7th IFAC Conference on Manufacturing Modelling, Management, and Control
International Federation of Automatic Control
June 19-21, 2013. Saint Petersburg, Russia

Predictive production planning in an integrated pulp and paper mill ⋆

Gonçalo Figueira * Marcos Furlan ** Bernardo Almada-Lobo *

* INESC TEC - Faculty of Engineering, University of Porto, Rua Dr. Roberto Frias, s/n 4200-465 Porto, Portugal (e-mail: goncalo.figueira@fe.up.pt).
** ICMC - University of São Paulo, Av. Trabalhador São-carlense, 400, 13560-970 São Carlos-SP

Abstract: Disturbance Management is a major issue in process industries like the pulp and paper (P&P) industry. In this paper, a case study in an integrated P&P mill is examined. Production plans for the whole mill need not only to be optimized concerning company's indicators, but also to be robust so that disturbances can be avoided. We present a simulation-optimization approach that generates plans, correctly weighting their quality (regarding various indicators) and robustness. A discrete-event simulation model replicates the dynamics of implementation and adaptation of production plans in practice. The simulation model gives then feedback to optimization, in order to enhance the analytical model, which is thus able to generate robust plans.

Keywords: Production planning and scheduling; Disturbance management; Pulp and paper industry; Simulation-optimization; Mixed integer programming; Variable neighbourhood search; Discrete-event simulation.

1. INTRODUCTION

Pulp and paper (P&P) industry converts fibrous raw materials into pulp, paper and paperboard. In a first step raw materials are processed into pulp (in the digester) and in a second step paper and paper products are produced out of this pulp (Schumacher and Sathaye, 1999). These two steps can be processed on separate plants or combined into an integrated mill.

P&P mills are complex systems, consisting of several processes with storage tanks in between. Disturbances in one process tend to propagate to other processes, resulting in production losses, unnecessary rate changes (which further cause quality disturbances and undesired wear on equipment) and increase of the environmental load of the mill. The propagation of disturbances can lead to enforced shut-downs. The start-up of the process after this kind of disturbance can also be problematic. It is clear that all these factors can cause considerable economic losses (Leiviska, 1996).

Disturbance management is a production control problem, typically neglected in the planning process. Still, some works in the literature have attempted to link both activities. The pioneer work provided by Alsholm and Pettersson (1969) and Pettersson (1970) tried to keep an appropriate level in each tank, in order to avoid the propagation of disturbances. For that purpose, in the objective function of the mathematical model that optimizes the system, the divergence from the desired target levels of the tanks was penalized. However, it was found difficult to correctly weight those penalties against the remaining criteria.

Leiviska and Niemi (1988) proposed a different approach. Instead of mathematical programming, a combination of simulation and heuristics was adopted. Simulation was used for testing production plans. Heuristic procedures were then executed for correcting and improving those plans. To avoid disturbances, a recommended region was defined for each tank. Then, corrective actions were carried out (by the heuristics), only if the tank level went out of that region. The drawback of this approach is that certain plans may need a drastic reformulation and hence, those corrective actions could not be enough. Moreover, the heuristic behaviour does not provide an optimal exploration of production plans.

Mercangoz and Doyle-III (2008) have built an optimization model from process variables with economic significance and the available degrees of freedom in the process control structure. The optimization results were then utilized to change the operating conditions of the benchmark problem.

The last two approaches start from the control point of view, simulating and refining the initial plans. However, it is also important that the planning process can be conducted considering the operational issues related to disturbances, so that plans are evaluated and selected based on realistic criteria. The challenge is then to create optimization methods that can on the one hand, generate good quality plans and on the other hand, consider robustness as an important criterion to avoid disturbances.

⋆ The first author is grateful to the Portuguese Foundation for Science and Technology for awarding him the Ph.D. Grant SFRH/BD/80109/2011.

Robustness measures also need to be properly weighted, in order not to fall in suboptimal solutions.

Uncertainty is addressed in optimization by one of three main approaches:

- Stochastic Programming (SP);
- Fuzzy Sets (FS);
- Simulation-Optimization (S-O).

