IMPROVING THE USAGE OF PROCESS DATA COLLECTED IN PROCESS INDUSTRY AND POWER PLANTS

Christer Karlsson, Anders Avelin, and Erik Dahlquist

Mälardalen University, Process Diagnostics Group,
P.O. Box 833 SE-721 23 Vasteras, Sweden

christer.karlsson@mdh.se (Christer Karlsson)

Abstract

Five years after the introduction of two classes of advanced control systems for continuous pulp digesters, only one of the classes is still in use after 5 years of operation. The five systems in current use are simple Excel sheets which the operators report are very satisfactory. The use of the six model-based systems has been discontinued even though they had the potential to increase the production by 2% compared to 0.5% for the Excel sheet based system. The current challenge is to develop a system that combines the benefits of the model-based system with the robustness of the Excel sheet based system. It is vital for the system to be robust in the sense that it is transparent and easy for the operator to maintain. Robustness is essential in many parts of the system, including measurement, process model validation, the ability of the model to adapt to changes in the process, optimisation algorithms, and of course the model itself. The optimisation algorithm here is a model predictive control algorithm that returns set-points for the PI controllers. The challenges when constructing such a system are in instilling operator confidence, filtering of misleading measured data, adaptation of process parameters when the process parameters change, and combined validation of measurements and process models. These challenges are met by using a combination of physical and statistical models and methods based on them such as model predictive control (MPC) and parameter estimation. The model should be maintained by a qualified engineer who should be able to explain the system to the operator so that it is understood and confidence is maintained in the system.

Keywords: statistical models, physical models, robustness, adaptation, process industry

Presenting Author's biography

Christer Karlsson (PhD Student and Licentiate Engineer, Master of Science in Energy Engineering) focuses on handling degeneration in first principle process models and on-line sensor measurements. Has developed applications for diagnostics on steam turbines, gas turbines, heat and power boilers using first principle models, artificial neural networks, and Bayesian networks in combination with optimization algorithms.

1 Introduction

The amount of data collected during a process has increased tremendously over the last twenty years. Both the number of signals monitored and the frequency of collection have increased. However, the application of the data to process control remains predominantly manual, although displays of trends in the data are often available on monitors. The use of these data in automatic control is mostly as single loop PI or PID controls, and the use of the data in diagnostics is rudimentary at present. With respect to production planning systems, these are at most tools for supporting the manual control, with a few exceptions. The reason for this is that there is an insufficient level of confidence in the systems among the operators. If the system is prone to faults or if there is no compensation for sensor drift, control cannot be left fully to the system, and thus they are often not used, even though they would be likely to improve performance.

An example of this is the case where an automation engineer participated in eleven installations of advanced control systems for continuous pulp digesters. Five of these used simple Excel sheets to indicate how to make the set points. The other six were based on more sophisticated mathematical models. Five years after installation, the five simple Excel sheets were still in use and the operators were very satisfied with their performance. Not one of the model-based control systems were in use at the end of the five year period. This was because the operators did not really understand the system, and were not capable of maintaining it. The engineer concluded that had the model-based systems been used throughout, production could have increased by 2 %, compared to the 0.5 % that was achieved with the Excel sheet based system.

Both the Excel sheet and the model based systems have attractive features. The Excel sheet based system is easy to maintain without a great deal of training. However, the model-based system has greater potential to increase productivity. We wish to combine the robustness and simplicity of the Excel sheet based system with the additional benefits of the model-based system. This paper will discuss the construction of such a model that may be applied in the power plant and process industries. The use of data and new possibilities in this area are highlighted, new areas of research are presented, as well as potential obstacles we foresee in following these avenues of research.

1.1 The Excel sheet based method

The planning tool in the Excel sheet based method for a continuous digester uses a lookup table method (Fig. 1). In this method, a recipe is applied to the input, which produces set points that produce the desired output. This method is simple and direct and therefore easy for the operator to understand.

