Hierarchical Scheduling and Utility Disturbance Management in the Process Industry

Lindholm, Anna

2013

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Lindholm, A. (2013). *Hierarchical Scheduling and Utility Disturbance Management in the Process Industry*. [Doctoral Thesis (compilation), Department of Automatic Control]. Department of Automatic Control, Lund Institute of Technology, Lund University.

Total number of authors:
1

Hierarchical Scheduling and Utility Disturbance Management in the Process Industry

ANNA LINDHOLM
DEPARTMENT OF AUTOMATIC CONTROL | LUND UNIVERSITY

Hierarchical Scheduling and Utility Disturbance Management in the Process Industry

Anna Lindholm

LUND UNIVERSITY

Department of Automatic Control

Department of Automatic Control
Lund University
Box 118
SE-221 00 LUND
Sweden

PhD Thesis
ISRN LUTFD2/TFRT--1097--SE
ISSN 0280–5316
ISBN 978-91-7473-678-6

Printed in Sweden by Media-Tryck.
Lund 2013

Abstract

This thesis deals with control of production at large-scale process industrial sites in the presence of disturbances. The main focus is on disturbances in the supply of utilities such as steam, cooling water and electricity. A general method for reducing the revenue loss due to disturbances in utilities is introduced, which may provide both proactive and reactive disturbance management strategies. Utility availability and area availability are introduced as performance indicators. These measures are used to obtain quick estimations of the revenue losses related to each utility. To obtain reactive strategies for utility disturbance management, a simple model of how utilities affect production in an area, and how utilities are shared between areas, is introduced. The modeling approach is utilized to formulate the production control problem at disturbances in utilities as an optimization problem. Measurement data are used to obtain empirical models of utility disturbances at an industrial site, which may be used as input to the optimization. Since production control closely relates to production scheduling, the integration of production scheduling based on orders and forecasts with production control at disturbances in utilities is studied in the final part of the thesis.

Acknowledgments

First of all I would like to thank my supervisor Charlotta Johnsson for her encouragement and good discussions during my years at the department. Her enthusiasm is highly contagious, which makes it a delight to work with her. This also holds for my co-supervisor Tore Hägglund, who always manages to get me to feel better about my research, written papers, or life in general, after each meeting. Inspiring people seems to be a theme for the people involved in the Process Industry Centre which my research project was a part of; every conference and meeting has generated a lot of useful and interesting discussions as well as nice dinners and other events. I am especially grateful for the good collaboration with Krister Forsman, Nils-Petter Nytzén, Hampus Carlsson and Jesper Jönsson at Perstorp.

I would also like to thank all colleagues at the department for helping to create such a beneficial environment for Ph.D. studies. Especially, I would like to thank Pontus Giselsson for good collaboration, Kristian Soltesz and Karl Berntorp for their helpful comments on this thesis, and Olof Sörnmo, Martin Hast, Andreas Stolt and Josefin Berner for allowing me to visit your room and distract you when I got tired of writing on the thesis. Thanks also to all Ph.D. students that I have been travelling around the world with for conferences, teaching and courses. You made these trips truly enjoyable! I'll never forget Corso Como, Livin' on a Prayer, the pool at Örenäs or Bohemian Rhapsody... Eva Westin and Ingrid Nilsson also deserve a special acknowledgment for their true concern and aim to keep everyone's spirit up at the department. While working at the department, all practical issues such as computers and administration works smoothly, which I am truly grateful for. A special thanks goes to you who have helped me with typesetting and administrative issues related to my thesis. Finally, I would like to thank my fantastic friends for putting up with all my whining about work- and non-work related issues during the past year, and my family for their constant support and encouragement.

Anna

Contents

Preface

Outline and Contributions of the Thesis

The thesis consists of five introductory chapters and four papers. This section describes the content of the introductory chapters and the contributions of each paper.

Chapter 1 – Introduction

This chapter provides a motivation for the research presented in the thesis. The used research methodology is described, and an overview is given of the process industry today, the research centre that the research is part of, and the company that has been involved as a collaborating partner.

Chapter 2 – Utility Disturbance Management

The challenges related to the area of utility disturbance management are discussed in this chapter, as well as some terminology for disturbance management and for statistical models of utility disturbances. This provides the background for Papers I-IV.

Chapter 3 – Production Planning, Scheduling, and Control

In this chapter, the areas of production planning, scheduling, and control are defined, and the challenges within these areas are discussed. This serves as background for Paper IV.

Chapter 4 – Optimization

Mathematical optimization is a main part of Papers II and IV. In this chapter, a short introduction to optimization is given, and the types of optimization problems that arise in Papers II and IV are discussed in more detail.

Chapter 5 – Discussion

In this chapter, the contributions of the thesis are summarized and possible future work directions are outlined.

Paper I

Lindholm, A. and C. Johnsson (2013): "Plant-wide utility disturbance management in the process industry." *Computers and Chemical Engineering*, **49**, pp. 146–157.

This paper presents a general method for reducing the loss due to utility disturbances, denoted the utility disturbance management (UDM) method. The method is presented step by step and is applied to a real industrial case at site Stenungsund, operated by Perstorp. Common utilities in the process industry and their usage are also described, and measures of utility and area availability are defined. A simple on/off modeling approach is suggested, and utilized for ordering utilities according to the revenue loss they cause. A matrix representation of a site is introduced to facilitate the calculations related to use of the UDM method.

The UDM method was outlined by A. Lindholm and C. Johnsson. A. Lindholm performed the literature review of the research area and wrote the summary on utilities using insights from visits at process industrial sites. C. Johnsson contributed with the idea of the matrix representation of a site, which was developed and extended by A. Lindholm. The case study was performed by A. Lindholm.

Paper II

Lindholm, A. and P. Giselsson (2013): "Minimization of economical losses due to utility disturbances in the process industry." *Journal of Process Control*, **23**, pp. 767–777.

This paper suggests a method for modeling utilities at process industrial sites. The modeling approach assumes a linear relation between assignment of a utility to an area and production in the area, and considers both utilities of continuous type and on/off type utilities such as electricity. The modeling approach is used for formulating an optimization problem for minimization of economical losses at disturbances in the supply of utilities. The resulting optimization problem becomes a mixed-integer quadratic program (MIQP). A quadratic-program (QP) relaxation for these types of problems is suggested to shorten the solution time. Two examples are given, and the solutions and solution times of the MIQP and QP problems are compared.

A. Lindholm contributed with the main idea for the utility modeling and optimization problem formulation. The MIQP optimization problem was formulated by A. Lindholm and P. Giselsson, and the QP reformulation of the problem was performed by P. Giselsson. The examples were constructed by A. Lindholm.

Paper III

Lindholm, A. and K. Forsman (2013): "Empirical models for utility disturbances in the process industry." Submitted to *The 19th World Congress of the International Federation of Automatic Control*. Cape Town, South Africa.

This paper introduces a general method for identifying empirical models of disturbances in the supply of utilities. The method includes prefiltering of the measurement data to exclude very infrequent disturbances, and investigates both periodic and nonperiodic disturbances. Statistical models are produced for measures such as duration and severity of a disturbance, and for time between failures. The method is presented step by step, and a case study is performed using industrial data from Perstorp.

A. Lindholm contributed with the main idea of the method and did the initial analysis of the disturbance data. K. Forsman assisted with valuable insights based on his experience from the process industry and helped constructing the final statistical models of the disturbance measures. The case study was performed by A. Lindholm using industrial data provided by K. Forsman.

Paper IV

Lindholm, A. and N. Nytzén (2013): "Hierarchical scheduling and disturbance management in the process industry." Submitted to *Computers and Chemical Engineering.*

The paper introduces a hierarchical structure for scheduling and disturbance management. Production scheduling (PS) with a horizon of one month and detailed production scheduling (DPS) with a horizon of one day is discussed, and a method for integrating these two activities is suggested. A list of specifications provided by Perstorp is given and used as a starting point for formulating the PS and DPS activities as optimization problems. It is suggested to solve the optimization problems in receding horizon fashion, and to enable updates of the input information at each time step. An example is given, which is constructed to resemble a real industrial site.

The initial idea for the hierarchical structure was provided by A. Lindholm together with C. Johnsson, N. Quttineh, H. Lidestam, M. Henningsson, J. Wikner, O. Tang, N. Nytzén, and K. Forsman (see citation in the paper). The scheduling procedure was further developed by A. Lindholm. The specification list was provided by N. Nytzén, and the mathematical formulation of the scheduling problems was performed by A. Lindholm using valuable inputs from P. Giselsson, N. Quttineh, H. Lidestam, C. Johnsson, and K. Forsman (see citations in the paper). The example was constructed by A. Lindholm.

Additional Publications

Lindholm, A. (2011): "Utility disturbance management in the process industry." Licentiate Thesis ISRN LUTFD2/TFRT–3253–SE. Department of Automatic Control, Lund University, Sweden.

The licentiate thesis includes Paper I and initial ideas for Paper II.

Lindholm, A., K. Forsman and C. Johnsson (2010): "Buffer management strategies for improving plant availability." In *Proceedings of Reglermöte 2010*, Lund, Sweden.

Lindholm, A., K. Forsman and C. Johnsson (2010): "A general method for defining and structuring buffer management problems." In *Proceedings of the 2010 American Control Conference*, pp. 4397–4402. Baltimore, ML.

Lindholm, A., H. Carlsson and C. Johnsson (2010): "Availability estimations for utilities in the process industry." In *Proceedings of the 16th Nordic Process Control Workshop.* Lund, Sweden.

Lindholm, A. (2011): "A method for improving plant availability with respect to utilities using buffer tanks." In *Proceedings of the 31st IASTED Conference on Modelling, Identification and Control*, pp. 378–383. Innsbruck, Austria.

Lindholm, A., H. Carlsson and C. Johnsson (2011): "Estimation of revenue loss due to disturbances on utilities in the process industry." In *Proceedings of the 22nd Annual Conference of the Production and Operations Management Society.* Reno, NV.

Lindholm, A., H. Carlsson and C. Johnsson (2011): "A general method for handling disturbances on utilities in the process industry." In *Proceedings of the 18th IFAC World Congress*, pp. 2761–2766. Milano, Italy.

Lindholm, A., A. Widd, A. Sootla and A. Sahlberg (2011): "Utvärdering av förståelse på skriftlig tentamen." In *3:e Utvecklingskonferensen för Sveriges ingenjörsutbildningar.* Linköping, Sweden.

Lindholm, A., C. Johnsson, T. Hägglund and H. Carlsson (2012): "Reducing revenue loss due to disturbances in utilities using buffer tanks – A case study at Perstorp." In *Proceedings of the Conference on Foundations of Computer-Aided Process Operations.* Savannah, GA.

Lindholm, A., C. Johnsson, T. Hägglund and H. Carlsson (2012): "Reducing revenue loss due to disturbances in utilities using buffer tanks – A case study at Perstorp." In *Proceedings of the 17th Nordic Process Control Workshop.* Lyngby, Denmark.

Lindholm, A. and C. Johnsson (2012): "A tool for utility disturbance management." In *Proceedings of the 14th IFAC Symposium on Information Control Problems in Manufacturing*, pp. 122–127. Bucharest, Romania.

Lindholm, A. and P. Giselsson (2012): "Formulating an optimization problem for minimization of losses due to utilities." In *Proceedings of the 8th IFAC Symposium on Advanced Control of Chemical Processes*, pp. 567–572. Singapore.

Lindholm, A., C. Johnsson, N. Quttineh, H. Lidestam, M. Henningsson, J. Wikner, O. Tang, N. Nytzén and K. Forsman (2013): "Hierarchical scheduling and utility disturbance management in the process industry." In *Proceedings of the IFAC Conference on Manufacturing Modelling, Management and Control*. Saint Petersburg, Russia.

Lindholm, A., P. Giselsson, N. Quttineh, H. Lidestam, C. Johnsson and K. Forsman (2013): "Production scheduling in the process industry." In *Proceedings of the 22nd International Conference on Production Research*. Iguassu Falls, Brazil.

Lindholm, A. and N. Nytzén (2013): "Hierarchical production scheduling in the process industry." In *Proceedings of the 18th Nordic Process Control Workshop*. Oulu, Finland.

1

Introduction

The research included in this thesis focuses on control of production at large-scale process industrial sites in the presence of disturbances. The research has been performed within the Process Industry Centre (PIC) at Lund University in close collaboration with process industrial companies, in particular with Perstorp. An inductive research approach has been used, where observations and scenarios at industrial sites have been studied and used to form new theories and methods. This approach is encouraged in the description of the grand challenges in process control, [Craig et al., 2011], where it is stated that research activities in process control need to be more focused on addressing industrial problems, rather than solving purely theoretical problems.

Initially, the research project focused on disturbances in utilities, such as steam and cooling water. A general method for finding both proactive and reactive disturbance management strategies was introduced. Since the reactive approaches involve control of the production at a site, the integration of utility disturbance management with production planning and scheduling was studied during the final phase of the research project. To provide a background and motivation for the research presented in the thesis, the process industry, PIC, and the company Perstorp are described in this chapter.

Figure 1.1 A process industrial site located in Stenungsund, Sweden.

1.1 The Process Industry

Process industries are industries in which raw materials are physically or chemically transformed, or where energy and material streams interact and transform each other [Craig et al., 2011]. In the report [IVA-M 353, 2006] by The Royal Swedish Academy of Engineering Sciences, IVA, the process industry is defined to consist of

- Pulp and paper industries
- Chemical and plastics industries
  (including petroleum and pharmaceutical production)
- Mining industries
- Iron and steel industries
- Food processing industries

The production in a process industrial company typically consists of refinement and processing of some raw materials, such as wood, crude oil, or natural gas. The industries are typically both raw material and capital intensive. The products are often raw materials for other producing companies and not products to be sold directly to end users. The production is usually automated to a high degree. In Sweden, the process industry is highly export oriented and stands for around 30% of the total export [IVA-M 353, 2006].

During the last decades, the chemical process industry has become a global marketplace with strong competition between manufacturers [Tousain, 2002]. The global competition means that the processing equipment is not the determining factor. Instead, know-how and process understanding become key factors for success. Market saturation and downward price pressures require chemical process industrial companies to continuously improve the operational efficiency and profitability to remain competitive [Bakhrankova, 2010]. The process industry is dominated by raw material and energy costs, which makes it essential to minimize these costs by efficient use of resources. The Swedish chemical industry has grown significantly during the past decade, and now constitutes about 14% of the total export[1] in Sweden [IVA-M 353, 2006].

[1] including plastics and pharmaceutical production

1.2 Process Industry Centre

The research presented in this thesis was performed within the framework of PIC, a Process Industry Centre located at Lund University and Linköping University [PIC, 2013]. The overall goal of PIC is to provide knowledge for the Swedish process industry to ensure future success, through research and competence development. The academic disciplines of Chemical Engineering, Automatic Control and Production Economics form the centre together with several industrial partners from the process industry.

In the research program, methods and tools for modeling, optimization and control of industrial processes are developed in order to improve production systems with respect to flexibility, controllability, and availability. The methods and the tools are developed from specific solutions to process control problems suggested by the industrial partners. The goal is to make the results from PIC industrially relevant, not only for the participating industries, but on a wide scale in process management, operation and automation. There are currently 17 industrial partners, including Siemens, Perstorp, Novo Nordisk, ABB, and K.A. Rasmussen.

In the competence development program, the main goal is to increase the competence level of process management, process optimization and control in industry as well as in academy. The goal is achieved through a master program in process management, industrial courses, PhD courses, Master theses, and conferences.

1.3 Perstorp

The research presented in this thesis was conducted in close collaboration with Perstorp, a world-leading company within several sectors of the specialty chemicals market [Perstorp, 2013]. Perstorp mainly produces additives for other chemical industries, such as additives for paints, coatings, and plastic processing. Some of their main product groups are polyols, organic acids, and esters, and their products can be found in for example automotive, food, packaging, and electronics applications. Perstorp is controlled by the French company PAI Partners and has about 1500 employees in 22 countries, and an annual turnover of more than 10 billion SEK in 2012. The company has 10 production sites around the world, which are shown in Figure 1.2. The production sites typically run in a continuous mode, without any product changes or grade changes. The aim of Perstorp is to run its production sites in a well-defined way even when there are site-wide disturbances such as disruptions in a utility or raw material.

Figure 1.2 Perstorp sites around the world.

2

Utility Disturbance Management

Utility disturbance management is the main topic of this thesis. In this chapter, some of the main challenges related to this area are discussed, and some terminology that is useful when reading Papers I-IV is introduced.

2.1 Challenges Related to Utility Disturbance Management

Utilities are support processes that are utilized in production, but are not part of the final product. At process industrial sites, utilities such as steam, cooling water, and electricity are very commonly used. Moreover, the utility costs often represent a large part of the total operating cost [Iyer and Grossmann, 1998]. Disturbances in the supply of utilities may also cause the production in affected areas to slow down or stop, which implies great loss of revenue. The production areas at a site are often connected by the flow of raw materials and products, and they also often share the same utilities. Thus, utility disturbance management becomes an intricate topic, since the production in each area at a site is affected both by a utility disturbance directly, and by the production in its neighboring areas, which may or may not be affected by the utility disturbance. One way to view this is that a site is a complex network of areas that interact with each other. These *process manufacturing networks* are mentioned as one of the grand challenges for control in [Marquardt and Frankl, 2011].

The great losses associated with poor utility disturbance management clearly motivates research in this area. To the best of my knowledge, not much work has been done regarding management of utility disturbances. A review of the related research areas is given in Paper I.

2.2 Proactive and Reactive Disturbance Management

Two main approaches of utility disturbance management can be identified: *proactive disturbance management* and *reactive disturbance management.* Proactive disturbance management aims to prevent future disturbance occurrences, whereas reactive disturbance management is strategies for handling disturbances when they occur. Paper I outlines a general method for finding both proactive and reactive utility disturbance management strategies, but focuses on proactive approaches. In Paper II, an optimization problem for optimal assignment of utilities to different areas during disturbances is formulated, which may be used both for proactive and reactive disturbance management. The reactive approaches involve decisions on how to share the available utility resources among the areas that require the utility. If the modeling approach in Paper II is used, this becomes equivalent to how to control the production in the areas affected by the disturbance. This means that utility disturbance management becomes closely connected to production planning and production scheduling, since all of these areas concern production control at the site/area level of an enterprise. Production scheduling for batch processes in the presence of utility constraints has previously been studied by, among others, [Kondili et al., 1993] and [Sundaramoorthy et al., 2009]. The integration of utility disturbance management with production scheduling for continuous processes is studied in Paper IV.

2.3 Utility Disturbance Models

To develop good reactive utility disturbance management strategies, the characteristics of disturbances in the supply of the utilities must be known, which is usually not the case at the time of the disturbance [Lee and Weekman, 1976]. To the best of my knowledge, there have been few attempts to model utility disturbances in the process industry. For other types of disturbances, characteristics such as *repair time* and *time to failure* have been modeled. The characteristics of the distributions that are commonly used to model these measures are discussed below.

The *lognormal distribution* is often used to model repair time for applications with repairable units [Rausand and Høyland, 2004]. The lognormal distribution has two parameters, μ and σ, and the probability density function (pdf) given μ and σ is

$$f(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} x^{-1} e^{-(\ln x-\mu)^2/(2\sigma^2)}, \quad 0 < x < \infty \tag{2.1}$$

with $\sigma > 0$. Some examples of probability density functions of lognormal distributions are given in Figure 2.1. The fact that the lognormal distribution

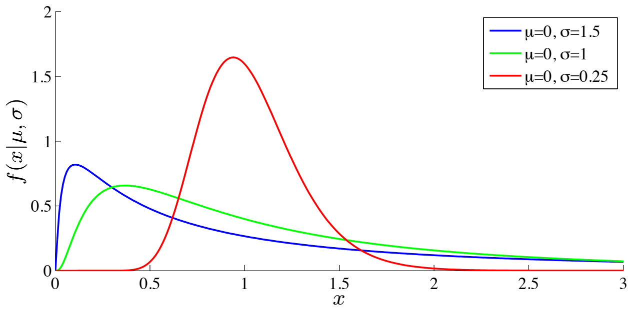

Figure 2.1 Lognormal probability density functions.

is commonly used for modeling repair time can be explained by the shape of the lognormal pdf, especially for distributions with larger σ, which makes the distribution more skewed. Consider the pdf for $\mu = 0$, $\sigma = 1.5$ in Figure 2.1 as a model for repair rate, where the x-axis indicates the elapsed repair time. As seen in Figure 2.1, the probability of completing the repair action within a short time interval is high. However, when the repair has been going on for a long time, the repair rate decreases with the repair time. This captures the situation of serious problems; for example that there are no spare parts available on the site [Rausand and Høyland, 2004].

The *Weibull distribution* is often used to model time to failure or the life span of a component [Pal et al., 2005]. The distribution has one scale parameter, a, and one shape parameter, b, and the pdf is given by

$$f(x|a,b) = \frac{b}{a^b} x^{b-1} e^{-(x/a)^b}, \quad 0 < x < \infty \tag{2.2}$$

where $a, b > 0$. Some examples of probability density functions of Weibull distributions are given in Figure 2.2.

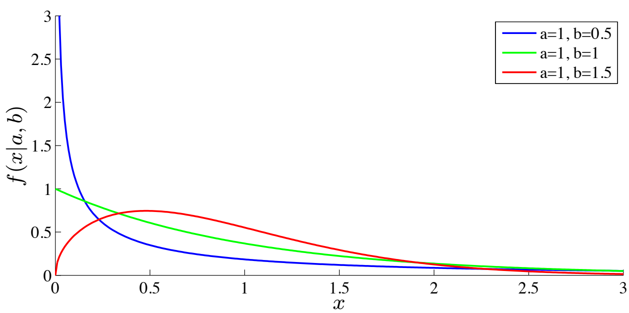

Figure 2.2 Weibull probability density functions.

A measure that is often used in the area of reliability engineering and risk assessment is the *failure rate*, which is the frequency with which a system or component fails. The failure rate function, $\lambda(t)$, is given by

$$\lambda(t) = \frac{f(x)}{1 - F(x)} \tag{2.3}$$

where $f(x)$ is the pdf and $F(x)$ is the cumulative distribution function (cdf). The cdf relates to the pdf as

$$F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t \tag{2.4}$$

For the Weibull distributions in Figure 2.2, the failure rate functions are depicted in Figure 2.3. As seen in Figure 2.3, the shape parameter, b, defines the characteristics of the failure rate for the Weibull distribution. If b is less than one, the failure rate decreases over time. For the life span of a component, this means that failures are more likely to occur early in the life of the component [Minitab, 2011]. For time to failure (also denoted time between failures), this indicates that failures are more likely to occur shortly after the previous failure than after a longer time. If b is greater than one, this indicates that the failure rate increases with time. These distributions are commonly used to model wear-out failures [Minitab, 2011]. For $b = 1$, the Weibull distribution becomes an exponential distribution, and the failure rate is constant over time. For time between utility failures it is in many cases reasonable to assume that the failure rate decreases over time, since poor operation of a utility may generate several consecutive failures.

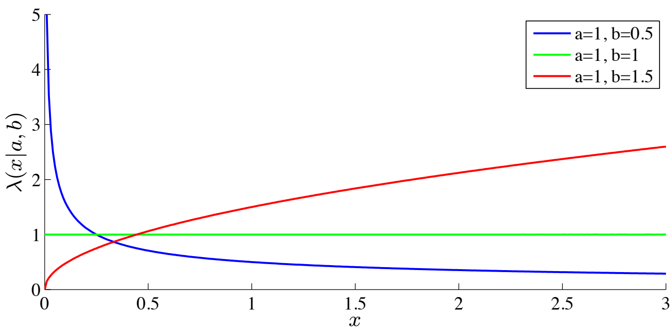

Figure 2.3 Failure rate functions for Weibull distributions.

Paper III presents a general method for finding utility disturbance models given measurement data. In this paper, the lognormal distribution is used to model the duration of utility disturbances, and the Weibull distribution for modeling the time between utility disturbances. The obtained disturbance models may be used as input for reactive disturbance management, for example as initial predictions of utility disturbances when they occur. The models presented in Paper III may be used for reactive utility disturbance management as presented in Paper II and IV.

3

Production Planning, Scheduling, and Control

The strong global competition in the process industry that has evolved during the past decades requires flexibility and reduction of production costs to remain competitive [Backx et al., 1998]. Planning, scheduling, and control are functions that have large economic impact on process industry operations [Shobrys and White, 2002]. In this section, the activities of planning, scheduling, and control are defined, and challenges related to these activities are discussed.

3.1 Production Planning

Production planning focuses on providing production goals over a time scale of weeks to months [Shobrys and White, 2002; Huang, 2010]. A more general definition is given by [APICS, 2013] as: "the process of setting goals for the organization and choosing various ways to use the organization's resources to achieve the goals". The questions that are addressed by the planning layer are, e.g., which products to make, which feedstocks to buy and how much to produce of each product [Huang, 2010]. Typical planning practice varies by industry, but most tools are still primarily spreadsheet based [Shobrys and White, 2002].

3.2 Production Scheduling

Scheduling defines the specific activities to be performed and the timing of these events over a time scale of days to weeks [Shobrys and White, 2002; Huang, 2010]. The scheduling activity also has to continuously compare what was expected with what actually happened. Normally, the scheduling activity passes the desired production rates and product quality limits to the control functions [Shobrys and White, 2002]. The tools that are used for scheduling today are commonly spreadsheet based [Shobrys and White, 2002].

In Paper IV, two production scheduling activities at different time scales are considered; *production scheduling* and *detailed production scheduling*. This choice of names complies with the standard [ISA-95.00.01, 2009]. In Paper IV, it is suggested for the production scheduling to operate over a horizon of one month with a time step of one day, and the detailed production scheduling to operate over a horizon of one day with a time step of one hour. The production scheduling activity can thus be seen as similar to the notation of *scheduling* in other papers, e.g., [Shobrys and White, 2002] and [Engell and Harjunkoski, 2012].

3.3 Process Control

Process control is defined as the technologies required to design and implement control systems in the process industry, and the goal is to bring about and maintain the conditions of a process at desired or optimal values [Craig et al., 2011]. Process control may be both basic control functions that are used to improve the dynamic response and reject disturbances, and advanced control for improving the economic performance of a process [Craig et al., 2011]. Thus, the time scale of process control may vary. In [Shobrys and White, 2002], it is stated that process control systems address real time execution over time horizons of seconds to minutes. For advanced process control or process control that aims for economic optimization, the time scales may be longer, e.g., hours. Advanced process control may be obtained either by the coordination of several PID controlled loops, or by model-based control, e.g., model predictive control (MPC). More information on MPC is provided in Chapter 4 of this thesis. Other advanced process control strategies that have been successfully used in the chemical industry are real-time optimization (RTO), statistical process control (SPC), and plant-wide loop oscillation detection [Craig et al., 2011]. The detailed production scheduling activity used in Paper IV may be seen as similar to the notation of *advanced process control* in other papers, e.g., [Shobrys and White, 2002] and [Engell and Harjunkoski, 2012].

3.4 Integration of Planning, Scheduling, and Control

Planning, scheduling, and control of process systems are activities that are closely linked. Integration pyramids are commonly used to show the hierarchy of these activities within an enterprise. These pyramids exist in many versions, which might be explained by the fact that people working in the field of process control come from many different areas [Tousain, 2002]. The integration pyramid that is used in [Craig et al., 2011] is depicted in Figure 3.1. Integration of the levels are crucial, and cross-layer integrated control solutions are needed [Craig et al., 2011]. Some work has been done on integrating the planning and scheduling activities, either by combining them and solving the planning and scheduling problem simultaneously, or by various decomposition techniques. An extensive review is provided in [Grossmann and Furman, 2009]. [Maravelias and Sung, 2009] present the major modeling approaches for the integration of planning and scheduling and discuss strategies for solving the problem. The topic of integrating planning and scheduling with control, on the other hand, is a topic that still has not received much attention in the literature [Craig et al., 2011; Grossmann, 2012]. The integration of the activities may be performed either using a monolithic approach, where both the scheduling and control problem characteristics are fully represented and the problems are solved simultaneously, or using a hierarchical approach, where the scheduling and control problems are solved separately but the activities communicate in some way. The hierarchical approach currently seems to be the only realistic approach for industrial-size problems [Engell and Harjunkoski, 2012]. In Paper IV, a hierarchical approach for integrating the production scheduling and detailed production scheduling activities is introduced.

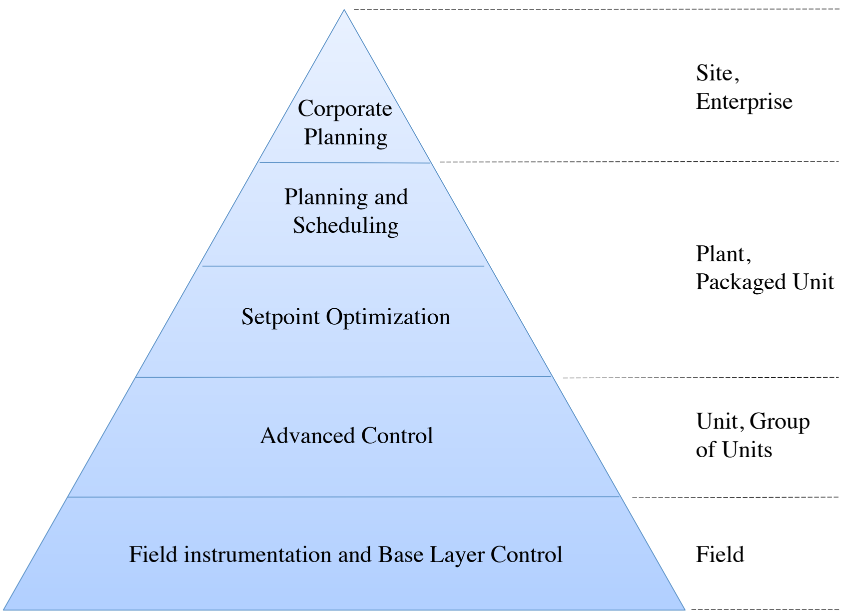

Figure 3.1 Integration pyramid for an enterprise.

3.5 Challenges Related to Planning, Scheduling, and Control

The main challenges related to production planning, scheduling, and control lie within the integration of the activities. The ability to achieve integration of the activities has in the past been limited by technical issues such as limited computer capabilities, lacking data capture, and poor user interfaces [Shobrys and White, 2002]. Today, the technical issues have been addressed, but other challenges remain in order to achieve an integrated planning, scheduling, and control system that is industrially useful. Four challenges are mentioned in [Grossmann and Furman, 2009]: the *modeling challenge*, the *multiscale challenge*, the *uncertainty challenge*, and the *algorithmic and computational challenge*. The modeling challenge involves the problem of how to develop models that capture the complexity of the operations in planning, scheduling, and control. Since the activities operate at different time scales, the models have to be coordinated over a time horizon that can reach from seconds to years. This is denoted the multiscale challenge. The uncertainty challenge is the challenge of how to handle the effect of uncertainties such as equipment breakdown or varying demands. For example, if there is a short-term limit in a multi-product plant that restricts the production below the desired amount, it must be decided which product that should be reduced first [Shobrys and White, 2002]. Creating models that account for the issues mentioned above may lead to challenging optimization problems. The algorithmic and computational challenge deals with the issues of how to efficiently solve these problems.

Apart from these challenges, there are challenges that relate more to the actual implementation of the techniques in the industry. Two challenges that are mentioned in [Shobrys and White, 2002] are the changing of human behavior to get acceptance of the more sophisticated tools, and the changing of the organizational structure to enable integration. It is mentioned in [Engell and Harjunkoski, 2012] that if the solutions are not accepted by the operators, they are not used in the long term. [Engell and Harjunkoski, 2012] also mentions the challenge of integration of different software systems, and the challenge of the design of the integrated systems, such as the communication between different functions. Paper IV in this thesis aims to present an industrially relevant structure for formulating and integrating the scheduling and detailed production scheduling problems. The level of detail of the models presented in the paper only takes the most important issues into account, but the aim is to present a general structure for the models and the integration, such that more details may be added to the models at a later stage.

4

Optimization

Increased pressure to reduce costs motivates the development of optimization tools for obtaining better solutions for plant operation in chemical plants [Grossmann and Biegler, 1995]. Optimization is a main part of the methods for reactive utility disturbance management and production scheduling that are presented in this thesis. In this chapter, the types of optimization problems that appear in Paper II and IV are described.

4.1 Optimization Problems

A mathematical optimization problem can be written on the form

$$\begin{array}{lll} \text{minimize} & f_0(x) & \\ \text{subject to} & g_i(x) \leq 0, & i = 1, \dots, d \\ & h_j(x) = 0, & j = 1, \dots, p \end{array} \tag{4.1}$$

where $x = (x_1, \dots, x_n)$ is the *optimization variable* and $f_0 : \mathbb{R}^n \to \mathbb{R}$ the *objective function*. The functions $g_i : \mathbb{R}^n \to \mathbb{R},\ i = 1, \dots, d$ define the *inequality constraints* and $h_j : \mathbb{R}^n \to \mathbb{R},\ j = 1, \dots, p$ the *equality constraints*. The types of optimization problems that are handled in this thesis are special cases of the optimization problem (4.1). These are described in the subsections below. In the thesis, all optimization problems are solved with the solvers in the CPLEX software package [CPLEX, 2009].

Linear Program

If the objective function $f_0(x)$ and constraint functions $g_i(x)$ and $h_j(x)$ in (4.1) are affine, the optimization problem (4.1) is called a *linear program* (LP). LPs are *convex optimization problems* (see [Boyd and Vandenberghe, 2004]) and there exists efficient methods for solving them. There is a large number of both free and commercial solvers available for solving these problems. Linear programming is widely used in a lot of applications; transportation, energy, telecommunications, and manufacturing, to mention some. In this thesis, linear programming is used to solve a steady-state production scheduling problem in Paper II.

Mixed-integer linear program If the objective function $f_0(x)$ and constraint functions $g_i(x)$ and $h_j(x)$ in (4.1) are affine like for the LP, but at least one element of x is required to be an integer, (4.1) is called a *mixed-integer linear program* (MILP). In contrast to linear programming, MILPs may be hard to solve since they are $\mathcal{NP}$ complete, which means that in the worst case, the solution time grows exponentially with the number of variables [Raman and Grossmann, 1991]. There exists good solvers for these types of problems, but most of them are commercial. Most of the commercial solvers use a combination of cutting plane methods and branch-and-bound (see e.g. [Nemhauser and Wolsey, 1988]) to solve the problems.