Whereas the first two include the stochastic elements in the optimization model (as a set of different scenarios or uncertain parameters), the last one uses a simulation model to replicate the stochastic behaviour of the system. Simulation provides considerably more detail and flexibility in modelling the system, but it usually results in higher computational time. Therefore, the framework that combines simulation and optimization has to be carefully designed, so that the algorithm can achieve a reasonable efficiency.

In this paper, we tackle the short-term production planning and scheduling of an integrated P&P mill, considering all types of disturbances that may occur. We propose a simulation-optimization approach, where optimization generates production plans and simulation implements them. All the relevant dynamics of the implementation and adaptation of production plans in practice are modelled. However, simulation does not serve as a simple evaluation function of the optimization. That would result in tremendous inefficiency, given the incredible large amount of solutions to evaluate. Instead, it is used to implement just particular plans and the obtained feedback is employed in the enhancement of the optimization model. The latter is thus able to find good and robust production plans.

The production planning and scheduling of integrated P&P mills was first proposed by Santos and Almada-Lobo (2012) and later approached by Figueira et al. (2013). The latter developed a hybrid variable neighbourhood search (VNS), which is able to provide good quality solutions for real-world instances in feasible time. This solution method is here extended to integrate a discrete-event simulation model within an appropriate simulation-optimization framework.

We are the first, to the best of our knowledge, to study the dynamics of the implementation and adaptation of production plans in an integrated P&P mill subject to disturbances. Most of the papers that simulate production flow do not consider random disturbances and focus only on pulp mills. In addition, we provide a simulation-optimization framework that generates good and robust production plans (where quality and robustness are correctly weighted) for real-world instances and within feasible time.

The paper is organized as follows: We start by reviewing in Section 2 the relevant literature in simulation-optimization, particularly the frameworks applied to production and supply chain planning. Section 3 briefly describes the production process, as well as how practitioners deal with planning activities and disturbance management. In Section 4 the simulation model is explored, first in a pure conceptual point of view and then in a more implementation perspective. Section 5 then presents our simulation-optimization framework, including a scheme to sample the uncertain parameters that reduces the number of required replications. Finally, we present general results and conclusions in Section 6.

2. LITERATURE REVIEW IN SIMULATION-OPTIMIZATION

As stated by Figueira and Almada-Lobo (2013), there are several ways of combining simulation and optimization. These methods can be divided into three main categories:

- Landscape Perception (LP), which corresponds to "simulation optimization" (the former working as a black-box evaluation function of the latter or creating a surrogate model);
- Analytical Model Enhancement (AME), related to "simulation for optimization" (the former computing parameters of the latter);
- Solution Generation (SG), frequently labelled as "optimization-based simulation" (the latter creating a solution, based on the output of the former).

The last category implies that optimization does not need any simulation feedback. This is not our case, since the generation of plans by optimization needs the feedback from their implementation, which is performed by simulation.

The first stream of methods may use simulation as the evaluation function of optimization. This framework could be used in our case. However, the amount of solutions to evaluate would be enormous and thus, the efficiency of the method would be strongly affected. Furthermore, it would only be applied to integer representations of the solution. Indeed, in order to consider stochastic elements in the optimization of the continuous variables, another S-O framework would be necessary. The issue is that all the linear relationships (between input and output variables) which are known would be lost.

That does not happen in AME. Figueira and Almada-Lobo (2013) review some of these methods. Three of them appear to fit well in our case. The first is ROSA, which consists on running recursively a relatively easy (typically linear) analytical model and a (more detailed) simulation model. Almeder et al. (2009) applies this approach for a supply chain planning problem under uncertainty. The uncertainty is related to both production and distribution lead-times. The authors propose the iterative ROSA procedure, considering either a linear program or a mixed integer program. The idea is to leave all the integer variables that do not have a major impact on performance to the simulation model. The latter is able to deal with nonlinearities, stochasticity and complex structure with no mathematical sophistication. The correct lead-times can thus be computed by simulation. The results from multiple replications are aggregated (not by a simple average, but using quantiles) and employed in the refinement of the analytical model parameters. The procedure stops when both models agree on the lead-times and performance measures.