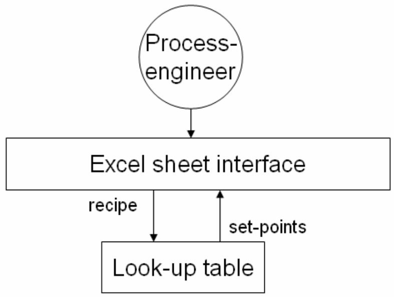

Fig 1. Excel sheet based method.

1.2 The model-based method

The model based method (Fig. 2) was troublesome to maintain in the long run, and required more input from the engineer than the Excel sheet based method. This method needs both a recipe and a number of additional factors depending on the desired quality of the results. The system used a predictor to optimize the set-points.

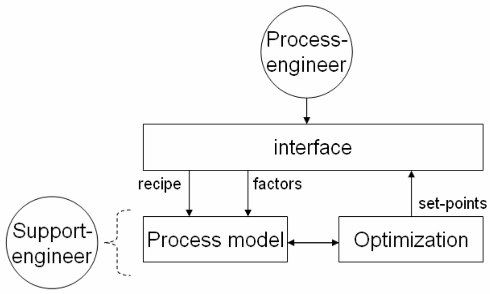

Fig 2. Model-based method.

The integration of support during the start-up of this system has been shown to be very important.

1.3 Making it easier and more complex

It is possible to reconcile the apparently conflicting need for simplicity in the interface and complexity in the model. If the Excel based method interface is retained and some of the more complex and user-unfriendly parts of the model based method are automated and better supported, the model based method is made easier and the Excel sheet method is made more sophisticated. This additional complexity however only occurs in the background and the operator still sees the same interface. Additional training in running the system and understanding the basics of the model in the background would also be needed to maintain the feeling of having control of the system. Robustness is maintained by gaining the trust of the process engineers as well as putting in place a support agreement with the developer of the system. It is also important to highlight the value of the process models and to be prepared to absorb the cost of keeping them up to date. Process models quickly become obsolete when not updated, a fact that is not always recognised.

2 Robustness in a wider sense

Robustness requires the inclusion of all the important variables in the system, and that these can be measured - either directly or indirectly. These measurements should then be used to adapt the statistical, physical or combined statistical and physical model to the real system. This adaptation has to be performed in such a way that only relevant data are considered by the model. If model-based control is going to be used, it is essential to have full control of all the measurements. If a sensor is faulty it should be ignored, or the reading adjusted to show an estimated value to the control system. This challenge can be resolved with data reconciliation and process performance monitoring using physical models. A view of the relationships between different steps is shown in Fig. 3. The system illustrated here is the production optimization Excel sheet based system and the underlying structure of the parts needed for the set-point optimization.

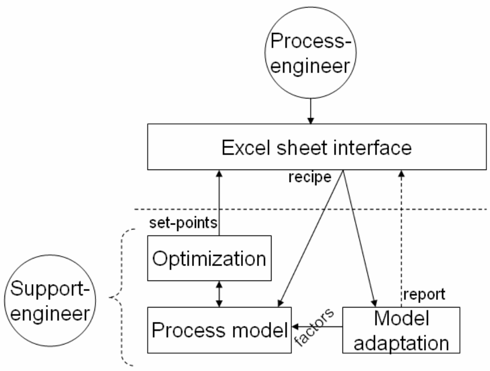

Fig 3. Combined method.

The successful application of the model depends on support for the process engineer from the support engineer, who should be an expert on the process model and the whole system. The day to day operation of the system should not require such support.

