Solution Methodology for Scheduling Problems in Batch Plants

Samuel Moniz,[*,†,§] Ana Paula Barbosa-Póvoa,[†] Jorge Pinho de Sousa,[‡,§] and Pedro Duarte[‖]

[†]Centro de Estudos de Gestão, Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisboa, Portugal
[‡]INESC TEC (formerly INESC Porto), 4200-465 Porto, Portugal
[§]Faculdade de Engenharia da Universidade do Porto, 4200-465 Porto, Portugal
[‖]Hovione FarmaCiencia SA, 2674-506 Loures, Portugal

S *Supporting Information*

ABSTRACT: This paper proposes a solution methodology for the production scheduling of batch plants. The methodology is defined by an integrated approach that simultaneously considers the representation of the scheduling problem, the optimization model, and the decision-making process. A problem representation and a mixed integer linear programing (MILP) model are developed and applied to solve a real world scheduling problem from the chemical—pharmaceutical industry. The main advantage of this approach is that it includes a general process representation that can be used across several departments of the company. Moreover, we also discuss development and implementation challenges of optimization methods for the process industry, and we provide some guidelines to mitigate existing problematic issues in this domain.

1. INTRODUCTION

Decision-making in the process industry tends to be inherently complex, since it may involve strategic, tactical, and operational decisions in very dynamic manufacturing systems. In particular, planning and scheduling decisions have a huge importance due to their interdependency with other functions, such as sales, procurement, production execution, and control. Hence, the integration of optimization methods to support these decisions caught the interest of many industrial companies. Model-based applications are seen as a way to improve competitiveness, to increase profitability, and also to reshape the product portfolio and to facilitate product and process innovations.[1]

In the past years, many academic and industrial efforts have been done to develop and implement model-based applications in manufacturing systems.[2] The major achievements in the area include exact, nonexact and hybrid methods, and also conceptual frameworks, ontologies, and problem representations. Exact methods include mixed integer linear programming (MILP) and mixed integer nonlinear programming (MINLP). Nonexact methods are usually based on heuristics, meta-heuristics, and artificial intelligence approaches. Hybrid methods combine the previous methods so as to build more efficient approaches. On the other hand, conceptual frameworks aim at defining the scope of the different problems addressed by the process systems engineering (PSE) community and aim at proposing general integration schemes. Ontologies attempt to clarify concepts and their relations. Finally, general problem representations attempt to provide unified and unambiguous views of planning and scheduling problems.

Although these developments clearly represent a huge progress on the integration of optimization methods with the decision-making processes, there are some open issues that have recurrently been reported by the literature. The most common ones are related to the computational performance, modeling uncertainty, multiscale optimization, or the modeling task itself. The modeling challenge is surely a complex issue, since it deals with the design of models targeting their integration with the companies decision-making processes.[3]

In this paper, we propose a methodology for the integration of scheduling model-based approaches with the decision-making processes, which was tested in real settings. In particular, we address a scheduling problem in the chemical—pharmaceutical industry from an integrated perspective. Issues related to the problem description, modeling, and implementation of scheduling models in batch plants are discussed and considered in the proposed methodology. This work was strongly motivated from the need of solving in an integrated way and in close collaboration with a company, their day-to-day scheduling problem. This need is addressed in this paper, and the methodology adopted is characterized, which has proven to result when solving real problems.

The rest of the paper is structured as follows: Background presents a literature review on conceptual frameworks and optimization methods. Methodology Conceptual Framework and Scheduling Methodology—Application describe the proposed methodology. We start by defining the concepts used, and then we apply the solution methodology in a real world scheduling problem from the chemical—pharmaceutical industry. In Implementation, we discuss the challenges related to the adoption of optimization methods by the industry and we present some implementation guidelines. Finally, in Conclusions, we present concluding remarks.

2. BACKGROUND

2.1. Conceptual Frameworks. Planning and scheduling are surely two critical activities performed by industrial companies. They involve the allocation of limited resources to operations

that occur in given time windows. Pinedo[4] defined scheduling as a decision-making process that deals with the allocation of resources to tasks over time, considering one or more objectives. Stephanopoulos and Reklaitis[5] defined process planning and scheduling as a subarea of process systems engineering (PSE) that deals with models, methods, and tools for supporting technical decisions related to the safety, efficiency, and reliability of the execution of the manufacturing functions of a process industry enterprise. These definitions are wide enough to incorporate relevant interactions with strategic areas, such as sales and forecasting and with operational areas such as production execution, control, and dispatching. Several authors have explored this area and have proposed conceptual frameworks (depicted in Appendix) that we will briefly describe for a better understanding of the planning and scheduling functions.

Considering a logistics perspective, Meyr et al.[6] presented the supply chain planning matrix, where planning activities are categorized in terms of time horizon and process: (a) the long-term planning, called strategic network planning, deals with the structure of the supply chain; (b) the midterm planning is responsible for the determination of production targets, distribution of the production, and capacity management; and (c) the short-term planning accounts for production planning and scheduling (i.e., operational decisions such as lot-sizing and tasks sequencing).

On the process operations, Bassett et al.[7] presented a decision-making hierarchy that integrates planning, scheduling, and control. The perspective supported by the authors is that model-based methodologies offer the most effective framework for integrating all these decisions. Nevertheless, due to the variety and scope of strategic and operational decisions, the authors claim that no single model would be sufficient to handle all aspects. Pinedo[4] proposed a similar framework for generic manufacturing systems. Scheduling is positioned between production planning and shop-floor control. The decision-making process is clearly hierarchical and allows bidirectional information flows. The planning process starts with a master production planning for determining the production quantities and due dates and to do the initial assessment of the production capacity. This data goes into the materials requirements planning (MRP) that is responsible for launching orders and ensuring that the materials required for production are available. The scheduling function receives the orders from the MRP and performs the sequencing. Orders are then dispatched following the production execution. The closed-loop information flow reinforces the possibility to revise the scheduling, the MRP, or the master production planning whenever necessary, and therefore, the system accuracy.