Another approach is based on the use of simulation to estimate functions (IFEBM). This estimation is conducted only in the beginning of the procedure. Then, those functions are embedded into an analytical model, which is solved afterwards. The model is typically nonlinear and hence harder to solve. Nevertheless, it only needs to be solved once and avoids the possible convergence issues of the previous approach. Asmundsson et al. (2006) implement this approach for dealing with "clearing functions" in a production planning problem. These functions represent the nonlinear relationship between workload and throughput of production resources.

Finally, Jung et al. (2004) propose a third methodology (OBSIR). The authors tackle a supply chain planning problem with a rolling-horizon approach. In each step, the analytical model is solved with updated information. The complete procedure is performed multiple times by a simulation model. The purpose of this procedure is to find the appropriate safety stock levels, which ensure no stockouts. The determination of the safety stock levels is thus defined as an outer optimization problem (LP framework).

3. PRODUCTION AND PLANNING PROCESSES

Figure 1 exhibits a schematic (simplified) view of the integrated P&P production process. The three major production stages are identified: pulp mill (or fibre line), paper mill and chemical recovery.

The whole process begins with the wood being debarked and reduced to small chips, which are later cooked inside one or several (batch or continuous) digesters in aqueous solutions with reagents at high temperature and pressure (chemical pulping). The virgin pulp is then stocked in tanks and later pulled by the paper machine, where it is mixed with recycled pulp, in different proportions according to the desired paper receipts. The final paper is characterized not only by the pulp mixture, but also by the paper grammage (measured in g/m^2). For the sake of simplicity, the term *grade* is used for the combination of both. The configuration of the machine to produce a new paper grade is sequence-dependent. Each setup leads to a loss in the production process in terms of time and quantity of a lower quality paper (as the machine is never idle, even during the changeover). In the pulping process a by-product is obtained: the weak black liquor. This product is later concentrated and burnt to produce steam and hence, generate electrical energy.

In many companies, the short-term production planning process is completely manual and typically follows a hierarchical scheme. The top-level focuses on the paper machine, dealing with the sizing and scheduling of paper runs, which are programmed to occur in cycles. One cycle corresponds to a certain period of time (e.g. one week), where a given sequence of grades is followed in order to fulfil current demand and backlog. The sequencing attempts to minimize setups and hence it usually consists of an ascending (or descending) order of grades, followed by a descending (or ascending) order. The base-level of the hierarchical scheme tries to schedule and control all the other production resources, subject to the input from the top-level.

Various iterations between the two levels may be needed to find a feasible solution. Indeed, the pulp quality strongly depends on the digester's stability and thus, its synchro-

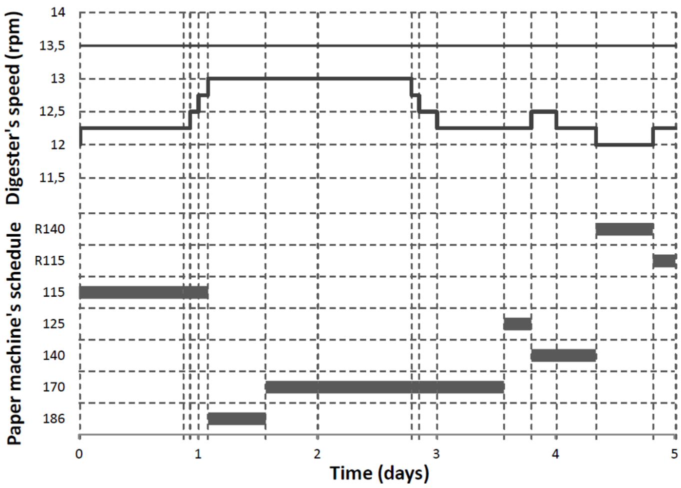

Fig. 2. Illustrative production plan that exhibits the desired synchronization between the digester and the paper machine (Figueira et al., 2013).

nization with the paper machine is crucial. In fact, different paper grades pull more or less pulp from the digester, forcing the change of its rate. The rate is constrained to a maximum variation, due to technological limitations of the equipment. Therefore, the close interrelation between these two stages must be carefully managed.