As it is difficult to have full trust in something that is not fully understood, it is important for the operators as well as the process engineers to understand the workings behind the control function. This understanding can be gained in a training simulator system. Lack of trust reduces the motivation to use the system. The most successful systems are those where the operator feels fully in control. This feeling of being in control was a factor in the success of the Excel based system. The operators of the system had a good understanding of the way it worked and therefore felt comfortable enough to continue using it. A good control system should therefore have a structure that includes robustness and is easy to understand, thereby increasing operator confidence. When it comes to process production optimization, all aspects must be considered, including the requirement for maintenance, and the risk of unforeseen problems arising. Any alarm filtering or maintenance on demand systems that are fitted, while they can solve some important problems, also have to be integrated and optimized so that they do not cause new ones.

3 System parts

The process models are at the heart of the system. Maintenance costs for the process models may occur at a number of points in the system such as control and adaptation of process models. The validation step and relationships between the system parts are shown in Fig 4.

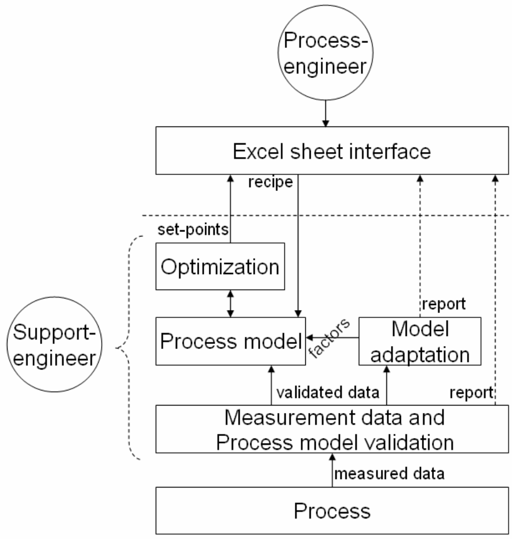

Fig. 4. Combined and extended system.

The system in Fig. 3 consists of a number of different parts that each solves a specific task. The system in Fig. 4 adds measurement data and process model validation to deal with measurement errors and process model errors. The next section deals with these functions, beginning with the measurement data from the process and ending with the Excel sheet interface, and demonstrates their necessity in a future robust control system.

3.1 Measurement collection and quality of measurements

Computers and computer networks have made it possible to collect huge amounts of data from industrial processes. Sensors have been developed to measure new properties directly or indirectly. Storage capacity of hard disks and removable media such as CDs and DVDs is increasing rapidly. Together, these developments have increased the amount of data collected in the process industry. With the available technology it is possible to store all sensor data that is produced from the moment the data collection system is installed. However, only a selected range of measurement tags are collected and stored in practice. In many cases the collected data is stored using a moving time window. It can range from seconds to years, depending on the type of measurement data and what it is used for.

Measurements used for monitoring and those used for reports to authorities (such as environmental reports and measured properties used for taxation) are stored for years. Sensor data for process control and low priority alarms are not usually stored. Filters are applied to reduce noise by high-pass or low-pass filters or averaging functions. Dead bands are also commonly used so that data is only recorded when large enough changes in the measured values occur. This effectively decreases the amount of data stored when the process deviates from the steady state. This also has the effect of reducing the strain on the data network during the steady state operation of the system. It is also common practice to compress data using averaging functions. This can be done in stages. For example, average data on a minute-by-minute basis can be stored as hourly averages when the data is more than a week old. It is important to realise that all these data reduction methods are irreversible. A simple calculation reveals that one year's data of 1000 tags with minute average values can be compressed using reversible methods such as zip files, to 12 GB of stored data. This kind of data in combination with a well structured process log with causes for downtime and process disturbance is the key to building useful statistical models and is valuable for diagnostics.

3.2 Measurement data and process model validation

The error in output from a process model can be divided into measurement faults and process model faults. Measurements error sources can be described as faults in the sensor or in the positioning of the sensor. The model fault is described either as faults in the structure of the model, which means that not all streams transporting mass and heat have been accounted for correctly, or as parameters not correctly tuned. The process model must be protected from erroneous data and bad tuning in order to avoid model faults and inaccurate descriptions of the process. Examples of techniques to deal with the problem of errors in measurement data and process parameters are described below.