On the production execution, Harjunkoski et al.[8] and Engell and Harjunkoski[9] presented the automation pyramid for discussing the integration of planning, scheduling, and control layers. The bottom level of the pyramid is composed by the control systems/sensors and is mainly related to hardware/software components. The middle level is the manufacturing execution system (MES) and deals with more advanced production control algorithms, scheduling, maintenance, inventory control, quality assurance, materials, and energy control. The top of the pyramid is in general based on the enterprise resource planning (ERP) system and is concerned with the long-term strategic and tactical planning decisions, performing business-related functions such as available-to-promise (ATP) checks. According to the authors, these levels are not fully standardized and their integration heavily depends on the characteristics of each company.

Standards are also being used for the definition of concepts and to perform the integration of the various subsystems of manufacturing environments. Two standards from ANSI/ISA (S88 and S95) are often referred in the literature. ANSI/ISA-88[10] provides models for integrating information related to the control of batch processes, and ANSI/ISA-95[11] has models for the integration of enterprise and control systems. The purdue reference model, presented in the S95 standard, describes the main components of an enterprise system, their functionalities, and interactions.

From a functional point-of-view, decision-making processes require infrastructures capable to effectively support information gathering, data integration, and models development, as mentioned by Venkatasubramanian et al.[12] These authors propose an information centric infrastructure based on an ontology to support product and process development of active-pharmaceutical ingredients (API). This approach provides a coherent knowledge base that can be used by software tools and models to promote information sharing. On the same line, Muñoz et al.[13] developed an ontology for batch processes, considering the scheduling and the control levels in a closed loop. Although the results presented by these authors are very interesting, substantial challenges will surely arise when implementing these frameworks in industrial facilities.

Klatt and Marquardt[1] presented an overview of methods and tools developed in the context of PSE. The authors argue that the development of user-friendly tools for industrial practitioners is still necessary. With a similar opinion as Bassett et al.,[7] Klatt and Marquardt[1] state that emphasis should be put on model-based applications and in the development of methodologies in which the economic impact and advantages are obvious at first glance.

Although the considerable achievements done by academia in the development of new scheduling formulations and encouraging solution approaches, the adoption of optimization planning tools by the industry is still poor.[14] Reasons for this are related to the way the information context is considered by these tools and are associated with an inadequate definition of the business process workflows. Stephanopoulos and Reklaitis[5] recognized that there are important research opportunities in the development of high level but flexible representations of the scheduling problems and innovative graphical representations, which would promote the adoption of advanced planning systems.

2.2. Models for Scheduling. Significant academic achievements have been done concerning modeling and solving batch planning and scheduling problems. Some relevant recent reviews on this topic provide a rather comprehensive survey on the domain.[2,15−19]

Mendez et al.[2] classified scheduling problems according to the network of processing tasks. Thus, we may have sequential and network processes. In sequential processes, the task-batch entity is preserved, thus batch mixing and splitting are not allowed. On the contrary, network processes have arbitrary topologies and allow batch mixing and splitting. Looking just at models suitable for network processes, we may have continuous-time formulations based on unit specific events[20−23] or based on global events.[24−27] Relevant contributions have also been made in what concerns discrete-time models,[28−34] and on the comparison of discrete-time and continuous-time models see refs 35−37.

2.3. Solving Real-World Scheduling Problems. Several works addressing real world scheduling problems and at different types of industries can be identified in the literature, where different models and solution approaches have been proposed.

In the pharmaceutical industry, Amaro and Barbosa-Póvoa[38] proposed a sequential modeling approach for the planning and scheduling of supply chains. The solution approach is applied to a real pharmaceutical supply chain producing several products such as injection drugs, tablets, and oral suspensions. Multistage multiproduct scheduling problems have been tackled by Kopanos et al.,[39] Stefansson et al.,[40] and Castro et al.[41] Kopanos et al.[39] and Castro et al.[41] have used similar decomposition strategies that reduce the computational complexity of the scheduling problem by scheduling orders sequentially and then improve the schedule by applying reordering procedures. Kopanos et al.[39] proposed general precedence and unit-specific general precedence models, while Castro et al.[41] proposed a unit-specific continuous-time formulation. Stefansson et al.[40] compared a discrete-time formulation based on refs 28 and 29 and a general precedence continuous-time formulation based on ref 42 in a scheduling problem of a secondary pharmaceutical production system. To tackle the combinatorial complexity of the MILP models, the authors have applied a decomposition algorithm that prioritizes the scheduling of the bottleneck stage. Susarla and Karimi[43] developed a unit slot continuous-time model to the campaign planning problem of the pharmaceutical industry, giving emphasis to the decision-making process. They studied several real scenarios considering different resources allocation profiles, safety stock limits, minimum campaign lengths, maintenance actions, and sequence-dependent changeovers. They remark that although planning is usually performed by the planning department, it is a collaborative process that seeks data from several other departments (sales, procurement, laboratory, maintenance, and higher management).

Addressing the steel making process scheduling, Harjunkoski and Grossmann[44] developed a decomposition algorithm that relies on splitting the original problem into smaller subproblems by exploring its special structure. The algorithm involves three MILP models and one LP model solved in a sequential solution strategy. Later, addressing also a scheduling problem of a steel plant, Harjunkoski et al.[45] discussed the implementation issues and benefits of planning and scheduling optimization. The authors state that reusability, flexibility, and configurability are relevant aspects that must be considered when encapsulating mathematical models in software applications to be used in industrial environments.

For the production scheduling in the polymer industry, Schulz et al.[46] formulated a discrete-time model and a continuous-time nonlinear model (MINLP) for a real case of a chemical batch plant producing expandable polystyrene. Algorithms have been developed to produce solutions in reasonable time. Considering the same scheduling problem, Wang et al.[47] have applied a genetic algorithm and Till et al.[48] addressed uncertainty by using a two-stage stochastic integer programming model and proposing a hybrid evolutionary algorithm to solve the stochastic problem. Castro et al.[49] explored the optimal periodic schedule of a resource constrained industrial problem of the pulp industry, through the use of discrete-time and continuous-time Resource-Task Network (RTN) based mathematical formulations. Adequate solution strategies were proposed for both formulations. While the exact optimal solution to the problem was achieved using the discrete-time formulation, the same was not true for the continuous-time formulation.