Facing several constraints, conflicting objectives and eventually multiple and shifting bottlenecks, planners are led to place feasibility over optimality. In addition, the production system is subject to disturbances, such as: paper breaks in the paper machine (which are particularly frequent in lower grades and changeovers), process delays, unsteady productivities (due to unstable pulp quality) and various types of breakdowns. Therefore, planners need also to consider slacks in their plans, in order to accommodate those disturbances and avoid tanks overflow. However, those slacks are arbitrarily defined, instead of being included in the optimization process.

In short, there is the need for decision support systems that can on the one hand optimize integrated P&P mills, correctly weighting the different company's KPIs (key performance indicators) and on the other hand, generate robust production plans that can accommodate a variety of disturbances.

In our case the production system includes a single digester and a single paper machine. An illustrative production plan is depicted in Figure 2, exhibiting the desired synchronization between the digester and the paper machine.

4. SIMULATION MODEL

4.1 General Approach

In this study the simulation model is used to analyse the dynamics of the implementation and adaptation of production plans in practice. In dynamic models, time can be represented in a continuous or discrete manner. The P&P production process is mostly continuous, since it involves continuous materials, whose state changes continuously (depending on chemical reactions and mixture properties). However, at the level of detail required (and desirable) by

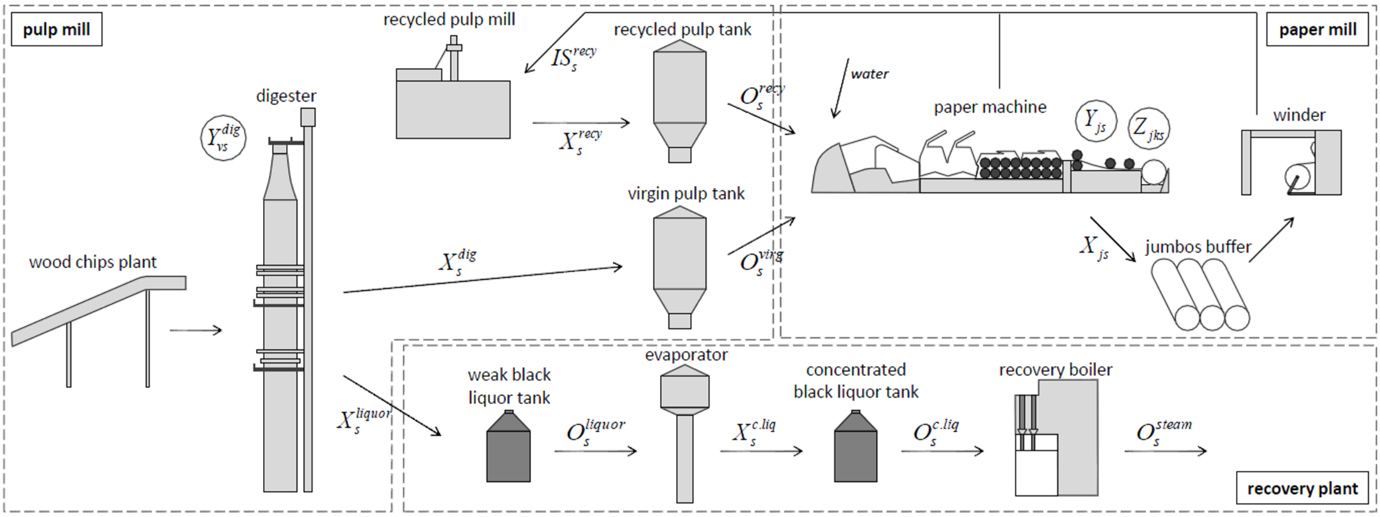

Fig. 1. The integrated pulp, paper and recovery plant (Figueira et al., 2013).

the planning process, the system can be seen as changing its state at discrete points in time and hence, modelled by a discrete-event approach.

In our model, the events are production orders to be implemented. These orders then change the state of the different production resources (entities), which is accomplished when the production task (activity) is executed. The state of an entity may consist of a given quantity of a specific paper grade being produced (in the case of the paper machine), a specified production rate during a certain period of time (e.g. in the digester) or simply a given inventory level (in the case of a tank).