Mixed-integer linear programming is commonly used for solving production planning and scheduling problems, and within application areas such as telecommunications and cellular networks. Despite the computational complexity, there are examples of real world problems that have been solved using mixed-integer linear programming. One example is the combined operational and strategic planning problem for a chemical industry that is presented in [Kallrath, 2002]. In this thesis, mixed-integer linear programming is used to solve production scheduling problems in Paper IV.

Quadratic Program

The optimization problem (4.1) is called a *quadratic program* (QP) if the objective function $f_0(x)$ is quadratic, the constraint functions $g_i(x)$ and $h_j(x)$ are affine, and $x \in \mathbb{R}^n$. QPs are, like LPs, convex optimization problems, and there exist many free and commercial solvers for these problems. Some applications where quadratic programming is used are portfolio optimization in finance, power generation optimization for electrical utilities, and design optimization in engineering.

Mixed-integer quadratic program If the objective function $f_0(x)$ is quadratic and the constraint functions $g_i(x)$ and $h_j(x)$ in (4.1) are affine like for the QP, but at least one element of x is required to be an integer, (4.1) is called a *mixed-integer quadratic program* (MIQP). MIQPs are, as MILPs, $\mathcal{NP}$ complete to solve, which may make such formulations unsuitable for large problems. An option if the solution time becomes unacceptably long is relaxing the problem to a QP, which can be done in various ways. Some suggestions of convex relaxations of MIQPs are given in [Axehill et al., 2010]. In this thesis, MIQPs are used to formulate the production scheduling and disturbance management problems in Papers II and IV. Paper II also suggests a QP relaxation of the problem to enable solution of larger problems.

4.2 Model Predictive Control

The basic concept of *model predictive control* (MPC) is to forecast the behavior of a system using a dynamic model, and to optimize the forecast to produce the best decision [Rawlings and Mayne, 2009]. Consider a model of a plant formulated as a difference equation

$$x_{t+1} = f(x_t, u_t) \tag{4.2}$$

where $x_t \in \mathbb{R}^n$ is the state of the plant at time t, $u_t \in \mathbb{R}^m$ is the input signal vector to the plant at time t, and $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ describes the plant dynamics. Models of this type can be included into the equality constraints of an optimization problem on the form (4.1) to optimize plant operation. An optimization problem may be formulated to, e.g., minimize losses at a site for chemical production. In that case, the objective function describes how the states and inputs of the plant contribute to the losses. The states could for example be the levels of the buffer tanks between areas at a site, and the inputs could be the production rates in the areas. Constraints may include capacity limitations on the buffer tank levels and production rates. If the optimization problem should be solved once over the time horizon N, (4.1) translates to

$$\begin{array}{ll} \text{minimize} & f_0(\mathbf{x}, \mathbf{u}) \\ \text{subject to} & g_i(\mathbf{x}, \mathbf{u}) \leq 0, \quad i = 1, \ldots, d \\ & h_j(\mathbf{x}, \mathbf{u}) = 0, \quad j = 1, \ldots, p \end{array} \tag{4.3}$$

where $\mathbf{x} = [x_t^T, x_{t+1}^T, \ldots, x_{t+N-1}^T]$ are the states, and $\mathbf{u} = [u_t^T, u_{t+1}^T, \ldots, u_{t+N-1}^T]$ are the inputs of the plant over the time horizon N. The model (4.2) is included in the equality constraints, $h_j(\mathbf{x}, \mathbf{u})$, and the initial conditions x_t are assumed to be given from measurements. Solving (4.3) once generates a solution trajectory with suggested inputs to the plant, $\mathbf{u}_t^\star$, from the current time t to the future time $t+N-1$, and a predicted behavior of the system, $\mathbf{x}_t^\star$. This is denoted *fixed-horizon optimization*. A drawback with fixed-horizon optimization is that if something unexpected happens at some time over the future interval $[t, t+N-1]$ that was not included in the model, the solution $\mathbf{u}_t^\star$ may no longer be optimal. A solution to this problem is to perform the optimization in *receding horizon*. This means that at time t and for the current state of the system, x_t, the optimization problem (4.3) is solved over a fixed future interval $[t, t+N-1]$ (with horizon N). The first step in the resulting optimal control trajectory $\mathbf{u}_t^\star$ is applied as the input signal to the system. Then, the state x_{t+1} that is reached at time $t+1$ is measured, and the fixed horizon optimization is repeated over the future interval $[t+1, t+N]$.

Model predictive control exists in many variants. Linear model predictive control and mixed-integer predictive control are described below. For more information on model predictive control and its variants, extensions and applications, see e.g., [Maciejowski, 2002] or [Rawlings and Mayne, 2009].

Linear Model Predictive Control

The most commonly used variant of MPC is linear MPC, where the constraints and the dynamics of the model are linear, and where a quadratic objective function is used. This implies that the optimization problem (4.3), which is to be solved at each time step, becomes a quadratic program. Since there exists efficient solvers for these problems, linear MPC has become very successful for industrial applications in a wide range of areas [Qin and Badgwell, 1997]. In this thesis, linear MPC is used in Paper II to minimize economical losses due to utility disturbances at a process industrial site.

Mixed-Integer Predictive Control

If some plant dynamics or constraints are defined by on/off characteristics, the optimization variables in the optimization problem (4.3) may be both continuous and binary. If the model and the constraints are linear, the system to be optimized is denoted a *mixed logical dynamical* (MLD) system. In [Bemporad and Morari, 1999], MLD systems are defined and an MPC framework for these systems is introduced, denoted *mixed integer predictive control* (MIPC). Extending the ordinary linear MPC to MIPC means that a MIQP has to be solved [Axehill, 2005]. Since these problems are $\mathcal{NP}$ complete to solve, industrial applications of MIPC are limited to systems of reasonable size. In this thesis, MIPC is used in Paper IV to solve a production scheduling problem.

5

Discussion

5.1 Summary

The thesis presents methods that could be used decrease the costs related to disturbances in the supply of utilities at industrial sites. A general method for decreasing the revenue loss due to utility disturbances is presented, and used together with production models of different level of detail. A simple on/off modeling approach is used to quickly assess which utilities that cause the greatest losses at a site. To simplify the related calculations, a matrix representation of a site is introduced.

The on/off model yields proactive disturbance management strategies only. To obtain reactive strategies, a modeling approach for modeling the effect of utility disturbances on production is presented. The utility model is used to formulate an optimization problem for minimization of losses at utility disturbances, where also production rate, inventory, and market constraints are taken into consideration. The solution to the optimization problem may be used as decision support for how the production should be controlled, and how inventories should be used, to minimize the economical effects of a disturbance.

The optimization problem formulation requires a prediction of the occurring utility disturbance to give decision support on how to control the production. A general method for identifying statistical disturbance models of utility disturbances from measurement data is presented to facilitate generation of the predicted trajectories.

Since the reactive utility disturbance management strategies involve how to control the production at a site, it is important to connect this with the other production planning and production scheduling functions at the site. A two-level hierarchical structure for production scheduling and utility disturbance management is presented, and the integration of the two levels is discussed. One optimization problem structure for each level is suggested, taking into account several constraints and specifications that are present at real process industrial sites.

Industrial data and site structures have been used to ensure that the methods presented in the thesis are useful for the process industry.

5.2 Future Work

The research presented in this thesis may be extended in several directions. One possibility is to strive for more accurate modeling of the effects of utility disturbances on production. A link is needed between the utility disturbance models in terms of deviation from normal operation, such as a drop in the steam pressure or a peak in the cooling temperature, and the effect of these disturbances on production in an area. These models are hard to obtain for integrated sites since the production areas are connected, and it is not clear directly from measurements what effects originate from the utility disturbance and what are indirect effects due to the area interconnections. To obtain good utility models, experience from process operators and physical laws for utilities probably have to be utilized in addition to the measurement data.

Another possible research direction is to extend the production scheduling and detailed production scheduling models to take more specifications into account and become more industrially relevant. This is currently done within the Process Industry Centre, described in Chapter 1. A major aspect to take into account is the transportation of products from the site via, e.g., trucks or boats, which connects the production scheduling problem to the supply chain. It would also be interesting to apply the integrated scheduling solutions to various industrial sites and compare the solutions with current practice. A related study within the Process Industry Centre is the study of inventory management at process industrial sites. The aim is to also interconnect this research with the research on production scheduling and control.

Bibliography

APICS (2013). *The Association for Operations Management.* [online] `http://www.apics.org/`, June 2013.

Axehill, D., L. Vandenberghe, and A. Hansson (2010). “Convex relaxations for mixed integer predictive control”. *Automatica* **46**:9, pp. 1540–1545.

Axehill, D. (2005). *Applications of integer quadratic programming in control and communication.* Licentiate thesis, Linköping University, Sweden.

Backx, T., O. Bosgra, and W. Marquardt (1998). “Towards intentional dynamics in supply chain conscious process operations”. In: *Proceedings of the international conference on Foundations Of Computer-Aided Process Operations.* Snowbird, UT.

Bakhrankova, K. (2010). “Decision support system for continuous production”. *Industrial Management & Data Systems* **110**:4, pp. 591–610.

Bemporad, A. and M. Morari (1999). “Control of systems integrating logic, dynamics, and constraints”. *Automatica* **35**:3, pp. 407–427.

Boyd, S. and L. Vandenberghe (2004). *Convex Optimization.* Cambridge University Press, USA.

CPLEX (2009). *IBM ILOG CPLEX V12.1: User’s manual for CPLEX.* International Business Machines Corporation.

Craig, I., C. Aldrich, R. Braatz, F. Cuzzola, E. Domlan, S. Engell, J. Hahn, V. Havlena, A. Horch, B. Huang, M. Khanbaghi, A. Konstantellos, W. Marquardt, T. McAvoy, T. Parsini, S. Pistikopoulos, T. Samad, S. Skogestad, N. Thornhill, and J. Yu (2011). “Control in the process industries”. In: *The impact of control technology.* Ed. by T. Samad and A. M. Annaswamy. Available at `www.ieeecss.org`. IEEE control systems society.

Engell, S. and H. Harjunkoski (2012). “Optimal operation: scheduling, advanced control and their integration”. In: *Proceedings of the international conference on Foundations Of Computer-Aided Process Operations.* Savannah, GA.

Grossmann, I. E. (2012). "Advances in mathematical programming models for enterprise-wide optimization". In: *Proceedings of the international conference on Foundations Of Computer-Aided Process Operations*. Savannah, GA.

Grossmann, I. E. and L. T. Biegler (1995). "Optimizing chemical processes". *Chemtech* **25**:12, pp. 27–35.

Grossmann, I. E. and K. C. Furman (2009). "Challenges in enterprise-wide optimization for the process industries". In: *Optimization and Logistics Challenges in the Enterprise*. Vol. 30. Springer Optimization and Its Applications. Springer, USA, pp. 3–59.

Huang, R. (2010). *Nonlinear model predictive control and dynamic real time optimization for large-scale processes*. PhD thesis. Carnegie Mellon University, PA.

ISA-95.00.01 (2009). *Enterprise-Control System Integration, Part 1: Models and Terminology*. Standard. Instrument Society of America, pp. 17–28.

IVA-M 353 (2006). *Ökad konkurrenskraft för svensk processindustri*. Report. The Royal Swedish Academy of Engineering Sciences.

Iyer, R. R. and I. E. Grossmann (1998). "Synthesis and operational planning of utility systems for multiperiod operation". *Computers & Chemical Engineering* **22**:7, pp. 979–993.

Kallrath, J. (2002). "Combined strategic and operational planning – an MILP success story in chemical industry". *OR Spectrum* **24**:3, pp. 315–341.

Kondili, E., C. C. Pantelides, and R. W. H. Sargent (1993). "A general algorithm for short-term scheduling of batch operations – I. MILP formulation". *Computers & Chemical Engineering* **17**:2, pp. 211–227.

Lee, W. and V. W. Weekman (1976). "Advanced control practice in the chemical process industry: A view from industry". *AIChE Journal* **22**:1, pp. 27–38.

Maciejowski, J. M. (2002). *Predictive Control with Constraints*. Prentice Hall, UK.

Maravelias, C. T. and C. Sung (2009). "Integration of production planning and scheduling: overview, challenges and opportunities". *Computers & Chemical Engineering* **33**:12, pp. 1919–1930.

Marquardt, W. and K. Frankl (2011). "Process manufacturing networks". In: *The impact of control technology*. Ed. by T. Samad and A. M. Annaswamy. Available at `www.ieeecss.org`. IEEE control systems society.

Minitab (2011). *Technical support document: distribution models for reliability data*. [online] `http://www.minitab.com/support/documentation/Answers/Reliability_Distribution.pdf`, June 2013.

Nemhauser, G. L. and L. A. Wolsey (1988). *Integer and combinatorial optimization.* Vol. 18. Wiley, USA.

Pal, N., C. Jin, and W. K. Lim (2005). *Handbook of Exponential and Related Distributions for Engineers and Scientists.* Chapman and Hall/CRC, UK.

Perstorp (2013). [online] `http://www.perstorp.com/`, June 2013.

PIC (2013). *Process Industry Centre – providing knowledge for the process industry.* [online] `http://www.processindustrielltcentrum.se/`, June 2013.

Qin, S. J. and T. A. Badgwell (1997). "An overview of industrial model predictive control technology". In: *AIChE Symposium Series.* Vol. 93. 316, pp. 232–256.

Raman, R. and I. E. Grossmann (1991). "Relation between MILP modelling and logical inference for chemical process synthesis". *Computers & Chemical Engineering* **15**:2, pp. 73–84.

Rausand, M. and A. Høyland (2004). *System Reliability Theory: Models, Statistical Methods, and Applications.* John Wiley & Sons, Inc., USA.

Rawlings, J. B. and D. Q. Mayne (2009). *Model Predictive Control: Theory and Design.* Nob Hill Publishing, USA.

Shobrys, D. E. and D. C. White (2002). "Planning, scheduling and control systems: why cannot they work together". *Computers & Chemical Engineering* **26**:2, pp. 149–160.

Sundaramoorthy, A., C. T. Maravelias, and P. Prasad (2009). "Scheduling of multistage batch processes under utility constraints". *Industrial & Engineering Chemistry Research* **48**:13, pp. 6050–6058.

Tousain, R. L. (2002). *Dynamic Optimization in Business-wide Process Control.* PhD thesis. Technical University Delft, The Netherlands.

Paper I

Plant-wide Utility Disturbance Management in the Process Industry

Anna Lindholm, Charlotta Johnsson

Abstract

Utilities, such as steam and cooling water, are often shared between production areas at large-scale sites. A disturbance in the supply of a utility is therefore likely to affect a large part of a site, and cause great loss of revenue. This study focuses on identifying disturbances in utilities and estimating the economical effects of such disturbances. A general method for reducing the loss of revenue due to utility disturbances, the utility disturbance management (UDM) method, is presented. Modeling of the effects of utility disturbances on production is needed to complete all steps of the method. In this paper, a simple on/off modeling approach is suggested to quickly obtain key performance indicators that may be used for decision support for proactive disturbance management. A matrix representation of a site and its utilities is introduced to simplify the computations. The UDM method is applied to an industrial case at Perstorp, Sweden.

1. Introduction

In the chemical process industry, companies must continuously improve their operational efficiency and profitability to remain competitive [Bakhrankova, 2010; Bansal et al., 2005]. This means it is of great importance to minimize losses in revenues due to e.g. disturbances in operation. Plant-wide disturbances cause considerable revenue losses at industrial sites [Thornhill et al., 2002; Bauer et al., 2007]. Some of these plant-wide disturbances are caused by utilities, such as steam or cooling water, which are used at most industrial large-scale production sites. At a disturbance in the supply of a utility, the production in all areas that use the utility is affected. Furthermore, areas are often connected by the product flow at the site, which makes the consequences of utility disturbances hard to predict. Two examples of utility disturbances at an industrial site are given in Figures 1 and 2. Both figures show a pressure drop in the middle-pressure steam net and the production in four areas that all require middle-pressure steam during the same time-period. The green dashed line shows the ideal steam pressure of 14 bar.

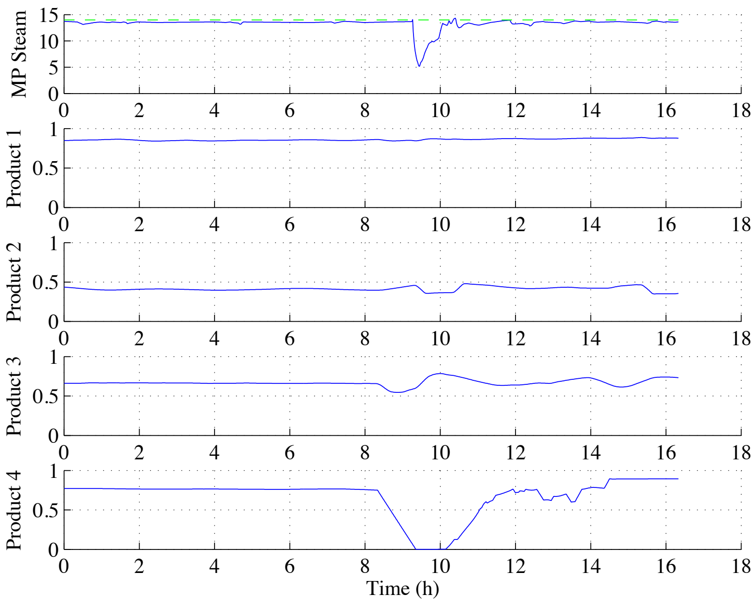

Figure 1. Example 1 of a pressure drop in the middle-pressure steam net.

In Figure 1 it can be seen that the production of product 1 is not affected by the disturbance, whereas mainly the production of product 4 has to handle the variations. The reason that the production of product 4 is reduced already before the pressure drop is detected is that the steam boiler fails before the pressure drops, and thus the operators can start to react to the disturbance before it is visible in the measurement data of the pressure. In Figure 2, the operators at the site handled the disturbance by starting to reduce the production of product 1 immediately as the pressure in the steam net drops. In this case, the production of products 3 and 4 is back to normal shortly after the disturbance, whereas the production of products 1 and 2 is reduced for a longer time. Consequently, the same type of disturbance is not handled using the same product flow control strategy each time. Furthermore, since utilities are shared by production areas, it is not evident which utility that causes the greatest economical loss. This motivates the need for a method that orders utilities according to the loss of revenue they cause, and suggests strategies for handling utility disturbances so that the loss of revenue is minimized. Economical loss is a useful measure, since it is a measure that is easily understandable by the managers at a site.

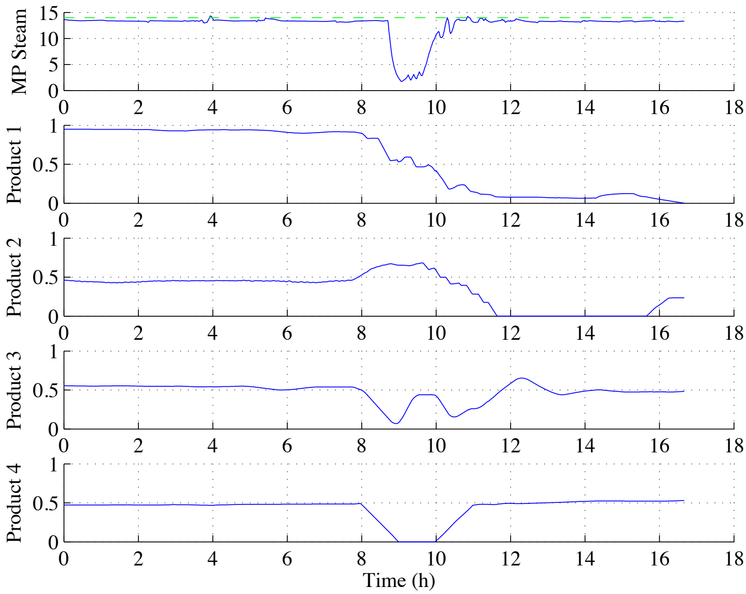

Figure 2. Example 2 of a pressure drop in the middle-pressure steam net.

2. Related Research Areas

To the best of our knowledge, managing disturbances in the supply of utilities at the site level is an unexplored topic which does not quite fit into any current research area. To show this, and to distinguish the contributions of this paper, a few related research areas are discussed below.

Disturbances in utilities are often plant-wide disturbances. Detection and diagnosis of plant-wide disturbances is discussed by, among others, [Bauer et al., 2007], [Thornhill and Horch, 2007] and [Thornhill et al., 2002]. However, these studies do not discuss how plant-wide disturbances should be handled when they occur. The area of plant-wide control has been studied by some researchers, among others in [Skogestad, 2004], [Downs and Skogestad, 2011], [Luyben et al., 1997] and [Zheng et al., 1999]. For the research on plant-wide control, the main concern is the choice of control structure. A well-known example is the Tennessee Eastman challenge problem, presented in [Downs and Vogel, 1993], which is a reactor/separator/recycle system. A related area of research is control of mixed logical dynamical (MLD) systems, studied in e.g. [Bemporad and Morari, 1999]. MLD systems may be used to model the connection of unit operations and can, in addition to the physical laws, also include logic rules and operating constraints. The applications of plant-wide control and control of MLD systems are mostly for the connection of unit operations, i.e. on the area level according to the standard [ISA-95.00.01, 2009]. The current study focuses on the connection of production areas, i.e. on disturbances in production at the site level according to [ISA-95.00.01, 2009]. However, the research areas have a key idea in common: The objective is to divert process variability away from critical locations to locations where it does less damage [Qin, 1998; Luyben et al., 1999]. If the production control problem is formulated as an MLD system, this would fit into the "Continuous production" modeling approach (see Section 8) when using the general framework presented in this paper.

Some studies related to this work have been performed at the enterprise level, mainly within the domain of production and operations management. In [Grossmann, 2005], [Varma et al., 2007] and [Grossmann and Furman, 2009], Enterprise-wide optimization (EWO) for the process industries is discussed, which includes planning, scheduling, real-time optimization and inventory control. However, the main focus of EWO is on planning and scheduling at the enterprise-level, and not on production control for disturbances at the site-level.

Regarding disturbances in utilities specifically, some studies have been done on the synthesis of utilities to satisfy the demand, e.g. by [Papoulias and Grossmann, 1983], [Maia et al., 1995], [Maia and Qassim, 1997] and [Iyer and Grossmann, 1998]. However, none of these studies treat the problem of how to control the production at utility disturbances.

In this paper, a general method for reducing the economical effects of disturbances in the supply of utilities is introduced, denoted the utility disturbance management (UDM) method. For completing all steps of the method, a model of the site is needed, that describes how utility disturbances affect production at different areas. The level of detail of the model determines the level of detail of the obtained disturbance management strategies. Detailed modeling enables development of online disturbance management strategies, but these models are often hard and time-consuming to obtain. In this paper, a simple on/off production modeling approach is used together with the UDM method to quickly obtain key performance indicators that may be used for decision support for proactive disturbance management for utility disturbances. A matrix representation of a site and its utilities is introduced that makes this possible, and an industrial case study at Perstorp is included to demonstrate how the method is applied at an industrial site.

3. Terminology

3.1 Equipment Hierarchy

The role based equipment hierarchy of an enterprise is defined in the standard [ISA-95.00.01, 2009]. For continuous production, there are five levels in the hierarchy; the enterprise, site, area, production unit, and unit levels. A plant is a concept that is not well defined with respect to the role based equipment hierarchy. In some contexts, a plant denotes a site, and in other contexts an area is intended. To avoid confusion regarding this term, the words site and area are used exclusively in this paper. This study belongs at the site level, i.e. considers the connection of production areas which produce one or more products each. Details such as single production units (e.g. reactors, distillation columns) in the areas are not considered. The objective is to find methods for minimizing the losses for an entire site due to disturbances in utilities affecting one or more areas.

At each level in the role based equipment hierarchy, a flowchart can be made showing the product flow through the enterprise, site or area. This is illustrated in Figure 3. In these flowcharts, the dependence by a site/area/production unit on other sites/areas/production units can be seen.

3.2 Disturbances

There are several different definitions of the term “disturbance” in different areas of research, and also within research areas. Also, in many studies, a definition is omitted. A definition that fits quite well for this paper is the definition of disturbances in logistics networks given in [Hinrichs et al., 2005]: “Disturbances are all events which prevent a process to deliver an economic result”. Using this definition, failures at the unit or production unit level in

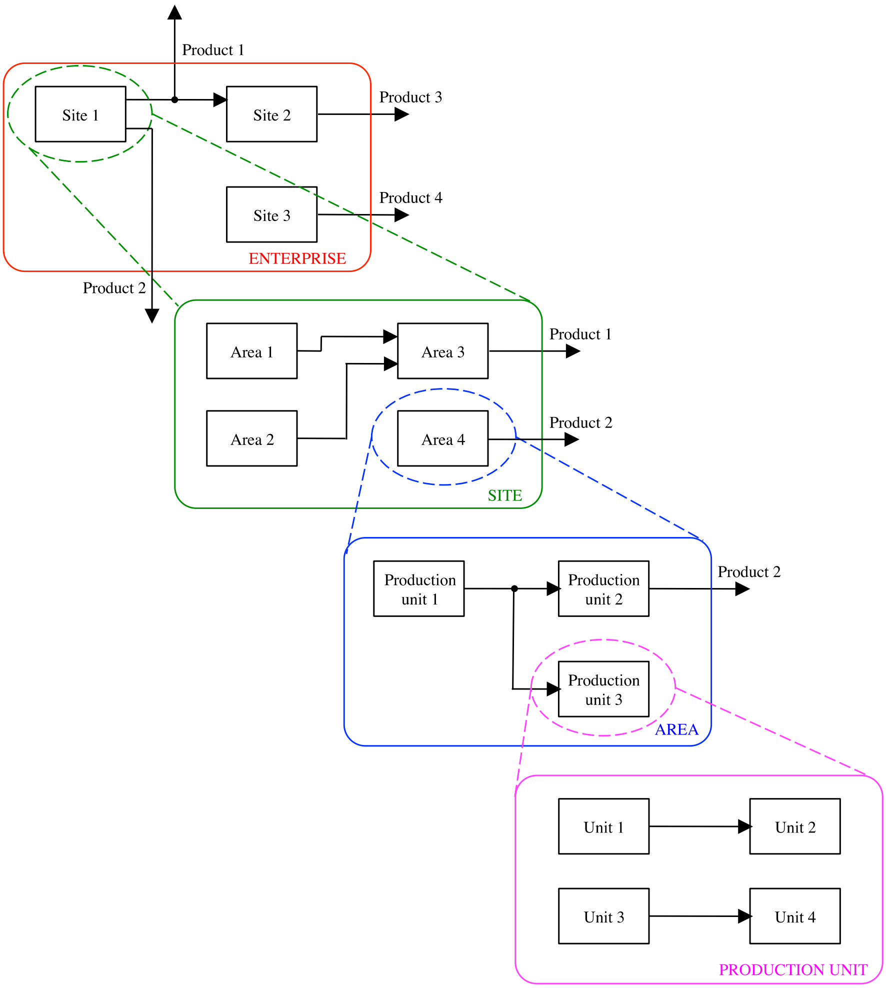

Figure 3. Flowcharts at different levels in the equipment hierarchy.

the role based equipment hierarchy will not be regarded as disturbances, if their effects do not propagate to the area level of the hierarchy. Examples of disturbances that fit into the definition are such as a pressure drop in the steam net, and variations in cooling water temperature; disturbances that affect the production in one or more areas at a site. These kinds of disturbances are the focus of this paper. Disturbances that affect entire areas or sites in an enterprise are sometimes also denoted plant-wide disturbances.

Two main approaches of disturbance management could be identified. One is disturbance management strategies that are aiming to prevent future disturbance occurrences, which is denoted *proactive disturbance management*. The other is disturbance management strategies for handling disturbances when they occur, which is denoted *reactive disturbance management*. These terms are also used in [Barroso et al., 2010], where they are defined for the supply chain, and in [Monostori et al., 1998], where they are used in the context of scheduling.

According to the standard [ISA-95.00.01, 2009] the production capability of an enterprise can divided into personnel, material, and equipment. Utilities, which are further discussed in Section 4, can also be seen as part of the material class. The cause of a disturbance may be any of these resources. What methods that can be used for disturbance management depends on the cause of the disturbance. For the material class, a proactive measure that can be taken to reduce the number of disturbances is to include more redundancy, such as investing in an extra steam boiler or cooling fan. However, these changes are often expensive since they require redesign of the system [Greenberg, 1991]. Investing in buffer tanks for raw materials or utilities may also in the same manner reduce the effects of disturbances caused by material. Reactive disturbance management strategies that may be used is to divide the resources of raw materials or utilities at a disturbance in a clever way at the site, i.e. to transfer the variability to a location where it does as little damage as possible. Buffer tanks may also be used for this purpose, since the effects of disturbances caused by material can be reduced by optimal use of the buffer volumes at the occurrence of a disturbance.

4. Utilities

4.1 Utilities in the Process Industry

Utilities are support processes that are utilized in production, but are not part of the final product. Utilities could either be specific to an area or be shared between production areas. In [Brennan, 1998], some utilities are described. Here, devices for combustion of tail gas and the vacuum system utility have been added to expand the list of common utilities in the process industry. Examples of uses of these utilities are described below:

Steam

The steam net is commonly used for heating, for example heating of a reactor at start-up, or for supplying energy, for example for distillation or endothermic reactions. There could be several steam nets at the same site, for example one net with high-pressure steam and one with low-pressure steam.

Cooling water

The cooling water system is used for example for cooling at exothermic reactions and in the condensing phase of distillation. Cooling fans, cooling a local cooling coil, are sometimes used for extra cooling in a certain area at the site.

Electricity

Electricity is needed in order for the instruments and pumps to operate. Electricity of different voltages could be required.

Fuel

Fuel, typically gas, oil or coal, may be needed for furnaces, kilns and steam boilers to operate. It may also be required for start-up of certain units or areas.

Water treatment

The water treatment utility, or effluent treatment, is used for purification of process water, precipitation and ground water.

Combustion of tail gas

A flare is a safety device used for combustion of tail gas at unforeseen events. There could also be other equipment for combustion of tail gas. These devices are often used for the combustion of tail gas during normal operation, and might be local utilities, i.e. utilities that only operate at a single area.

Nitrogen

Nitrogen is needed to maintain pressure in vessels by pushing away oxygen to prevent oxidation.

Water

Feed water, which consists of varying proportion of recovered condensed water (return water) and purified fresh water (make-up water) is needed for the boilers to be able to produce steam. Water is also required for washing and for the fire protection water system.

Compressed air

Compressed air could be both process air and instrument air. Instrument air is needed for the pneumatic instruments to work. Instrument air might be a local utility; e.g. every area has its own instrument air system.

Vacuum system

Vacuum is used to lower the boiling point of a liquid to facilitate distillation and to remove gas produced in reactions. The vacuum system might be a local utility; i.e. every area has its own vacuum system.

4.2 Utility Disturbance Limits

Utilities most often affect production only when their supply is interrupted or does not meet the specifications. For example, the cooling water utility does often not affect production until the temperature of the cooling water is over some temperature limit. One possible way of identifying when a utility suffers a disturbance is thus as when the measurement of a utility parameter, such as temperature or pressure, goes outside a limit at which the poor operation of the utility will have negative consequences for the production at the site. The consequences could be of different severity depending on how large the deviations from the limits are; a very high cooling water temperature could for example affect the production more than just a little too high temperature. The suggestion is to set the limit so that it represents the limit for when maximum production can no longer be maintained because of poor operation of the utility. An example of the operation of the cooling water utility at an industrial site is given in Figure 4. The suggested disturbance limit is marked with a dashed red line.

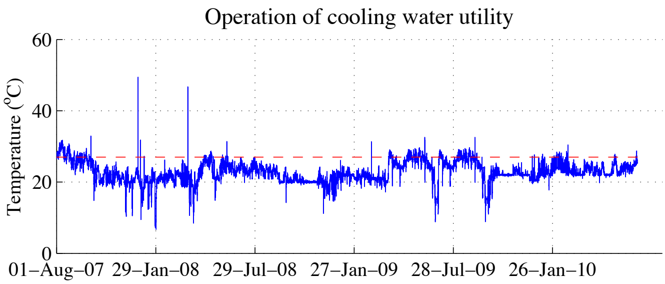

Figure 4. Cooling water temperature and disturbance limit.

5. A Matrix Representation of Utilities at a Site

To enable efficient computation of key performance indicators such as utility availability and area availability (described in Sections 6 and 7), a matrix representation of a site and its utilities is introduced in this section. A simple site with four areas and five utilities is used as an example throughout the paper, to show how the matrix representation is utilized. A flowchart of the product flow at the example site is shown in Figure 5.

The example site requires five utilities: steam (st), cooling water (cw), electricity (el), feed water (fw) and instrument air (ia). The corresponding utility parameters, that determine if the utility works correctly, are 'pressure' for steam, feed water and instrument air, and 'temperature' for cooling wa-

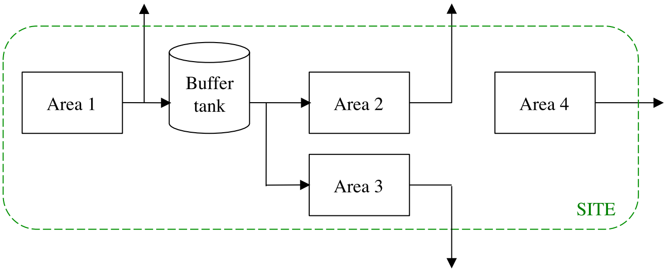

Figure 5. Product flow at the example site.

ter. Electricity can only be operating or not operating, i.e. 'on' or 'off'. The measurements of the utility parameters during 10 hours with the sampling interval 1 hour are given below

$$
\begin{array}{ll}
\mathrm{st} = & [42 \quad 38 \quad 34 \quad 32 \quad 35 \quad 41 \quad 40 \quad 36 \quad 34 \quad 37] \\
\mathrm{cw} = & [25 \quad 24 \quad 24 \quad 26 \quad 28 \quad 30 \quad 27 \quad 25 \quad 24 \quad 25] \\
\mathrm{el} = & [1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 0 \quad 1 \quad 1 \quad 1] \\
\mathrm{fw} = & [22 \quad 19 \quad 18 \quad 20 \quad 22 \quad 21 \quad 21 \quad 21 \quad 21 \quad 21] \\
\mathrm{ia} = & [1 \quad 2 \quad 1 \quad 1 \quad 3 \quad 2 \quad 1 \quad 0 \quad 0 \quad 1]
\end{array}
$$

The data set does not include any planned stops. The disturbance limits that apply for these utilities at the site are:

$$
\begin{array}{ll}
\text{Steam :} & \text{pressure} < 35 \text{ bar} \\
\text{Cooling water :} & \text{temperature} > 27^\circ\mathrm{C} \\
\text{Electricity :} & \text{on/off} \\
\text{Feed water :} & \text{pressure} < 20 \text{ bar} \\
\text{Instrument air :} & \text{pressure} \leq 0 \text{ bar}
\end{array}
$$

Table 1 shows the utilities that are required by each area in the example.