Considering enterprise-wide optimization (EWO) in complex production systems, Wassick[50] presented the case of Dow Chemical Company and discussed opportunities for the integration of design, planning, and scheduling optimization models in the industry. The problem of waste disposal scheduling is solved using the RTN discrete-time formulation. Moreover, the author presented some useful considerations on modeling and implementation. So, to the company, the choice of the discrete-time RTN relied on the simplicity and generality of the formulation due to a uniform treatment of all resources. A creative definition of the production resources allows solving a variety of scheduling problems without changing the model (and the code). Simple linear representations of the processes are adequate for long time frames or more strategic decisions, but for short time frames or operational decisions, it becomes necessary to account for the nonlinearities of the chemical processes. For this author, the greatest modeling challenge concerns capturing complex operating policies. In these cases, it is recommended to negotiate simplifications with the decision makers, instead of dealing with complicated constraints. Moreover, the process design, planning and scheduling integration, and the representation of uncertainty and risk, should be viewed as critical. An interesting case-specific aspect about the implementation is that during the first year of operation, the scheduling model was used together with the existing scheduling procedure, in order to compare both methods and to gain confidence in the model.

In general, Applequist et al.[51] pointed out four practical issues that make planning and scheduling problems particularly difficult to address, namely, (a) social considerations: manufacturing is considered a cooperative activity in the company, and the "strategies and scheduling" function is viewed as having the responsibility to orchestrate this cooperation; (b) dynamic nature: the active environment of the manufacturing system requires flexible and scalable planning and scheduling tools that must be able to adapt to different production scenarios; (c) information intensity: even relatively small planning and scheduling problems require a considerable amount of data, and this creates additional complications concerning data management; and (d) intrinsic combinatorial character: leading to significant mathematical challenges to solve these problems. The relevance of these issues, and the practical need to address them, are the main motivation for this work.

2.4. Integration of Optimization Models in the Industry. As referred, it is clear that there is still a lot of work to do concerning the integration of optimization models in the decision-making processes of the companies. The literature addressing models and solution approaches for planning and scheduling problems is mainly focused on time performance and comparison between models. When addressing the quality of the solutions, few confront the solutions obtained by the models with the solutions obtained by the planners, and just a small number of works address practical issues associated with modeling and implementation of optimization methods in industrial companies. Models are rarely evaluated in the context where the information is available, thus they simply do not consider the internal decision-making processes of the company. This interaction is missing and it would surely provide valuable information for improvement of the optimization methods.

Although we also consider that, in practical terms, time to obtain solutions is the most critical issue of many models solving scheduling problems, research should also focus in other aspects that are determinant to integrate optimization methods in industrial practices. The development of systematic approaches for structuring the scheduling problems and the inclusion of the decision-making processes into models that can rapidly be understood by the industrial practitioners are essential aspects that have been somehow neglected.

3. METHODOLOGY CONCEPTUAL FRAMEWORK

In this section, we discuss the components of the methodology and then, in Scheduling Methodology—Application, we present the resolution of a real world problem from the chemical–pharmaceutical industry to demonstrate its applicability.

3.1. Key Components. The methodology is defined by three main components, as shown in Figure 1. The problem representation
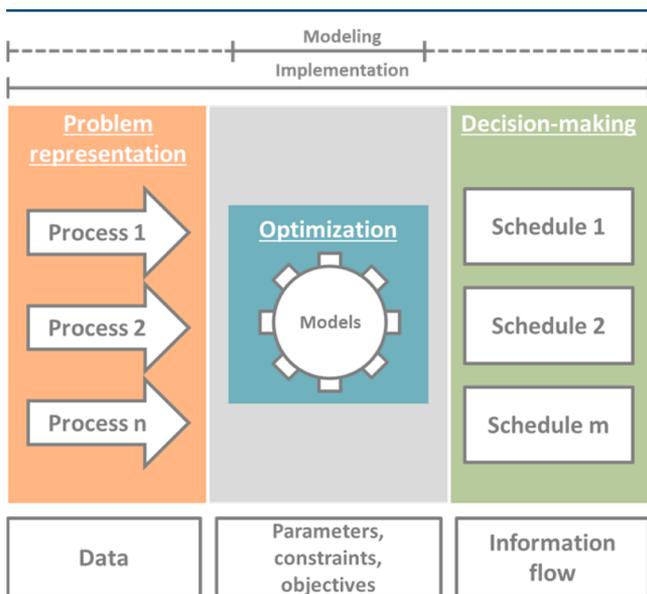

Figure 1. Conceptual framework for the solution methodology.

component that is related to the interpretation of the scheduling problem, and it is used as an interface with the decision-makers and to capture data for the optimization model. The optimization component that deals with case-specific models and solution approaches developed to solve the scheduling problem. Finally, the decision-making component that has to do with the analysis of the solution pool provided by the optimization component and involves the visualization and user-interface interactivity required to support the decision-making process.

We consider two main sets of tasks: implementation and modeling. Implementation includes all tasks required to place the application running in the company, and it is typically assigned to IT consultants. Its scope must be wide enough to incorporate all components that in the end will constitute the decision-making tool. Modeling concerns the definition of the model and is usually a task performed by academics and researchers. The scope of modeling is in many situations limited to the development and test of models. Thus, it disregards the context where the data is created and gathered and mainly, how decisions are made. In our view, when addressing real-world optimization problems, the scope of the modeling task must be broadened to include the data context and to encompass the information flow of the decision-making processes. The modeling task must be extended to define more complex interactions between the representation of the problem and the decision-making process (see Figure 1). This will surely require a deeper collaboration between academics, industrial practitioners, and IT consultants.