3.3 Treatment of measurements and data reconciliation

Apparent faults must be removed from measurement data so that they do not corrupt the process model calculations and results. A detection of a fault is often used to trigger an alarm and is an indication that something has gone wrong. Detection of faults can be achieved by measurement of threshold values or by observers noting discrepancies between the measurement and predicted values. Isolation of measurement faults in the process industry is not done automatically. It has to be carried out by experienced maintenance personnel, operators or engineers. However, fault isolation could be automated to some degree with the use of robust process models. Self diagnosis, which includes fault isolation in sensors is common for an increasing number of products, for example in cars and printers, for a limited number of faults. These products are mass produced and therefore the unit cost for developing such a diagnostic system is low.

There is great potential for using process models for multiple applications, but the quality of the models are limited by the quality of the input measured variables. Several techniques are used for signal validation. Some of these use models to predict signal values and compare predicted values with measured values to validate the measurement. When the model is developed and fully working, the next step is to tune the constants in the equations. The model is tuned with process data for a number of different operational conditions by measuring the signal as a function of time. If the deviation between the model prediction and the measured value is above a certain level, the model is adjusted iteratively until the prediction is good enough for the operating conditions. Thereafter the conditions are changed and the model is tuned for the new conditions. The tuning proceeds systematically according to a pre-prepared scheme. This can be performed using different algorithms, but they are basically different versions of the same principles.

The example presented in Fig. 5 shows the deviation between predicted and measured values for five different sensors. Where the deviation remains constant over the observation period as shown for the sensors Fl2, Fl3, Fl4, and Fl5, there is no warning. For sensor Fl1 the deviation changes and the system indicates that the condition of the sensor should be examined.

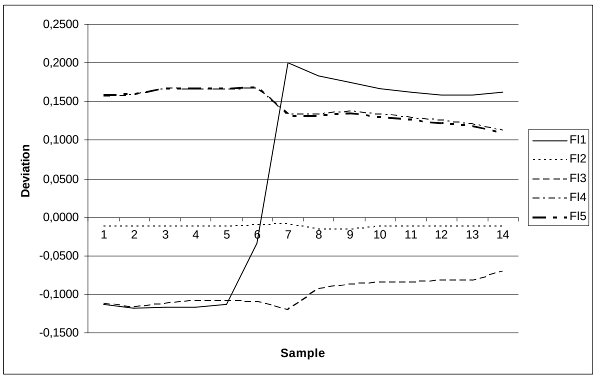

Fig. 5 Deviation between the model prediction and measured values for a number of related sensors [1].

Setting up a mass balance over a process is an efficient tool to detect leaks and flow sensors that are out of calibration. The flow and the sensors are described in a network (also called a graph) made up of nodes and connecting arcs. A method called data reconciliation, which makes use of redundant flow sensors can be used to reduce noise. The method solves a weighted optimization problem where all sensors are used in order to reduce sensor noise, and the constraints involve the graph model [2, 3]. This method requires that there are no gross errors in the measured flows. Examples of implementations and tests have been described for a refinery [4] and chemical process [5]. In the refinery example it was concluded that leaks and badly calibrated equipment were found, and substantial financial savings were achieved mainly by better measurement of mass flow of produced refinery products. The data reconciliation method has also been investigated by for a flue gas channel in a heat and power plant [6]. The authors and their co-researchers are developing methods for fault isolation in gas and steam turbines [7] and for measuring moisture content in flue gas using combined statistical and physical models. The flue gas moisture content measurement has been implemented in a heat and power plant [8].

3.4 Combined process model and sensor validation

Measurement validation requires a robust and correct process model and process model validation requires a reliable set of measured data. The combination of simultaneous process and model validation is a complex problem that has not yet been solved. The authors are making continuous efforts to solve such problems [9]. One way to validate a process model is to estimate the parameters while running the model on a reliable data set and analyzing the changes in the parameters.