4.2 Stochastic Parameters

Uncertainty and disturbances are included in the simulation model as stochastic parameters. Two different types of parameters are considered:

- productivity losses in the digester, paper machine, recycled pulp mill, evaporator and recovery boiler (rp_i^{dig}, rp_i^{pm}, rp_i^{recy}, rp_i^{evap}, rp_i^{boil}, respectively - where i is a portion of a time period, not necessarily equal to the planning subperiods of the optimization model);
- time between failures and time to repair in the digester and paper machine (tbf^{dig}, ttr^{dig}, tbf^{pm} and ttr^{pm}, respectively).

The distributions of all parameters are estimated based on real data from the company. Some parameters require multiple distributions. For instance, the rate of failures in the paper machine, λ_{pm}, is higher for lower grades and also during grade changeovers, since there is a higher probability of paper breaking. Therefore, a distribution was estimated for each grade and changeover.

In the presence of stochastic parameters, the simulation model has to run multiple times and the results have to be aggregated, so that one can have a certain confidence on them. The confidence level is naturally correlated to the number of simulation replications. However, there are techniques to reduce the variance of results without increasing the number of simulations (see Subsection 5.4). This is particularly important for simulation-optimization algorithms, where the simulation is executed iteratively.

4.3 Simulation Process

The model generally described in the beginning of this section is a special case of discrete-event simulation, since there are no moving entities (only fixed resources, whose state evolves over time). Thus, it might be difficult to implement it in a conventional simulation package. Moreover, there are two important requirements of our model:

- flexibility in the generation of pseudo-random numbers (necessary for variance reduction techniques);
- the highest performance possible, since it is to be included in an optimization algorithm.

Therefore, it was found to be most appropriate an implementation in a general purpose programming language, like C++.

The simulation process is as follows. Production orders are given, one at a time, first to the paper machine, then to the digester, recycled pulp mill, evaporator and finally to the recovery boiler. Then, it restarts with the following order to the paper machine. The amounts to be produced on it are to be kept and hence, their completion times vary (because of uncertainty factors). The plan for the other resources uses the those time periods (in order to keep synchronization), for which given production rates are to be followed. These production rates can however be adapted, according to changes in the state of the system.

After all simulation runs have been executed, results concerning the violation of the tanks are aggregated. Tank violations are computed as the area above the maximum limit and below the minimum limit. Figure 3 shows an example of a tank violation. The measure of the total violation is then fed-back to the optimization and the analytical model is refined accordingly.

4.4 Adaptation of Production Plans

The simulation process previously described, which implements the production orders initially planned, can be

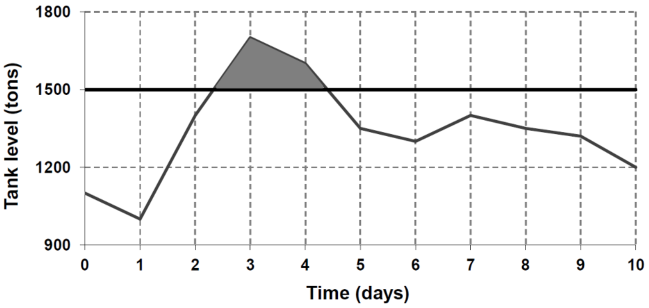

Fig. 3. Tank upper limit violation (shaded area).

alternated with an adaptation procedure. The aim is to replicate the activities that take place in practice, where some adaptations are performed such that the current state of the system (influenced by disturbances and uncertainty) is kept close to the original plan. In this way, the actual behaviour of the mill and its control actions are considered (or at least approximated) and hence, the real impact of production plans is assessed. The adaptation of each order is executed by simple heuristics which are solved before its implementation.

5. SIMULATION-OPTIMIZATION APPROACH

5.1 Analytical Model

The analytical model is presented in detail by Santos and Almada-Lobo (2012) and Figueira et al. (2013), the latter containing small refinements. The formulation is an extension of the general lotsizing and scheduling problem (GLSP), originally proposed by Fleischmann and Meyr (1997), with sequence-dependent setup times and costs and multiple stages. In this model the length of each subperiod s is a decision variable N_s. The time grid that results from these variables is common for all resources and production stages (similar modelling techniques are used by Camargo et al. (2012)).