Note that although the focus of this methodology is on scheduling problems, we think that it can also be applied to other types of problems.

The components and their interactions will be discussed in detail in the following subsections.

3.2. Representation of the Scheduling Problem. As did Bassett et al.,[7] we also view scheduling as an integration activity. Accordingly, scheduling problems should be represented so that different types of knowledge can be captured in a coherent way. For that, different scheduling views may be necessary in order to ensure a comprehensive representation of the problem. Such representation can then have several layers (or views) and must be able to be integrated with any model or solution approach. Grossmann et al.[52] argue that the application of mathematical programming approaches to process design, and synthesis problems require the development of superstructures for the representation of the alternatives, regardless of the detail of the model. We can say that this reasoning is also valid to planning or scheduling problems, since these problems use similar super-structures.

Although there is a general consensus that models/solution approaches must be adapted to the specific features of each case, we believe that it is possible to develop general representations of the scheduling problems that could be used later by different models. A good example of this is the state-task network (STN) of Kondili et al.[28] and the RTN of Pantelides[30] representations that are being applied to represent a variety of scheduling problems and are used by many different formulations. For example, extensions of STN to account for design and operational decisions were developed by Barbosa-Póvoa and Macchietto,[31] and Amaro and Barbosa-Póvoa[53] developed the chain-STN to solve supply chain problems. In this way, scheduling representations could evolve independently from the model formulations and provide a coherent representation of the problems. This research direction has been recently followed by Maravelias.[54] This author proposes a framework for the description of scheduling problems in chemical industries based on the characteristics of the processes.

Furthermore, the representation of the scheduling problem is typically based on the process structure, in which the level of abstraction is a critical issue. High detail representations of the processes may allow the development of more detailed models and reach theoretically optimal solutions but may result in models that are computationally intractable. In practice, this approach requires the involvement of industrial and academic specialists and the integration of different types of knowledge, which will easily turn into a very time-consuming task. On the other hand, less detailed processes result into more simple models that are easier to solve and to manage but may result in infeasible schedules. In summary, a careful exploitation of the problem structure is required in order to keep the balance between these trade-offs, and here a close collaboration between academics and industrial planners must exist.

In the PSE community, planning and scheduling problems appear closely connected to process development and design problems.[15] The process design focus is to define the characteristics of products, the production tasks, and the specifications of processing units. The planning and scheduling problem often take the design into account and seek the effective use of the enterprise resources.[5] In this work, we have developed a comprehensive representation of the scheduling problem that captures the characteristics of the processes and available equipment, defining superstructures with all possible production alternatives. The representation of the scheduling problem is then a key part of the scheduling methodology that integrates with the optimization and the decision-making components.

3.3. Optimization. A single model would hardly be sufficient to address all types of planning and scheduling decisions. Thus, if

a model-based approach is followed then the methodology must be able to include several models, in which the links between those models play a crucial role. Furthermore, since the decision-making processes vary from company-to-company, case-specific models may also be developed.

Concerning the computational complexity, many scheduling models solving real-world problems are considered too large to be solved to optimality in affordable time. This is due to the combinatorial nature of the problems, associated with binary decisions such as task-unit assignments, tasks sequencing, changeovers, and storage tasks. Problems with a significant number of tasks and processing units and considering long scheduling horizons tend to be difficult to solve with exact methods. In these cases, alternative solution approaches can be applied to obtain satisfactory solutions in reasonable time. A discussion on scheduling models and solution approaches has been presented in Background.

In summary, models and solution approaches should be built taking into account the characteristics of the problems and the decision-making processes of the companies, thus defining concise methodologies that integrate mathematical approaches with existing decision-making procedures and resulting in solvable models that represent adequately the real problem.

3.4. Decision-Making. The planning information flow presented by Pinedo[4] and the Purdue reference model are two comprehensive frameworks where the complexity of the planning and scheduling becomes more evident. To address this complexity, the development and implementation of decision-support tools should start by addressing the core decision-making processes of the company. The approach may vary from company-to-company but should always involve academics and industrial practitioners.

The scheduling process must be supported by adequate information flows between those participating in the decision-making process. Assuming that scheduling is performed collaboratively, the scheduling methodology (problem representation and optimization, decision-making) should ensure that the necessary data is available and up to date before being used by the optimization model. In this way, the methodology must be integrated transversally in the company, since it is common that several functions in the company can use that information and benefit from it. To clearly explain this methodology, the application to a case study is detailed below.

4. SCHEDULING METHODOLOGY—APPLICATION

The proposed conceptual framework is now applied to a real case-study from the chemical—pharmaceutical industry. The main goal of this exercise is to demonstrate how the components (representation of the scheduling problem, optimization, and decision making) can be designed in order to implement a decision-support methodology for solving scheduling problems. For that purpose, we briefly describe the context of the scheduling problem and a typical scheduling decision-making process. We then present the methodology that was implemented in our case study and discuss the main decisions involved in that process. The optimization model and results are also presented.

4.1. Scheduling Problem in Chemical-Pharmaceutical Industry. The chemical—pharmaceutical industry is responsible for the development and manufacturing of fine chemicals called active pharmaceutical ingredients (API). Manufacturing such products involves complex and long processes that are executed under a close supervision of the regulatory authorities, with

responsiveness of the manufacturing system and cost reduction being two critical aspects.

Production may simultaneously include products that are under development and products that are in commercialization, and the plant resources may be shared between these products. Products are associated with recipes that describe in detail the production processes. Recipes are defined by a network of production tasks that must be executed to manufacture a product[55] and include process data such as materials consumption and production proportions, tasks processing times, and characteristics of the units required by each task. Recipes differ from product-to-product (i.e., tasks sequencing and material flows are product specific). To manufacture a single product, several days of effective production time may be required, with tasks processing times varying between 1 h and 2 days.

Cleaning of processing units, pipelines, and other resources are needed to avoid cross contamination of the products. Therefore, changeovers between lots of the same product and between lots of different products are present and may impose significant downtime periods.