Table 1. Utilities required at each area in the example.

	Area 1	Area 2	Area 3	Area 4
Steam	x		x	
Cooling water		x	x	
Electricity	x	x	x	x
Feed water	x		x	
Instrument air	x		x	x

5.1 Site Structure

As described in Section 3, a site consists of one or more production areas. Some areas produce intermediate products that are refined to end products in other areas, so that the product of an area is the raw material to one or more other areas. This gives interdependence of areas that can be described by an *area dependence matrix*, where a one at row i and column j means that area i obtains raw material from area j. A zero means that the areas are independent. All areas are assumed to be dependent on themselves, which gives ones on the diagonal of the matrix. The area dependence matrix, A_d, will have the size $n_a \times n_a$, where n_a is the number of areas at the site. For the example site described in this section, areas 2 and 3 are dependent on raw materials from area 1, whereas area 4 is independent with respect to the product flow. This gives the area dependence matrix

$$A_d = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

5.2 Utility Measurement Data

Measurement data of utility parameters can be compared to the critical limits for each utility to form vectors where, for each sample, the value is one if the utility works properly, and zero otherwise. These vectors could be row-stacked to obtain the *utility operation matrix* U of size $n_u \times n_s$, where n_u is the number of utilities, and n_s the number of samples. The sampling interval is denoted t_s. The measurements and utility disturbance limits given for the example in this section gives the utility operation matrix

$$U = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

5.3 Utility Requirements

Every production area requires a specific set of utilities in order to operate correctly. The set of utilities each area requires can be presented in an *area–utility matrix*, where a one at row i and column j means that area i requires utility j. This matrix is denoted A_u and has n_a rows and n_u columns. For the example in this section, the area–utility matrix

$$A_u = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix}$$

is obtained using Table 1. The rows in the area–utility matrix corresponds to the columns in Table 1.

6. Utility Availability

A key performance indicator that could be used for determining how often a utility suffers a disturbance is the availability of the utility. The availability of a production unit is according to the standard [ISO/WD-22400-2, 2011] the ratio between the actual production time and the planned allocation time, where the planned allocation time is the time in which the unit can be used (the operation time) minus the planned downtime. For utilities, the suggestion is to define availability as the fraction of time all utility parameters are inside their limits. This represents the fraction of time when there is a possibility for maximum production, assuming that all utility disturbance limits have been correctly set. Utility availability can be computed if measurements of all utility parameters are available. Planned stops should not be included in the availability computations. The utility availabilities for all utilities at a site are easily obtained using the matrix representation described in Section 5. The utility availabilities are obtained as a column vector, U_{av}, by taking the row-sum of the utility matrix and dividing by the number of samples, or equivalently

$$U_{av} = U \cdot \mathbf{1}/n_s \tag{1}$$

where U is the utility operation matrix and $\mathbf{1}$ denotes a column vector of ones. For the example in Section 5, we get

$$U_{av} = \begin{bmatrix} 0.7 & 0.8 & 0.9 & 0.8 & 0.8 \end{bmatrix}^T$$

with the same ordering of the utilities as in the example.

6.1 Utility Dependence

Some utilities are dependent on other utilities, which may have as consequence that a disturbance on one utility also shows up in the measurements of other utilities. For example, if feed water is not available, steam could not be produced, and the steam utility is defined as unavailable even though there is nothing wrong with the steam boiler. This should be considered a feed water failure and not a steam failure, since feed water is the root cause of the disturbance. Utility interdependence may be visualized in a flowchart. An example of a utility dependence flowchart for electricity, cooling water, instrument air, feed water and steam is given in Figure 6.

Utility dependence may also be represented by a matrix, U_d, where a one at row i, column j, means that utility i is dependent on the operation of utility j. A zero means that the utilities operate independently of each other.

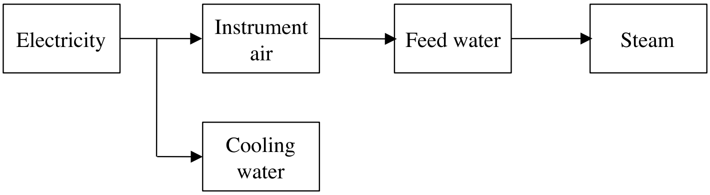

Figure 6. Utility dependence flowchart for five utilities.

All utilities are assumed to be dependent on the operation of themselves, which gives ones on the diagonal of the matrix. U_d is denoted the *utility dependence matrix.*

Utility dependence can be taken into account when calculating utility availability. Since the measurements do not reflect the root cause of a disturbance, utilities that are dependent on other utilities might appear to have lower availability than they actually have, if utility dependence is not taken into account. If utility dependence is considered, the utility operation matrix becomes

$$U_{ud} = \text{sign}\left(U + \text{sign}\left((I - U_d)(U - \mathbf{1}\mathbf{1}^T)\right)\right) \tag{2}$$

where $\mathbf{1}$ denotes a column vector of ones, and I is the identity matrix. U is the utility operation matrix without consideration to utility dependence, as defined in Section 5. The utility availabilities when considering utility dependence is given by (1) using U_{ud} as the utility operation matrix. For the utilities in the example in Section 5, the utility dependence matrix becomes

$$U_d = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix}$$

if the utilities are dependent according to Figure 6. The ordering of the utilities is the same as in Table 1. If utility dependence is considered, we get the utility operation matrix

$$U_{ud} = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

which gives the utility availabilities

$$U_{av}^{ud} = U \cdot \mathbf{1}/n_s = \begin{bmatrix} 0.9 & 0.8 & 0.9 & 0.8 & 0.8 \end{bmatrix}^T$$

7. Area Availability

A simple estimate of the availability of each area, with respect to utilities, is as the fraction of time all utilities needed at the area are available; i.e. the intersection of the operation of all concerned utilities. The measure of area availability should be interpreted as the fraction of time an area has a possibility of operating at maximum production rate, with respect to utilities. Area availability computed without considering the connection of areas at a site is denoted *direct area availability*. A column vector containing the direct area availabilities of all areas at a site, A_{av}^{dir}, is obtained by

$$A_{av}^{dir} = A_{dir} \cdot \mathbf{1}/n_s \tag{3}$$

where

$$A_{dir} = \mathbf{1}\mathbf{1}^T + \text{sign}\left(A_u(U - \mathbf{1}\mathbf{1}^T)\right) \tag{4}$$

A_u is the area–utility matrix and U is the utility operation matrix, as defined in Section 5. A_{dir} is denoted the *direct area operation matrix*. Note that the utility operation matrix without consideration to utility dependence should be used, since the measure of area availability should describe how the areas actually have operated during the time-period. For the example in Section 5 we get

$$A_{av}^{dir} = \begin{bmatrix} 0.4 & 0.7 & 0.2 & 0.7 \end{bmatrix}^T$$

7.1 Area Dependence

If an area is unavailable, this may also affect other areas at the site, since areas could be dependent on obtaining raw materials from, or delivering product to, other areas. The area interdependence can be seen in a flowchart of the product flow at the site. To include area dependence in the measure of area availability, *total area availability* is introduced. Total area availability is defined as the fraction of time all the required utilities and all areas that the area is dependent on are available. Thus, total area availability contains both the direct effects of a utility disturbance, and the indirect effects because of area interdependence. The total area availabilities of all areas are given by the column vector A_{av}^{tot}, which is computed as

$$A_{av}^{tot} = A_{tot} \cdot \mathbf{1}/n_s \tag{5}$$

where

$$A_{tot} = \mathbf{1}\mathbf{1}^T + \text{sign}\left(A_d(A_{dir} - \mathbf{1}\mathbf{1}^T)\right) \tag{6}$$

A_{dir} is the direct area operation matrix and A_d is the area dependence matrix, as defined in Section 5. A_{tot} is denoted the *total area operation matrix*. For the example in Section 5 we get

$$A_{av}^{tot} = \begin{bmatrix} 0.4 & 0.2 & 0.2 & 0.7 \end{bmatrix}^T$$

8. Production Modeling Approaches

Utilities are often shared between the production areas at a site. Three approaches for modeling the production at a site with respect to utilities are suggested in this paper, here listed according to level of detail of the obtained model. The model should describe how the production in all areas is affected at utility disturbances.

1. **On/off production without buffer tanks**
   Utilities and areas are considered to be either operating or not operating, i.e. 'on' or 'off'. An area operates at maximum production speed when all its required utilities are available, and does not operate when any of its required utilities are unavailable. It is assumed that there are no buffer tanks between the areas at the site. This means that if an area is unavailable, downstream areas of that area will also be unavailable.

2. **On/off production including buffer tanks**
   The same modeling approach as approach 1, but buffer tanks between areas are included in the model. The buffer tanks act as delays from when an area upstream/downstream of the tank stops producing until its downstream/upstream areas have to be shut down.

3. **Continuous production**
   Utility operation and production are considered to be continuous. Areas can operate at any production rate below the maximum limit determined by the operation of utilities.

9. Estimation of Revenue Loss due to Utility Disturbances

The on/off production modeling approach without buffer tanks, described in Section 8, enables estimation of revenue losses due to disturbances in utilities using the matrix representation introduced in Section 5. The on/off model does not adequately reflect how the actual production is managed for most sites, but is because of its simplicity useful for obtaining quick estimates of production losses due to disturbances. The estimates may be used as decision support for proactive disturbance management.

9.1 Estimation of Revenue Loss for each Product

With on/off production modeling without buffer tanks, areas are assumed to be operating at maximum production speed when available, and not at all when not available. Thus, the loss of revenue in an area can be estimated as the area availability times the profit the area would have raised if it would have operated at full speed. The profit could be estimated as the flow to the market at maximum production times the contribution margin for the product and the length of the time period, i.e. the number of samples, n_s times the sampling interval, t_s. The estimates of the direct and total revenue losses, J_p^{dir} and J_p^{tot}, are obtained as column vectors:

$$J_p^{dir} = \left(\mathbf{1} - A_{av}^{dir}\right) \circ q^m \circ p n_s t_s \tag{7}$$

$$J_p^{tot} = \left(\mathbf{1} - A_{av}^{tot}\right) \circ q^m \circ p n_s t_s \tag{8}$$

where q^m is a column vector containing the flows to the market of all products at maximum production, p a column vector containing the contribution margins of all products, and $\circ$ denotes the Hadamard product. The area availabilities are obtained from (3) and (5).

9.2 Estimation of Revenue Loss due to each Utility

For on/off production modeling without buffer tanks, the utility availabilities may be used to estimate the direct and total revenue loss caused by each utility. The loss can be estimated as the sum of the revenue loss the utility causes in each area where the utility is required. Thus, information on the connection of areas and which utilities that are required at each area is needed to estimate the revenue loss due to each utility. The direct and total revenue loss due to each utility (column vectors J_u^{dir} and J_u^{tot}) become

$$J_u^{dir} = \operatorname{diag}\left[\mathbf{1} - U_{av}^{ud}\right] \cdot (A_u)^T (q^m \circ p) n_s t_s \tag{9}$$

$$J_u^{tot} = \operatorname{diag}\left[\mathbf{1} - U_{av}^{ud}\right] \cdot \operatorname{sign}\left(A_d A_u\right)^T (q^m \circ p) n_s t_s \tag{10}$$

with the same notation as in Section 9.1. A_u is the utility operation matrix and A_d the area dependence matrix, as defined in Section 5, and the utility availabilities are obtained from (1) with utility dependence taken into account according to (2).

10. Utility Disturbance Management Method

The general method for reducing the revenue loss due to disturbances in utilities is denoted the utility disturbance management (UDM) method. The strategies for reducing the loss may be both proactive and reactive disturbance management strategies, depending on which production modeling approach that is selected. Some suggestions of modeling approaches are given in Section 8. The accuracy of the strategies for reducing the revenue loss depends on the level of detail of the model of the site. A less detailed model will only give simple strategies for reducing the revenue loss, whereas a more detailed model is required to obtain advanced disturbance management strategies. The choice of modeling approach depends on how detailed strategies that are required, and how much time and effort that is available for modeling.

10.1 The UDM Method – Step by Step

The UDM method is divided into four steps. The first three steps concern acquiring the necessary information about the site and estimating the revenue loss that is caused by each utility during a certain time period, and the last reducing the revenue loss due to future disturbances in utilities. Only the last two steps of the method are model specific, i.e. depend on which production modeling approach that is used. The four steps, with sub-steps, are listed below.

Step 1: get information on site-structure and utilities

a) Depict the overall structure of the site

b) List all utilities used at the site

c) Determine which utilities that are required at each area

d) Draw a utility dependence flowchart

e) Define disturbance limits for each utility

f) Get relevant measurement data

g) List all planned stops during the time period

Step 2: compute utility and area availabilities

a) Compute utility availabilities

b) Compute direct and total area availabilities

Step 3: estimate revenue loss due to disturbances in utilities

a) Select site model

b) Estimate flow to the market of each product

c) Get contribution margins for each product

d) Estimate revenue loss for each product

e) Estimate revenue loss due to each utility

Step 4: reduce revenue loss due to future disturbances in utilities

The four steps of the UDM method are described in the following four subsections. Step 3 d), 3 e), and 4 are dependent on the choice of site model and could not described in detail for the general case. If on/off production modeling without buffer tanks is applied, these steps are easily completed using the matrix representation introduced in Section 5. Step 3 d) and 3 e) with on/off production modeling are handled in Section 9.

Step 1: get information on site-structure and utilities

a) *Depict the overall structure of the site*
Determine the role based equipment hierarchy (see Section 3). Draw the overall structure of the site by highlighting its production areas and the physical connections between them. The structure could be represented by a flowchart on the site level. The interdependence of areas at a site may also be represented by the area dependence matrix, as described in Section 5.

b) *List all utilities used at the site*
List the utilities that are required for production at the site. Common utilities within the process industry are listed in Section 4.

c) *Determine which utilities that are required at each area*
Determine which utilities that are required at each area. This can be summarized in a table or in the area–utility matrix, as described in Section 5.

d) *Draw a utility dependence flowchart*
Some utilities might be dependent on the operation of other utilities. Determine the hierarchy of the utilities and draw a utility dependence flowchart. An example is given in Figure 6 in Section 6. Utility dependence can also be represented in the utility dependence matrix, as described in Section 5.

e) *Define disturbance limits for each utility*
Determine the utility parameters that describe the operation of each utility. Determine the critical limits for each utility parameter, i.e. determine the limits for when disturbances in the utility parameters have negative impact on production. This is described more thoroughly in Section 4.

f) *Get relevant measurement data*
Decide which time period to consider. Get measurement data for all utility parameters for all utilities for this time period.

g) *List all planned stops during the time period*
Find the time and duration of all planned stops at the site during the considered time period. Data from these periods should not be included in the availability computations.

Step 2: compute utility and area availabilities

a) *Compute utility availabilities*
Compute the utility availabilities, i.e. the fraction of time each utility operates correctly, using the disturbance limits (step 1 e)). Take utility dependence (step 1 d)) and planned stops (step 1 g)) into account. Computation of utility availability using the matrix representation, and further information about utility availability is provided in Section 6.

b) *Compute direct and total area availabilities*
Compute the area availability for each area, i.e. the fraction of time each area has access to all required utilities, and thereby has the possibility of operating at its maximum production rate. Do the computations both without and with consideration to area interdependence (direct and total availability). The measure of area availability is introduced in Section 7, where also computation of area availability using the matrix representation is handled.

Step 3: estimate revenue loss due to disturbances in utilities

a) *Select site model*
For completing step 3 of the method, a production model of the site is required. Three approaches for modeling the production at a site are suggested in Section 8. The model of the site has to describe how the production of each area is affected when a utility is unavailable.

b) *Estimate flow to the market of each product*
Estimate the flow to the market of each product during the selected time-period. The flows may vary over the time-period.

c) *Get contribution margins for each product*
Get the contribution margins, i.e. the marginal profit per volume, for each product. This information can often be acquired from sales personnel at the site. The margins may vary over time.

d) *Estimate revenue loss for each product*
Compute the direct and total revenue loss for each product due to all disturbances in all utilities during the selected time-period, given the modeling approach selected (step 3 a)). Production revenue loss is obtained when an area cannot operate at the desired production rate. For constant contribution margins and flows to the market, the revenue loss is given by the contribution margin (step 3 c)) multiplied by the duration when production was reduced, and by the difference of the desired production rate and the actual production rate during the time

period (step 3 b)). The direct revenue loss is the loss of revenue due to poor operation of the utilities in that area. The total revenue loss is obtained if also the effects of utility disturbances in areas that the area is dependent on are considered.

e) *Estimate revenue loss due to each utility*
Compute the direct and total revenue loss caused by disturbances in each utility separately, given the modeling approach selected (step 3 a)). Use the estimation of the flows to the market (step 3 b)) and the contribution margins (step 3 c)). The direct revenue loss is the loss each utility causes directly, because of reduced production in the areas that require the utility. The total revenue loss also includes the revenue loss due to reduced production in areas that are dependent on the areas that require the utility.

Step 4: reduce revenue loss due to future disturbances in utilities

This step is dependent on which modeling approach for modeling the site that was chosen (step 3 a)). In general, the step consists of finding proactive and/or reactive disturbance management strategies (see Section 3) for reducing the revenue loss due to future disturbances in utilities, based on the information acquired in step 3 d) and 3 e).

11. Case study: Application of the UDM Method

The case study is performed at Perstorp, at their site in Stenungsund, Sweden. Perstorp is a worldwide enterprise that is a world leader in several sectors of the specialty chemicals market. Their products can be found in for example automotive, food, packaging, and electronics applications. Site Stenungsund is located on the Swedish west coast, approximately 50 km north of Gothenburg. The main products of the site are aldehydes, organic acids, alcohols, and plasticizers [Perstorp, 2012].

11.1 Step 1: Get Information on Site-structure and Utilities

a) *Depict the overall structure of the site*
Site Stenungsund is one of 13 sites owned by the enterprise Perstorp. The site consists of 10 production areas. The products of the 10 areas at the site are here denoted product 1-10 for area 1-10 respectively. Internal buffer tanks exist for products 1-5. Their location and the interdependence of areas can be seen in the flowchart of the product flow at the site in Figure 7.

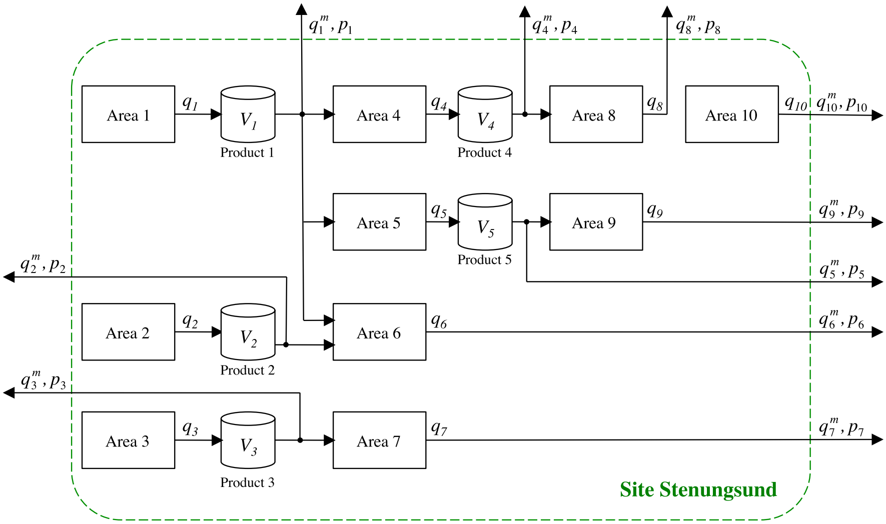

Figure 7. Flowchart of the product flow at site Stenungsund.

The area dependence matrix becomes

$$A_d = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{11}$$

b) *List all utilities used at the site*

Site Stenungsund has all utilities listed in Section 4.

Steam There are two steam nets at the site, one with high-pressure steam, ideally 41 bar, and one with middle-pressure steam, ideally 14 bar.

Cooling water There is one cooling water network that supplies all areas with cooling water. Areas 1, 2, 3 and 7 have cooling fans in addition to ordinary cooling water.

Electricity Site Stenungsund uses both 130 kV and 40 kV electricity. A list of disturbances in electricity is provided by the supplier.

Fuel Fuel is not included as a utility in this case study. However, the effects of fuel being unavailable will show up in measurements of other utilities, for example steam.

Water treatment For the water treatment utility (WTU) there are monthly limits for how large amounts of suspended material (SUSP) and dissolved organic carbon (DOC) that is allowed in the outgoing water, that should not be exceeded. There are also more strict limits for how much of these substances that are allowed in the outgoing water each year. If the yearly limits are exceeded, the site has to be shut down immediately. The monthly and yearly limits are individual for each production site. In this case study, only the yearly limits are considered.

Combustion of tail gas The site contains a flare and there are also three areas, areas 7, 8 and 9, that have devices for local combustion of tail gas at normal operation. However, measurements are only available for the combustion devices at area 7 and area 9, why the combustion device at area 8 will not be considered in this case study.

Nitrogen Nitrogen is used to maintain pressure in vessels.

Water Both feed water, washing water and fire protection water are used at the site, but only feed water will be considered in the case study.

Compressed air Compressed air could be both process air and instrument air. At site Stenungsund, only instrument air is used.

Vacuum system The vacuum systems are individual for each area.

c) *Determine which utilities that are required at each area*
A table showing which utilities that are needed at each area is presented in Table 2. Some utilities have been divided into sub-utilities to give a more specific view of what causes the greatest revenue losses; the steam utility has been divided into high-pressure (HP) steam and middle-pressure (MP) steam, the cooling water utility into cooling water and cooling fans and combustion of tail gas into flare and devices for combustion of tail gas at normal operation (here denoted 'combustion devices').

Table 2. Utilities required at areas at site Stenungsund.

	1	2	3	4	5	6	7	8	9	10
HP steam							x	x	x	x
MP steam	x	x	x	x	x	x	x		x	
Cooling water	x	x	x	x	x	x	x	x	x	x
Cooling fan 1	x									
Cooling fan 2		x								
Cooling fan 3			x							
Cooling fan 7							x			
Electricity	x	x	x	x	x	x	x	x	x	x
Water treatment	x	x	x	x	x	x		x	x	
Flare	x	x	x	x	x	x				x
Combustion device 7							x			
Combustion device 9									x	
Nitrogen	x	x	x	x	x	x	x	x	x	x
Feed water	x	x	x	x	x			x		
Instrument air	x	x	x	x	x	x	x	x	x	x
Vacuum system	x	x	x	x	x	x	x	x	x	x

The area–utility matrix becomes

$$A_u = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \end{bmatrix} \tag{12}$$

where utilities are ordered: HP steam, MP steam, cooling water, cooling fan area 1, cooling fan area 2, cooling fan area 3, cooling fan area 7, electricity, water treatment, combustion device area 7, combustion device area 9, nitrogen, feed water, instrument air.

The ten vacuum systems and the flare have been left out in the A_u matrix to reduce the size of the matrix and make the problem more transparent. Both the vacuum systems and the flare have been available 100% of the time during the considered time-period.

d) *Draw a utility dependence flowchart*

The interdependence of utilities at site Stenungsund is shown in Figure 8.

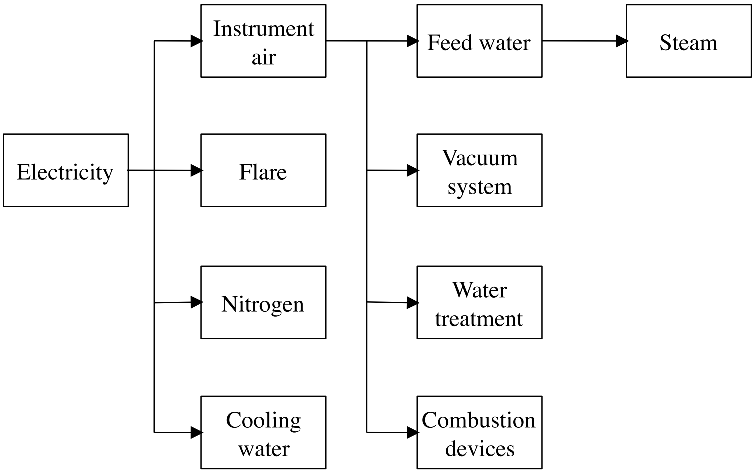

Figure 8. Utility dependence flowchart for site Stenungsund.

This gives the utility dependence matrix

$$U_d = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{13}$$

e) *Define the disturbance limits for each utility*
The disturbance limits for disturbances in utilities at site Stenungsund are listed below.

Steam

- Pressure in high-pressure steam net below 33 bar
- Pressure in high-pressure steam net over 45 bar
- Pressure in middle-pressure steam net below 12 bar

Cooling water

- Cooling water temperature higher than 27°C
- Temperature of water cooled by cooling fans in areas 1, 2 or 3 higher than 70°C
- Temperature of water cooled by cooling fan in area 7 higher than 65°C
- Loss of cooling water flow

Electricity

- Voltage below 99% of normal voltage for 40 kV electricity
- Voltage below 99% of normal voltage for 130 kV electricity
- Loss of low voltage electricity
- Loss of electricity

Water treatment

- Amount of SUSP in outgoing water more than 4000 kg a year
- Amount of DOC in outgoing water more than 4000 kg a year

Combustion of tail gas

- Flare flame goes out.
- Failure of combustion device at area 7
- Failure of combustion device at area 9

Nitrogen

- Pressure in main nitrogen pipe less than 21 bar

Water

- Pressure in main feed water pipe less than 20 bar

Compressed air

- Zero pressure of instrument air

Vacuum system

- Loss of vacuum system in any area

f) *Get relevant measurement data*
The time period from August 1, 2007 to July 1, 2010 is considered. Data for all utility parameters that are considered is available for this time period, with a sampling interval of 1 minute. Measurement data is compared to the disturbance limits (step 1 e)) to produce the utility operation matrix, U. The size of the utility operation matrix becomes $14 \times 1\ 535\ 040$.

g) *List all planned stops during the time period*
There has been one planned stop during the time period, from September 15 to October 8, 2009. Data from this time period is removed from the utility operation matrix, U. The utility operation matrix is now of size $14 \times 1\ 501\ 921$.

11.2 Step 2: Compute Utility and Area Availabilities

a) *Compute utility availabilities*
Since all needed measurements are available, availabilities for all utilities can be computed according to (1). Utility dependence according to (13) is taken into account. The results presented in Table 3 are obtained for the selected time-period.

b) *Compute direct and total area availabilities*
The direct and total area availabilities are computed according to (3) and (5) respectively, using (11), (12), and the utility operation matrix. The resulting area availabilities are listed in Table 4.

11.3 Step 3: Estimate Revenue Loss due to Disturbances in Utilities

a) *Select site model*
The on/off production modeling approach without buffer tanks is used for modeling the site. Only downstream effects of disturbances are considered, since the site to a large extent is not market limited. The products of upstream areas may therefore be sold on the market when downstream areas are unavailable.

Table 3. Utility availabilities at site Stenungsund.

Utility	**Availability** (%)
Flare	100
Vacuum systems	100
Water treatment	100
Instrument air	100
Cooling fan area 7	100
Nitrogen	100
Electricity	99
Feed water	99
HP steam	99
Cooling fan area 1	97
Cooling fan area 2	97
Cooling fan area 3	97
MP steam	97
Combustion area 9	96
Combustion area 7	94
Cooling water	92

Table 4. Availabilities of areas at site Stenungsund.

Area	**Direct availability** (%)	**Total availability** (%)
1	84	84
2	84	84
3	84	84
4	87	84
5	87	84
6	87	84
7	82	80
8	89	84
9	84	81
10	90	90

b) *Estimate flow to the market of each product*

The maximum production rates of each product at the site are available, but not the corresponding inflows to all areas. The inflows are estimated from the maximum production in the areas via a conversion factor, denoted y_{ij} for the conversion between product i and j. The conversion factors have been obtained from personnel at the site. An estimation of the flows to the market becomes

$$q_1^m = \max(0, q_1 - q_4 y_{14} - q_5 y_{15} - q_6 y_{16}) \tag{14}$$
$$q_2^m = \max(0, q_2 - q_6 y_{26}) \tag{15}$$
$$q_3^m = \max(0, q_3 - q_7 y_{37}) \tag{16}$$
$$q_4^m = \max(0, q_4 - q_8 y_{48}) \tag{17}$$
$$q_5^m = \max(0, q_5 - q_9 y_{59}) \tag{18}$$
$$q_i^m = q_i, \quad i = 6, 7, 8, 9, 10 \tag{19}$$

where q_i is the maximum production rate of area i in the unit volume/time. The flows to market are stored in the array

$$q^m = \begin{bmatrix} q_1^m & q_2^m & q_3^m & q_4^m & q_5^m & q_6^m & q_7^m & q_8^m & q_9^m & q_{10}^m \end{bmatrix}^T$$

c) *Get contribution margins for each product*

Contribution margins for all products at the site are available and are stored in the array

$$p = \begin{bmatrix} p_1 & p_2 & p_3 & p_4 & p_5 & p_6 & p_7 & p_8 & p_9 & p_{10} \end{bmatrix}^T$$

where p_i is the contribution margin of area i in unit profit/volume.

d) *Estimate revenue loss for each product*

The direct and total revenue loss corresponding to each product is estimated using (7) and (8) respectively. The calculations are performed in milliseconds in Matlab on a desktop computer. In Table 5, the products are ordered according to the direct and total loss of revenue they stand for, in descending order.

e) *Estimate revenue loss due to each utility*

The direct and total revenue loss that is caused by each utility is computed using (9) and (10) respectively. The calculations are performed in milliseconds in Matlab on a desktop computer. In Table 6, the utilities are ordered according to the direct and total revenue loss they cause, in descending order.

Table 5. Products ordered according to the revenue loss they cause.

Direct loss	Total loss
Product 9	Product 9
Product 1	Product 1
Product 6	Product 6
Product 7	Product 7
Product 8	Product 8
Product 10	Product 4
Product 4	Product 10
Product 3	Product 3
Product 2	Product 2
Product 5	Product 5

Table 6. Utilities ordered according to the revenue loss they cause.

Direct loss	Total loss
Cooling water	Cooling water
MP steam	MP steam
Combustion device 9	Cooling fan 1
Combustion device 7	Feed water
Cooling fan 1	Combustion device 9
Electricity	Combustion device 7
HP steam	Electricity
Feed water	HP steam
Nitrogen	Cooling fan 2
Cooling fan 3	Cooling fan 3
Cooling fan 2	Nitrogen
Instrument air	Instrument air
Cooling fan 7	Cooling fan 7
Flare	Flare
Vacuum system	Vacuum system
Water treatment	Water treatment

11.4 Step 4: Reduce Revenue Loss due to Future Disturbances in Utilities

From step 3 e), it can be concluded that the cooling water utility seems to have caused the greatest revenue loss over the selected time period. Thus, if proactive disturbance management should be applied, it should be most profitable to try to improve the cooling water availability. Even a small increase of the availability might give a large reduction of the total revenue loss that this utility causes at the site. For example, to include redundancy in the cooling water system or to use cooling fans for additional cooling in critical areas should reduce the revenue loss due to disturbances in utilities considerably. However, the cost of improvements for different utilities may vary and might also have to be considered.

With on/off production modeling, we get no reactive disturbance management strategies that tell us how the production should be controlled at the occurrence of a utility disturbance. If this is desirable, on/off production including buffer tanks or continuous production modeling has to be used.

An interesting result that can be seen in step 3 d) in the case study is that the area with the lowest direct and total area availability is area 7, whereas the product that stands for the greatest direct and total revenue loss is product 9, produced in area 9. Thus it can be seen that it cannot be concluded directly from the area availabilities which product that stands for the greatest revenue loss. This depends on the flows to the market of each product as well as the contribution margins. For the total revenue loss, also the indirect effects of disturbances are important. In the same manner, the utility with the lowest availability is not necessarily the utility that causes the greatest revenue loss.

12. Discussions

12.1 Simplicity vs. Accuracy

The level of detail of the model of the site determines the level of detail of the strategies for minimizing the revenue loss that are given by the UDM method. When choosing modeling approach for the site, there is a trade-off between simplicity and accuracy of the results. This makes different modeling approaches suitable for different situations. For cases when modeling effort and time is limited, the on/off approach without buffer tanks might be a good choice, whereas if a detailed reactive strategy for how to handle disturbances in utilities should be developed, continuous production modeling might be required. The suggestion in [Morris, 1967] is to start with a simple model, and step by step work towards more elaborate models. In this case, this means starting with the simple on/off approach to get a broad view of which utilities that causes the greatest revenue losses at the site, and then work successively towards the more elaborate modeling approaches to be able to develop better strategies for utility disturbance management.

12.2 Recomputation of Revenue Losses

One advantage with the UDM method is that the estimate of the revenue loss due to disturbances in utilities can be computed over any period of time. After having used the method one time, the needed information about the site is acquired, and thus the method can be applied for other time periods at a minimal cost. It might be useful to recompute the results after a time period where adjustments have been made, for example in terms of maintenance or acquisition of additional equipment, to see if the adjustments had the desired effect on the revenue losses.

13. Conclusions

A general method for reducing the revenue loss due to disturbances in utilities was presented, denoted the utility disturbance management (UDM) method. In the method, both direct effects on areas due to disturbances in utilities, and indirect effects because of the connections of production areas at the site was investigated. The UDM method is easy to apply to any site by following the step-by-step instructions. A model of the site with respect to utilities is required to complete all steps of the method. Some modeling approaches of different level of detail that may be used were described. A framework for representing a site and its utilities by matrices was also presented. This matrix representation simplifies the computations associated with the UDM method, and enables quick estimation of revenue losses caused by utilities via the on/off modeling approach. The UDM method was applied to an industrial site at Perstorp. In this case study, on/off production modeling without buffer tanks was used. Thanks to the UDM method, Perstorp now has the knowledge of which utilities that cause the greatest revenue losses at site Stenungsund. Efforts can now be focused on reducing the effects of disturbances in these utilities.

Acknowledgments

The research is performed within the framework of the Process Industrial Centre at Lund University (PIC-LU) supported by the Swedish Foundation for Strategic Research (SSF).