In addition to the binary variables regarding the sequence-dependent setups, this problem presents the difficulty of the digester's rotation speed (V_s^{dig}) being a decision variable that is constrained to a maximum variation ($\Delta \cdot \Phi$) in each subperiod (to ensure the stability and smoothness of the cooking process). Hence, it has to be explicitly represented in the model. In order to avoid a non-linear formulation, which would result from the multiplication of V_s^{dig} by N_s, a grid of discrete values for the digester's speed (sp_v) had to be defined, with a binary variable (Y_{vs}^{dig}) associated to each of them.

The hybrid method developed by Figueira et al. (2013) thus combines a VNS (which manages the setup-related variables), a specific heuristic to determine the digester's production speed rates and an exact method to optimize the production and flow movement decisions. The heuristic proposed was designed such that the final digester's speeds respected the discrete values of the defined grid, in order to compare the complete solution procedure to exact methods. Here, we determine the speeds in a more plain and accurate way, but not respecting the grid. After solving the first linear program (which relaxes the speed constraints), we fix the subperiods N_s and solve another linear program (now determining V_s^{dig}).

5.2 Parameters Refinement

The aforementioned analytical model had to suffer some modifications, in order to include the stochastic elements that allow to consider disturbances. The stochastic parameters concerning productivity losses, presented in Subsection 4.2, were hence introduced in the corresponding capacity constraints. Failures, however, despite of their important impact in the production system, were not represented analytically, since it would require unnecessary mathematical sophistication. Instead, their effect is aggregated with productivity losses.

These parameters can be refined such that the model accurately represents the system in average terms. Still, considering its variability is imperative. Almeder et al. (2009) have dealt with this issue, aggregating lead-times using quantiles. This strategy is effective when the disturbance affects only one of the tails of the probability distribution. In other words, considering an upper bound (pessimistic scenario) of the lead-time will guarantee with a certain confidence level that the system will work without any disturbance. However, in our case disturbances may appear due to either underestimations or overestimations of the production lead-times, which could result in underflow or overflow of tank levels, respectively.

Variability issues are therefore approached here in a different way. In order to avoid disturbances, we penalize in the objective function the proximity of tank levels to the corresponding limits. As opposed to previous works in the literature, the penalty weights are refined according to the output of a simulation model, more precisely the total expected violation. Thus, both quality (with respect to company's KPIs) and robustness of production plans are correctly weighted in our solution method.

5.3 S-O Framework

Different S-O structures were presented for AME approaches in Section 2. ROSA is a recursive method, where optimization and simulation run alternately (classified as ASO by Figueira and Almada-Lobo (2013)). IFEBM is a pure sequential procedure, with optimization following an initial simulation study (SSO). In OBSIR optimization is solved multiple times within simulation (SOBI).

Our paper proposes a framework that has not yet been studied. The idea is to use simulation to refine the analytical model, while the optimization problem is being solved. The interaction between simulation and optimization can thus be classified as AME-OSBI. Our purpose is to explore simultaneously the convergence to a solution (in the optimization process) and the refinement of the analytical model (performed using simulation). In this way, we reduce both the computational time of optimization, since the problem is solved only once, and that related to simulation, which profits from executing a gradual refinement. Our Progressive Model Refinement framework is based on the assumption that in the beginning of the search procedure it is not necessary to have a very accurate model for selecting improvement moves. As the search progresses, however, the importance of refining the model increases. Therefore, convergence to a solution is performed with increasing accuracy.

5.4 Variance Reduction

In order to increase the confidence on the simulation results without increasing the number of simulations (which results in higher computational effort), variance reduction (VR) techniques are applied. We combine here three of these techniques: stratified sampling (SS), Latin hypercube sampling (LHS) (McKay et al., 2000) and antithetic variates (AV) (Hammersley and Morton, 1956).

The first two techniques are similar and consist in defining n_i strata for each parameter i, all with the same probability of occurrence. Then, in SS the same number of samples (typically one) are generated from every stratum, whereas in LHS not all strata are covered (only one stratum per variable is chosen). SS thus seems to be appropriate when the combination of samples is critical, as for the time between failures. LHS is applied to the other parameters. Both schemes are then combined, i.e. for each SS sample of the time between failures, an LHS sample is generated for all the remaining parameters.