Often the chemical—pharmaceutical industry relies on general purpose batch plants with multipurpose processing units between which connections are usually not fixed. Instead, units are organized in such a way that almost all connectivity options are possible. Operations flexibility is achieved through multipurpose units, capable of executing a variety of chemical tasks, as well as, through the connections between units, that can be changed when there is a change from a product to another product (i.e., connections between units can change with the production demand).

In such plants, the most common units are reactors, filters, and dryers of different volumes, packaging rooms, and auxiliary equipment like condensers, temperature systems, cleaning in place (CIP), vacuum pumps, etc. that may be attached to the processing units, thus changing their configuration with additional characteristics important for the task-unit assignment. For instance, a reactor is defined by its maximum and minimum volume, type (glass lined or stainless steel) and also by the agitation system, the temperature system, CIP, etc. The material flows are established through a complex system of pipelines and mobile vessels. Furthermore, people are critical resources and are usually considered in the medium and short-term scheduling. This happens because tasks require specialized manpower to execute or control the production.

The planning and scheduling functions are typically a responsibility of the planning department of the company. However, other areas, such as sales, procurement, and R&D, are also involved. These areas contribute with data inputs relevant for planning and scheduling and may as well do the analysis of the schedules.

Planning is in general done for a time horizon of up to 12 months and involves the determination of the production quantities and a preliminary allocation of the processing units to products. But because production capacity is defined at the level of the processing units, planning is referred here as medium-term scheduling. The medium-term scheduling tends to be stable, at least for the next months, and is revised every month or whenever an unexpected event that has impact on the plan appears.

There is also the short-term scheduling that has a time horizon of up to 2 weeks and is revised on a daily basis. Data from the medium-term term scheduling is used as a reference for building the short-term scheduling, namely in what concerns recipes, inventory, and product demand. Decisions at this level refer to
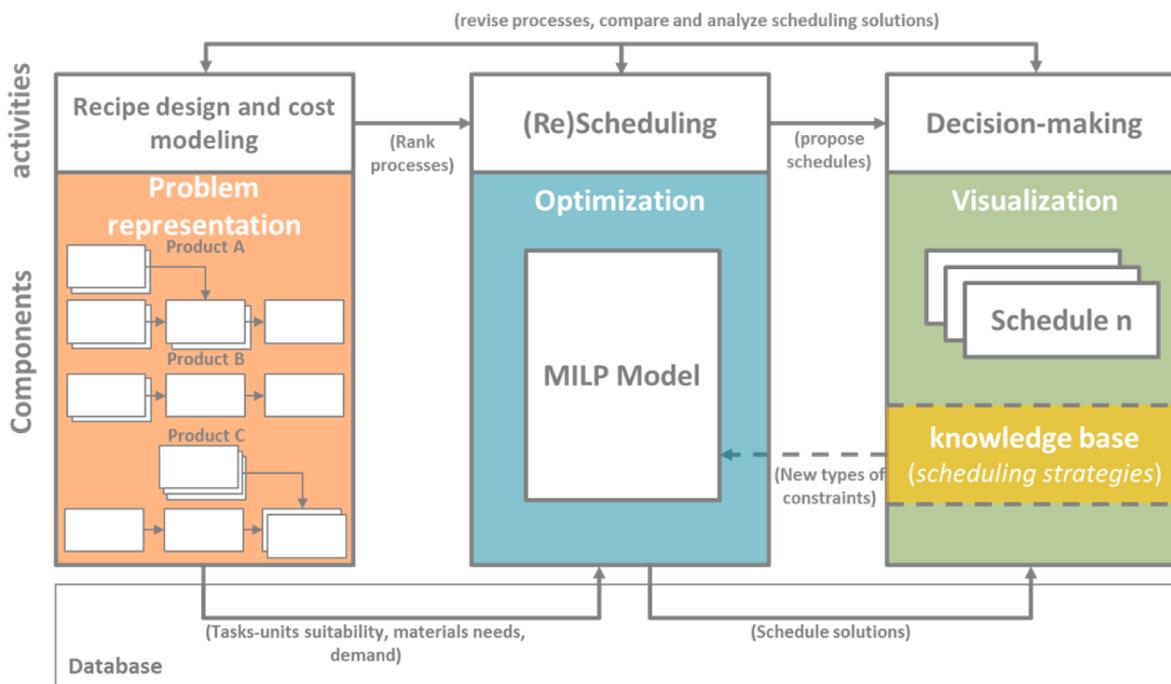
Figure 2. Proposed scheduling methodology.

the assignment of tasks to units and to the determination of the exact time when tasks are going to be executed. Industrial planners are therefore challenged to obtain the "best" set of processing units to manufacture each product and to obtain effective task sequencing, taking into account objectives related to cost and total production time. Since recipes may have a large number of tasks and tasks may be processed through multiple units of different capacities, where sequence-dependent change-overs must be respected, scheduling decisions become extremely complex.

The status of the tasks execution is continuously checked and potential delays are evaluated. Products that are under development add more complexity to the scheduling function, since they regularly impose the revision of the schedule. Schedule deviations that do not have any further impact in the plan are promptly solved, while significant delays trigger the revision of the medium-term schedule.

Although we assume here that medium and short-term scheduling are two distinct problems, they are linked because they constrain each other. Both scheduling problems should therefore use a unique problem representation, defined by the methodology proposed in this work (see Proposed Scheduling Methodology). This will ensure that with respect to data, both problems use the same structures, although the detail level of the models can be different. Our focus is on the short-term scheduling problem, as defined in detail in the next section.

4.2. Problem Statement. The problem consists then in finding the optimal scheduling of multipurpose batch plants, in which products have arbitrary network structures.

Given (a) the recipes of the products (materials flows and proportions, tasks processing times and characteristics); (b) the processing units (including all characteristics that define the task-unit suitability); (c) the resources availability (intermediaries, final products and processing units for every time interval); (d) the demand (quantities and delivery due dates); (e) the minimum and maximum allowed lot sizes; (f) the scheduling time

horizon; (g) the costs (storage, changeover, and missed deliveries); and (h) the economic value of the products.