References

Bakhrankova, K. (2010). "Decision support system for continuous production". *Industrial Management & Data Systems* **110**:4, pp. 591–610.

Bansal, M., A. Adhitya, R. Srinivasan, and I. A. Karimi (2005). "An online decision support framework for managing abnormal supply chain events". In: *Proceedings of the European Symposium on Computer Aided Process Engineering-15*. Barcelona, Spain, pp. 985–990.

Barroso, A. P., V. H. Machado, A. R. Barros, and V. C. Machado (2010). "Toward a resilient supply chain with supply disturbances". In: *Proceedings of the IEEE International Conference on Industrial Engineering and Engineering Management*. Macau, China, pp. 245–249.

Bauer, M., J. W. Cox, M. H. Caveness, J. J. Downs, and N. F. Thornhill (2007). "Nearest neighbors methods for root cause analysis of plantwide disturbances". *Industrial & Engineering Chemistry Research* **46**, pp. 5977–5984.

Bemporad, A. and M. Morari (1999). "Control of systems integrating logic, dynamics, and constraints". *Automatica* **35**:3, pp. 407–427.

Brennan, D. (1998). *Process Industry Economics: An International Perspective*. Institution of Chemical Engineers, Instrument Society of America.

Downs, J. J. and S. Skogestad (2011). "An industrial and academic perspective on plantwide control". *Annual Reviews in Control* **35**:1, pp. 99–110.

Downs, J. J. and E. F. Vogel (1993). "A plant-wide industrial process control problem". *Computers & Chemical Engineering* **17**:3, pp. 245–255.

Greenberg, H. R. (1991). *Risk Assessment and Risk Management for the Chemical Process Industry*. John Wiley & Sons, Inc., USA.

Grossmann, I. E. (2005). "Enterprise-wide optimization: a new frontier in process systems engineering". *AIChE Journal* **51**:7, pp. 1846–1857.

Grossmann, I. E. and K. C. Furman (2009). "Challenges in enterprise-wide optimization for the process industries". In: *Optimization and Logistics Challenges in the Enterprise*. Vol. 30. Springer Optimization and Its Applications. Springer, USA, pp. 3–59.

Hinrichs, J., J. Rittscher, F. Laakmann, and B. Hellingrath (2005). "Disturbance management in large logistics networks". In: *Proceedings of the International Conference on Industrial Logistics*. Montevideo, Uruguay, pp. 141–152.

ISA-95.00.01 (2009). *Enterprise-Control System Integration, Part 1: Models and Terminology*. Standard. Instrument Society of America, pp. 17–28.

ISO/WD-22400-2 (2011). *Manufacturing operations management - Key performance indicators - Part 2: Definitions and descriptions of KPIs*. Standard. (draft). International Organization for Standardization.

Iyer, R. R. and I. E. Grossmann (1998). “Synthesis and operational planning of utility systems for multiperiod operation”. *Computers & Chemical Engineering* **22**:7, pp. 979–993.

Luyben, M. L., B. D. Tyreus, and W. L. Luyben (1997). “Plantwide control design procedure”. *AIChE Journal* **43**:12, pp. 3161–3174.

Luyben, W. L., B. D. Tyreus, and M. L. Luyben (1999). *Plantwide process control*. McGraw-Hill, USA.

Maia, L. O. A. and R. Y. Qassim (1997). “Synthesis of utility systems with variable demands using simulated annealing”. *Computers & Chemical Engineering* **21**:9, pp. 947–950.

Maia, L. O. A., L. A. V. de Carvalho, and R. Y. Qassim (1995). “Synthesis of utility systems by simulated annealing”. *Computers & Chemical Engineering* **19**:4, pp. 481–488.

Monostori, L., E. Szelke, and B. Kádár (1998). “Management of changes and disturbances in manufacturing systems”. *Annual Reviews in Control* **22**, pp. 85–97.

Morris, W. T. (1967). “On the art of modeling”. *Management Science* **13**:12, pp. 707–717.

Papoulias, S. A. and I. E. Grossmann (1983). “A structural optimization approach in process synthesis – I: utility systems”. *Computers & Chemical Engineering* **7**:6, pp. 695–706.

Perstorp (2012). [online] `http://www.perstorp.com/`, January 2012.

Qin, S. (1998). “Control performance monitoring – A review and assessment”. *Computers & Chemical Engineering* **23**, pp. 173–186.

Skogestad, S. (2004). “Control structure design for complete chemical plants”. *Computers & Chemical Engineering* **28**, pp. 219–234.

Thornhill, N. F. and A. Horch (2007). “Advances and new directions in plant-wide disturbance detection and diagnosis”. *Control Engineering Practice* **15**:10, pp. 1196–1206.

Thornhill, N. F., C. Xia, J. Howell, J. Cox, and M. Paulonis (2002). “Analysis of plant-wide disturbances through data-driven techniques and process understanding”. In: *Proceedings of the 15th IFAC World Congress*. Barcelona, Spain.

Varma, V. A., G. V. Reklaitis, G. E. Blau, and J. F. Pekny (2007). “Enterprise-wide modeling & optimization – An overview of emerging research challenges and opportunities”. *Computers & Chemical Engineering* **31**:5-6, pp. 692–711.

Zheng, A., R. V. Mahajanam, and J. M. Douglas (1999). “Hierarchical procedure for plantwide control system synthesis”. *AIChE Journal* **45**:6, pp. 1255–1265.

Paper II

Minimization of Economical Losses due to Utility Disturbances in the Process Industry

Anna Lindholm, Pontus Giselsson

Abstract

A process industrial site may consist of several production areas, some producing intermediate products for further refinement in other areas, and some producing end products. The areas may share the same utilities, such as steam and cooling water, which means that the areas could be connected both by the flow of products through the site and by the use of the same utilities. Management of utility disturbances thus becomes an interesting topic. In this paper, a simple approach for modeling utilities is suggested and used to formulate a mixed-integer quadratic program (MIQP) that aims at minimizing the total economic loss at the site, due to utility disturbances. The optimization problem is reformulated as an ordinary quadratic program (QP), where auxiliary variables are utilized to avoid the use of integer variables. For suitable choices of the optimization weights, the solutions to the MIQP and the QP are in many cases equal. Two examples are given, where one is a small example inspired by a real site at the specialty chemicals company Perstorp, and the second is a larger problem that aims to show the advantage of the QP formulation when the number of areas, and thus the number of integer variables, becomes large.

1. Introduction

At chemical production sites, utilities such as steam and cooling water are often shared between production areas. The production areas could also be connected, since intermediate products from some areas might be refined to end products in other areas at the site. These kinds of sites, which may be seen as a network of integrated areas, may be enormously complex [Wassick, 2009]. Most studies on plant-wide control, among others [Downs and Vogel, 1993] and [Zheng et al., 1999], focus on individual production units, such as reactors and distillation columns, and the connections of these. In this paper, the dynamics of the production units within an area are ignored, and only the inputs and outputs of the interconnected areas are considered. Modeling of a site at a similar abstraction level is performed in [Wassick, 2009] and [Hackman and Leachman, 1989]. However, in these papers the focus is on production planning and scheduling on a longer time horizon, and not on real-time disturbance management as in the current study.

Some utility disturbances may result in that the utility operates at a lower rate. One example is a cooling water disturbance, where the temperature becomes a few degrees too high. Another is a disturbance in the steam net, where the steam pressure becomes reduced. The decision on how much of the available cooling or heating effect that should be supplied to each area may be formulated as an optimization problem, which most often does not have a trivial solution for an integrated site. The objective of the optimization is to find the product control strategy that minimizes the total economic loss at a utility disturbance. Buffer tanks or inventories at the site should also be used appropriately to achieve this. Synthesis of utilities to satisfy a given demand, as studied by among others [Papoulias and Grossmann, 1983], [Hui and Natori, 1996] and [Iyer and Grossmann, 1998], is not handled in this paper. Here it is assumed that the operation of utilities is given, and the objective is to control the production optimally for this operation of utilities. Furthermore, the focus of the previously mentioned studies is mainly on utilities for heat and power production. The present study enables consideration of other utilities, such as nitrogen and instrument air, as well. In order to optimize the supply of utilities to each area, the effects of utility disturbances on production has to be modeled. Modeling of utilities may be very complex, but if only the utilities' effect on production is relevant, a detailed model might not be needed. In this paper, a simple modeling approach is suggested based on the assumption that utility resources may be interpreted as a volume, or power, which is shared by all production areas that require them. A linear relation between assigned utility volume and maximum possible production in an area is assumed. This representation is used when formulating the optimization problem for minimization of losses at utility disturbances. The formulation of the optimization problem is similar to the formulation in [Kondili et al., 1993] for scheduling of batch operations, but in [Kondili et al., 1993] the focus is on scheduling rather than real-time disturbance management.

The optimization problem is formulated as a mixed-integer quadratic program (MIQP). The integer variables are needed to model that an area could have a nonzero minimum rate at which it can operate, such that production between zero and the minimum rate is impossible. However, when the number of areas becomes large, this might not be a suitable formulation since mixed-integer quadratic programming problems are $\mathcal{NP}$ complete, which means that in the worst case, the solution time grows exponentially with the number of variables [Raman and Grossmann, 1991]. Although algorithms have been proposed that solve medium- and large-scale application problems [Floudas, 1995], there could still be advantages of relaxing the problem to a quadratic program (QP), if the solutions becomes similar to the solutions of the MIQP. The problems of the structure described in this paper may be reformulated as QPs that give the same solutions as the MIQPs as long as the QP solutions satisfy the constraints of the original MIQP. The MIQP problem is formulated in Section 5 and the problem is reformulated as a QP in Section 6. The formulation of the optimization problem builds on the QP formulation introduced in [Lindholm and Giselsson, 2012].

The results from the optimization may be used either for proactive or reactive disturbance management, as described in [Lindholm and Johnsson, 2013]. Studying the solution to the problem for typical utility disturbances gives proactive advice on how to control the production at these types of disturbances. If online reactive disturbance management is desired, the optimization problem can be solved in receding horizon fashion. This enables online advice to plant operators, given an estimated disturbance trajectory some steps ahead. The optimization may also be connected to other economic optimization at the site, e.g., monthly production scheduling. A straightforward approach for doing this is to let the production scheduling activity set the reference values for production rates, sales, and inventory levels for the utility disturbance management optimization. Solving the MIQP formulation of the problem in receding horizon fashion gives mixed-integer predictive control (MIPC), as introduced in [Bemporad and Morari, 1999]. The optimization model may also be used to evaluate how planned production disturbances, e.g., an area running at reduced production rate due to maintenance, should be handled. In this case, the optimization results can give advice also on how to prepare for the disturbance in terms of building up suitable buffer volumes in the tanks. An example of this is provided in Section 7.

The research presented in this paper is conducted within the Process Industrial Centre at Lund University (PIC-LU), an organization where academia and industry work together on real process industrial problems. The work on utility disturbance management is conducted in close collaboration with the companies in the research centre, in particular with the enterprise Perstorp, that is a world leader in several sectors of the specialty chemicals market. For more information about Perstorp, see [Perstorp, 2013]. The first example problem presented in Section 7 is inspired by Perstorp's site at Stenungsund, Sweden.

2. Hierarchy Models

2.1 Role-based Equipment Hierarchy

According to the standard ISA-95 [ISA-95.00.01, 2009], a site consists of one or more production areas, where each area produces either end products or intermediates. The intermediate products may either be sold on the market or refined to end products in other areas at the site. Buffer tanks could be placed between areas to serve as inventory of the intermediates or as buffer tanks with the purpose to allow independent operation of upstream and downstream areas. Thus, the production at a site may be viewed as a network of areas, with intermediate buffer tanks between some areas. An example is given in Figure 1. In this paper, it is assumed that each area produces one product, which may be stored in a dedicated buffer tank. Furthermore,

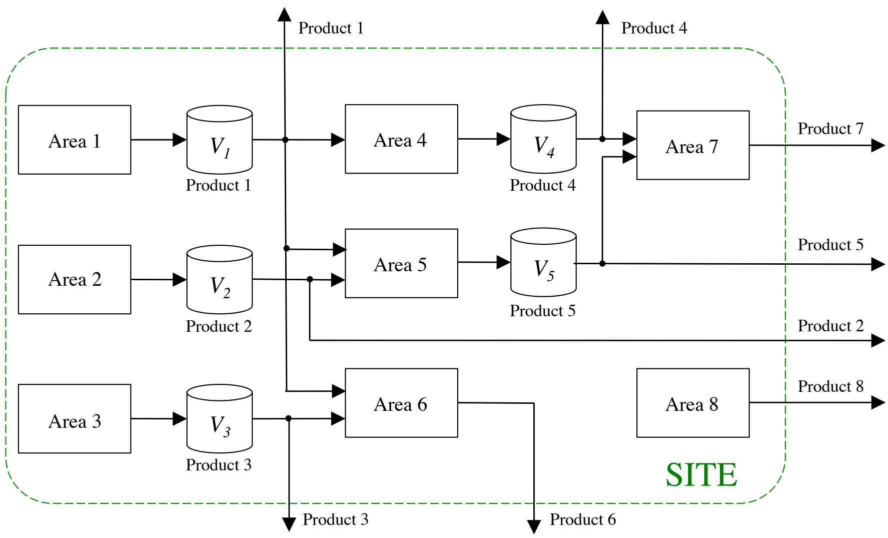

Figure 1. Example site.

the dynamics of the production units within the areas are ignored, i.e., it is assumed that the production in an area is directly proportional to the inflows to the area:

$$q_{jk}^{\text{in}}(t) = q_k(t)a_{jk} \tag{1}$$

where $q_{jk}^{\text{in}}(t)$ is the inflow of product j to area k at time t, $q_k(t)$ the production in area k at time t, and a_{jk} is called the conversion factor between product j and product k. This assumption is reasonable since the area dynamics are usually fast compared to the dynamics of the production network.

2.2 Functional Hierarchy

In ISA-95 [ISA-95.00.01, 2009], a simplified version of the Purdue Hierarchy Model [Williams, 1991] is given, which divides the functional hierarchy of an enterprise as three major levels. Level 4 handles business planning and logistics, level 3 manufacturing operations and control, and level 2, 1, and 0 the basic process control of a plant (batch/continuous or discrete). The work presented in this paper focuses on level 3 of the functional hierarchy, more specifically on the detailed production scheduling activity.

3. Utility Disturbance Management

Utilities are support processes that are utilized in production, but that are not part of the final product. Common utilities include steam, cooling water, electricity, compressed air, and water treatment. The functions of these utilities are described in [Brennan, 1998] and [Lindholm and Johnsson, 2013]. Some utilities operate continuously, such as steam, cooling water, feed water and vacuum systems, whereas some utilities have on/off characteristics. Examples of such utilities are electricity and nitrogen.

A general method for utility disturbance management, the UDM method, was introduced in [Lindholm and Johnsson, 2013]. The method aims to reduce the revenue loss due to disturbances in utilities, and the workflow of the method is presented step by step in the paper. For completing all steps of the method, a model of the site is needed that describes how utility disturbances affect production. Three modeling approaches are suggested in [Lindholm and Johnsson, 2013]: On/off production modeling without buffer tanks, on/off production modeling with buffer tanks and continuous production modeling. For the first two approaches, production areas are assumed to be either operating at maximum production speed or at zero production speed, which enables quick estimation of revenue losses. Use of the UDM method with these simple modeling approaches is suitable when quick results are important, since the time to obtain the model and make the calculations is negligible. The disturbance management strategies that are obtained are primarily proactive approaches, which suggest, e.g., where maintenance activities should be focused for maximum reduction of revenue loss due to utility disturbances. If more elaborate reactive disturbance management strategies are required, such as strategies for how to control the production at utility disturbances, continuous production modeling has to be performed. For the continuous production modeling approach, areas and utilities are considered to be continuous, and an area can operate at any production rate that is possible with respect to utility and production constraints. In this paper, a simple approach for continuous production modeling is suggested. There are two relations that have to be modeled; the effect of disturbances in utilities on production, and the effect of reduced production in one area on other areas. These two parts of the modeling procedure are discussed in Section 4.

4. Modeling of Utilities

4.1 Effects of Utility Disturbances on Production

The mapping from measurements of utility properties to the constraints they impose on production is not trivial, and might look different for different utilities. Thus, operation outside the normal limits for a utility might give very different effects on the production of the areas that require the utility. When a utility operates poorly, the capacity of the production areas that require the utility is affected. How much the capacity is affected depends on how severe the disturbance is, e.g., how much the utility operation deviates from normal operation. The measurements related to utilities are often parameters like temperature, flow or pressure of the utility. A first approach for modeling the effect of utilities on production is to assume a simple relation between the utility parameters (temperature, pressure, etc.) and production in an area, inspired by the physical properties of the utility, and by measurement data. To show the complexity of the modeling, cooling water is chosen as an example. In Figure 2, the maximum hourly production of an area at an industrial site is plotted as a function of the cooling water temperature. Too cold cooling water temperature is not a problem at the site. Despite of this, it can be seen in Figure 2 that the maximum production that has been achieved decreases with the cooling water temperature for temperatures below approximately 22°C. This shows that the measurement data that is available is actually part of a feedback loop. When the area produces at low production speed, for reasons other than cooling water disturbances, less cooling effect will be needed, and the cooling water temperature will be low. Let us assume that the maximum production rate could be kept even for cooling water temperatures below 22°C. The problem then becomes to determine how the production capacity is affected by temperatures above the limit marked in Figure 2, where the maximum production decreases with

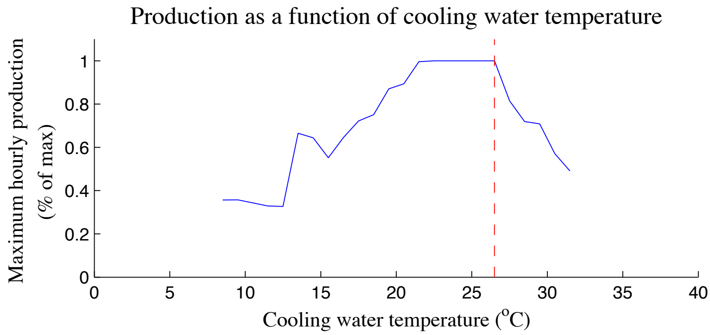

Figure 2. Maximum hourly production at an industrial site as a function of cooling water temperature.

increasing cooling water temperature. Inspired by the measurement data in Figure 2 and the assumption that too cold cooling water does not affect the production capacity, three suggestions for the utility-production relation in are formed as shown in Figure 3. The relation could be, e.g., linear over the critical limit (as in the leftmost subfigure), piecewise linear (as in the middle subfigure), or nonlinear (as in the rightmost subfigure).

In this paper, a slightly different interpretation is made. Instead of trying to model the production as a function of cooling water temperature, or other utility parameters acquired from measurement data, utilities are interpreted as volumes, or power, which all areas that require the utilities have to share. This interpretation makes sense for example for cooling water and steam utilities, where all areas that require these utilities have to split the total cooling or heating power. The relation that has to be modeled when using this framework is how the production in an area that requires the utility depends on the amount of the utility it is assigned. This reformulation of the modeling problem disconnects the optimization problem for determining the optimal supply of utilities to each area from the physical modeling from measurement data. A first modeling step has to be added, where the relationship between the utility measurements and the volume interpretation is handled, if the op-

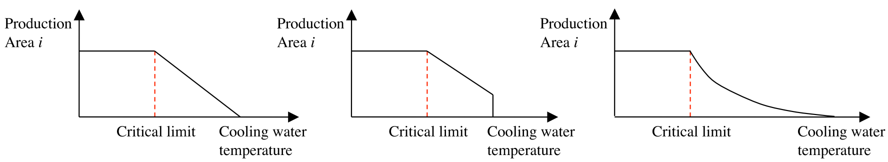

Figure 3. Possible models for the effects of cooling water temperature on production.

timization should be connected to the real site via the utility measurements. Assuming a linear relation between the amount of utility assigned to an area and the maximum production in that area, the constraint on the production rate in area j due to a utility i can be expressed as

$$q_j(t) \leq c_{ij}u_{ij}(t) + m_{ij} \tag{2}$$

where $q_j(t)$ is the production in area j at time t, $u_{ij}(t)$ the assignment of utility i to area j at time t, and c_{ij} and m_{ij} are constants specific for utility i, area j.

If $c_{ij} > 0$, (2) should correspond well to many utilities with continuous characteristics. For example, for cooling water: the cooling water utility produces a certain cooling power, that is shared between production areas that are connected to the cooling water system. If an area is assigned more cooling water power, it should be able to produce at a higher production speed, within the normal range of production rates. The constraint in (2) is presented graphically for $m = 0$ and some $c_{ij} > 0$ in Figure 4.

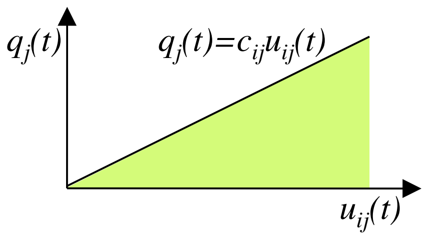

Figure 4. A continuous utility's constraints on production.

If $c_{ij} = 0$, (2) corresponds to representation of a utility with on/off characteristics, where the area can produce at some maximum speed, $m_{ij} = q_j^{\max}$, if the supply of utility is greater than zero, and not at all when it does not get assigned any amount of the utility. This is represented in Figure 5.

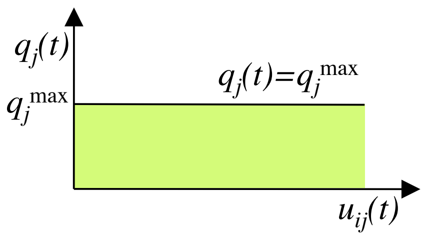

Figure 5. An on/off utility's constraints on production.

In reality, there could be a minimum amount of a utility that is required for a production area to be able to operate, here denoted $u_{ij}^{\min}$. Also, there could be an upper limit, $u_{ij}^{\max}$, such that supplying more utility than $u_{ij}^{\max}$ does not permit higher production than the maximum production if $u_{ij}^{\max}$ is assigned to the area. This modification to the constraint in (2) could be captured by setting maximum and minimum constraints on the production rates. If these constraints are taken into account, the representation of the continuous type and on/off type utilities become as in Figure 6.

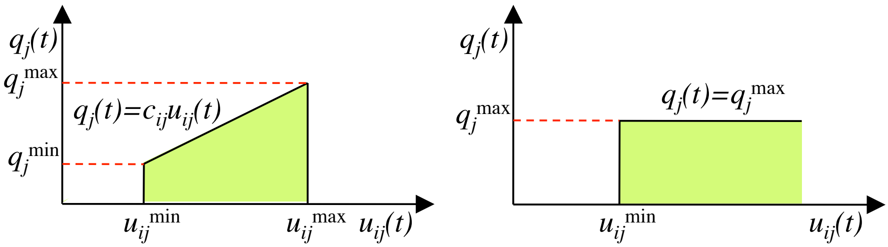

Figure 6. Utility representations with production constraints.

4.2 Connections of Areas via the Product Flow

As mentioned previously, utilities are often shared between several production areas at a site. Combining the effects of disturbances in utilities with the area connections might be cumbersome, which is here illustrated by an example: Assume that two areas, area 1 and 2, are connected by the product flow, as in Figure 7, and that the cooling water utility is required by both areas. Assume that area 1 is forced to reduce its production to 80% and area 2 to 70% because of a disturbance in the cooling water utility. Area 2 is affected by the disturbance both directly, and indirectly because of the disturbance in area 1, which delivers raw material to area 2. What limits the production of area 2 could be either the lack of raw materials (due to reduced production in area 1), or the cooling water temperature.

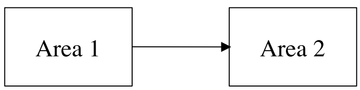

Figure 7. Two areas connected by the product flow.

The sharing of utilities means that a shutdown of one area that requires the utility may enable another area, which also requires the utility, to run instead. For example, the cooling water at a site gives a certain cooling effect that is shared by all areas that use cooling water. One area that is important to keep running might be able to get its required cooling effect at the expense of shutting down one or more other areas, or running them at reduced speed.

These discussions illustrate the complexity of the problem, and shows the suitability for formulating this as an optimization problem. The volume representation of utilities makes it possible to represent the constraints due to a shared utility i as

$$\sum_{j \in \mathcal{M}_i} u_{ij}(t) \leq U_i(t) \tag{3}$$

where $u_{ij}(t)$ is the amount of utility i that is assigned to area j at time t, $U_i(t)$ is the total amount of utility i available at time t, and $\mathcal{M}_i$ is the set of areas that require utility i. This is used to formulate the optimization problem in Section 5.

5. Formulating the Optimization Problem

In this section, an optimization problem is formulated in order to determine the optimal supply of utilities to each production area at an integrated site at the occurrence of utility disturbances. Since utilities often are shared between production areas, and production areas are coupled by the product flow at the site, the solution is not trivial for most setups. In this paper, solutions that minimize economical losses are searched for. The optimization considers the contribution margins of each product and the deviation from the reference levels in the inventories to determine the economically optimal solution. Another aspect that has to be considered is the shutdown and start-up of production areas, which often is very expensive at process industrial sites. Solutions where areas are shut down and started up frequently should be avoided to as large extent as possible. To model that an area has a nonzero minimum rate at which it can operate, unless it is entirely shut down, binary variables are used. These binary variables describe if the area operates or not in each time step. A quadratic objective function is used and the constraints that are defined are affine, which makes the optimization problem a mixed-integer quadratic program (MIQP).

The inputs to the optimization problem are the reference values for the production rates, sales, and buffer tank levels, and the utility disturbance trajectories. The utility disturbance trajectories are estimates of the total available amounts of all utilities, $U_i(t)$, at each time t. The defined optimization problem is solved in receding horizon, which means that the utility disturbance trajectories have to be estimated over the prediction horizon. Disturbances at a real site are often such that nothing is known about the disturbance in advance, but the optimization may also be used e.g. for planned maintenance of utilities, or for planned production disturbances, where the production rate in one or more areas is reduced during some time period. The results from the receding horizon optimization may be used as online advice to operators or, if possible, applied directly to form closed loop mixed-integer predictive control (MIPC). In the following subsections, the optimization problem is formulated step by step.

5.1 Required Site Information

The information about the site that is required for formulation of the optimization problem is the contribution margins and the maximum and minimum quantity of sales of all products (if such limits exist), the maximum and minimum production rates of all areas, and maximum, minimum and reference levels for all buffer tanks between areas at the site. Information of which utilities that exist at the site, and at which areas they are used, is also required, as well as how the utilities are shared among the areas at normal (often maximum) production at the site. The costs for having areas shut down and for shutting down and starting-up each area could also be useful, since these costs may be used for choosing the weights in the objective function of the optimization problem.

5.2 Mass Balance Constraints

The model of the site is given by the connections of its production areas. An example of what the site structure could look like is given in Figure 1. The connections of areas are represented by the mass balances at the internal buffer tanks, i.e.,

$$V_j(t+1) = V_j(t) + q_j(t) - q_j^m(t) - \sum_{k \in \mathcal{D}_j} q_k(t) a_{jk}, \quad j \in \mathcal{B} \tag{4}$$

where $V_j(t)$ is the volume in the buffer tank for product j at time t, $q_j(t)$ the production of product j at time t, $q_j^m(t)$ the flow to the market of product j at time t, and a_{jk} the conversion factor between product j and k (see (1)). $\mathcal{B}$ is the set of areas/products that have a dedicated buffer tank and $\mathcal{D}_j$ is the set of areas directly downstream of area j.

5.3 Utility Constraints

At a disturbance in the supply of a utility, the available amount of the utility may not be enough to supply all areas with the amount they need for maximum production. This is described by (3) in Section 4. If (2) holds for all areas j and utilities i, (3) become time-varying constraints of the production rates of all areas that share a utility, since $U_i(t)$ varies over time.

For continuous utilities, it can be assumed that equality holds in (2) since it would not be optimal for an area to not use all its assigned utility volume to produce its product, as the production in other areas might be limited by the same utility. Since $c_{ij} > 0$ for continuous utilities, (3) can then be rewritten using (2) as

$$\sum_{j \in \mathcal{M}_i} k_{ij} q_j(t) - d_{ij} \leq U_i(t), \quad i \in \mathcal{K} \tag{5}$$

where $k_{ij} = 1/c_{ij}$ and $d_{ij} = m_{ij}/c_{ij}$ are positive constants for utility i, area j. $\mathcal{K}$ is the set of utilities of continuous type and $\mathcal{M}_i$ is the set of areas that require utility i.

For on/off utilities, where $c_{ij} = 0$, (2) becomes equivalent to the maximum constraint on the production rate. The areas that require the utility can operate at maximum speed at time t if the utility is available at time t, and can not operate if the utility is unavailable. The constraints on production due to on/off utilities may thus be written as

$$q_j(t) \leq q_j^{\max} \qquad \text{if } U_i(t) = 1, \quad j \in \mathcal{M}_i, \ i \in \mathcal{L} \tag{6}$$

$$q_j(t) = 0 \qquad \text{if } U_i(t) = 0, \quad j \in \mathcal{M}_i, \ i \in \mathcal{L} \tag{7}$$

where $\mathcal{L}$ is the set of on/off type utilities. $U_i(t)$ is always zero or one for on/off type utilities.

5.4 Buffer Tank Level Constraints

The maximum and minimum limitations on the buffer tank volumes give constraints

$$V_j^{\min} \leq V_j(t) \leq V_j^{\max}, \quad j \in \mathcal{B} \tag{8}$$

The maximum and minimum limits, $V_j^{\max}$ and $V_j^{\min}$, might correspond to the entire buffer tank, or they could correspond to a part of the tank that is reserved for handling disturbances in utilities.

5.5 Production Rate Constraints

The production rate constraint may be formulated using binary variables, to model that an area can not operate at a rate lower than the minimum production rate, unless it is completely shut down. With binary variables $w_j(t) \in \{0, 1\}$ that are equal to one if area j runs at time t and zero otherwise, the production rate constraints become

$$w_j(t) q_j^{\min} \leq q_j(t) \leq w_j(t) q_j^{\max}, \quad j \in \mathcal{A} \tag{9}$$

where $q_j^{\max}$ and $q_j^{\min}$ are the maximum and minimum rates at which area j can operate, and $\mathcal{A}$ the set of areas at the site. The constraint is formulated in the same manner as in [Kondili et al., 1993], and ensures that an area can not operate at a rate between 0 and $q_j^{\min}$. This formulation also enables penalization of start-up/shutdown of areas. An area has been shut down from time t to $t+1$ if $w(t) = 1$ and $w(t+1) = 0$. If there are planned production disturbances at the site that should be taken into account in the optimization, this may be expressed by a time-varying maximum production rate for the affected areas.

5.6 Market Constraints

The amount that can be sold of each product at each time instance may also be limited. This gives constraints

$$q_j^{m,\min} \leq q_j^m(t) \leq q_j^{m,\max}, \quad j \in \mathcal{A} \tag{10}$$

where $q_j^{m,\min}$ and $q_j^{m,\max}$ are minimum and maximum limitations for the flow to the market from area j. The limits at a real site may be time-varying, but in this study they are assumed to be constant for simplicity.

5.7 Steady-state Optimization

Before the dynamic optimization problem can be cast, reference values need to be computed. The production rates, q, and the flows to the market, q^m, that give the optimal profit may be determined by a steady-state optimization, i.e., by assuming that there are no disturbances in utilities, and no buffer tanks between areas. The optimal production rates and flows to the market and the optimal profit from the steady-state optimization are used as reference values for the dynamic optimization, where the objective is to minimize the economical effects of disturbances in the supply of utilities. The reference values may also be obtained from some other higher-level optimization at the site.

The profit is given by the sum of the contribution margin multiplied by the quantity of the sales for all products. If m is a column vector containing the contribution margins of all products and there are no disturbances and no buffer tanks between areas, the optimal profit can thus be determined by the mixed-integer linear program (MILP)

$$\begin{aligned} \text{maximize} \quad & m^T q^m \\ \text{subject to} \quad & (4) \text{ with } V_j(t+1) = V_j(t) \\ & (9) - (10) \end{aligned} \tag{11}$$

The variables are q, q^m, and w, which are all column vectors with one element for each area. The solution to the steady-state optimization problem is denoted q_{ref}, q_{ref}^m, with the optimal profit p_{ref}.

5.8 Dynamic Optimization

The solution to the steady-state optimization problem in Section 5.7 gives reference values for the dynamic optimization, which aims to optimize the economical performance of the entire site. A quadratic objective function that maximizes the contribution, minimizes the deviation from the reference values and penalizes start-up and shutdown of areas is suggested. Since the inequality constraints and equality constraints stated in the previous subsections are affine, the problem becomes a mixed-integer quadratic program (MIQP). The suggested objective function is

$$J\left(\mathbf{q}, \mathbf{q^m}, \mathbf{V}, \mathbf{w}\right) = \sum_{\tau=0}^{N-1} J_t\left(q(\tau), q^m(\tau), V(\tau), w(\tau)\right) \tag{12}$$

with variables

$$\begin{aligned}
\mathbf{q} &= \begin{bmatrix} q(0)^T & \dots & q(N-1)^T \end{bmatrix}^T \\
\mathbf{q}^m &= \begin{bmatrix} q^m(0)^T & \dots & q^m(N-1)^T \end{bmatrix}^T \\
\mathbf{V} &= \begin{bmatrix} V(0)^T & \dots & V(N-1)^T \end{bmatrix}^T \\
\mathbf{w} &= \begin{bmatrix} w(0)^T & \dots & w(N-1)^T \end{bmatrix}^T
\end{aligned}$$

where

$$\begin{aligned}
J_t\left(q(t), q^m(t), V(t), w(t)\right) = &\left(m^T q^m(t) - p_{\text{ref}}\right)^2 Q_p + \Delta V^T(t) Q \Delta V(t) \\
&+ \Delta q^T(t) R_1 \Delta q(t) + \Delta {q^m}^T(t) R_2 \Delta q^m(t) \\
&+ \sum_{j\in\mathcal{A}} g_j\left(1 - w_j(t)\right) \\
&+ \sum_{j\in\mathcal{A}} k_j\left(w_j(t+1) - w_j(t)\right)^2
\end{aligned}$$

with

$$\begin{aligned}
\Delta V(t) &= V(t) - V_{\text{ref}} \\
\Delta q(t) &= q(t) - q_{\text{ref}} \\
\Delta q^m(t) &= q^m(t) - q^m_{\text{ref}}
\end{aligned}$$

$Q_p > 0$ is a scalar weight, Q is positive semidefinite and R_1, R_2 positive definite weighting matrices, and g_j and k_j are scalar non-negative weights for all j. Thus, the problem can be solved using convex methods and branch-and-bound. The reference values q_{ref}, q^m_{ref}, and p_{ref} are given from the steady-state optimization, (11), and V_{ref} are the given reference levels for the buffer tanks. The objective function consists of four major parts:

- $(m^T q^m(t) - p_{\text{ref}})^2 Q_p$

  To penalize deviation from the reference profit (maximize the contribution).