AV is combined with both SS and LHS. The idea is to use symmetric random seeds in the sampling process. This means that for every sample path obtained, its antithetic path is also considered, thus requiring the definition of an even number of strata in SS and LHS.

6. RESULTS AND CONCLUSIONS

In this paper we have proposed a novel simulation-optimization framework for tackling a production planning problem subject to disturbances. A case study of a real pulp and paper mill was examined. The specificities of the production system required new ways to refine the analytical model. Moreover, the complexity of the model made the utilization of recursive approaches impractical. Therefore, a progressive refinement was employed.

With our approach we are able to generate production plans for the whole mill which consider all the relevant KPIs, as well as robustness against disturbances. Both aspects are accurately weighted due to the refinement strategy of the solution method. Experimental tests have revealed that our algorithm outperforms the original method given by Figueira et al. (2013), which does not consider uncertainty. In fact, as opposed to the latter, it is able to effectively mitigate disturbances while generating production plans that succeed in all KPIs.

Our framework can be applied to a large variety of problems that are subject to uncertainty, particularly those represented by difficult models. The methodology followed in this case study may also give insights to other problems in the P&P industry and other similar process industries.

REFERENCES

Almeder, C., Preusser, M., and Hartl, R. (2009). Simulation and optimization of supply chains: alternative or complementary approaches? *OR Spectrum*, 31, 95–119.

Alsholm, O. and Pettersson, B. (1969). Integration of pulp mill control with the production planning of the paper mill. *Trans. of the BP&BMA Symposium*, 209–231.

Asmundsson, J., Rardin, R., and Uzsoy, R. (2006). Tractable nonlinear production planning models for semiconductor wafer fabrication facilities. *Semiconductor Manufacturing, IEEE Transactions on*, 19(1), 95–111.

Camargo, V., Toledo, F., and Almada-Lobo, B. (2012). Three time-based scale formulations for the two-stage lot sizing and scheduling in process industries. *Journal of the Operational Research Society*, 63, 1613–1630.

Figueira, G. and Almada-Lobo, B. (2013). Hybrid simulation-optimization methods: A taxonomy and discussion. Submitted for publication (unpublished).

Figueira, G., Santos, M., and Almada-Lobo, B. (2013). A hybrid vns approach for the short-term production planning and scheduling: A case study in the pulp and paper industry. *Computers & Operations Research*. DOI: 10.1016/j.cor.2013.01.015.

Fleischmann, B. and Meyr, H. (1997). The general lotsizing and scheduling problem. *OR Spectrum*, 19, 11–21.

Hammersley, J.M. and Morton, K.W. (1956). A new monte carlo technique: antithetic variates. *Mathematical Proceedings of the Cambridge Philosophical Society*, 52(03), 449–475.

Jung, J.Y., Blau, G., Pekny, J.F., Reklaitis, G.V., and Eversdyk, D. (2004). A simulation based optimization approach to supply chain management under demand uncertainty. *Computers & Chemical Engineering*, 28(10), 2087–2106.

Leiviska, K. (1996). Simulation in pulp and paper industry. Technical report, University of Oulu, Control Engineering Laboratory.

Leiviska, K. and Niemi, T. (1988). Production scheduling in pulp mills. In *Pulp and Paper Symposium 'Paskov 88, Mars*, 29–31. Ostrava, Czechoslovakia.

McKay, M.D., Beckman, R.J., and Conover, W.J. (2000). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 42(1), 55–61.

Mercangoz, M. and Doyle-III, F.J. (2008). Real-time optimization of the pulp mill benchmark problem. *Computers & Chemical Engineering*, 32(45), 789 – 804.

Pettersson, B. (1970). Optimal production schemes coordinating subprocesses in a complex pulp and paper mill. *Pulp and Paper Magazine of Canada*, 71(5), 59–63.

Santos, M.O. and Almada-Lobo, B. (2012). Integrated pulp and paper mill planning and scheduling. *Computers & Industrial Engineering*, 63(1), 1–12.

Schumacher, K. and Sathaye, J. (1999). India's pulp and paper industry: Productivity and energy efficiency. Technical report, LBNL-41843.