Our goal is to obtain optimal production schedules by determining: (a) the task-unit suitability for a given process; (b) the task-unit assignment of the production schedules; (c) the lot sizes and product deliveries; (d) the materials inventory levels; and (e) the optimal process schedule.

In this context, we have used profit maximization and cost minimization objectives, defining a short-term scheduling problem that can be solved by a linear model, with deterministic data.

4.3. Proposed Scheduling Methodology. The conceptual framework depicted in Figure 1 gave origin to the scheduling methodology shown in Figure 2. The three components of the methodology (problem representation, optimization, and decision making) are now framed by the associated activities [recipe design and cost modeling, (re)scheduling, and decision making] that interact and provide/receive data to/from the optimization model. These activities were identified in our case study as being core activities that have a huge relevance in the scheduling decision-making process. In the following subsections each component will be addressed in detail.

Note that this methodology views the scheduling activity as an interactive and collaborative process that may involve several departments of a company. Thus, the involved departments may provide data to the process and revise and analyze scheduling solutions.

In order to test and validate the proposed methodology, we have, during one year, performed meetings on a regular basis with process engineers and planners. Insights from industrial practitioners revealed to be very useful in redefining the components of the methodology and the integration requirements between those components.

4.3.1. Representation of the Scheduling Problem. This component aims at providing a standard representation of the processes in such a way that they can be readily understood by all the participants in the scheduling problem. Having this goal in mind, we developed novel representation of the processes in

CIP:	YES
Sampling:	YES
Vacuum:	NO
Stirring:	GOOD
Stirring vol.[L]:	60
Temp. system:	Type 1

1st step →

Reaction
Material: acid
Duration [H] : 8
Volume [L] : 5000

2nd step ← {U1,U3,U4}

Charact.\Unit	U1	U2	U3	U4	U5
Type	acid(res)	basic(res)	all	acid(res)	basic(res)
CIP	•	•			•
Sampling	•		•		
Vacuum		•		•	•
Stirring vol.	50	20	100	40	110
Max vol.	10000	3000	6000	5000	2000
Temp. control	Type 1	Type 2	Type 2	Type 1	Type 3
Cost	1000	350	1750	900	1100
...					
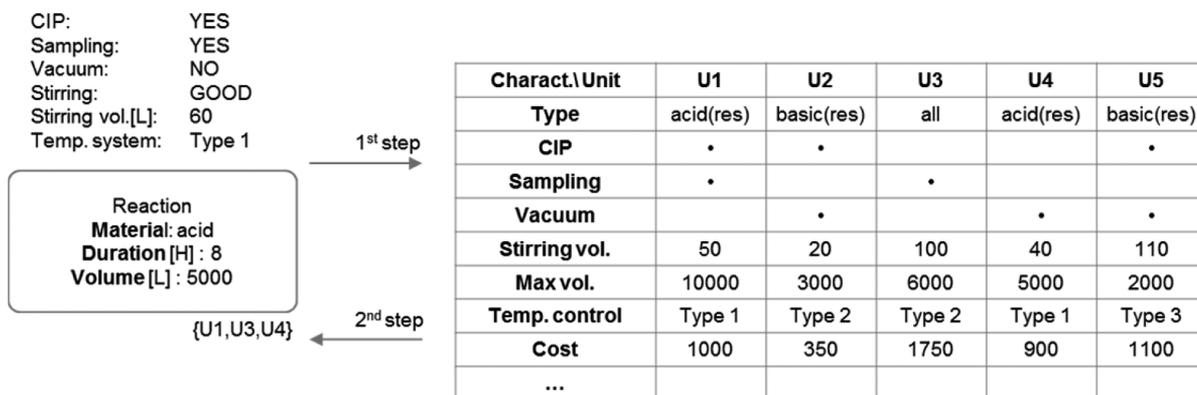
Figure 3. Mapping between tasks and processing units characteristics.
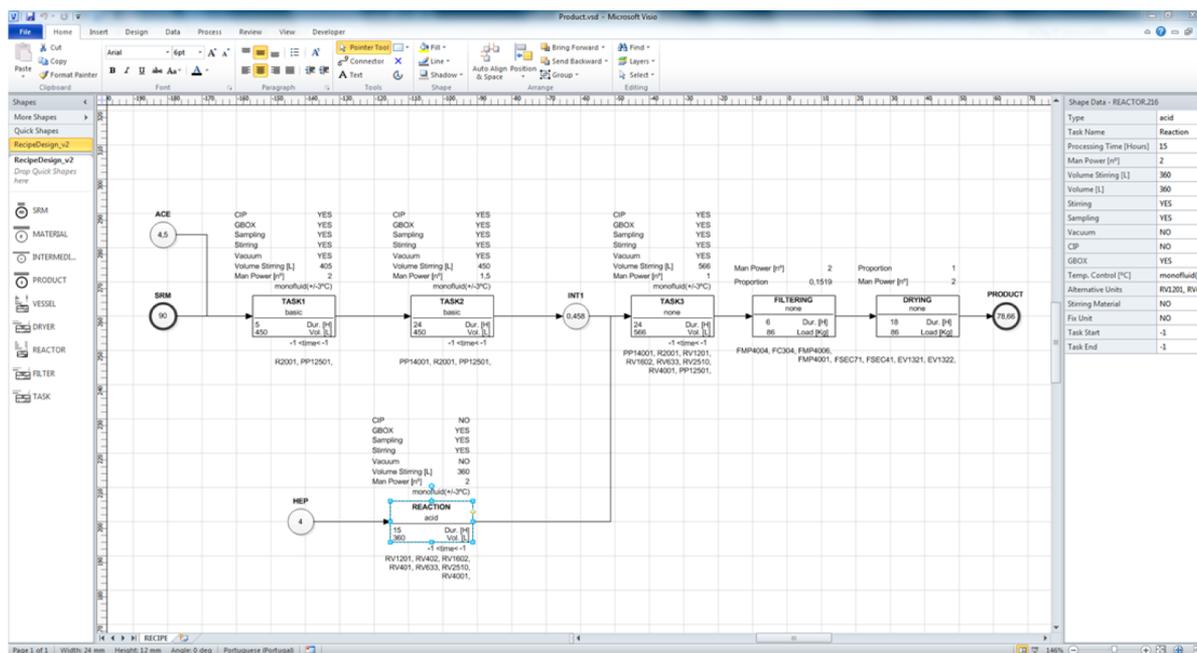

Figure 4. Example of a process representation using the recipe design tool.

which emphasis is given to the definition of the task characteristics that will determine the task-unit assignment.