- $\Delta V^T(t) Q \Delta V(t)$

  To penalize deviations from reference buffer tank levels, to avoid solutions where all inventories are sold.

- $\Delta q^T(t) R_1 \Delta q(t)$ and $\Delta q^{m\,T}(t) R_2 \Delta q^m(t)$

  To penalize deviations from nominal production and sales.

- $\sum_j g_j\,(1 - w_j(t))$ and $\sum_j k_j\,(w_j(t+1) - w_j(t))^2$

  To inflict extra cost to area shutdown. g_j is the cost to have area j shut down, and k_j the cost for shutting down/starting up area j.

The optimization problem becomes

$$\begin{array}{lll} \text{minimize} & (12) & \\ \text{subject to} & (4), & t = 0, \ldots, N-2 \\ & (5) - (10), & t = 0, \ldots, N-1 \end{array} \qquad (13)$$

with variables $\mathbf{q}$, $\mathbf{q^m}$, $\mathbf{V}$, and $\mathbf{w}$. The problem may be solved, e.g., using CPLEX MIQP solver, either offline, or online in receding horizon fashion.

The utility disturbance trajectories, $U_i(t)$, are inputs to the optimization that come in at constraints (5) and (6) – (7). Information about utility disturbances is most often not known in advance, which means that $U_i(t)$ is equal to the maximum available amount for all times t in the prediction horizon until the disturbance occurs. However, the optimization can also be used when the disturbance is known in advance. Planned disturbances in production may also be studied by modifying the maximum production rates for the considered areas at constraint (9), for the time steps when reduced production is planned.

6. Reformulation as a Quadratic Program (QP)

The optimization problem (13) is a MIQP and is therefore $\mathcal{NP}$ complete to solve. As the problem size grows, this may result in prohibitively long execution times. Thus, efficient relaxations that can be solved using quadratic programming may be desirable. These relaxations are also advantageous since there are many free and commercial solvers for these problems, as opposed to for MIQPs, where most good solvers are commercial. Some suggestions of convex relaxations of mixed-integer quadratic programs are given in [Axehill et al., 2010]. In this paper, we present a QP relaxation that is suitable for MIQP problems of the structure presented in Section 5. The suggested QP relaxation gives the same solution as the MIQP, provided that the optimal QP solution satisfies the constraints in (13). To this end, we introduce auxiliary variables s and exchange the mixed-integer constraints in (9) by

$$q_j^{\min} + s_j(t) \le q_j(t) \le q_j^{\max}, \qquad j \in \mathcal{A} \tag{14}$$

$$-q_j^{\min} \le s_j(t) \le 0, \qquad j \in \mathcal{A} \tag{15}$$

where $s_j(t)$ is the auxiliary variable for area j at time t.

This formulation does not guarantee that production rates in $(0,\ q_j^{\min})$ is avoided. However, for solutions for which this is avoided, the cost can be chosen to give the same solution as in the MIQP formulation in (13). We choose

$$\sum_{j\in\mathcal{A}} -\frac{g_j}{q_j^{\min}} s_j(t) + \sum_{j\in\mathcal{A}} \frac{k_j}{(q_j^{\min})^2}(s_j(t) - s_j(t+1))^2 \tag{16}$$

to replace the cost for the binary variables in the MIQP case (the last part of the MIQP objective function in Section 5.8). We have that $s_j(t) = -q_j^{\min}$ corresponds to $w_j(t) = 0$, i.e., that area j is stopped, and $s_j(t) = 0$ corresponds to $w_j(t) = 1$, i.e., that area j is running. The objective function for the relaxed QP case becomes

$$J(\mathbf{q}, \mathbf{q}^m, \mathbf{V}, \mathbf{s}) = \sum_{\tau=0}^{N-1} J_t\left(q(\tau), q^m(\tau), V(\tau), s(\tau)\right) \tag{17}$$

with variables

$$\begin{aligned}
\mathbf{q} &= \begin{bmatrix} q(0)^T & \dots & q(N-1)^T \end{bmatrix}^T \\
\mathbf{q}^m &= \begin{bmatrix} q^m(0)^T & \dots & q^m(N-1)^T \end{bmatrix}^T \\
\mathbf{V} &= \begin{bmatrix} V(0)^T & \dots & V(N-1)^T \end{bmatrix}^T \\
\mathbf{s} &= \begin{bmatrix} s(0)^T & \dots & s(N-1)^T \end{bmatrix}^T
\end{aligned}$$

and

$$\begin{aligned} J_t\left(q(t), q^m(t), V(t), s(t)\right) = &\left(m^T q^m(t) - p_{\mathrm{ref}}\right)^2 Q_p + \Delta V^T(t) Q \Delta V(t) \\ &+ \Delta q^T(t) R_1 \Delta q(t) + \Delta q^{mT}(t) R_2 \Delta q^m(t) \\ &+ \sum_{j \in \mathcal{A}} -\frac{g_j}{q_j^{\min}} s_j(t) \\ &+ \sum_{j \in \mathcal{A}} \frac{k_j}{(q_j^{\min})^2} \left(s_j(t+1) - s_j(t)\right)^2 \end{aligned}$$

with

$$\begin{aligned} \Delta V(t) &= V(t) - V_{\mathrm{ref}} \\ \Delta q(t) &= q(t) - q_{\mathrm{ref}} \\ \Delta q^m(t) &= q^m(t) - q^m_{\mathrm{ref}} \end{aligned}$$

and all matrices the same as in the MIQP case. The optimization problem becomes

$$\begin{aligned} &\text{minimize} \;\; (17) && \\ &\text{subject to} \;\; (4), && t = 0, \ldots, N-2 \\ &\qquad (5)-(8),\ (10),\ (14)-(15), && t = 0, \ldots, N-1 \end{aligned} \tag{18}$$

with variables $\mathbf{q}$, $\mathbf{q^m}$, $\mathbf{V}$, and $\mathbf{s}$. The choice of optimization weights for the auxiliary variables according to (16) gives that the QP and MIQP formulations are equivalent as long as the optimal solution for the relaxed problem satisfies $s_j(t) = 0$ or $s_j(t) = -q_j^{\min}$ for all j and t. Since the flows and tank levels are penalized quadratically and the auxiliary variables are non-positive and penalized linearly in the cost function, the QP reformulation is closely related to ℓ_1-norm regularization and basis pursuit [Boyd and Vandenberghe, 2004, §6.5.4, §11.4.1]. Basis pursuit and ℓ_1-norm regularization are known to generate sparse solutions [Boyd and Vandenberghe, 2004, §6.5.4], i.e., solutions that are prone to satisfy $s_j(t) = 0$ in our formulation. This behavior can be interpreted as that an additional step-cost is inflicted when $s_j(t) < 0$. Thus, for a large enough linear cost, we are likely to get solutions that satisfy either $s_j(t) = 0$ or $s_j(t) = -q_j^{\min}$. This behavior is also confirmed by numerical examples.

The QP optimization problem has a structure that makes it possible to solve it in a distributed fashion. Therefore, the solution method presented in [Giselsson et al., 2013] may be used to solve this problem efficiently. Otherwise, any quadratic programming solver such as CPLEX or GUROBI may be used. In this paper, CPLEX is used to solve both the MIQP and QP problems.

6.1 Two-phase Optimization

It cannot be guaranteed a priori that the solution to (18) satisfies the original MIQP constraints, i.e., that either $s_j(t) = -q_j^{\min}$ or $s_j(t) = 0$ for all j, t. To avoid solutions that are infeasible for the MIQP formulation, the following simple two-phase procedure can be used.

ALGORITHM 1
Two-phase optimization

---

1. Solve the relaxed problem (18)
2. **If** $s_j(t) \neq 0$ and $s_j(t) \neq -q_j^{\min}$

    If $|s_j(t) - 0| < |s_j(t) + q_j^{\min}|$

        Add constraint: $s_j(t) = 0$

    Else

        Add constraint: $s_j(t) = -q_j^{\min}$

    End

   End
3. Solve (18) with additional constraints □

Using this two-phase procedure the original MIQP constraints will hold after the second optimization in step 3 of Algorithm 1.

7. Examples

In this section, the optimization problem formulation presented in Section 5 is used to formulate and solve two specific problems. The solutions to the QP relaxations of the problems, as suggested in Section 6, are also presented and the results are compared to the original MIQP formulations. The first problem is a small example with six production areas that is inspired by a real industrial case. This example aims to show how the general structure for formulating these kinds of optimization problems is used in practice. The second problem is a larger problem with 30 production areas. Here, the formulation of the problem is handled in less detail, and the goal is instead to show more clearly the differences in solution time of the solutions to the MIQP and QP formulations.

7.1 Example I

This example is inspired by Perstorp's site in Stenungsund, Sweden, that produces mainly aldehydes, organic acids, alcohols and plasticizers. The production data, and to some extent the site structure, has been modified in the paper due to secrecy matters.

Required site information The site that is considered is the site with six production areas given in Figure 8. Table 1 summarizes the maximum and minimum production rates of all areas, $q^{\max}$ and $q^{\min}$, the contribution margins of all products, m, the maximum and minimum volume of all buffer tanks, $V^{\max}$ and $V^{\min}$, and the reference volumes for the buffer tanks, V_{ref}. It is assumed that the plant is not market limited, such that there is no upper bound on the sales of the products. A lower bound of zero is used for the flows to the market in the optimization.

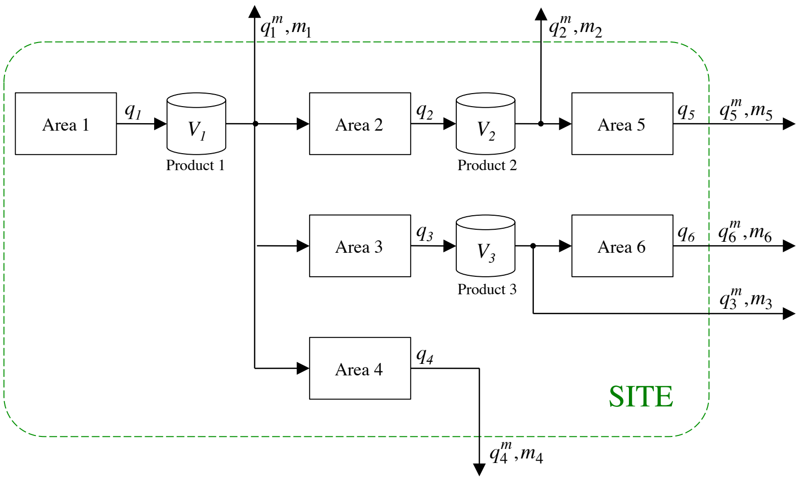

Figure 8. Site considered in example I.

Table 1. Production data.

	$q^{\min}$	$q^{\max}$	m	$V^{\min}$	$V^{\max}$	V_{ref}
Product 1	0.10	1	0.4	0	0.5	0.5
Product 2	0.05	0.5	0.7	0	0.5	0.5
Product 3	0.02	0.2	0.1	0	0.5	0.5
Product 4	0.01	0.1	0.5	-	-	-
Product 5	0.02	0.2	0.8	-	-	-
Product 6	0.02	0.2	1.0	-	-	-

Three utilities are considered in the example; high pressure (HP) steam, middle pressure (MP) steam, and cooling water. Table 2 shows which utilities that are required at each area. It is assumed that, at maximum production, the utilities are shared equally between the areas that require them.

Table 2. Utilities required at each area.

Area →	1	2	3	4	5	6
HP steam	x		x			
MP steam		x		x		x
Cooling water	x	x	x	x	x	x

Mass balance constraints The mass balances at all buffer tanks at the site give

$$V_1(t+1) = V_1(t) + q_1(t) - q_1^m(t) - q_2(t) - q_3(t) - q_4(t) \tag{19}$$
$$V_2(t+1) = V_2(t) + q_2(t) - q_2^m(t) - q_5(t) \tag{20}$$
$$V_3(t+1) = V_3(t) + q_3(t) - q_3^m(t) - q_6(t) \tag{21}$$

As can be seen in the equations, all conversion factors, a_{jk}, are chosen to be equal to one in the example, for simplicity.

Utility constraints It is assumed that the total amount of each utility is equal to one ($U_1 = U_2 = U_3 = 1$) when it operates at maximum capacity. The utilities in the example are of continuous type (see Section 4), and it is assumed that zero assignment of a utility to an area gives zero production, i.e., $c_{ij}, k_{ij} > 0$ and $m_{ij} = d_{ij} = 0$ for all utilities i and areas j. The time-varying constraints on the production rates due to shared utilities are obtained from (5) using Table 2. We get

$$k_{11}q_1(t) + k_{13}q_3(t) \leq U_1(t) \tag{22}$$
$$k_{22}q_2(t) + k_{24}q_4(t) + k_{26}q_6(t) \leq U_2(t) \tag{23}$$
$$\sum_{j=1}^{6} k_{3j}q_j(t) \leq U_3(t) \tag{24}$$

where $U_i(t)$ is equal to one if utility i operates at maximum capacity at time t, and less than one otherwise. Since the utilities in the example are shared equally at maximum production, we get

$$k_{11} = \frac{1}{2q_1^{\max}}, \quad k_{13} = \frac{1}{2q_3^{\max}} \tag{25}$$

$$k_{22} = \frac{1}{3q_2^{\max}}, \quad k_{24} = \frac{1}{3q_4^{\max}}, \quad k_{26} = \frac{1}{3q_6^{\max}} \tag{26}$$

$$k_{3j} = \frac{1}{6q_j^{\max}}, \quad j = 1, \ldots, 6 \tag{27}$$

Buffer tank level constraints Upper and lower level constraints are given by (8) for the three buffer tanks, $j = 1, 2, 3$. The limits $V^{\max}$ and $V^{\min}$ are given in Table 1.

Production rate constraints Maximum and minimum limitations on production rates give constraints according to (9) for the six production areas, $j = 1, \ldots, 6$. The limits $q^{\max}$ and $q^{\min}$ are given in Table 1.

Market constraints The flows to the market are not bounded above, because of the assumption that the plant is not market limited. The only constraint on the flows to the market is thus

$$q_j^m(t) \geq 0, \quad j = 1, \ldots, 6 \tag{28}$$

Steady-state optimization Since there are no buffer tanks for the end products, the flows to the market from the end product areas are the same as the production in these areas. Thus, the flows to the market from end product areas may be omitted from the optimization. Merging the production of all areas, and the flows to the market of intermediate products to one decision variable vector, we get

$$\bar{q} = \begin{bmatrix} q_1 & q_2 & q_3 & q_4 & q_5 & q_6 & q_1^m & q_2^m & q_3^m \end{bmatrix}^T$$

The extended contribution margin vector becomes

$$\bar{m} = \begin{bmatrix} 0 & 0 & 0 & m_4 & m_5 & m_6 & m_1 & m_2 & m_3 \end{bmatrix}^T$$

The steady-state solution that maximizes the objective function of (11) subject to constraints (19) – (21) with $V_j(t+1) = V_j(t)$, and (9), (28) becomes

$$\bar{q}_{\text{ref}} = \begin{bmatrix} 1 & 0.5 & 0.2 & 0.1 & 0.2 & 0.2 & 0.2 & 0.3 & 0 \end{bmatrix}^T$$

with the optimal profit $p_{\text{ref}} = 0.7$. $w_j = 1$ for all j in the steady state solution, which means that all areas operate in steady state.

Dynamic optimization The MIQP optimization problem is given by

$$
\begin{aligned}
&\text{minimize} && (12) && \\
&\text{subject to} && (19)-(21), && t=0,\ldots,N-2 \\
& && (8),(9),(22)-(28), && t=0,\ldots,N-1
\end{aligned}
\tag{29}
$$

with variables $\mathbf{q}$, $\mathbf{q^m}$, $\mathbf{V}$, and $\mathbf{w}$. The optimization weights are chosen as $Q_p = 100$, $R = \text{diag}\left(\begin{bmatrix}0.1 & 0.1 & 0.1 & 10 & 10 & 10 & 10 & 10 & 10\end{bmatrix}\right)$, $Q = I_3$, $g = -100 \cdot \mathbf{1}$, $k = \mathbf{1}$, where R is the merged weight of R_1 and R_2 for $\bar{q}$, and $\mathbf{1}$ denotes a column vector of ones. The problem is solved in receding horizon with $N = 10$.

Results The disturbance trajectory that is evaluated is a middle-pressure steam disturbance, where the available middle-pressure steam is 30% of the maximum amount during 10 time units. The disturbance enters the system at $t = 5$, and before this time nothing is known about the disturbance. The solution trajectories are given in Figure 9. To give a clearer view of how the disturbance is handled, the production and sales of each product around the time of the disturbance are shown in Figure 10 together with the reference values and the minimum production rates.

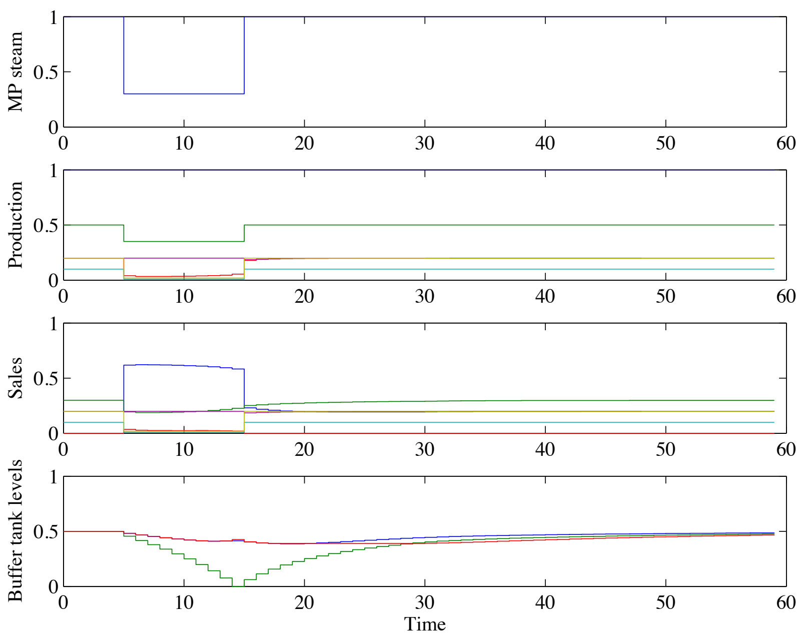

Figure 9. Optimal trajectories for MP steam disturbance.

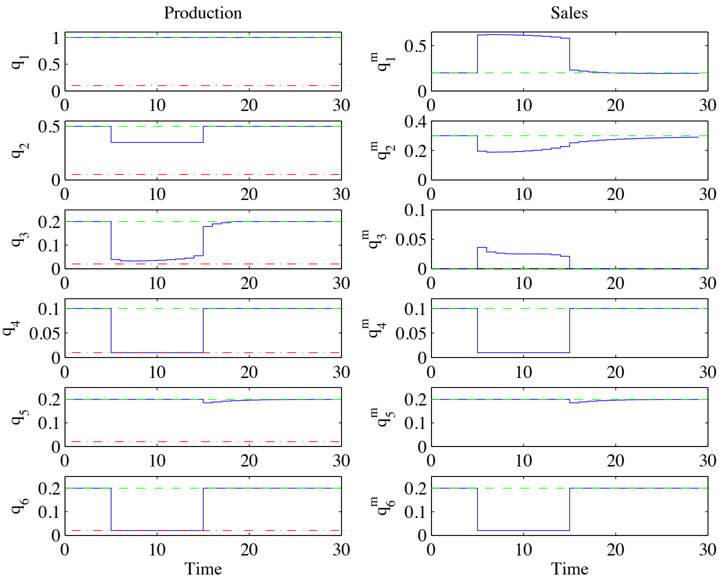

Figure 10. Optimal production and sales trajectories at MP steam disturbance (solid lines). The steady-state solution is shown as dashed lines and the minimum production rates as and dash-dotted lines.

According to Table 2, middle-pressure steam is required at areas 2, 4, and 6. In Figure 10 it can be seen that area 4 and 6, which are both end product areas, are run at the minimum rate during the disturbance to avoid the costs for having an area shut down. The production in area 2 has not been reduced as much, because it is an intermediate area that needs to operate in order for the profitable downstream area 5 to be able to run. Since area 6 runs at reduced rate, area 3 does not have to run at full production speed, since it is less profitable to produce and sell product 3 than to sell product 1 directly. The solution shows that the optimal product flow control at a disturbance that only affects some production areas is not trivial to find without use of optimization.

QP reformulation If the MIQP problem is relaxed to a QP according to Section 6, the optimization weights on the slack variables become

$$\sum_{j=1}^{6} -\frac{100}{q_j^{\min}} s_j(t) + \sum_{j=1}^{6} \frac{1}{(q_j^{\min})^2} (s_j(t) - s_j(t+1))^2$$

The solution to the QP reformulation of the problem is as similar to the MIQP solution that the differences compared to Figures 9 and 10 can not be seen with the naked eye. None of the areas operate at a rate between zero and the minimum rate. This means that the QP solution is feasible also for the MIQP case, which means that the solutions to the MIQP and QP are actually equal. Thus, the use of integer variables may be avoided for this example, by expressing the MIQP problem as an equivalent QP. The time to solve the QP is not much smaller than the time to solve the MIQP, which means that no significant speedup is achieved for this small example.

7.2 Example II

This example is a site with 30 interconnected areas with 12 end products, 18 intermediate products, and 18 buffer tanks between areas. The example aims to show the differences in solution time between the original MIQP formulation of the problem and the suggested QP relaxation. In the example, both a utility disturbance and a planned production disturbance are studied.

Required site information The site that is studied is depicted in Figure 11. The maximum production rates are marked in the figure. The minimum production rates are 10% of the maximum rates. The maximum buffer tank volumes are equal to 20 and the minimum volumes are zero for all 18 buffer tanks. The reference levels for all tanks are set to 50% of the total volume, i.e., 10. The plant is assumed not to be market limited. The contribution margins are evenly spaced numbers from 0.1 to 0.5 for the 30 areas, where the contribution margin increases with increasing area number. One utility is considered, that is used at all areas with even numbers. The utility

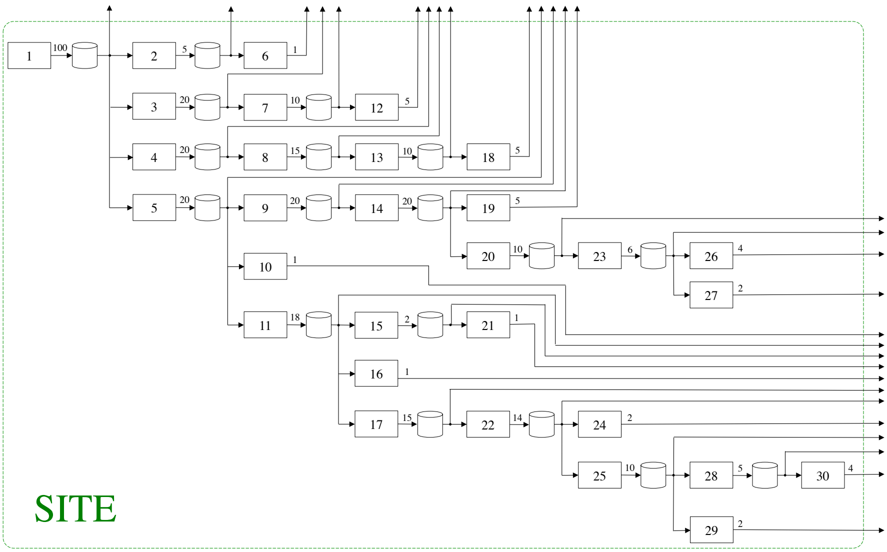

Figure 11. Site considered in example II.

is of continuous type and is shared equally at maximum production among the areas that require the utility.

Constraints The constraints are defined in the same manner as in example I. We have mass balance constraints according to (4) and the site structure in Figure 11, utility constraints according to (5), buffer tank level constraints according to (8), constraints on the flows to the market to be greater than zero, and production rate constraints according to (9).

Steady-state optimization The steady state optimization gives the reference values for production rates and sales, $\bar{q}$, to be used by the dynamic optimization. The optimal steady state operation in the example is to let all areas run at the maximum rate.

Dynamic optimization The dynamic optimization problem is given by (29) with the constraints mentioned in this subsection. The weights are chosen as $Q_p = 100$, $R = 10I_{48}$, $Q = I_{18}$, $g = 10(q^{\max})^2$, $k = \mathbf{0}$, where R is the merged weight of R_1 and R_2 for $\bar{q}$ and $\mathbf{0}$ denotes a vector with zeros. The problem is solved in receding horizon for different horizons N.

Results Two disturbances are evaluated in this example. First, at time $t = 5$, a utility disturbance equal to the one in example I affects the system, i.e., the utility that affects all areas with even numbers operates at 30% during 10 time units. Then, at time $t = 75$, there is planned production disturbance, where the production in area 11 is reduced to its minimum rate during 5 time units. The trajectories for $N = 30$ are shown in Figure 12. In the figure, the difference of the reaction to utility disturbances and planned production disturbances can be seen. For planned production disturbances, the system can start preparing for the disturbance in advance by building up suitable levels in the buffer tanks.

QP reformulation When relaxing the MIQP problem to a QP according to Section 6, it turns out that the QP solution satisfies the original MIQP constraints. This implies that the solutions to the MIQP and QP are equal. In Figure 13, the solution time to solve the MIQP and the QP is plotted as a function of the horizon, N. In the figure it can be seen that the maximum solution time to solve the MIQP increases rapidly for long horizons, i.e., when the number of integer variables becomes large. For the QP, the solution time increases much more slowly with the horizon. However, the solution times for both the MIQP and QP are small for this example, and the problems are solved in seconds even for long prediction horizons. A problem that can not be seen in the figure is that the solver often runs into numerical problems when solving the MIQP, which may lead to problems in the receding horizon optimization. For other site structures, the MIQP problem could not even be solved by CPLEX in a reasonable time.

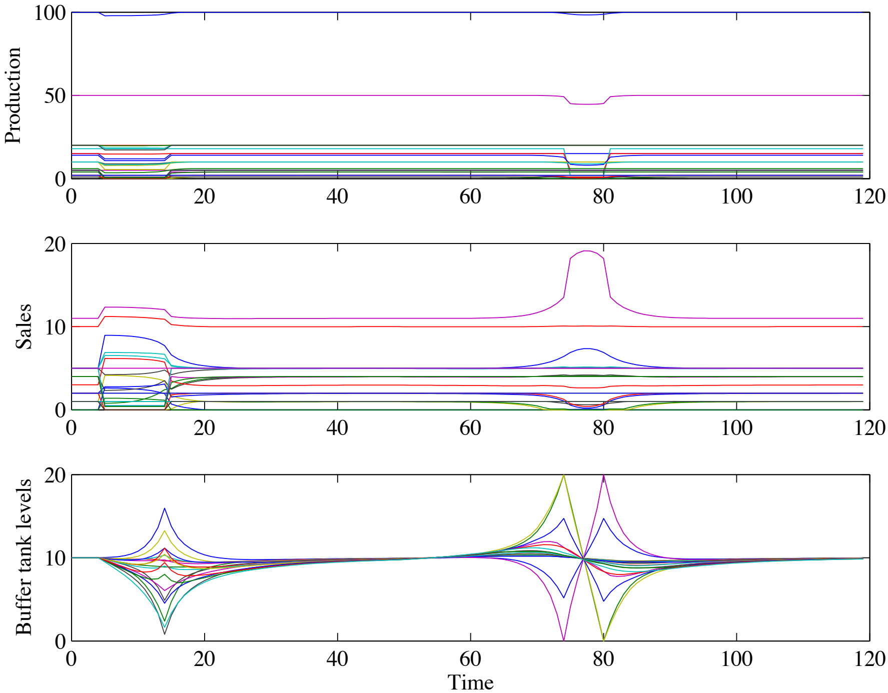

Figure 12. Optimal trajectories in example II.

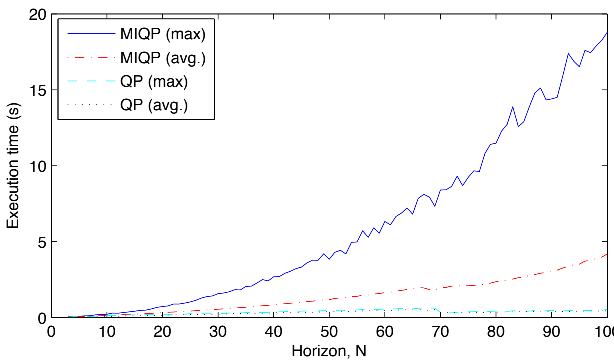

Figure 13. Solution time to solve MIQP and QP as a function of the horizon.

8. Conclusions

A simple modeling approach for modeling the effects of utility disturbances on production was introduced, that can represent both continuous and on/off type utilities. This representation allows formulation of an optimization problem that aims to find the production control strategy that minimizes the revenue loss due to disturbances in utilities. The formulation is presented step by step in the paper and aims to be general, such that it can be applied to any process industrial site. The optimization results can be used to analyze the effects of different plant-wide disturbances in utilities offline, to improve proactive disturbance management for utility disturbances. The problem may also be solved in receding horizon fashion to enable online advice to site operators on how to control the production at utility disturbances, or to use for closed-loop model predictive control (MPC) or mixed-integer predictive control (MIPC). In this case, the disturbance trajectories have to be estimated over the prediction horizon. These predicted disturbance trajectories may be updated in each time step, which may be useful if new information about the disturbances becomes available.

The formulation of the optimization problem resulted in a mixed-integer quadratic program (MIQP). To avoid long solution times when the number of areas, and thereby the number if integer variables, becomes large, the MIQP is reformulated as an ordinary quadratic program (QP). The suggested QP relaxation gives the same solution as the MIQP if the QP solution satisfies the original MIQP constraints.

The research presented in this paper fits into the framework presented in [Lindholm and Johnsson, 2013] and aims to be general, such that the methodology may be used at any process industrial site. Applications such as the pulp and paper industry could also be considered, with utilities such as electricity and waste water treatment. In that case, units in the production line are considered rather than production areas.

Acknowledgments

The research is performed within the Process Industry Centre (PIC) supported by the Swedish Foundation for Strategic Research (SSF). The authors are members of the LCCC Linnaeus Center and the eLLIIT Excellence Center at Lund University.

References

Axehill, D., L. Vandenberghe, and A. Hansson (2010). “Convex relaxations for mixed integer predictive control”. *Automatica* **46**:9, pp. 1540–1545.

Bemporad, A. and M. Morari (1999). “Control of systems integrating logic, dynamics, and constraints”. *Automatica* **35**:3, pp. 407–427.

Boyd, S. and L. Vandenberghe (2004). *Convex Optimization.* Cambridge University Press, USA.

Brennan, D. (1998). *Process Industry Economics: An International Perspective.* Institution of Chemical Engineers, Instrument Society of America.

Downs, J. J. and E. F. Vogel (1993). “A plant-wide industrial process control problem”. *Computers & Chemical Engineering* **17**:3, pp. 245–255.

Floudas, C. A. (1995). *Nonlinear and mixed-integer optimization: fundamentals and applications.* Oxford University Press, USA.

Giselsson, P., M. D. Doan, T. Keviczky, B. De Schutter, and A. Rantzer (2013). “Accelerated gradient methods and dual decomposition in distributed model predictive control”. *Automatica* **49**:3, pp. 829–833.

Hackman, S. T. and R. C. Leachman (1989). “A general framework for modeling production”. *Management Science* **35**:4, pp. 478–495.

Hui, C.-W. and Y. Natori (1996). “An industrial application using mixed-integer programming technique: A multi-period utility system model”. *Computers & Chemical Engineering* **20** (Supplement 2), S1577–S1582.

ISA-95.00.01 (2009). *Enterprise-Control System Integration, Part 1: Models and Terminology.* Standard. Instrument Society of America, pp. 17–28.

Iyer, R. R. and I. E. Grossmann (1998). “Synthesis and operational planning of utility systems for multiperiod operation”. *Computers & Chemical Engineering* **22**:7, pp. 979–993.

Kondili, E., C. C. Pantelides, and R. W. H. Sargent (1993). “A general algorithm for short-term scheduling of batch operations – I. MILP formulation”. *Computers & Chemical Engineering* **17**:2, pp. 211–227.

Lindholm, A. and P. Giselsson (2012). “Formulating an optimization problem for minimization of losses due to utilities”. In: *Proceedings of the 8th IFAC Symposium on Advanced Control of Chemical Processes.* Singapore, pp. 567–572.

Lindholm, A. and C. Johnsson (2013). “Plant-wide utility disturbance management in the process industry”. *Computers & Chemical Engineering* **49**, pp. 146–157.

Papoulias, S. A. and I. E. Grossmann (1983). “A structural optimization approach in process synthesis – I: utility systems”. *Computers & Chemical Engineering* **7**:6, pp. 695–706.

Perstorp (2013). [online] http://www.perstorp.com/, February 2013.

Raman, R. and I. E. Grossmann (1991). "Relation between MILP modelling and logical inference for chemical process synthesis". *Computers & Chemical Engineering* **15**:2, pp. 73–84.

Wassick, J. M. (2009). "Enterprise-wide optimization in an integrated chemical complex". *Computers & Chemical Engineering* **33**:12, pp. 1950–1963.

Williams, T. J., ed. (1991). *A reference model for computer integrated manufacturing (CIM)*. Instrument Society of America, USA.

Zheng, A., R. V. Mahajanam, and J. M. Douglas (1999). "Hierarchical procedure for plantwide control system synthesis". *AIChE Journal* **45**:6, pp. 1255–1265.

Paper III

Empirical Models for Utility Disturbances in the Process Industry

Anna Lindholm, Krister Forsman

Abstract

In this paper, a general method for identifying common disturbances in the supply of utilities to a process industrial site is presented, with focus on the chemical process industry. The method aims at finding typical utility disturbance trajectories and define them by simple measures, such as duration and time between failures. Statistical models are suggested for each of the measures. The estimated disturbance trajectories may be used as inputs to optimization problems for determining the optimal supply of utilities to each area of a site during a disturbance. Use of such optimization procedures can improve both proactive and reactive disturbance management for utility disturbances. Industrial data for some utilities, among others steam and cooling water, have been used to retrieve disturbance models for utilities at a specific site. These models may also give some information on the general characteristics of utility disturbances.