3.5 Process models and model adaptation

As we are considering model-based approaches, first principles models are an essential part of the algorithm. These models are often combined with statistical models, for example where there is an insufficient number of on-line measurements. The continuous digester model has been published earlier [10] and is not further described here.

With the model adaptation we aim to update the parameters in the model continuously. This can be achieved by estimating the parameters within a time window under the supervision of the support engineer to ensure safe tuning. We propose that this task be semi-automated as there are numerous factors to consider when it comes to filtering data during tasks such as maintenance and start and shutdown procedures.

3.6 Optimization and control

Optimization can be realized using a Model Predictive Control (MPC) algorithm. This is a multivariable model predictive control technology [11]. A combined statistical and physical process model is one way to control the process in a MPC. The main advantage of MPC is its ability to handle multivariable processes with strong interactions between process variables and with constraints involving both process and state variables. The potential of this control method is dependent on the quality of the process model and of the measured variables. With a good mass balance simulation of the process the MPC is less dependent on the measured variables [12]. The system discussed here is based on PI control loops and the set-points are set by operators with guidance from the Excel sheet. If the set-points were input directly into the MPC this stage would not be necessary, but the demand on robustness would be much higher than before. Therefore, such closure of the loop can only be carried out with caution and when the process is not disturbed by actions not taken into account by the system, such as maintenance.

4 Examples of combined models within different applications

Process industry uses a number emerging techniques on a test scale and others in day-to-day work. For example, MPC is used by ABB in their ProcessIT platform. Processes that have very high downtime costs and high impact on the environment when the process goes out of control - such as nuclear power plants and large chemical industries - already use model-based early detection systems. These methods are now becoming available for processes with lower downtime costs such as pulp and paper, and power plants.

An example of the use of simulation models in connection with optimization solvers was implemented in the DOTS project, a €5 million EU joint project between Sweden, Finland, Germany and France [13]. Here the physical model interacted with different optimization solvers. This was done both to make direct use of the dynamic simulator to find the optimum process pass for the next minutes to hours and to verify that the solution was feasible if a conventional optimization was used for a simplified problem formulation.

One of the drivers for new techniques and methods is the shift of maintenance methods from repairing broken equipment to time scheduled maintenance and the possibility of condition based maintenance (CBM). Monitoring the condition of process equipment becomes more important as the process is automated. The increasing complexity of the equipment means that many more personnel would be needed to continuously track performance and condition. Knowing the equipment condition also gives the advantages of potentially increasing up-time, narrowing margins to full capacity of a process, and makes it easier to detect differences between normal and abnormal operation. In short, knowing the condition of the equipment confers the ability to utilize the plant better.

5 Challenges when implementing combined statistical and physical models

Many industrial processes are non-steady state. In a paper machine the wire gets clogged and sensors are fouled. In a roll mill rolls are torn down. Generally the production switches between different product grades or capacities on a frequent basis, and so on. This makes it difficult to use the exact same model all the time. As the operational conditions are very different, statistical models are normally difficult to implement, whereas physical models tuned by process data are potentially very useful. To achieve this, parameters in the physical models need to be adapted to compensate for the varying conditions. The adaptation can be done by continuous retuning of the model or by switching between different values of predefined parameters. This is easy in theory, but in reality there are many issues to consider. Firstly, the relevant conditions that need to be adapted to must be identified. Otherwise the adaptation may destabilize the model prediction rather than improve it.

The adaptation can be both of parameters in a process model and of tuning parameters in a control algorithm. In both cases the collected data has to be pre-treated in some way. The data have to be evaluated so that reliable data is used and data that is invalid for whatever reason is rejected. This is not a trivial task. Parameter estimation is the preferred method for adaptation. It may be wise to filter signals with different sampling frequency and then update parallel models. It is advisable to keep any old models for comparison. It is often advisable to use an old model until the cause of drift in the model prediction is understood to avoid problems when the cause of drift is removed. Once this has been done, the adapted models are likely to give significantly improved performance over a longer time period. It may even be possible or desirable to include the cause of the drift as a new variable. If for instance the drift in paper quality is due to clogging of the wire, time for replacement or cleaning can be included as a time function, even though it would not be measured as a 'real variable'.