In a way similar to the STN, we design recipes through a network of tasks and states. As shown in Figure 3, tasks (represented as rectangles) are defined through an object that considers all the characteristics (e.g., volume, task duration, need for CIP, sampling, vacuum pump, etc.) that are relevant to the determination of the suitable processing units. For example, in Figure 3, if the task has an acid material, only units U1, U3, and U4 can be used. But if CIP and sampling systems are also required, only U1 can be used. The states define the materials and are represented by circles.

Processes are represented by the definition of tasks, material states, and material flows using directed arrows, in a prototype developed in Microsoft Visio. Predefined objects are available to support the process design. So, as a first step, the description of the process is done taking into account just the characteristics of the tasks, and in a second step these characteristics are automatically mapped into the characteristics of the existent processing units for the determination of suitable units. More advanced rules can be used to account for approximate (roughly defined) characteristics such as the need for very good, good, and

normal stirring. Additionally, the cost modeling of the processes can also be performed, taking into account the resources involved. This can, for example, be used for ranking the alternative processes based on their cost.

Having defined the tasks and associated tasks to units, a global process representation is obtained (see Figure 4). The recipe design tool shown in Figure 4 is a prototype developed in Microsoft Visio that allows an immediate assessment of the process concerning the determination of the units capable of manufacturing it. In the left-hand side of the screen is shown the library of standard objects, in the middle we can see the process (in this case, we have a process with 6 tasks), and in the right the characteristics of the selected task are presented. The processing units suitable for each task are depicted just below the rectangle and were automatically determined through the task-unit matching characteristics as explained above.

All the data of this process is saved in a database that can be used later by any optimization system. Thus, the problem representation and optimization components are indeed independent.

Challenges here related to the data management, since many data inputs are often required to describe a single process, and to

produce a representation of processes that can be used by all participants. The tests performed with the company demonstrated that the developed interface (as shown in Figure 4) is user-friendly and can be used by all involved departments and that the design of the processes is quite fast. These characteristics are fundamental for the planners since they are determinant to an effective usage of the tool.

4.3.2. Optimization. After having the processes represented in the recipe design tool, data structures are automatically generated and can be used by the scheduling model. With this approach, recipe design decisions can promptly be revised and integrated with the scheduling model.

The development of the model is case specific. However, the data structures facilitate the development and testing of new models in the company and ensure data consistency. In this work, we have applied a MILP model; however, other models could also be used. Alternative models, constraints, and objective functions that have been considered under this methodology can be found in Moniz et al.[56]

The optimization stage is related to the scheduling function, but in real manufacturing systems, this function is mainly used for rescheduling. In fact, when this function is performed, either processing units are executing tasks or planned orders have already been allocated for the near future. Then when new orders arrive, planners may have to revise the current schedule, compare and analyze scheduling solutions, or even evaluate alternative processes. Scheduling responsiveness is ensured here by an immediate assessment of the alternative processes and by integration with the optimization model.

The solutions delivered by the model can be quantified not only by the value of the objective function but also by the computation of several key performance indicators (KPIs) done in the postprocessing phase of CPLEX. For example, we may have KPIs for measuring the free capacity and volume usage of the reactors or the missing deliveries of a schedule.

The scheduling model used in the developed framework is based on the formulation proposed by Moniz et al.[56] The present model extends the previous work by considering explicitly the lots production through the definition of lot size continuous variables and lot production binary variables. Also lots backlog is now modeled, an aspect that is quite important in the industrial reality. This is a MILP model where time is uniformly discretized along the scheduling horizon of interest. One particular characteristic of this model is that the material balance constraints consider explicitly the inventory carried out by each task and production lots. Production lots refer to the amount of stable intermediary or final product manufactured through a known set of tasks, units, and materials, so as to keep record of lots blending operations, thus ensuring lots traceability. By following this approach, new types of constraints for modeling temporary storage in the processing units and sequence-dependent changeovers can be derived.

The indices, sets, parameters, and decision variables used by the formulation are described in Appendix.

Model. The scheduling problem considers a scheduling horizon with length T, divided into time intervals $t \in H$ of equal and fixed duration. Scheduling decisions are made through the N_{klt} task-unit assignment/sequencing binary variables that are equal to 1 if task k of lot l starts at time interval t; Y_l lot production binary variables, equal to 1 if l lot is produced; ξ_{klt} task batch size continuous variables that determine the batch size of task k of lot l at time interval t; β_l lot size continuous variables that state the amount of product manufactured in lot l; the R_{krlt}^p and R_{krlt}^c

continuous variables that define the materials r production (p) and consumption (c) for each task k of lot l and time interval t; R_{krlt} continuous variables that give the resultant materials r availability; $\prod_{krlt}$ continuous variables for product r deliveries, given by task k, lot l and time interval t; and B_{rt} continuous variables of backlogged demand, determined for product r and time interval t.