Submitted to *The 19th World Congress of the International Federation of Automatic Control*, Cape Town, South Africa, 2013.

1. Introduction

Utilities such as steam and cooling water are commonly used at process industrial sites, and the utility costs often represent a large part of the total operating cost [Iyer and Grossmann, 1998]. Disturbances in the supply of utilities may also cause the production to slow down or stop, which implies great loss of revenue. Since utilities are often shared between production areas, the problem becomes even more intricate. At the occurrence of a disturbance, critical information about the disturbance is often unknown [Lee and Weekman, 1976]. The survey of utility disturbances that is presented in this paper aims to find common disturbance characteristics for each utility, which may give operators at a site some insight in what could be likely scenarios when a disturbance occurs. The objective is to present a general method for finding typical disturbances for each utility, given historical measurement data. The method is by no means the single approach to obtain these typical disturbance trajectories, and each of the steps of the method could be developed and augmented in order to ensure more accurate results for all cases. However, we do believe that there is a need for such a method, since little work seems to have been done on statistical modeling of utility disturbance characteristics within the chemical process industry. A related study for the pulp and paper industry has been performed in [Khanbaghi et al., 1997], where statistical analysis of paper break data from a typical paper mill is used to form a Markov chain model for the paper-break process. Modeling the process as a continuous Markov chain implies that the failure rate is constant, due to the memoryless character of the exponential distribution. In [Ogawa, 2003], the nature of the input disturbances to a broke storage tank at a pulp and paper production site are investigated, and it is found that break and normal durations can be modeled by exponential distributions. However, the exponential distribution with its constant failure rate does not seem to fit as well for utility disturbances. This is further discussed in Section 2.3 of this paper.

The method proposed in this paper is applied to industrial data from Perstorp, a world leading company within several sectors of the specialty chemicals market [Perstorp, 2013]. The disturbance characteristics for the utilities at Perstorp may be useful also for other companies, since the utilities used at the Perstorp site are very common also at other industrial sites. The estimated disturbance trajectories can be used as input to optimization problems for finding the optimal supply of a utility to different production areas at the occurrence of a disturbance. A method for formulating such optimization problems is presented in [Lindholm and Giselsson, 2013]. The results from the optimization may be used by process operators to improve disturbance management for utility disturbances.

The objective of the method presented in this paper is to find statistical models for common disturbances in utilities. As a consequence, unlikely but possibly serious disturbances are disregarded. For that topic, we refer to literature on risk management, e.g. [Greenberg, 1991].

2. A General Method for Finding Empirical Utility Disturbance Models

A utility disturbance is according to [Lindholm and Johnsson, 2013] defined as when the measurement of a utility parameter, such as temperature or pressure of the utility, goes outside the limits for normal operation[1]. This definition enables to detect disturbances directly from measurement data, which gives the possibility to estimate typical utility disturbance trajectories for each utility. These trajectories may be used to evaluate how such disturbances should be handled, e.g. by the approach introduced in [Lindholm and Giselsson, 2013]. A method to find these typical disturbance trajectories from measurement data is presented here. The method consists of three steps:

1. Prefilter the utility measurement data
2. Identify type(s) of disturbances for the utility
3. Find one or more typical disturbance(s) for the utility

Step 3 is divided into different cases based on which type(s) of disturbances that were identified in step 2.

2.1 Step 1: Prefilter the Utility Measurement Data

The utility measurement data could be measurements of e.g. pressure, flow, or temperature of the utility. Since the definition of utility disturbances in [Lindholm and Johnsson, 2013] is used, the detection of a disturbance is entirely defined by the disturbance limit that is set for the utility. Special care should be taken to two cases:

Very short disturbances To avoid the detection of many short disturbances when a measurement oscillates around the disturbance limit, the data may have to be prefiltered before continuing to the second and third step of the method. The situation is illustrated in Figure 1, where measurements of a utility parameter are given for approximately two hours of operation, during which one major utility disturbance occurred. As seen in the figure, many consecutive disturbance limit crossings give very short disturbance durations and time between failures. The disturbance limit is marked with a red dashed line in the figure. A simple filtering strategy is to simply merge

[1] Hereafter denoted *disturbance limits*.

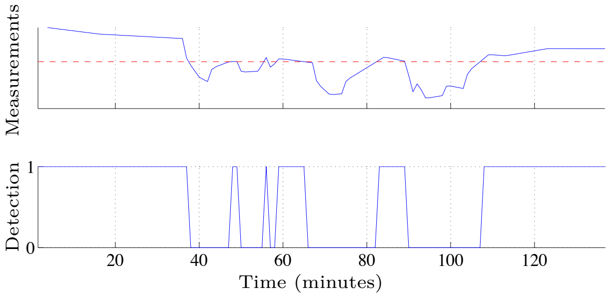

Figure 1. Measurement data and disturbance detection.

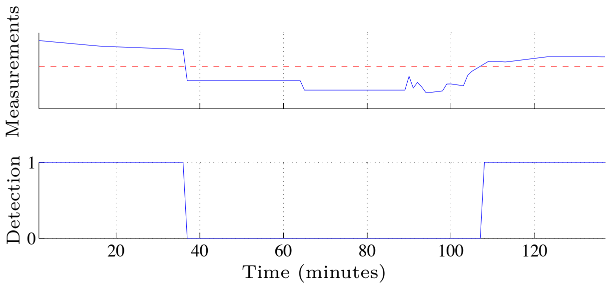

Figure 2. Measurement data and disturbance detection after filtering.

disturbances where the time between failures is shorter than some limit, e.g., 10 minutes. For the example in Figure 1, this gives the result shown in Figure 2. Another option is to use a *spike-pass filter*, which is a nonlinear filter using change detection to identify abrupt changes. Between abrupt changes, it is a simple low-pass filter. In this way, high frequency noise is reduced, while capturing sudden changes in the signal. The principle is described in [Gustafsson, 2000]. Spike-pass filtering gives a filtered signal that is similar to the resulting signal when the simple approach described above is used. Simple low-pass filtering can also be an alternative for the data prefiltering. The parameters of the filter should, independently of the choice of filter, be tuned in such a way that the result of the filtering agrees with the intuitive notion of what is noise and what is disturbances. There is obviously no general choice of filtering parameters that works for any type of data, from any plant.

Very long disturbances Another filtering of the data that may have to be performed before continuing to steps 2 and 3 is to remove disturbances of very long durations, since these disturbances typically correspond to unlikely situations. One way to identify if these disturbances are very unlikely to occur is to look at the percentage of the number of disturbances that are of shorter duration than a threshold. This is visualized in Figure 3. The suggestion

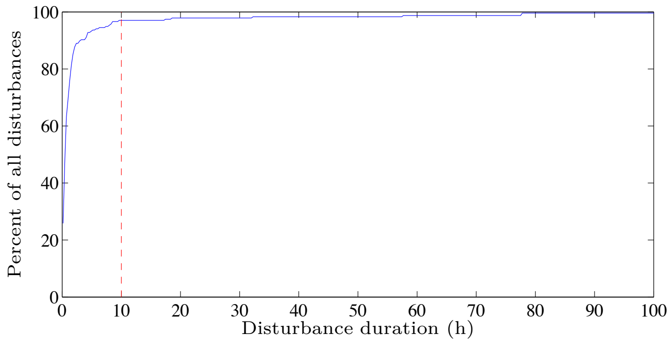

Figure 3. Percentage of disturbances of duration shorter than a threshold.

is to identify a threshold for the duration, for which most disturbances are of shorter durations. This may be done by studying figures like Figure 3 and see if a 'knee' can be identified. If no knee can be identified, include all disturbance data. For the example in Figure 3, one possible threshold seems to be at approximately 10 hours, which is marked by a red dashed line in the figure. The data used in the figure are real industrial data for the middle-pressure steam utility from a site operated by Perstorp. This data will be used as an example for nonperiodic disturbances throughout remainder of the section.

2.2 Step 2: Identify Type(s) of Disturbances for the Utility

To begin with, the characteristics of the disturbances have to be evaluated. Two main types of utility disturbances may be identified: *periodic* and *nonperiodic*.

Periodic disturbances Periodic disturbances can be characterized by a shape (e.g. a sinusoid or a square wave) and a period time/frequency. A sinusoidal disturbance may be described as $A\sin(\omega t)$, where A is the amplitude and ω the angular frequency. A utility that could typically suffer from periodic disturbances is the cooling water utility, because of daily and yearly variations of the outdoor temperature.

Nonperiodic disturbances The idea is to characterize nonperiodic disturbances by three parameters: duration, severity, and time between failures. The severity is evaluated as the depth of the drop or the height of the peak in the measurements, and the time between failures is defined as the duration from the end of one disturbance until the beginning of the next disturbance. The disturbance measures are visualized in Figure 4.

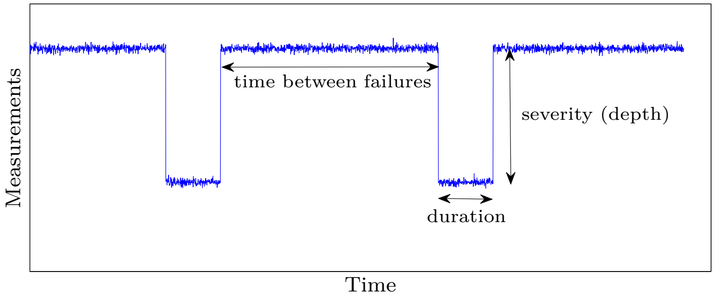

Figure 4. Disturbance measures for nonperiodic disturbances.

2.3 Step 3: Find One or More Typical Disturbance(s) for the Utility

This step is divided into two parts: One for periodic disturbances and one for nonperiodic disturbances. There could be several typical both periodic and nonperiodic disturbances for a utility. If the results at one of the substeps seems to be inconsistent, two or more typical disturbances can be identified, with different amplitude and/or period time for periodic disturbances, or with different disturbance duration and/or severity for nonperiodic disturbances.

Periodic disturbances The existence of periodic disturbances may be seen by studying the measurement data, to see if a shape could be identified. There could also be physical indications that there should exist periodic disturbances, such as daily or yearly variations of the outdoor temperature. When the shape has been identified, identify the amplitude and period time, and possibly other parameters that define the disturbance trajectory.

An example is taken from a site operated by Perstorp. Figure 5 shows the cooling water temperature together with the identified disturbance limit of 27°C, and the outdoor temperature during almost three years. The measurement data imply that there are both daily and yearly variations of the temperature, and the variations seem to be approximately sinusoidal-shaped. The temperature $T(t)$ may thus be described as

$$T(t) = T_{\text{norm}} + A_d \sin\left(\frac{2\pi}{24 \cdot 60} t\right) + A_y \sin\left(\frac{2\pi}{24 \cdot 60 \cdot 365} t\right) \tag{1}$$

where $T_{\text{norm}} = 24°\text{C}$ is the normal cooling water temperature, A_d the daily temperature variations, and A_y the yearly temperature variations. The daily and yearly variations could be studied separately to find the amplitudes A_d and A_y. Given the data in Figure 5, amplitudes of $A_d = 1.5°\text{C}$ and $A_y = 4°\text{C}$ seem appropriate. The resulting disturbance trajectories for the daily and yearly temperature variations are shown together with the measurement data in Figure 6.

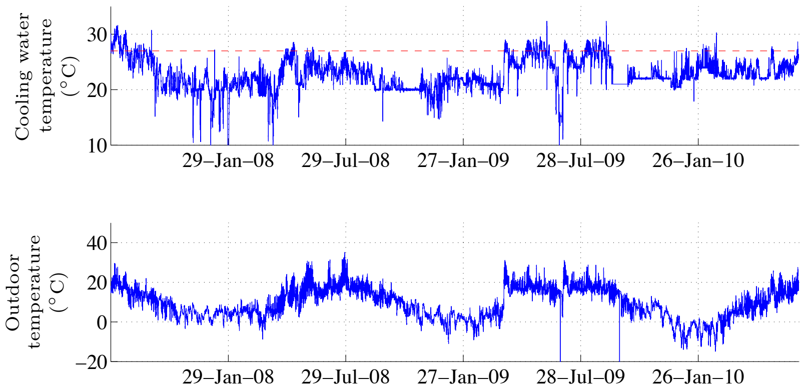

Figure 5. Temperature of cooling water and outdoor temperature at an industrial site.

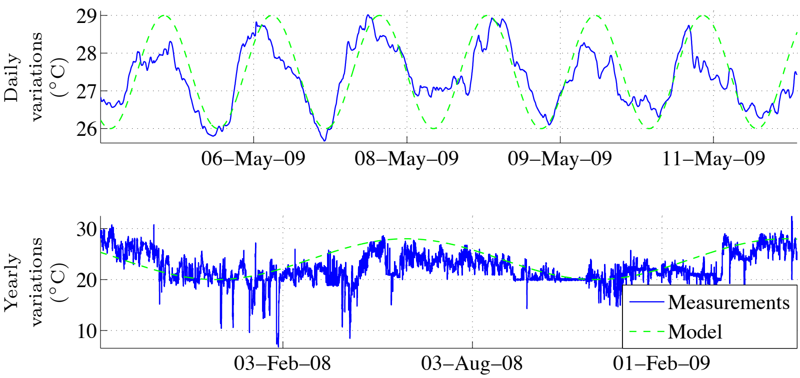

Figure 6. Model for daily and yearly variations of cooling water temperature.

Nonperiodic disturbances For nonperiodic disturbances, the procedure is to find the duration, severity, and time between failures for one or more typical utility disturbances. How to find these measures by forming statistical models is discussed below.

I. *Estimate most probable disturbance duration:* Some different probability distributions were evaluated as candidates for the distribution of the disturbance durations, among others the normal distribution, lognormal distribution, exponential distribution, and the Weibull distribution. Probability plots (see e.g. [Rice, 2007]) where used to assess the fit of data to the distributions. The lognormal distribution showed to give the best fit for the utility data from Perstorp, why this is our suggestion for the statistical model of the disturbance durations for utilities. The duration of a disturbance is related to the concept of *repair time* for applications with repairable units. Time to repair is according to e.g. [Hamada et al., 2008] and [Schroeder and Gibson, 2010] also well characterized by a lognormal distribution, which further motivates the choice of this distribution for the disturbance durations.

The suggestion is thus to fit a lognormal distribution to the disturbance duration data. The lognormal distribution has two parameters, μ and σ and the probability density function (pdf) is given by

$$f(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} x^{-1} e^{-(\ln x-\mu)^2/(2\sigma^2)} \tag{2}$$

where $0 < x < \infty$. Further information about the lognormal distribution and its applications is given e.g. in [Crow and Shimizu, 1988].

In the upper subfigure of Figure 7, a histogram of the disturbance durations for the middle-pressure steam utility at the Perstorp site is given together with the lognormal probability density function that gives the best fit (red line). The median disturbance duration is marked with a green dashed line in the figure, and the average disturbance duration with a blue dash-dotted line. As seen in the figure, the average duration is significantly longer than the median, because of some long and uncommon disturbances. The parameters of the distribution that gives the best fit are $\mu = 3.22$ and $\sigma = 1.44$. The median disturbance duration is $e^{\mu} = 25$ minutes and the average duration $e^{\mu+\sigma^2/2} = 70$ minutes. The cumulative probability density

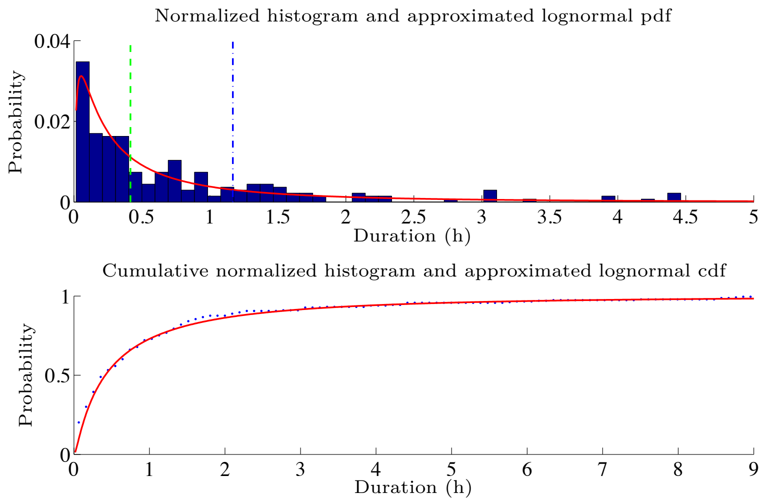

Figure 7. Statistical model for disturbance durations for middle-pressure steam.

function (cdf, red line) is plotted in the bottom subfigure of Figure 7 together with the measured disturbance durations (blue dots). According to this plot, the lognormal distribution seems to be a good fit for the disturbance durations.

II. *Estimate most probable severity of disturbance:* The maximum deviation in the measurements from the normal operating point is correlated to the duration, but the correlation is not very strong. In fact, the probability distribution of the deviation is not at all lognormal in the cases studied here. It does not fit any of the "standard" distributions (normal, lognormal, exponential, Beta, Weibull, or Gamma) to a statistically significant level. The suggestion is therefore to use a normalized histogram with an appropriate number of bins to produce an empirical probability density function for the severity of the disturbance. A normalized histogram of the maximum deviation from the nominal pressure at disturbances in the middle-pressure steam utility at the Perstorp site is shown in Figure 8. In the histogram, the median and average deviation are marked with a green dashed line and a blue dash-dotted line, respectively. For the example, the median is approximately 5 bar, whereas the average is around 6 bar.

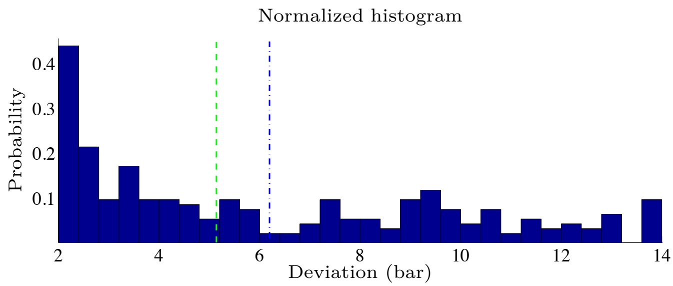

Figure 8. Normalized histogram of maximum deviation from nominal pressure for middle-pressure steam.

III. *Estimate most probable time between failures:* The Weibull distribution is often used to model the time to failure or life span of a component [Pal et al., 2005]. Some application areas include computing systems, for which among others [Zhang et al., 2005] and [Schroeder and Gibson, 2010] model the time between failures using the Weibull distribution. By evaluating the probability plots for some common distributions (as in step I), it was found that the Weibull distribution also seems to fit the time between utility failures well, based on the data from Perstorp. The distribution has two parameters, one scale parameter, a, and one shape parameter, b. The probability density function is given by

$$f(x|a,b) = \frac{b}{a^b} x^{b-1} e^{-(x/a)^b} \tag{3}$$

where $0 < x < \infty$, and $a, b > 0$. More information about the Weibull distribution can be found in, e.g., [Pal et al., 2005].

In the upper subfigure of Figure 9, the Weibull probability density function (red line) that gives the best fit for the middle-pressure steam utility at the Perstorp site is plotted together with a normalized histogram of the time between failures for the utility. The median and average time between failures are marked by a green dashed line and a blue dash-dotted line, respectively. The cumulative distribution function (red line) is plotted together with the measurements (blue dots) in the lower subfigure of Figure 9. These plots show that the Weibull distribution seems to be a good fit for the example data. The parameters of the Weibull distribution that gives the best fit are $a = 2500$ and $b = 0.46$, for disturbance data in minutes. The median time between failures for the distribution is $a(\ln 2)^{1/b} = 19$ hours, whereas the mean time between failures is $a\Gamma(1 + 1/b) = 4$ days for this data set, where Γ is the gamma function. Mean time between failures (MTBF) is a measure that is often used for different kinds of repairable systems.

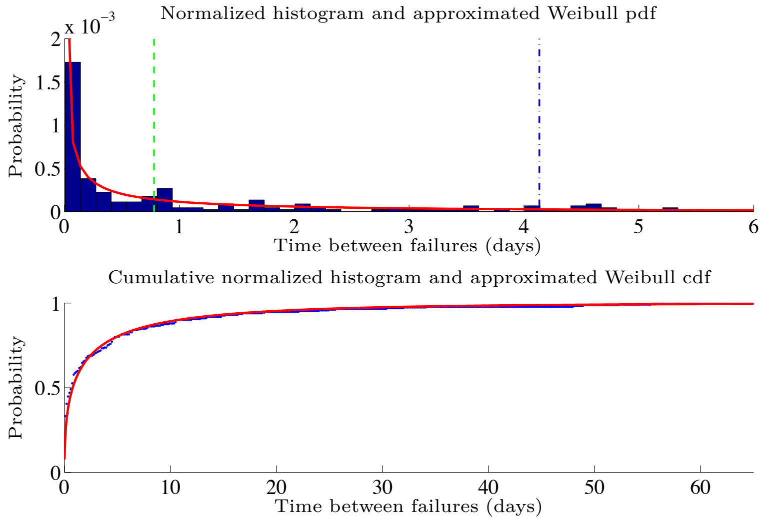

Figure 9. Statistical model for time between failures for middle-pressure steam.

2.4 Results

The typical disturbance trajectories are obtained after performing the three steps described in this section. For nonperiodic disturbances, distributions for the duration, severity, and time between failures of the disturbance were identified in substeps I-III of step 3 of the method. A typical disturbance trajectory is obtained as a realization using the statistical models of these measures. An alternative is to use the median or average for the distributions to generate a typical disturbance trajectory.

3. A Case Study

In this section, the general method described in Section 2 is used to identify typical disturbances for selected utilities at a site operated by Perstorp. The utilities that have been studied are

- Middle-pressure (MP) steam
- High-pressure (HP) steam
- Cooling water
- Nitrogen
- Combustion device

Nitrogen is used at the site to maintain pressure in vessels and the combustion device for combustion of tail gas. Other utilities are also used at the site, but for these utilities, the disturbances have been so few during the considered time period that the data set is not sufficiently large to produce reliable disturbance models. The data that are available span approximately three years of operation.

3.1 Step 1: Prefilter the Utility Measurement Data

The approach that was selected for prefiltering of the data was to merge all disturbances where the time between failures were shorter than 10 minutes. To remove very unlikely disturbances of long durations from the data sets, Figure 10 was studied. The figure shows the percentage of all disturbances for the utilities that are shorter than a threshold. MP steam was omitted in the figure, since this utility has already been treated in Figure 7. The red dashed lines in Figure 10 mark the thresholds that where set for the considered utilities: approximately 10 hours, 5 hours, 100 hours, 2 hours, and 21 hours for MP steam, HP steam, cooling water, nitrogen, and the combustion device, respectively.

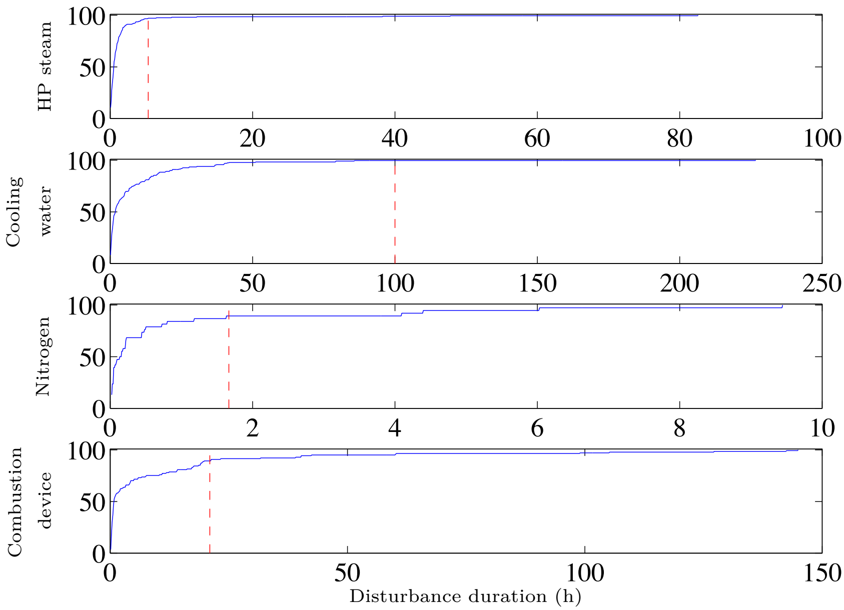

Figure 10. Percentage of disturbances of duration shorter than a threshold.

3.2 Step 2: Identify Type(s) of Disturbances for the Utilities

In this step, it was concluded that the cooling water utility was the only utility that indicated a periodical behavior.

3.3 Step 3: Find One or More Typical Disturbance(s) for the Utilities

One typical disturbance for each utility was identified; a periodic disturbance for cooling water, and nonperiodic disturbances for the MP steam, HP steam, nitrogen, and combustion device utilities. The procedure is described in more detail below. Here, MP steam and cooling water have been left out, since these utilities were used as examples in Section 2, and the details can be found there.

Periodic disturbances The identification of the characteristics of the periodic cooling water disturbance was handled as an example in Section 2, with the resulting disturbance trajectory structure (1) for the temperature.

Nonperiodic disturbances The three steps to identify typical nonperiodic utility disturbances are here carried out for high-pressure steam, nitrogen, and the combustion device simultaneously.

I. *Estimate most probable disturbance duration:* The estimated lognormal probability distributions for the disturbance durations are shown together with the normalized histograms in Figure 11 for the three utilities. The median disturbance durations are marked with green dashed lines and the average durations with blue dash-dotted lines. The parameters μ and σ of the best fit lognormal distributions are given in Table 1, for disturbance data in minutes.

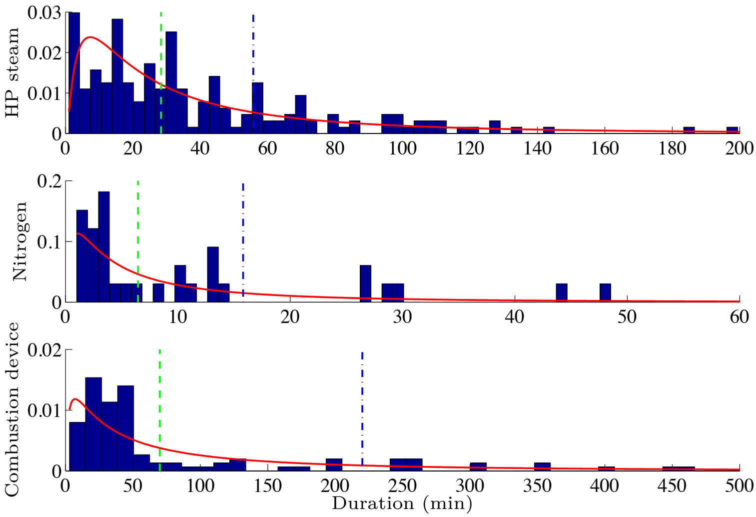

Figure 11. Statistical models for disturbance durations.

Table 1. Parameters of lognormal distribution for disturbance duration.

Utility	μ	σ
HP steam	3.34	1.16
Nitrogen	1.87	1.34
Combustion device	4.25	1.51

II. *Estimate most probable severity of disturbance:* Normalized histograms of the maximum deviation from the normal operating point at disturbances in the three utilities are shown in Figure 12. For HP steam and nitrogen, the measurement unit is pressure in bar, and for the combustion device it is flow in kg/h. For all three utilities, a disturbance is a drop in the pressure or flow. The median depths of the drops are marked with dashed green lines and the average with dotted blue lines.

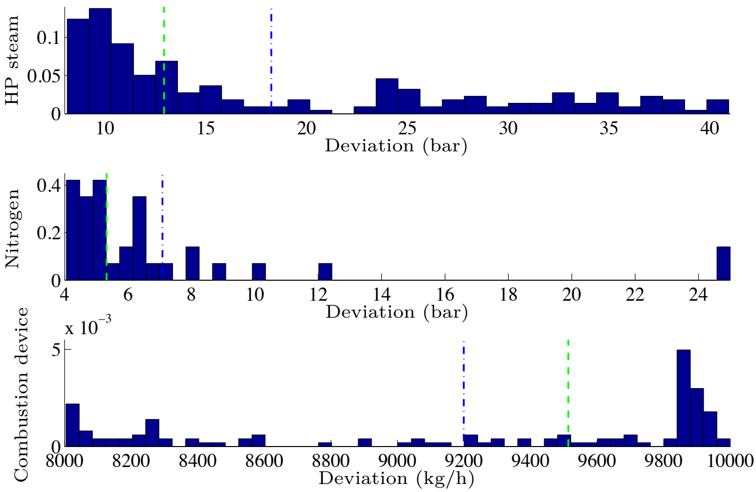

Figure 12. Normalized histograms of maximum deviation from normal operating point.

III. *Estimate most probable time between failures:* Figure 13 shows histograms of the time between failures for the three utilities together with the suggested Weibull distributions. The median time between failures is marked with dashed green lines and the average with blue dotted lines. The scale and shape parameters a and b of the best fit Weibull distributions are given in Table 2, for disturbance data in minutes.

Table 2. Parameters of Weibull distribution for time between failures.

Utility	a	b
HP steam	2676	0.45
Nitrogen	3875	0.32
Combustion device	2630	0.40

3.4 Results

The typical periodic disturbance for the cooling water utility is given by the temperature trajectory in °C:

$$T(t) = 24 + 1.5 \sin\left(\frac{2\pi}{24 \cdot 60} t\right) + 4 \sin\left(\frac{2\pi}{24 \cdot 60 \cdot 365} t\right) \tag{4}$$

The other utilities suffer from nonperiodic disturbances. The distributions of the duration, severity, and time between failures for the nonperiodic dis-

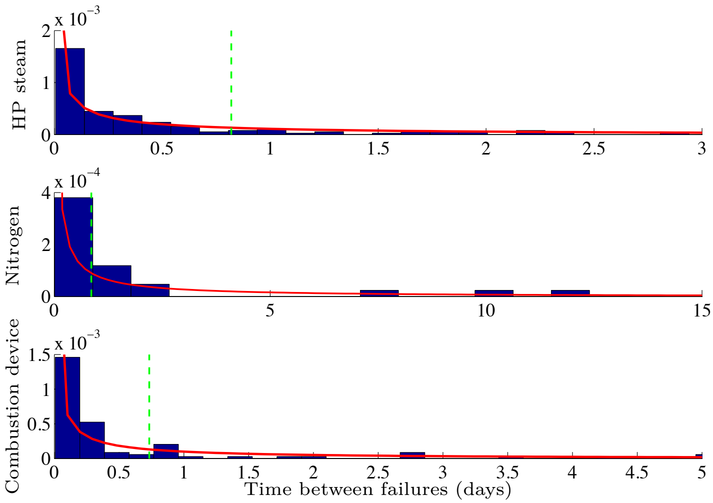

Figure 13. Statistical models for time between failures.

turbances were estimated in step 2 of the method. To get a feeling for what the typical disturbance looks like, the median of the disturbance duration (l), the median of the severity parameter (d), and the mean time between failures (tbf) are presented in Table 3.

Table 3. Typical utility disturbances at the Perstorp site.

Utility	l	d	tbf
MP steam	25 min	5 bar	4 days
HP steam	28 min	13 bar	5 days
Nitrogen	6 min	5 bar	19 days
Combustion device	70 min	9513 kg/h	6 days

3.5 Discussion

For the periodic cooling water disturbance, the suggested sinusoidal disturbance seems to capture the disturbance characteristics quite well, even though it can be seen in Figure 5 that the periodicity of the measurements of the outdoor temperature is more distinct than those of the cooling water temperature. This might be due to the feedback effect: if the plant is (for other reasons than cooling water problems) run at reduced speed, less cooling effect is needed and the cooling water temperature will be lower. This might explain the drops in the measurements of the cooling water temperature that can be seen in the figure.

For the nonperiodic disturbances, most figures seem to show similar results for the four utilities, with some exceptions. In step I, the probability distributions in Figure 11 seem to fit better for utilities where the data contain a larger number of failures, in this case for the steam utilities. This is also seen in step III of the distributions for the time between failures, but not as clearly. The weakest result is the estimation of the severity of the disturbances in step II. The histograms do not clearly show a peak or a tendency of the data to fit any of the standard distributions. For the nitrogen and combustion device utilities, it seems like there might be two types of disturbances; one smaller disturbance, and one with a greater deviation from the normal operating point. An attempt to improve the modeling could be to divide these into two separate cases and redo substeps I-III for each of the cases.

4. Conclusions

A general method for finding typical utility disturbance trajectories were presented and used to analyze disturbances in the supply of utilities at a site at Perstorp. The disturbances were categorized into periodic and nonperiodic disturbances, and characterized by simple measures: period time and amplitude for periodic disturbances, and duration, severity and time between failures for nonperiodic disturbances. The duration of a disturbance was modeled using the lognormal distribution, and the time between failures using the Weibull distribution. The resulting disturbance trajectories may be useful for both proactive and reactive disturbance management.

Acknowledgments

The research is performed within the Process Industry Centre (PIC) supported by the Swedish Foundation for Strategic Research (SSF). The first author is a member of the LCCC Linnaeus Center and the eLLIIT Excellence Center at Lund University.

References

Crow, E. L. and K. Shimizu (1988). *Lognormal Distributions: Theory and Applications*. Marcel Dekker, USA.

Greenberg, H. R. (1991). *Risk Assessment and Risk Management for the Chemical Process Industry*. John Wiley & Sons, Inc., USA.

Gustafsson, F. (2000). *Adaptive filtering and Change Detection*. Wiley, UK.

Hamada, M. S., A. Wilson, C. S. Reese, and H. Martz (2008). *Bayesian reliability*. Springer, USA.

Iyer, R. R. and I. E. Grossmann (1998). "Synthesis and operational planning of utility systems for multiperiod operation". *Computers & Chemical Engineering* **22**:7, pp. 979–993.

Khanbaghi, M., R. Malhame, M. Perrier, and A. Roche (1997). "A statistical model of paper breaks in an integrated tmp-newsprint mill". *Journal of pulp and paper science* **23**:6, J282–J288.

Lee, W. and V. W. Weekman (1976). "Advanced control practice in the chemical process industry: A view from industry". *AIChE Journal* **22**:1, pp. 27–38.

Lindholm, A. and P. Giselsson (2013). "Minimization of economical losses due to utility disturbances in the process industry". *Journal of Process Control* **23**:5, pp. 767–777.

