. and Xu, Z. "Design of a novel knowledge-based fault detection and isolation scheme", *IEEE Transactions on system, man and cybernetics, Part B, cybernetics*, IEEE, 34 (2004) 2, pp. 1089-1095.

[Zie06] Zielinski, M. "EDDL- more than just diagnostics?", *Automation magazine*, Emerson, December/January 2006, http://www.emersonprocess.com/home/library/articles/automation/automation0601_eddl.pdf

[Zuc03] Zucchero, N. "Technology soul mates: FDT and OPC", *InTech magazine*, ISA, January 2003, http://findarticles.com/p/articles/mi_qa3739/is_200301/ai_n9216540