It sometimes happens that models that work very well at first drift away and become almost useless after a period of a few months, due to changes in non-measured variables. An example of this was seen when using multivariate data analysis at Dynäs, a pulp and paper mill in Sweden. 15 different paper properties were predicted from models produced from process data and near infra red (NIR) spectra of pulp sampled at certain positions. During the first two to four weeks the model produced good predictions for many of the properties, but after four months many of the predictions were too poor to be used in the operations. For instance, the prediction power Q2 for Tensile index dropped from 0.82 to around 0.65 over this period. The main reason for this was a combination of drifting sensors and clogging of the paper machine wire. Some of these could have been compensated for by additional measurements, which would have resulted in robust soft sensors [14].

Robustness of the models can be closely related to the factors discussed above with respect to adaptation. It may sometimes be better to have a less accurate but more stable model to avoid instabilities in the process. This is because shutdowns are often more serious than sometimes suboptimal operating conditions. Robustness can sometimes be achieved by using mean values computed in a moving window approach, so that steep changes are smoothed out. However, this can also make it difficult to notice real process disturbances, which may cause significant negative effects on the overall process performance. An example of this was in a paper machine application, where the consistency at the head box was fluctuating heavily [15]. A decision was taken to strongly filter the signal. However, when the sensor was moved five meters downstream the noise level was significantly reduced. When the previously applied filtering was removed it was then possible to see strong spikes that turned out to be correlated to many of the paper breaks. When the cause of these spikes was eliminated (a poor control of switching between two water filters), the number of paper breaks was reduced by 29%, estimated to be worth US$7 million per year in increased production. In this case the filtering, whose purpose was to produce a robust application, instead masked a real problem, even though the signal looked better.

The combination of physical models and process data can produce more robust overall solutions that can be used for both diagnostics and control. One example of this is the use of a physical model to connect different physical properties in a complex process. This is implemented at Korsnäs pulp mill. Here we measure NIR spectra of ingoing wood chips. The fibers are followed through the mill using a physical model (using Modelica). A specific parameter value for chemical reactivity of the wood is used to predict the dissolution of lignin from fibers as a function of time, concentrations of chemicals, temperature, production capacity and other factors. By measuring the kappa number of the fibers produced as well as residual alkali and dissolved lignin of the extracted liquors and comparing these to the predicted values, the NIR spectra can be correlated to the reactivity of the incoming wood. Subsequently the NIR spectra can be used to predict the properties of the outgoing pulp, and for MPC on the continuous digester. Additionally, different diagnostics can be performed such as identification of hang ups and channelling. The diagnostics is also used to identify normal operations and the data that can be used to fine-tune the wood properties. This is an example of how a combination of physical models and statistical models with measured data can produce more robust controls as well as diagnostics [1].

In another application a soft sensor was made with a combined statistical and physical model to predict the ‘ring crush test’ on paper board. The physical model resulted in a very robust base for the system, and the parameter estimation resulted in good output values [16]. This application was implemented at several mills and gave good results for years without retuning.

6 Conclusions

It is important to be aware of the possible uses of the different data that are generated in a process, because the decisions that are made about filtering, compression etc., affect the possibilities of later analysis and of building data driven models in the future. The methods need to be more transparent to the user or completely automated. Examples have been described of the powerful applications based on combined statistical and physical process models.

Applied research that simplifies and lowers the threshold for acceptance of methods based on combined statistical and physical models is one of the keys to better use of the data collected in process industry and power plants. An alternative is to keep developing new methods that make use of existing process models, in order to avoid having to develop new models for each application and the costs for maintaining several distinct models. Adaptation of process models and robustness in sensor output data are increasingly important when an automated method is to be implemented.