Lindholm, A. and C. Johnsson (2013). "Plant-wide utility disturbance management in the process industry". *Computers & Chemical Engineering* **49**, pp. 146–157.

Ogawa, S. (2003). *Stochastic Averaging Level Control and Its Application to Broke Management in Paper Machines*. PhD thesis. The University of British Colombia, Canada.

Pal, N., C. Jin, and W. K. Lim (2005). *Handbook of Exponential and Related Distributions for Engineers and Scientists*. Chapman and Hall/CRC, UK.

Perstorp (2013). [online] `http://www.perstorp.com/`, June 2013.

Rice, J. A. (2007). *Mathematical statistics and data analysis*. Thomson, USA.

Schroeder, B. and G. A. Gibson (2010). "A large-scale study of failures in high-performance computing systems". *IEEE Transactions on Dependable and Secure Computing* **7**:4, pp. 337–350.

Zhang, Y., M. Squillante, A. Sivasubramaniam, and R. Sahoo (2005). "Performance implications of failures in large-scale cluster scheduling". In: *Job Scheduling Strategies for Parallel Processing*. Ed. by D. Feitelson, L. Rudolph, and U. Schwiegelshohn. Vol. 3277. Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 233–252.

Paper IV

Hierarchical Scheduling and Disturbance Management in the Process Industry

Anna Lindholm, Nils-Petter Nytzén

Abstract

The integration of scheduling and control in the process industry is a topic that has been frequently discussed during the recent years, but many challenges remain in order to obtain integrated solutions that may be implemented at large-scale industrial sites. This paper introduces a general framework for production scheduling (PS) and detailed production scheduling (DPS) using a two-level hierarchical approach. The PS activity generates a monthly production schedule based on information on orders and forecasts, and the DPS activity handles disturbances in production on an hourly basis. The focus is on disturbances in the supply of utilities, which often cause great losses at process industrial sites. The research has been conducted in close collaboration with Perstorp, a world-leading company within several sectors of the specialty chemicals market. A specification list provided by Perstorp has been used as a starting point for formulating the PS and DPS activities as optimization problems. An example that is inspired by a real industrial site is presented to show how the PS and DPS may operate and how the integration of these two functions behaves.

Submitted to *Computers and Chemical Engineering.*

1. Introduction

The chemical industry has during the past decades become a global marketplace with strong competition between manufacturers [Tousain, 2002], which motivates the need for optimizing the operational efficiency. Planning, scheduling, and control are key features that have large economic impact on process industry operations [Shobrys and White, 2002]. These areas are often not easily distinguishable, and border lines between the areas are often diffuse [Kallrath, 2002]. Common definitions are that planning is the activity to make production, distribution, and inventory plans, and scheduling to decide the timing of actions to execute the plan and make use of the available resources [Kallrath, 2002; Rawlings and Amrit, 2009; Huang, 2010; Engell and Harjunkoski, 2012]. The timescales in which the activities operate also vary. Usually, planning is said to work on a time scale of one or more months, and scheduling on a horizon of weeks. For control in the process industry, it is much harder to find a general definition both of the timescale and the activities to be performed, because of the many different interpretations of process control. In this paper, two activities are handled, which are denoted *production scheduling* (PS) and *detailed production scheduling* (DPS), in line with the definition in [ISA-95.00.03, 2009]. The PS operates on a horizon of one month, and the DPS on a horizon of one day. The activity of production scheduling is sometimes denoted *scheduling*, and detailed production scheduling is denoted *advanced control* in other papers, e.g. in [Shobrys and White, 2002] and [Engell and Harjunkoski, 2012].

The vague definitions of the activities to be performed at the planning, scheduling, and control level also makes it more difficult to define what is meant by the integration of these areas. Some work has been done on integrating planning and scheduling, either by combining them and solving the planning and scheduling problem simultaneously, or by various decomposition techniques. An extensive review is provided in [Grossmann and Furman, 2009]. The topic of integrating planning and scheduling with control, on the other hand, is a topic that still has not received much attention in the literature [Craig et al., 2011; Grossmann, 2012]. [Shobrys and White, 2002] and [Engell and Harjunkoski, 2012] provide a good view of the activities that have to be integrated and describe the current practice and challenges for integrating the planning, scheduling and control functions in the process industry. A lot of case-specific contributions regarding integration of scheduling and control have also been made, of which [Harjunkoski et al., 2009] provide an excellent overview. In [Tousain, 2002], a hierarchical approach for integrating scheduling with control is presented, but only a single plant/area is studied, and the focus is on multi-grade plants. In the current study, a hierarchical approach for integrating the PS and DPS activities is suggested. The focus is on one process industrial site with several connected production areas.

The focus in this paper is on production scheduling for chemical process industries with continuous production. Several models for scheduling for chemical sites have been proposed previously, but the majority of these, e.g. the models suggested in [Kondili et al., 1993], [Neumann et al., 2002], and [Maravelias and Grossmann, 2003], handle the scheduling of batch processes. The state-task network (STN) introduced in [Kondili et al., 1993] is also used by, among others, [Ierapetritou and Floudas, 1998] and [Shaik et al., 2009] to formulate production scheduling models for continuous production sites. However, these studies focus on the unit level of the equipment hierarchy rather than the area/site level that is relevant for the current study. General frameworks for chemical production scheduling are suggested by [Sundaramoorthy and Maravelias, 2011] and [Maravelias, 2012], but these frameworks focus on batch processes and are not as intuitive for sites with continuous production.

Perstorp is a world-leading company within several sectors of the specialty chemical market. The company has ten production sites around the world, where each production site is divided into about 5-10 production areas. The production sites typically run in a continuous mode, without any product changes or grade changes. The aim of Perstorp is to run its production sites in a well-defined way even when there are site-wide disturbances such as disruptions in a utility or raw material. In order to do so, decision makers at Perstorp have generated a specification list containing demands and desires for the production scheduling. This list is used as a starting point for finding models for the PS and DPS that are generic enough to be applied to all its production sites. The specifications are listed in Section 4, and formulated as optimization problems is Sections 5 and 6.

2. Hierarchy Models

To clarify at which levels of the physical and functional hierarchy of an enterprise the current study is focused, the role-based equipment hierarchy, functional hierarchy and scheduling hierarchy are defined in this section.

2.1 Role-based Equipment Hierarchy

According to the standard [ISA-95.00.01, 2009], there are five levels of the role-based equipment hierarchy of an enterprise with continuous production; the enterprise, site, area, production unit, and unit levels. Traditionally, the area of process control is focused on control of production units, e.g. reactors or distillation columns, or on control of some connected production units. This would correspond to the production unit level or area level of the equipment hierarchy. In this study, the focus is on the area and site levels of the hierarchy; on control of the production in the different areas of a site.

The areas at a process industrial site are often connected, such that one area produces raw materials for other areas. This is in [Wassick, 2009] denoted an *integrated site*, and in process flow scheduling (PFS) a *process train*. Changing the production rate in one area, e.g. due to a disturbance, may thus affect the production in several other areas at the site. An example of an integrated site with six production areas and three buffer tanks is given in Figure 1.

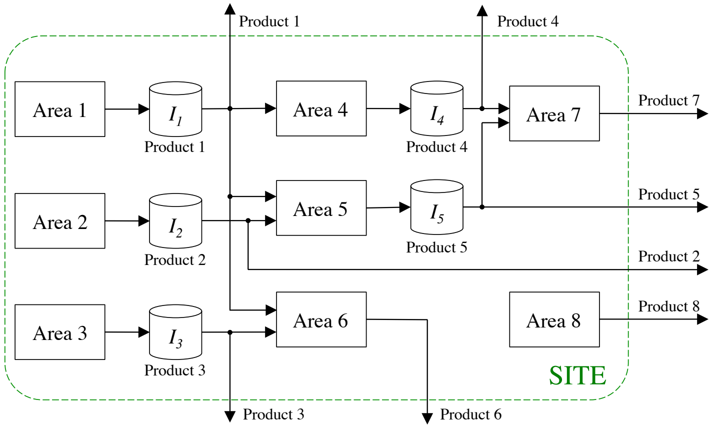

Figure 1. Example of an integrated site.

If modeling at the site/area level in the hierarchy should be performed, and the area dynamics are fast compared to the dynamics of the production network, the production in an area can be assumed to be directly proportional to the inflows to the area (i.e., the dynamics within the area are ignored). This assumption is also made in [Lindholm and Giselsson, 2013]. The assumption can be expressed as

$$q_{ijt}^{\text{in}} = q_{jt} a_{ij} \tag{1}$$

where q_{ijt}^{in} is the inflow of product i to area j at time t, q_{jt} the production in area j at time t, and a_{ij} is called the conversion factor between product i and product j. In Figure 5 in Section 5.2, the notation is shown in a flowchart of an example site.

2.2 Functional Hierarchy

The functions that are used for operating an enterprise are often viewed in a hierarchical structure. In papers that discuss the integration of different functions, such as production planning, scheduling, and control, 'integration pyramids' like the one in Figure 2 (left) are commonly used. These pyramids might look quite different, which is no surprise since the people working in the field of process control come from many different areas [Tousain, 2002].

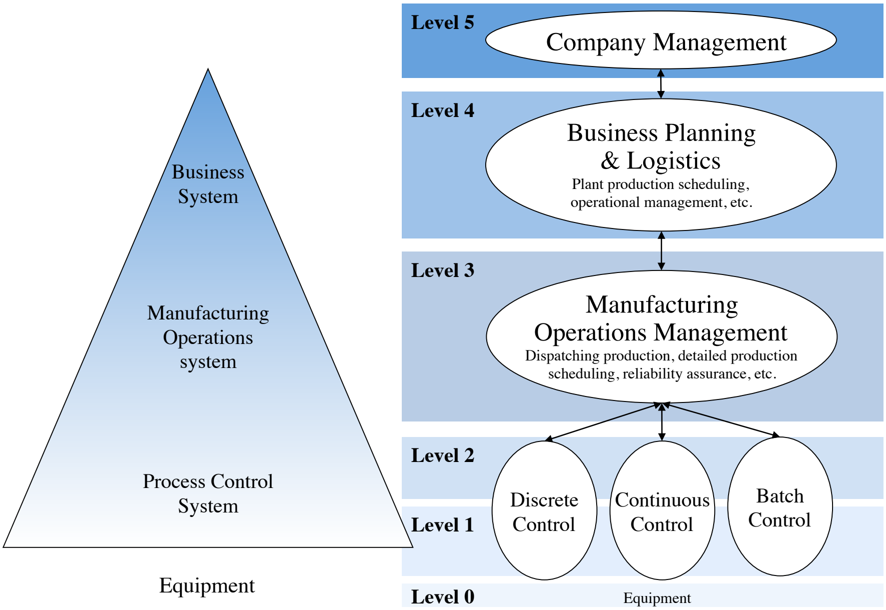

Figure 2. Functional hierarchy of an enterprise.

In this paper, we stick to the definition in the standard [ISA-95.00.01, 2009], as presented in Figure 2 (right). The levels represent activities at various timescales, where level 1-2 include activities with the shortest timescales such as sensing and reading (milliseconds, seconds, minutes), level 3 includes activities with a longer timescale such as scheduling (minutes, hours, weeks) and level 4 includes activities with an even longer timescale such as planning (days, weeks, months). The current study is focused on level 3 and 4 of the functional hierarchy.

2.3 Scheduling Hierarchy

The scheduling hierarchy that is presented in the appendix of standard [ISA-95.00.03, 2009], is depicted in Figure 3. This hierarchy is derived from common terms used in APICS dictionary [Blackstone and Cox, 2004] and [ISA-95.00.03, 2009]. The figure could be extended with a control layer at the very bottom, visualizing the fact that control and scheduling are tightly coupled. In this paper, we have adopted a hierarchical scheduling approach in which the short-term scheduling (referred to as 'detailed production scheduling') takes care of the fifth level in the scheduling hierarchy, whereas the long-term scheduling (referred to as 'production scheduling') takes care of the fourth level.

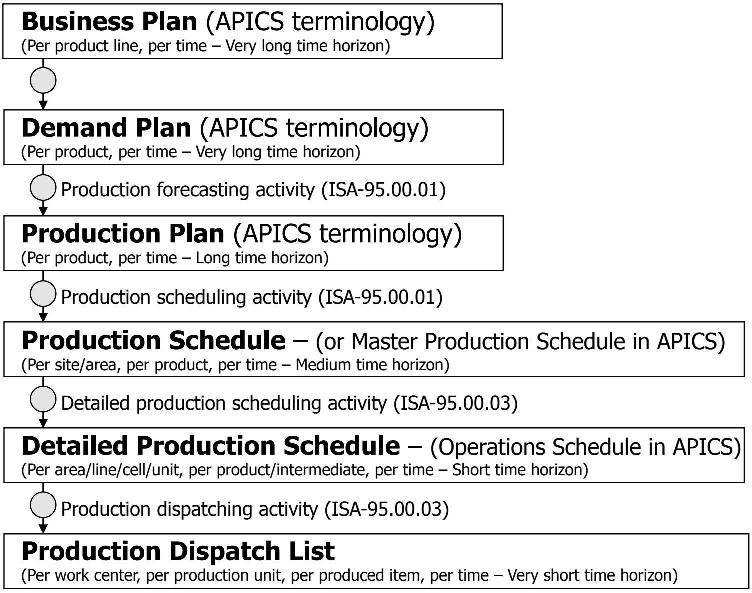

Figure 3. Scheduling hierarchy of an enterprise.

3. Suggested Scheduling Hierarchy

The purpose of the PS activity is to produce a production schedule. The production schedule suggests how much to produce of each product at each area at a specific site, and how the inventories of the products at the site should be used. The purpose when making the production schedule is to minimize the backlog of orders while considering production and inventory limitations. The suggestion is that the PS activity produces a production schedule one month ahead and updates the plan every day. This time step and horizon has been suggested after discussions with Perstorp. The information (input) needed for making the production schedule can for example be different kinds of capacity, levels of storage for different products, incoming orders, planned maintenance and transports.

The DPS activity should determine how the production should be controlled to handle disturbances on a time-scale of hours in an economically optimal way. The detailed production schedule consists of trajectories that suggest how much to produce and sell at each hour during the day, and how the buffer tanks at the site should be utilized. The suggestion is that this activity operates on a horizon of one day, and updates the schedule every hour. The inputs to the DPS are reference levels for the production, sales, and inventories of all products. Predicted disturbance trajectories are also needed.

If the DPS is performed in receding horizon as suggested, the disturbance predictions may be updated every hour.

The ideal solution for integrating PS with DPS would be a monolithic solution, where the characteristics of both problems are fully represented [Engell and Harjunkoski, 2012]. However, this will result in extremely challenging mixed-integer optimization problems that have to be solved online [Engell and Harjunkoski, 2012]. Therefore, the current apprehension among most authors in the field, among others [Tousain, 2002], [Kadam and Marquardt, 2007], and [Engell and Harjunkoski, 2012], seems to be that a hierarchical approach is currently the only realistic one to tackle industrial-size problems, even if this may lead to suboptimal solutions. In addition, the two activities are often handled at different levels of the functional hierarchy of a site. To make the scheduling solutions transparent and understandable for all users, a hierarchical approach is clearly advantageous. One of the greatest challenges today is getting people to accept and efficiently use new tools for performing the scheduling [Shobrys and White, 2002]. If the solutions are not accepted by the operators, they will not be used in the long term, even if benefits could be shown [Engell and Harjunkoski, 2012].

In this paper, a two-level approach for integrating PS and DPS is suggested, as shown in Figure 4. This structure was first introduced in [Lindholm

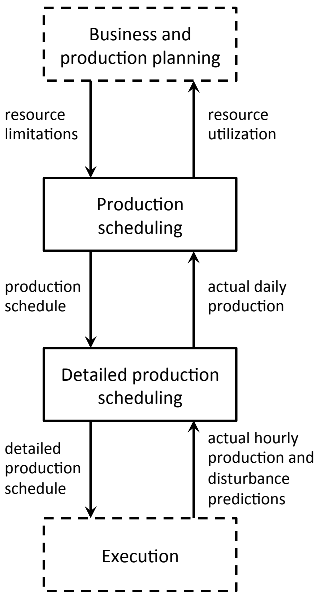

Figure 4. Suggested structure for scheduling and utility disturbance management.

et al., 2013a], and is based on the scheduling hierarchy presented in [ISA-95.00.03, 2009]. In the PS layer, a production schedule is set, which serves as a reference for the DPS. To produce the schedule, information about the current resource limitations at the site are acquired from the upper level in the hierarchy, 'Business and production planning'. The DPS activity gets its reference values for the production, sales, and inventory levels from the production schedule at the beginning of each day. This activity has an interface to the actual site, where the schedule is executed. This could be done e.g. by the operators at the site or using model-predictive control (MPC). Measurements from the site report to the DPS layer how the production was actually conducted each hour. The site also provides predicted disturbance trajectories that are used to produce the detailed production schedule. The information on the actual production and inventory usage is aggregated and reported to the PS activity at the end of each day. The resource utilization given the production schedule is reported back to the business and production planning level. The scheduling procedure is summarized below.

ALGORITHM 1
Scheduling procedure

For every day of the month

a) Get measurements of inventory levels, and yesterday's production and sales.
b) Get information on orders and forecasted orders.
c) Perform PS.
d) Report reference values to DPS.

   For every hour of the day

   i. Get measurements of inventory levels, and production and sales during the last hour.
   ii. Get predicted disturbance trajectories.
   iii. Perform DPS.

   End

End □

Our suggestion is to run both the PS and DPS in receding horizon, and produce a new schedule in every time step, but an alternative would be to redo the schedule only when needed, as in [Kadam and Marquardt, 2007]. An advantage of the suggested scheduling procedure compared to current practice at most process industrial companies is that the scheduling problems are considered dynamic problems, and not static open-loop problems. Thus, the schedules are revised as new information becomes available or production disturbances occur.

4. Scheduling Specifications

Together with Perstorp, a list of specifications for the PS and DPS activities was produced. The specfications that relate to the PS layer of the hierarchy in Figure 4 were first presented in [Lindholm et al., 2013b]. Here, the activities related to the DPS layer have been added. Of course, the list can be made much longer if more details and special cases are taken into account. To make the list more general, only the specifications that were thought to be most important are included. All of the specifications given in this section are then handled in the formulation of the optimization problems for the PS and DPS in Sections 5 and 6.

4.1 Interconnection of Production Areas

At integrated sites, as the one shown in Figure 1, production areas are connected via the flow of products at the site. Both the PS and DPS need to take this into account when producing the schedules.

4.2 Orders and Forecasts

Information on which ordered volumes that exist, and when they should be delivered, should be provided as an input for the PS. Both actual orders and forecasts should be handled. The actual orders are the most important to fulfill, which has to be accounted for in the scheduling. For the first days of the scheduling period, it is reasonable to assume that most of the orders are final, whereas for the last days, fewer orders have probably been placed and forecasted orders dominate. The PS should be performed such that the backlog of orders is as small as possible. The DPS does not need to get information on orders and forecasts, since it gets its reference values for the production and sales from the PS layer.

4.3 Production Rate Limitations

Production areas have a maximum capacity, which should be incorporated into the scheduling. Most areas also have a minimum rate at which they can operate, unless they are completely shut down. The minimum production rate limits exist due to physical limitations, e.g. a distillation column has a minimum rate at which it can operate. The maximum and minimum hourly production is considered for the DPS, whereas the aggregated maximum and minimum daily production is used for the PS.

4.4 Inventory Limitations

Most inventories have a maximum capacity. At some sites, this is not a strong limitation, since extra inventory capacity may be achieved by e.g. renting storage space in temporary warehouses. However, at chemical plants the inventories are often (liquid) buffer tanks, which means that the maximum and minimum limits are hard constraints. This has to be considered both by the PS and DPS.

In some cases, it could be desirable to keep the inventory levels at the site at certain reference levels, or to minimize inventories to reduce inventory costs. However, in many cases it might be a better approach to consider a reference interval, where it does not matter what the inventory level is, as long as it is kept between some minimum and maximum levels. These limits have to be within the hard constraints on the maximum and minimum inventory levels discussed previously. The reference interval should be considered by the PS. The DPS gets reference levels for the inventories from the production scheduler, and these levels should be kept if possible. However, if there are disturbances in production, the DPS should be able to utilize the buffer tanks to handle the disturbances.

4.5 Start-up Costs and Start-up Times

Shutting down and starting up areas is often very expensive and time-consuming, which should be taken into account when performing both the PS and DPS. The start-up costs originate from the cost of utilities and raw materials that are consumed during the start-up phase, e.g. for heating up reactors. The start-up time is the time it takes for the area to return to normal production rate after a shutdown. It is important to take into account that some areas take several days to start up, whereas others may be started after only a few hours. Thus, the PS needs to take the start-up times into account when performing the monthly production planning. For the DPS, it may be enough to consider the cost of starting up the area, and the cost of having an area shut down.

4.6 Market Conditions

The PS has to consider the profitability of the forecasted orders, in addition to the orders that have to be fulfilled. The PS should aim to maximize the total profit. When disturbances in production are present, the DPS should aim at minimizing the losses, or maximizing the contribution, when suggesting how to handle the disturbances.

4.7 Costs for Late Delivery

Late delivery of orders may lead to penalty costs. The aim of the PS should be to minimize the backlog of orders to avoid these costs.

4.8 Cost of Production Rate Changes

Changing the production rate of an area quickly may be hard, and a stable production rate is also often more economically profitable because of lower average utility and raw material consumption. Both the PS and DPS should take this into account, and penalize large production rate changes from one day to the next, and from one hour to the next, respectively.

4.9 Limited Availability of Utilities

The production at a site may require use of one or more utilities. The utilities might have continuous characteristics, such as steam, cooling water, or be of on/off type, such as electricity and nitrogen. If a utility is of continuous type, the amount of a utility that is assigned to an area will affect the maximum possible production rate of the area. The utilities might be shared between production areas. This means that at a disturbance in the supply of a utility, all areas may not be able to operate at full production speed. Utility disturbances are often in a time scale of minutes/hours, which means that they should be handled by the DPS.

5. Mathematical Formulation of the Production Scheduling

In this section, the specifications in Section 4 are formulated as an optimization problem to produce a production schedule. The optimization problem is meant to be solved in receding horizon fashion, such that the PS problem for one month ahead is solved every day. The horizon for the PS is denoted N_{PS}, and is thus equal to 30 days (one month).

The model is formulated for integrated sites, and it is assumed that each area produces one product, which is stored in one buffer tank. The set of production areas at the site is denoted $\mathcal{A}$.

5.1 Variables and Parameters

The variables that are used to formulate the PS problem are summarized in Table 1. All variables are continuous except for the binary variables $w_{i\tau}$, that represent the operational mode of the areas, and $s_{i\tau}$, that indicate if start-ups of areas have been performed.

The parameters that are used to formulate the problem are summarized in Table 2. The values of these parameters may be obtained from personnel at the site, and are constant for each optimization problem in the receding horizon formulation. In addition to these parameters, five penalty parameters; α_i, β_i, γ_i, δ_i and ϵ_i, are used in the objective function of the optimization problem. These parameters should be tuned to get the desired behavior of the PS.

Table 1. Variables in the PS model.

Variable	Description
$I_{i\tau}$	Inventory level of tank i at the end of day τ
$q_{i\tau}$	Production in area i during day τ
$q_{i\tau}^{O}$	Flow to the market of orders of product i during day τ
$q_{i\tau}^{F}$	Flow to the market of forecasted orders of product i during day τ
$B_{i\tau}$	Backlog of product i at the end of day τ
$w_{i\tau}$	Operational mode of area i (on/off) during day τ
$s_{i\tau}$	Shutdown/start-up variable for area i for day τ
$z_{i\tau}$	Auxiliary variable for buffer tank reference interval for tank i, day τ
$x_{i\tau}$	Auxiliary variable for production rate changes for area i, day τ

Table 2. Parameters in the PS model.

Parameter	Description
a_{ij}	Conversion factor between product i and j
$q_i^{\mathrm{PS,min}}$	Minimum daily production in area i
$q_i^{\mathrm{PS,max}}$	Maximum daily production in area i
$I_i^{\min}$	Minimum inventory level of tank i
$I_i^{\max}$	Maximum inventory level of tank i
I_i^{lb}	Lower bound of reference interval for tank i
I_i^{ub}	Upper bound of reference interval for tank i
n_i	Start-up time of area i (days)
d_i	Number of days that area i has to remain being shut down
m_i	Contribution margin of product i

Other information about the site that is required is how the production areas at the site are connected. In the formulation of the constraints, the set $\mathcal{D}_i$ is introduced, denoting the set of areas directly downstream of area i.

The external inputs to the model are the total order volumes, $O_{i\tau}$, and forecasted order volumes, $F_{i\tau}$, for all products $i \in \mathcal{A}$ and each day $\tau = 1, \ldots, N_{\mathrm{PS}}$ over the horizon. These may be seen as input trajectories to the problem that can be updated each day, since the problem is solved in receding horizon fashion. This should be useful since new information about orders and forecasted orders during the next 30 days may become available from one day to the next. For each instance of the PS problem, the remaining days that the areas have to be shut down, d_i, also have to be supplied.

5.2 Formulation of Constraints

The specifications from Section 4 are formulated as constraints below.

Interconnection of production areas The interconnection of production areas are expressed as mass balances at the buffer tanks at the site. If the approximation in (1) is used to model the product flow at a site, the mass balance equations become

$$I_{i\tau} = I_{i,\tau-1} + q_{i\tau} - q_{i\tau}^{O} - q_{i\tau}^{F} - \sum_{j \in \mathcal{D}_i} q_{jt} a_{ij}, \quad i \in \mathcal{A},\ \tau = 1, \ldots, N_{\text{PS}} \tag{2}$$

where the initial conditions I_{i0} are given from measurements. The notation is shown in Figure 5, which shows a site with three production areas.

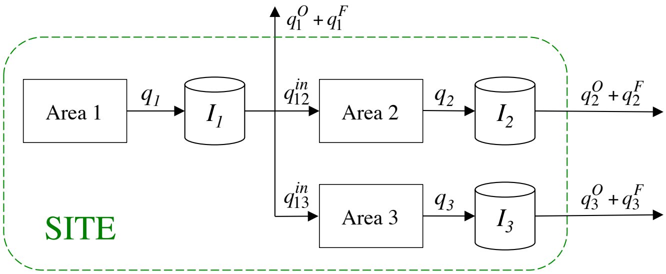

Figure 5. Example of a site with three areas.

Orders and forecasts The total order quantity of product i that should be delivered during day τ is denoted $O_{i\tau}$, and is given as an input to the optimization for all products at the site and all times over the horizon. A forecast of the anticipated orders over the horizon is also assumed to be given. The forecasted order quantity of product i during day τ is denoted $F_{i\tau}$.

Production rate limitations To avoid solutions where the areas operate at rates between zero and the minimum rate, binary variables $w_{i\tau}$ may be used. $w_{i\tau}$ is equal to one if area i operates at day τ, and zero otherwise. The production rate constraints are expressed as

$$w_{i\tau} q_i^{\text{PS,min}} \leq q_{i\tau} \leq w_{i\tau} q_i^{\text{PS,max}}, \quad i \in \mathcal{A},\ \tau = 1, \ldots, N_{\text{PS}} \tag{3}$$

This formulation is made in the same manner as the capacity limitation constraints in [Kondili et al., 1993]. The formulation enables penalization of start-up/shutdown of areas, which is handled under 'Start-up costs and start-up times'. If there are planned production disruptions at the site, such as planned maintenance, this may be expressed as a time-varying maximum production rate for the affected areas.

Inventory limitations The hard constraints on the inventory levels may simply be expressed as

$$I_i^{\min} \leq I_{i\tau} \leq I_i^{\max}, \quad i \in \mathcal{A},\ \tau = 1, \ldots, N_{\mathrm{PS}} \tag{4}$$

Auxiliary variables $z_{i\tau}$ may be used to put a linear penalty on deviating from the reference interval in a buffer tank. Constraints

$$I_i^{lb} - z_{i\tau} \leq I_{i\tau} \leq I_i^{ub} + z_{i\tau}, \quad i \in \mathcal{A},\ \tau = 1, \ldots, N_{\mathrm{PS}} \tag{5}$$

may be formulated, and the auxiliary variables penalized in the objective function to achieve this.

Start-up costs and start-up times A binary variable $s_{i\tau}$ can be introduced to keep track of when an area has been shut down. Inequality constraints

$$s_{i\tau} \geq w_{i,\tau-1} - w_{i\tau}, \quad i \in \mathcal{A},\ \tau = 1, \ldots, N_{\mathrm{PS}} \tag{6}$$

force $s_{i\tau}$ to be equal to one if area i has been shut down from day $\tau - 1$ to τ. Initial conditions w_{i0} are given from measurements. To penalize shutdown/start-up costs, $s_{i\tau}$ may be penalized in the objective function.

If area i is shut down it has to remain being shut down for n_i days, where n_i is the start-up time of area i. This gives the constraint

$$w_{i\tau} = 0, \quad i \in \mathcal{A},\ \tau = 1, \ldots, d_i \tag{7}$$

where d_i is the number of remaining days for area i to be shut down, which is decreased by one day at the end of each day. When area i is shut down, d_i becomes equal to n_i. The PS has to account for that an area could have to remain to be shut down for several days if it is shut down. This is captured by the constraint

$$\begin{aligned} w_{i,\tau+j} \leq 1 - s_{i\tau}, \quad i \in \mathcal{A},\ & \tau = d_i + 1, \ldots, N_{\mathrm{PS}} \\ & j = 0, \ldots, n_i - 1 \end{aligned} \tag{8}$$

Market conditions The forecasted orders that should be prioritized are the orders that give the highest profit. This may be achieved by penalizing $-\sum_\tau m_i q_{i\tau}^F$ in the objective function, where m_i is the contribution margin of product i. Another constraint imposed by the market is that the flows to the market may not be greater than the demand. For the flow to the market of orders, both the orders during the day and the backlog of orders from the previous day has to be taken into account. The constraint may be expressed as

$$q_{i\tau}^O \leq O_{i\tau} + B_{i,\tau-1}, \quad i \in \mathcal{A},\ \tau = 1, \ldots, N_{\mathrm{PS}} \tag{9}$$

where $B_{i,\tau-1}$ is the backlog of product i at the end of day $\tau - 1$, which is defined under 'Costs for late delivery'.

For forecasts, the constraint

$$q_{i\tau}^{F} \leq F_{i\tau} \quad i \in \mathcal{A},\ \tau = 1, \ldots, N_{\mathrm{PS}} \tag{10}$$

may be used to ensure that the PS does not plan to produce more of a product than can be sold.

Costs for late delivery The backlog of orders of product i at the end of day τ is given by

$$B_{i\tau} = B_{i,\tau-1} + O_{i\tau} - q_{i\tau}^{O}, \quad i \in \mathcal{A},\ \tau = 1, \ldots, N_{\mathrm{PS}} \tag{11}$$

where $B_{i\tau} \geq 0$ at all times, and initial conditions B_{i0} are given from measurements. To avoid late delivery of orders, the backlog can be penalized in the objective function. To ensure that the backlog of the most profitable products are handled first, the term may be weighted by factors proportional to the contribution margins, m_i.

Cost of production rate changes Auxiliary variables $x_{i\tau}$ may be used to penalize production rate changes. Constraints

$$q_{i\tau} - x_{i\tau} \leq q_{i,\tau-1} \leq q_{i\tau} + x_{i\tau}, \quad i \in \mathcal{A},\ \tau = 1, \ldots, N_{\mathrm{PS}} \tag{12}$$

are introduced, and $x_{i\tau}$ penalized in the objective function.

5.3 Formulation of Objective Function

As described previously, some variables have to be penalized in the objective function to handle the specifications from Section 4. The suggested objective function is

$$Z_{\mathrm{PS}} = \sum_{i \in \mathcal{A}} \sum_{\tau=1}^{N_{\mathrm{PS}}} \left[\alpha_i m_i B_{i\tau} - \beta_i m_i q_{i\tau}^{F} + \gamma_i s_{i\tau} + \delta_i z_{i\tau} + \epsilon_i x_{i\tau} \right] \tag{13}$$

The first term penalizes backlog of orders, where the weighting with the contribution margins ensures that backlog of the most profitable products is handled first. The second term encourages to plan the production for forecasted orders. The most profitable forecasted orders are prioritized, since the term is weighted with the contribution margins. The third term penalizes start-up/shutdown of areas. The weight on this variable, γ, should correspond to the cost of starting up the different production areas. The last two terms penalize the auxiliary variables $z_{i\tau}$ and $x_{i\tau}$, to account for the buffer tank reference interval and the cost of production rate changes, respectively.

5.4 Resulting Optimization Problem

An optimization problem that takes the specifications in Section 4 into account may be formulated as

$$\begin{aligned} &\text{minimize } (13) \\ &\text{subject to } (2) - (12) \\ &\text{and } I_{i\tau}, q_{i\tau}, q^O_{i\tau}, q^F_{i\tau}, B_{i\tau}, z_{i\tau}, x_{i\tau} \geq 0, \quad i \in \mathcal{A}, \tau = 1, \dots, N_{\text{PS}} \\ &\qquad w_{i\tau}, s_{i\tau} \in \{0, 1\}, \quad i \in \mathcal{A}, \tau = 1, \dots, N_{\text{PS}} \end{aligned} \tag{14}$$

with variables $I_{i\tau}, q_{i\tau}, q^O_{i\tau}, q^F_{i\tau}, B_{i\tau}, z_{i\tau}, x_{i\tau}, w_{i\tau}, s_{i\tau}$ for $i \in \mathcal{A}, \tau = 1, \dots, N_{\text{PS}}$. The problem is a mixed-integer linear program (MILP) since it consists of both binary and continuous variables and the objective function and the constraints are linear. Thus, the problem may be solved by any MILP solver, e.g. CPLEX. In total, there are $7|\mathcal{A}|N_{\text{PS}}$ continuous variables and $2|\mathcal{A}|N_{\text{PS}}$ binary variables. In the worst case, the solution time for MILPs grows exponentially with the number of variables, but good branch-and-bound solvers in general perform much better. A parallel branch-and-bound algorithm specifically developed for chemical production scheduling problems is presented in [Velez and Maravelias, 2013].

The optimization model may also be written as a model on linear state-space form

$$x(\tau + 1) = Ax(\tau) + Bu(\tau) + B_d d(\tau) \tag{15}$$

where the states x are given by the inventories and backlog of all products, and the remaining variables are the inputs, u. The disturbances, d, correspond to the orders and forecasted orders. To handle the start-up constraints, *lifting* according to [Subramanian et al., 2012] has to be performed. As described in [Subramanian et al., 2012], formulation of the scheduling problem as a state-space model can help bridge the gap between scheduling and control.