7 Acknowledgement

The authors like to acknowledge grants from the Knowledge Foundation and Swedish Energy Agency.

8 References

[1] Avelin A., Jansson J., Dahlquist E, “Use of Combined Physical and Statistical Models for On-line Applications in Pulp and Paper Industry”, *submitted to Journal of Mathematical and Computer Modelling of Dynamical Systems in August 2006.*

[2] Crowe C.M., Campos Y.A.G., Hrymak A., “Reconciliation of process flow rates by matrix projection. Part I: Linear case”, *American Institute of Chemical Engineers Journal,* vol. 29. pp 881-883, 1983.

[3] Crowe C.M., “Reconciliation of process flow rates by matrix projection. Part II: The nonlinear case”, *American Institute of Chemical Engineers Journal,* vol. 32. pp 616-623, 1986.

[4] Wang Y., Rong G., Wang S., “Linear dynamic data reconciliation: refinery application”, *6th IFAC Symposium on Dynamics and Control of Process Systems*, Korea, pp 650-655, 2001.

[5] Sanchez M. A., Bandoni A., Romagnoli J., “PLADAT – A package for Process Variable Classification and Plant Data Reconciliation.” *Journal of Computers and Chemical Engineering* (suppl. 1992), pages 499-506.

[6] Karlsson C., Widarsson B., Dotzauer E., “Data reconciliation and gross error detection for the flue gas channel in a heat and power plant”. *Conference on Probabilistic Methods Applied to Power Systems, PMAPS2004*, USA, Ames, 2004.

[7] Karlsson C., Arriagada J., Genrup M., “Detection and interactive isolation of faults in steam turbines for maintenance decision support”, *submitted to Journal of Modelling and Simulation, Practice and Theory in February 2007.*

[8] Avelin A, Widarsson B, Wester L. “Moisture Content Determination of Biomass fuel supported by Bayesian Networks”. *ASCC, Indonesia, Bali 2006*

[9] Karlsson C., Kvarnström A, Dotzauer E., ” Estimation of process model parameters and process measurements – a heat exchanger example”, *Conference on New Trends in Automation*, Sweden, Västerås, 2006.

[10] Avelin A, Jansson J, Dahlquist E. Use of Modelica For Multi Phase Flow In Complex systems, With application For Continous Pulp Digester. *APMMCT, Ukraine, Khmelnitsky, 2005*

[11] Morari M., Lee J., Garcia C.,”Model Predictive Control”, Book at Prentice Hall, Engelewoods Cliffs, 2000.

[12] Jansson J., Erik Dahlquist., “Model-based Control and Optimization in Pulp Industry”, *SIMS 2004*, Denmark, Copenhagen Sept 22-24, 2004.

[13] Dhak J., Dahlquist E., Holmström K., Ruiz J., Belle J., Goedsch F.: DEVELOPING A GENERIC METHOD FOR PAPER MILL OPTIMIZATION, *Conference on Control Systems* 2004, Canada, Quebec City, June 14-17, 2004.

[14] Liljenberg T., Backa S., Lindberg J. , Ekwall H., Dahlquist E.: *On-line characterization of pulp - Stock preparation department*; SPCI, Sweden, Stockholm, 1999

[15] Jansson J., Malardalen University, Dahlquist E., Malardalen University and ABB, Subramani S.R., ABB Singapore., “Application of Physical Models for Optimization and Control of Digesters in Pulp and Paper Industry”, *Asia Paper*, Singapore, April 22-26, 2002.

[16] Pettersson J., Dahlquist E., Warnqvist J., Carlsson M., “On-line strength prediction and optimization for multi-ply kraft liner”, *Conference on Control Systems, Sweden, Stockholm*, 2002.

ISBN 978-3-901608-32-2