Initial conditions The model needs initial conditions for the inventories, the backlog of orders, and the operational mode of the areas (I_{i0}, B_{i0}, w_{i0}). The initial inventory levels are given as measurements from the actual site. The initial backlog is computed from the orders and the actual production during the previous time step. To handle the start-up time constraints (7) and (8) in the receding horizon formulation, the remaining days that area i has to be shut down, d_i is also needed.

6. Mathematical Formulation of Detailed Production Scheduling

In this section, the specifications in Section 4 are formulated as an optimization problem to produce a detailed production schedule. The optimization problem is meant to be solved in receding horizon fashion, such that the PS problem for one day ahead is solved every hour. The horizon for the PS is denoted N_{DPS}, and is thus equal to 24 hours.

As for the PS model, it is assumed that each area produces one product, which is stored in one buffer tank. The set of production areas at the site is denoted $\mathcal{A}$. The site may use both continuous and on/off type utilities. The set of continuous utilities is denoted $\mathcal{K}$ and the set of on/off type utilities $\mathcal{L}$.

6.1 Variables and Parameters

The variables that are used to formulate the DPS problem are summarized in Table 3. All variables are continuous except for the binary variables w_{it} and s_{it}. The parameters that are used to formulate the problem are summarized in Table 4. The values of these parameters may be obtained from personnel at the site, and are constant for each optimization problem in the receding horizon formulation. Many parameters are the same as those for the PS formulation. In addition to the parameters in Table 4, five penalty parameters, ζ_i, η_i, θ_i, κ_i, and λ_i, are used in the objective function of the optimization problem. These should be tuned to get the desired behavior of the PS.

Table 3. Variables in the DPS model.

Variable	Description
I_{it}	Inventory level of tank i at the end of hour t
q_{it}	Production rate of area i during hour t
q_{it}^m	Flow to the market of product i during hour t
w_{it}	Operational mode of area i (on/off) during hour t
s_{it}	Shutdown/start-up variable for area i for hour t
x_{it}	Auxiliary variable for production rate changes for area i, hour t

Table 4. Parameters in the DPS model.

Parameter	Description
a_{ij}	Conversion factor between product i and j
$q_i^{\text{DPS,min}}$	Minimum hourly production in area i
$q_i^{\text{DPS,max}}$	Maximum hourly production in area i
I_i^{min}	Minimum inventory level of tank i
I_i^{max}	Maximum inventory level of tank i
n_i	Start-up time of area i (days)
d_i	Number of days that area i has to remain being shut down
q_i^D	Remaining volume to be sold of product i during the day
c_{ki}	Utility model constant for continuous utility k, area i (slope)
d_{ki}	Utility model constant for continuous utility k, area i (offset)
I_i^{ref}	Reference value for inventory level of tank i (from PS)
q_i^{ref}	Reference value for production of product i during each hour (from PS)
$q_i^{\text{m,ref}}$	Reference value for sales of product i during each hour (from PS)

Information on how the production areas at the site are connected, and at which areas the utilities are used is also required. The interconnection of production areas is handled by introducing $\mathcal{D}_i$ as the set of areas directly downstream of area i. For utility usage, the set $\mathcal{M}_k$ is used, denoting the set of areas that require utility k.

The external inputs to the model are the predicted utility disturbance trajectories, U_{kt}, for all utilities $k \in \mathcal{K} \cup \mathcal{L}$ at each time step $t = 1, \ldots, N_{\mathrm{DPS}}$ over the horizon. The volume interpretation of utilities introduced in [Lindholm and Giselsson, 2013] is used, which means that the utilities are modeled as a volume, or power, that all areas which require them have to share. U_{kt} are thus predictions of how much of utility k (e.g. in percent of the maximum available amount) that is going to be available during hour t. The predicted utility disturbance trajectories may be updated each hour, since the problem is solved in receding horizon fashion. This is useful since new information about the disturbances might have become available. Usually, disruptions in the utility supply are not known in advance, which means that the U_{kt} is equal to 100% for all k and t until the disturbance occurs.

The current production schedule, which specifies the desired daily production, sales, and inventory usage for the next 30 days, are given by the PS activity at the start of each day. The accumulated production for the current day (first time step in the production schedule) is divided evenly over the hours of the day to produce reference values I_i^{ref}, q_i^{ref} and $q_i^{\mathrm{m,ref}}$ for the DPS. The references for the next day is also used in the same manner, since the DPS operates in receding horizon, and uses reference values for the following 24 hours.

Feedback in terms of measurements from the actual site is also needed to keep track of the actual accumulated production during the receding horizon optimization. The volume that remains to be sent to the market (sold) during at hour t of the day if the production schedule should be followed perfectly is denoted q_{it}^D. The initial value of this parameter is given from the production schedule at the beginning of the day, and the parameter is then subtracted by the actual hourly production at the end of each hour.

6.2 Formulation of Constraints

The specifications from Section 4 are formulated as constraints below.

Interconnection of production areas As for the PS problem, the mass balance constraints are formulated as

$$I_{it} = I_{i,t-1} + q_{it} - q_{it}^m - \sum_{j \in \mathcal{D}_i} q_{jt} a_{ij}, \quad i \in \mathcal{A},\ t = 1, \ldots, N_{\mathrm{DPS}} \tag{16}$$

Production rate limitations The production rate constraints are formulated in the same manner as for the PS problem. The production rate constraints become

$$w_{it} q_i^{\text{DPS,min}} \le q_{it} \le w_{it} q_i^{\text{DPS,max}}, \quad i \in \mathcal{A},\ t = 1, \dots, N_{\text{DPS}} \tag{17}$$

Inventory limitations As for the PS model, the maximum and minimum limitations on the buffer tank levels are defined by

$$I_i^{\min} \le I_{it} \le I_i^{\max}, \quad i \in \mathcal{A},\ t = 1, \dots, N_{\text{DPS}} \tag{18}$$

Reference levels for the inventories are given by the production schedule. A penalty on deviating from these levels is imposed in the objective function of the DPS problem.

Start-up costs and start-up times In the same manner as for the PS problem, binary variables s_{it} are used to formulate the start-up constraints

$$s_{it} \ge w_{i,t-1} - w_{it}, \quad i \in \mathcal{A},\ t = 1, \dots, N_{\text{DPS}} \tag{19}$$

To penalize shutdown/start-up costs, s_{it} may be penalized in the objective function.

If an area is shut down, e.g. due to a complete utility failure, it needs to remain being shut during all remaining hours of the day if the start-up time for the area is one day or more. This can be captured by the constraint

$$w_{it} = 0 \quad \text{if } d_i > 0, \quad i \in \mathcal{A},\ t = 1, \dots, N_{\text{DPS}} \tag{20}$$

where d_i is the number of days that area i has to remain being shut down. d_i is set to the start-up time of area i, n_i, when area i is shut down ($s_{it} = 1$).

Market conditions The market conditions are captured by the production schedule, which is given from solution of the PS problem. The detailed production schedule should not plan to produce more of any of the products during the day than given by the production schedule. The constraint may be expressed as

$$\sum_{t=1}^{N_{\text{DPS}}} q_{it}^m \le q_{it}^D, \quad i \in \mathcal{A}, t = 1, \dots, N_{\text{DPS}} \tag{21}$$

where q_{i1}^D is equal to the planned sales given by the production schedule, $q_{i\tau}^O + q_{i\tau}^F$ for day τ (first instance in the production schedule). The parameter q_{it}^D is then updated at the end of each hour t by subtracting the actual production during the hour, that is given from measurements. If the production schedule is followed perfectly, $q_{iN_{DPS}}^D$ will be equal to zero after the last hour of the day. If q_{it}^D is greater than zero, this means that the backlog of product i will increase, which is a problem that will be handled by the PS activity.

Cost of production rate changes Auxiliary variables are used to penalize production rate changes, in the same manner as for the PS. The constraints become

$$q_{it} - x_{it} \leq q_{i,t-1} \leq q_{it} + x_{it}, \quad i \in \mathcal{A},\ t = 1, \ldots, N_{\text{DPS}} \tag{22}$$

and x_{it} is penalized in the objective function.

Limited availability of utilities If utilities are modeled according to [Lindholm and Giselsson, 2013], the constraints on the production rate due to utilities with continuous characteristics may be expressed as

$$\sum_{i \in \mathcal{M}_k} c_{ki} q_{it} - d_{ki} \leq U_{kt}, \quad k \in \mathcal{K},\ t = 1, \ldots, N_{\text{DPS}} \tag{23}$$

where c_{ki} and d_{ki} are positive constants that are specific for utility k, area i. A linear relationship between the assignment of a utility to an area and the production in the area is thus assumed.

For on/off utilities, the assumption in [Lindholm and Giselsson, 2013] is that the areas that require a utility with on/off characteristics can operate at maximum speed during hour t if the utility is available during hour t, and can not operate if the utility is unavailable. The constraints on production due to on/off type utilities become

$$q_{it} \leq q_i^{\max} \qquad \text{if } U_{kt} = 1, \quad i \in \mathcal{M}_k,\ k \in \mathcal{L},\ t = 1, \ldots, N_{\text{DPS}} \tag{24}$$

$$q_{it} = 0 \qquad \text{if } U_{kt} = 0, \quad i \in \mathcal{M}_k,\ k \in \mathcal{L},\ t = 1, \ldots, N_{\text{DPS}} \tag{25}$$

U_{kt} is always zero or one for on/off type utilities.

6.3 Formulation of Objective Function

To take the specifications from Section 4 into account, the following quadratic objective function is suggested:

$$\begin{aligned} Z_{\text{DPS}} = \sum_{i \in \mathcal{A}} \sum_{t=1}^{N_{\text{DPS}}} [(I_{it} - I_i^{\text{ref}})^2 \zeta_i + (q_{it} - q_i^{\text{ref}})^2 \eta_i + \\ + (q_{it}^m - q_i^{m,\text{ref}})^2 \theta_i + \kappa_i s_{it} + \lambda_i x_{it}] \end{aligned} \tag{26}$$

The first, second, and third terms are quadratic terms that penalize deviations from the reference values given from the solution of the PS problem. The fourth term penalizes start-ups, and the last production rate changes. These penalties are linear, to give large penalties for small deviations.

6.4 Resulting Optimization Problem

An optimization problem that together with (14) captures the specifications in Section 4 may be formulated as

$$\begin{aligned} &\text{minimize } (26) \\ &\text{subject to } (16) - (22) \\ &\text{and } I_{it}, q_{it}, q_{it}^m, x_{it} \geq 0, \quad i \in \mathcal{A}, t = 1, \ldots, N_{\text{DPS}} \\ &\qquad w_{it}, s_{it} \in \{0, 1\}, \quad i \in \mathcal{A}, t = 1, \ldots, N_{\text{DPS}} \end{aligned} \tag{27}$$

with variables $I_{it}, q_{it}, q_{it}^m, x_{it}, w_{it}, s_{it}$ for $i \in \mathcal{A}, t = 1, \ldots, N_{\text{DPS}}$. Since the inequality and equality constraints are linear and the cost function is quadratic, the problem becomes a mixed-integer quadratic program (MIQP), which may be solved e.g. using CPLEX MIQP solver. The problem has $4|\mathcal{A}|N_{\text{DPS}}$ continuous variables and $2|\mathcal{A}|N_{\text{DPS}}$ binary variables. As for MILPs, the solution time for MIQPs grows exponentially with the number of variables in the worst case, but good branch-and-bound solvers in general perform much better.

As for the PS problem, the optimization model may be written as a model on linear state-space form, in line with the scheduling model formulation in [Subramanian et al., 2012]. The states are then given by the inventories, and the remaining variables are inputs. The disturbances are given by the utility disturbance trajectories.

Initial conditions The model needs initial conditions for the inventories and for the binary operational mode variables (I_{i0} and w_{i0}). The inventory levels are given as measurements from the site. The operational mode of the areas at the start of the optimization is given from the measurements of the production rates at the site. The number of days that area i has to remain to be shut down, d_i is also needed.

7. An Example

The example provided in this section is constructed to resemble a real industrial site. Real data from Perstorp can unfortunately not be published due to secrecy matters. The planning horizon for the PS is one month ($N_{\text{PS}} = 30$ days) and for the DPS one day ($N_{\text{DPS}} = 24$ hours).

7.1 Background Data

The site structure is depicted in Figure 6. The site has six production areas, six products and six buffer tanks. The downstream areas are most profitable, and product 3 is a byproduct for which there is no demand. The contribution margins, m_i, maximum and minimum production rates, $q_i^{\max}$ and $q_i^{\min}$, and start-up times, n_i, for all production areas/products are given in Table 5.

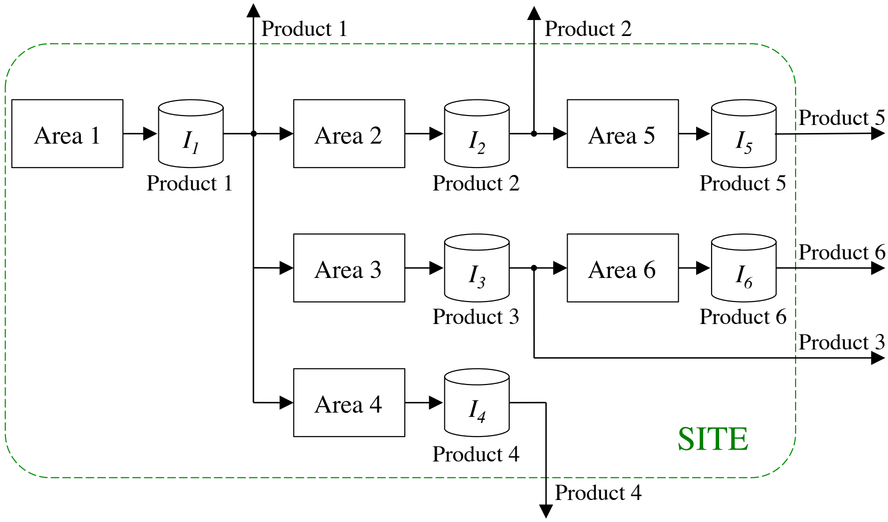

Figure 6. Site considered in the example.

Table 5. Production data.

	$q_i^{\min}$ (m^3/h)	$q_i^{\max}$ (m^3/h)	m_i (\$/m^3)	n_i (days)	$V_i^{\max}$ (m^3)
Product 1	0.1	1	0.4	1	10
Product 2	0.05	0.5	0.7	1	10
Product 3	0.02	0.2	0.1	3	10
Product 4	0.01	0.1	0.5	0	10
Product 5	0.02	0.2	0.8	0	10
Product 6	0.02	0.2	1	0	10

The buffer tanks are limited between zero and the maximum limit $V_i^{\max}$, and a reference interval between 60% and 80% of the buffer tank volume is desired for all buffer tanks.

To make the results easier to interpret, it is assumed that the forecasts are perfect, such that all forecasted orders become actual orders. The information on all orders of the month is given already at the first day of the month. This means that the flows to the market q_{it}^O and q_{it}^F can be merged into one flow, q_{it}^M, in the PS problem. Without loss of generality, it is also assumed that all the conversion factors a_{ij} are equal to one in the mass balance constraints, (2) and (16).

The site uses three utilities; electricity (utility 1), steam (utility 2), and cooling water (utility 3). Steam and cooling water are utilities with continuous characteristics, and electricity is an on/off type utility. Steam is required by areas 4 and 6, cooling water by areas 3, 5, and 6, and electricity by all areas.

7.2 Production Scheduling

The objective function for the PS problem is given by (13). At the site, the costs for shutting down/starting up areas are very high. Thus, a large penalty on shutdown/start-ups is used in the objective function (large γ_i). The cost of starting up/shutting down is approximately equal for all areas. It is more important to deliver orders on time (avoid backlog) and deliver forecasted orders than to avoid to deviate from the inventory reference interval or to avoid changing the production rate rapidly (α_i and β_i greater than δ_i and ϵ_i). The buffer tanks should be used to handle unexpected production disturbances (not too large δ_i and ϵ_i in relation to α_i and β_i). The penalty parameters for the PS optimization are thus chosen as

$$\alpha_i = 100,\ \beta_i = 10,\ \gamma_i = 1000,\ \delta_i = 1,\ \epsilon_i = 0.1, \quad i \in \mathcal{A}$$

The optimization problem is given by (14), with the parameter values stated in this section. The problem consists of $6 \cdot 6 \cdot 30 = 1080$ continuous variables (since q_{it}^O and q_{it}^F have been merged) and $2 \cdot 6 \cdot 30 = 360$ binary variables.

7.3 Detailed Production Scheduling

The objective function for the PS problem is given by (26). Since the costs for shutting down/starting up areas are very high, a large penalty on shutdown/start-up is used also for the DPS problem (large κ_i). Keeping to the economically optimal plan for the production and sales that is given by the PS is more important than keeping to the given reference levels, and avoiding rapid changes in the production rate (large η_i and θ_i in relation to ζ_i and λ_i). The penalty parameters for the DPS optimization are thus chosen as

$$\zeta_i = 0.01,\ \eta_i = 100,\ \theta_i = 100,\ \kappa_i = 1000,\ \lambda_i = 0.1, \quad i \in \mathcal{A}$$

The optimization problem is given by (27), with the parameter values stated in this section. The problem consists of $4 \cdot 6 \cdot 24 = 576$ continuous variables and $2 \cdot 6 \cdot 24 = 288$ binary variables.

7.4 Simulation

The PS and DPS are simulated during one month using the suggested models. An initial plan for the sales of the six products is given by the total order volumes for each day. The example is constructed as five periods with different levels of daily orders. The accumulated daily order volumes of each product in these periods are summarized in Table 6.

Table 6. Orders of each product in m^3/day.

Days $\rightarrow$	$1-4$	$5-10$	$11-15$	$16-23$	$24-30$
Product 1	4.80	2.03	5.96	4.80	4.11
Product 2	7.20	6.93	7.20	7.20	6.74
Product 3	0	0	0	0	0
Product 4	2.40	1.90	1.20	2.40	2.35
Product 5	4.80	1.23	4.80	4.80	4.40
Product 6	4.80	1.40	4.80	4.80	4.18

Three utility disturbances are simulated: An electricity failure at day 2, a steam disturbance at day 12, and a cooling water disturbance at day 18. The electricity failure is a complete failure, since it is an on/off utility. The failure starts at 18:00 day 2 and lasts for two hours ($U_{1t} = 0,\ t = 18, \ldots, 20$). During the steam disturbance, the available amount of steam is 30% of the maximum amount from the start of the disturbance at 04:00 day 12 and two hours ahead ($U_{2t} = 0.3,\ t = 4, \ldots, 6$). The cooling water disturbance starts at 05:00 day 18 and lasts for 15 hours. During the disturbance, 80% of the normal cooling capacity is available ($U_{13} = 0.8,\ t = 5, \ldots, 20$).

The scheduling is performed according to Algorithm 1. The resulting production, sales, and inventory trajectories are shown in Figure 7-9. The actual production, sales, and inventory usage at the site is assumed to be the same as given by the detailed production schedule for the example.

In Figures 7 and 8 it can be seen that the DPS solution follows the reference for the production rates and sales that is given by the PS solution closely, when there are no disturbances present. In Figure 8 it is that the backlog of products due to disturbances is reduced when this is possible. At the end of the month, 100% of the total ordered amount of each of the products has been delivered. Some orders are delivered one or a few days late. If this should be allowed or not depends on the contracts and relations with the customers at the actual site. It can be noted that areas 4 and 5 are shut down for one day due to the electricity failure, even if the start-up time of these areas are zero. This is because areas 1 and 2, which provide raw materials to areas 4 and 5, have to remain shut down during this time. In Figure 9 it is seen that the buffer tanks are utilized to fulfill the orders, at times where there has been disturbances in production. When it is possible, the buffer tank levels are returned to within the reference interval.

The solution time for solving one DPS problem is around 0.1 seconds, and the time for solving one PS problem is around 0.2 seconds using CPLEX on a desktop computer. Thus, the solution times can be considered small in relation to the time steps (hours, days).

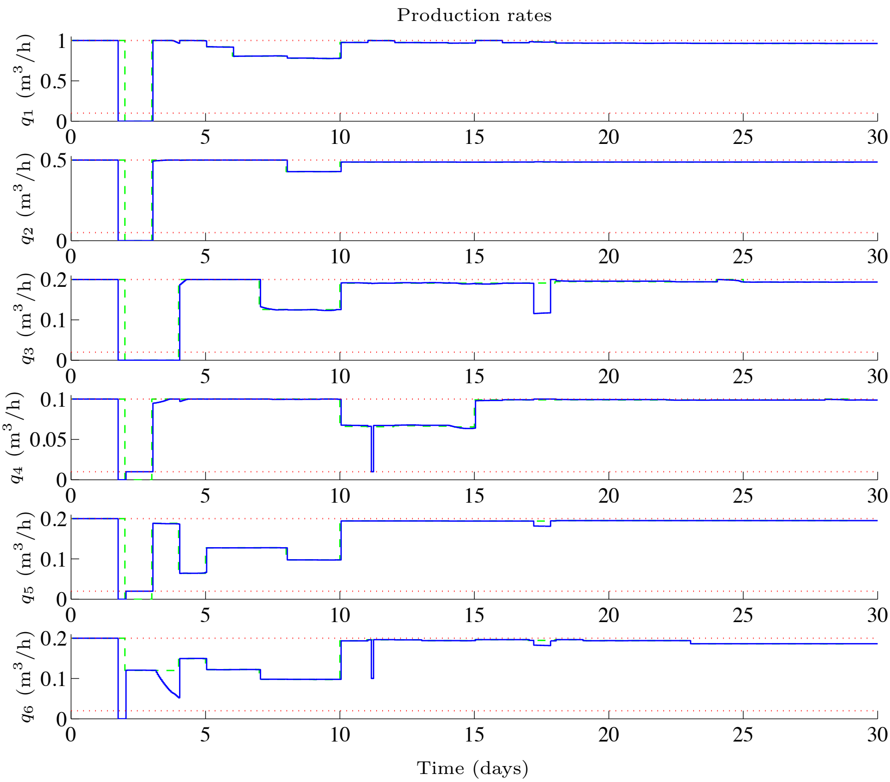

Figure 7. Production rate trajectories. The green dashed lines show the PS solution, blue solid lines mark the DPS solution, and the red dotted lines the maximum and minimum limits.

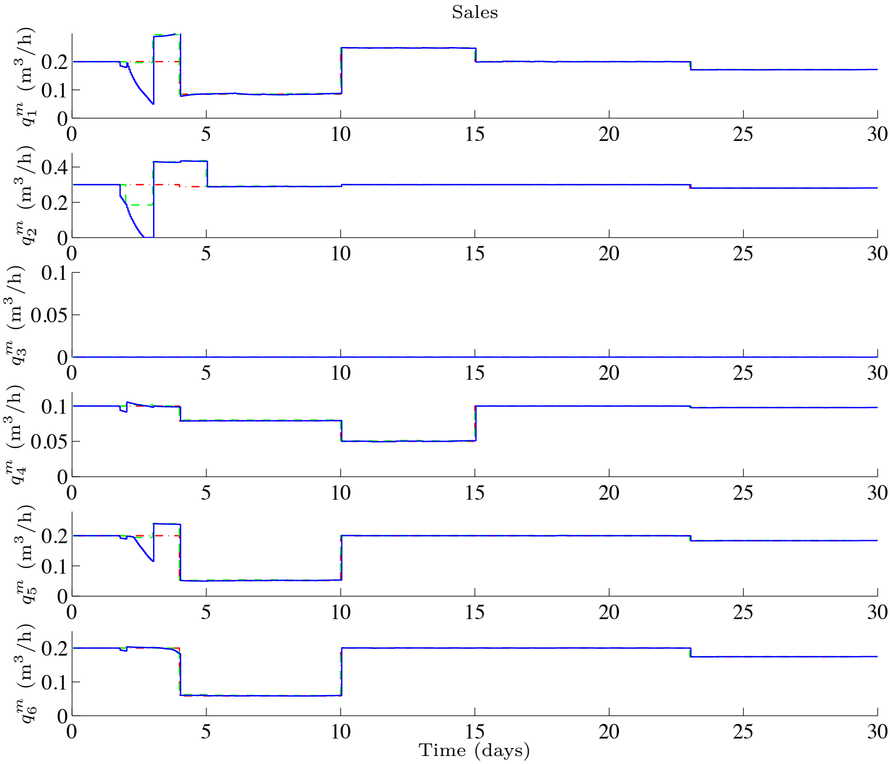

Figure 8. Sales trajectories. The red dash-dotted lines mark the orders, the green dashed lines the PS solution, and the blue solid lines the DPS solution.

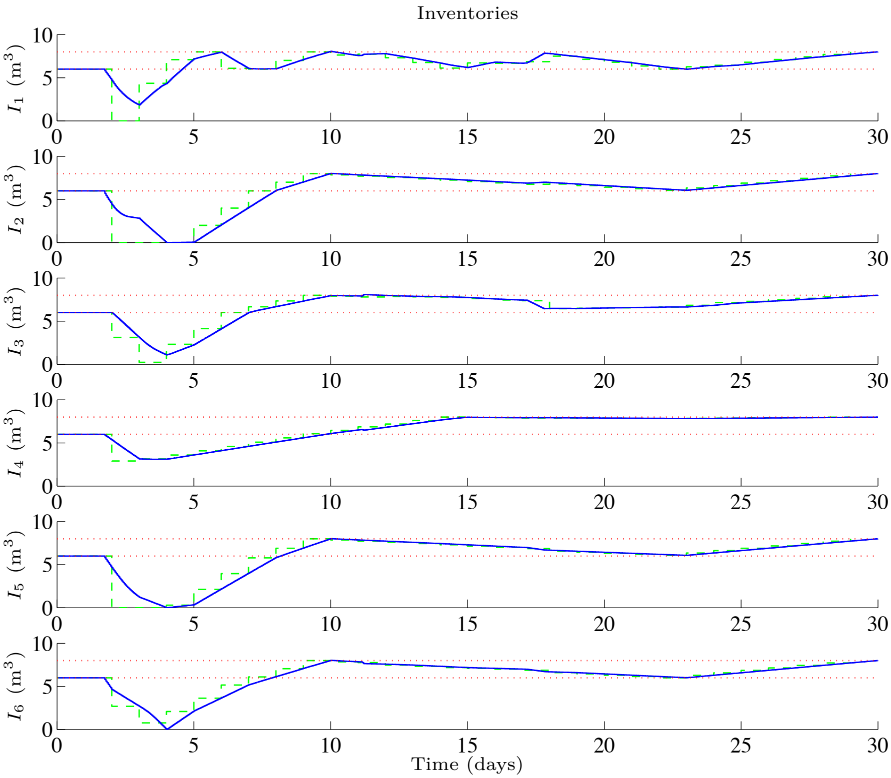

Figure 9. Inventory trajectories. The green dashed lines show the PS solution, and the blue solid lines the DPS solution. The reference intervals for the buffer tanks are marked with red dotted lines.

8. Discussion

With as simple models as in this paper, it should be possible to incorporate the two scheduling problems into a single optimization problem with a time scale of hours. However, it might be required to take more constraints into account, e.g. that there are different penalty costs for late delivery of different individual orders, or that it is not possible to make minor changes in the production rate. When such constraints are introduced, the complexity of the scheduling problems will quickly increase, and a monolithic solution would be impossible for industrial-size problems. At Linköping University, Sweden, the PS model has been extended to take further specifications into account. These studies have shown that even for small sites with only three production areas, the time for solving the PS problem may be as long as several hours. Therefore, the two-level approach seems more reasonable in the long run. If the dynamics and information used in the two problems have a significant difference in time scale, the solution when using the integrated approach should not differ much from the monolithic solution.

9. Conclusions

A hierarchical approach for integrating production scheduling (PS) and detailed production scheduling (DPS) was presented, which aims to be general; such that it can be used at any process industrial site. A specification list from Perstorp was used as a starting point for formulating the two scheduling problems, and the interaction and flow of information between the problems is described. The PS problem becomes a mixed-integer linear program (MILP) and the DPS problem a mixed-integer quadratic program (MIQP). Both problems are solved in receding horizon, the PS problem with a horizon of one month, and the DPS problem with a horizon of one day. An example inspired by a real site at Perstorp is used to show how the PS and DPS operate and interact.

The PS model is currently further developed at Linköping University, Sweden. The aim is to take more aspects of the PS into account to make the model more realistic. Another interesting future work direction would be to solve the monolithic scheduling problem, where the PS and DPS problems are merged into one problem, for a simple problem setup that can be solved in reasonable time, and compare this truly optimal solution to the solution when using the hierarchical scheduling approach.

Acknowledgments

The research is performed within the Process Industry Centre (PIC) supported by the Swedish Foundation for Strategic Research (SSF). The first author is a member of the LCCC Linnaeus Center and the eLLIIT Excellence Center at Lund University.

References

Blackstone, J. H. and J. F. Cox, eds. (2004). *APICS Dictionary, 11th edition.* The Association for Operations Management.

Craig, I., C. Aldrich, R. Braatz, F. Cuzzola, E. Domlan, S. Engell, J. Hahn, V. Havlena, A. Horch, B. Huang, M. Khanbaghi, A. Konstantellos, W. Marquardt, T. McAvoy, T. Parsini, S. Pistikopoulos, T. Samad, S. Skogestad, N. Thornhill, and J. Yu (2011). "Control in the process industries". In: *The impact of control technology.* Ed. by T. Samad and A. M. Annaswamy. Available at `www.ieeecss.org`. IEEE control systems society.

Engell, S. and H. Harjunkoski (2012). "Optimal operation: scheduling, advanced control and their integration". In: *Proceedings of the international conference on Foundations Of Computer-Aided Process Operations.* Savannah, GA.

Grossmann, I. E. (2012). "Advances in mathematical programming models for enterprise-wide optimization". In: *Proceedings of the international conference on Foundations Of Computer-Aided Process Operations.* Savannah, GA.

Grossmann, I. E. and K. C. Furman (2009). "Challenges in enterprise-wide optimization for the process industries". In: *Optimization and Logistics Challenges in the Enterprise.* Vol. 30. Springer Optimization and Its Applications. Springer, USA, pp. 3–59.

Harjunkoski, I., R. Nyström, and A. Horch (2009). "Integration of scheduling and control–Theory or practice?" *Computers & Chemical Engineering* **33**:12, pp. 1909–1918.

Huang, R. (2010). *Nonlinear model predictive control and dynamic real time optimization for large-scale processes.* PhD thesis. Carnegie Mellon University, PA.

Ierapetritou, M. G. and C. A. Floudas (1998). "Effective continuous-time formulation for short-term scheduling. 2. continuous and semicontinuous processes". *Industrial & engineering chemistry research* **37**:11, pp. 4360–4374.

ISA-95.00.01 (2009). *Enterprise-Control System Integration, Part 1: Models and Terminology.* Standard. Instrument Society of America, pp. 17–28.

ISA-95.00.03 (2009). *Enterprise-Control System Integration, Part 3: Models of Manufacturing Operations Management.* Standard. Instrument Society of America.

Kadam, J. and W. Marquardt (2007). "Integration of economical optimization and control for intentionally transient process operation". In: *Assessment and Future Directions of Nonlinear Model Predictive Control.*

Vol. 358. Lecture Notes in Control and Information Sciences. Springer Berlin Heidelberg, pp. 419–434.

Kallrath, J. (2002). "Planning and scheduling in the process industry". *OR Spectrum* **24**:3, pp. 219–250.

Kondili, E., C. C. Pantelides, and R. W. H. Sargent (1993). "A general algorithm for short-term scheduling of batch operations – I. MILP formulation". *Computers & Chemical Engineering* **17**:2, pp. 211–227.

Lindholm, A. and P. Giselsson (2013). "Minimization of economical losses due to utility disturbances in the process industry". *Journal of Process Control* **23**:5, pp. 767–777.

Lindholm, A., C. Johnsson, N.-H. Quttineh, H. Lidestam, M. Henningsson, J. Wikner, O. Tang, N.-P. Nytzén, and K. Forsman (2013a). "Hierarchical scheduling and utility disturbance management in the process industry". In: *Proceedings of the IFAC Conference on Manufacturing Modelling, Management and Control.* Saint Petersburg, Russia.

Lindholm, A., P. Giselsson, N.-H. Quttineh, H. Lidestam, C. Johnsson, and K. Forsman (2013b). "Production scheduling in the process industry". In: *Proceedings of the 22nd International Conference on Production Research.* Iguassu Falls, Brazil.

Maravelias, C. T. (2012). "General framework and modeling approach classification for chemical production scheduling". *AIChE Journal* **58**:6, pp. 1812–1828.

Maravelias, C. T. and I. E. Grossmann (2003). "New general continuous-time state–task network formulation for short-term scheduling of multipurpose batch plants". *Industrial & Engineering Chemistry Research* **42**:13, pp. 3056–3074.

Neumann, K., C. Schwindt, and N. Trautmann (2002). "Advanced production scheduling for batch plants in process industries". *OR Spectrum* 24, pp. 251–279.

Rawlings, J. B. and R. Amrit (2009). "Optimizing process economic performance using model predictive control". In: *Nonlinear Model Predictive Control: Towards New Challenging Applications.* Vol. 384. Lecture Notes in Control and Information Sciences, pp. 119–138.

Shaik, M. A., C. A. Floudas, J. Kallrath, and H. J. Pitz (2009). "Production scheduling of a large-scale industrial continuous plant: short-term and medium-term scheduling". *Computers & Chemical Engineering* **33**:3, pp. 670–686.

Shobrys, D. E. and D. C. White (2002). "Planning, scheduling and control systems: why cannot they work together". *Computers & Chemical Engineering* **26**:2, pp. 149–160.

Subramanian, K., C. T. Maravelias, and J. B. Rawlings (2012). "A state-space model for chemical production scheduling". *Computers & Chemical Engineering*.

Sundaramoorthy, A. and C. T. Maravelias (2011). "A general framework for process scheduling". *AIChE Journal* **57**:3, pp. 695–710.

Tousain, R. L. (2002). *Dynamic Optimization in Business-wide Process Control*. PhD thesis. Technical University Delft, The Netherlands.

Velez, S. and C. T. Maravelias (2013). "A branch-and-bound algorithm for the solution of chemical production scheduling MIP models using parallel computing". *Computers & Chemical Engineering* **55**, pp. 28–39.

Wassick, J. M. (2009). "Enterprise-wide optimization in an integrated chemical complex". *Computers & Chemical Engineering* **33**:12, pp. 1950–1963.