The model is defined by constraints 1 that express the fact that either processing units are allocated to production tasks or to storage operations. In other words, constrains 1 define the task-unit assignment and sequencing and the temporary storage in the processing units. In the chemical industry, it is common to find processes in which material storage may occur in the processing unit where the material has been produced. In these cases, units work temporarily as storage vessels until all material is consumed by subsequent tasks of the process. The first term of the constraints does the task-unit assignment and sequencing, while the second term indicates if the processing unit is performing storage ($R_{krlt} \geq 0$) for the intermediaries produced by task k and subject to non-intermediate storage policy (NIS) I_k^{NIS}.

Constraints 2 determine the amount of resource r (intermediaries and final products) produced, and constraints 3 give the amount of resource r consumed (intermediaries) by task k of lot l at each time interval t. Parameters $\nu_{kr\theta}^p$ and $\nu_{kr\theta}^c$ give the materials production and consumption proportions of the batch size of task k for resource r. Constraints 4 express the materials balance for each resource r (intermediary or final product), task k and lot l. The amount of resource r available R_{krlt} in each task k of lot l is equal to the amount stored in the previous time interval, plus the amount produced R_{krlt}^p, minus the amount consumed R_{krlt}^c, plus the amount that is delivered $\prod_{krlt}$. Note that $\prod_{krlt}$ takes negative values for product deliveries and that we assume no receipts of materials occur during the scheduling horizon. Constraints 5 bound the resource r availability to a maximum value given by parameter R_{rt}^{max} and are only defined for intermediaries subject to finite intermediate storage (FIS), zero-wait (ZW), and unlimited intermediate storage (UIS) policies. Concerning the materials temporarily held on the processing units (NIS), the amount of material that can be stored is bounded by the maximum capacity of the unit. Constraints 6 ensure that the tasks batch size ξ_{klt} is within the minimum V_{krl}^{min} and maximum V_{krl}^{max} capacities of resource r (processing unit) for task k and lot l. Constraints 7 impose that the total amount of product manufactured must be equal to the lot size β_l, and constraints 8 bound the lot size β_l between the minimum q_l^{min} and maximum q_l^{max} allowed size for lot l. Constraints 9 define that lot l can only be produced if lot $l-1$ has been produced.

Backlogged demand B_{rt} is defined by expressions 10, where B_{rt} will take a value greater than zero whenever a product delivery D_{rt} is not fulfilled, partially or totally.

Sequence-dependent changeovers are required whenever cleaning and units setup operations need to be performed, when changing the production to a new product or lot. Thus, constraints 11 state that if task k of lot l occurs at time interval t, then the first term is equal to one, and the second is forced to be zero for all tasks k' of lot l' and time intervals corresponding to $t - \tau_{k'} - \theta$.

Constraints 12 define that tasks must finish in the time horizon of interest. Constraints 13 impose that delivery variables $\prod_{krlt}$ cannot take values either for the time intervals outside the delivery time windows or for resources other than final products. And expressions 14 define the variables domain.

$$\sum_{k \in K_r} \sum_{l \in L_r} \sum_{t'=t-\tau_k+1}^{t} N_{klt'} + \sum_{k \in K_r^p} \sum_{r' \in I_k^{NIS}} \sum_{l \in L_r} \left(\frac{R_{kr'lt}}{V_{kr'l}^{max}} \right) \le 1$$

$$\forall\, r \in E,\, t \in H \tag{1}$$

$$R_{krlt}^{p} = \sum_{\theta=0}^{\tau_k} (\nu_{kr\theta}^{p} \xi_{kl,t-\theta}) \quad \forall\, r \in I \cup P,\, k \in K_r^p,\, l \in L_r,\, t$$

$$\in H \tag{2}$$

$$\sum_{k \in K_r^p} R_{krlt}^{c} = \sum_{k \in K_r^c} \sum_{\theta=0}^{\tau_k} (\nu_{kr\theta}^{c} \xi_{kl,t-\theta}) \quad \forall\, r \in I,\, l \in L_r,\, t \in H \tag{3}$$

$$R_{krlt} = (R_{krl}^{init}|_{t=0},\, R_{krl,t-1}|_{t>0}) + R_{krlt}^{p} - R_{krlt}^{c} + \Pi_{krlt}$$

$$\forall\, r \in I \cup P,\, k \in K_r^p,\, l \in L_r,\, t \in H \tag{4}$$

$$0 \le \sum_{k \in K_r^p} \sum_{l \in L_r} R_{krlt} \le R_{rt}^{max} \quad \forall\, r \in I \backslash I^{NIS} \cup P,\, t \in H \tag{5}$$

$$V_{krl}^{min} N_{klt} \le \xi_{klt} \le V_{krl}^{max} N_{klt} \quad \forall\, r \in E,\, k \in K_r,\, l \in L_k,\, t \in H \tag{6}$$

$$\sum_{k \in K_r^p} \sum_{t \in H} R_{krlt}^{p} = \beta_l \quad \forall\, r \in P,\, l \in L_r \tag{7}$$

$$Y_l q_l^{min} \le \beta_l \le q_l^{max} Y_l \quad \forall\, l \in L \tag{8}$$

$$Y_{l-1} \ge Y_l \quad \forall\, l \in L: l > 1 \tag{9}$$

$$B_{rt} = (B_{r,t-1}|_{t>0}) + D_{rt} + \sum_{l \in L_r} \sum_{k \in K_r^p} \Pi_{krlt} \quad \forall\, r \in P,\, t \in H \tag{10}$$

$$\sum_{k \in f_l^r} N_{klt} + \sum_{k' \in f_{l'}^r} N_{k'l't-\tau_{k'}-\theta} \le 1 \quad \forall\, r \in E,\, l,\, l'$$

$$\in L_r,\, \theta = 0,\, ..,\, c_{rl'l} - 1,\, t \in H \tag{11}$$

$$\sum_{t=T-\tau_k+1}^{T} N_{klt} = 0 \quad \forall\, k \in K,\, l \in L_k \tag{12}$$

$$\Pi_{krlt} = 0 \quad \forall\, r \in P,\, l \in L_r,\, d \in D_r,\, k \in K_r^p,\, t \in H \backslash DW_{rd}$$

$$\Pi_{krlt} = 0 \quad \forall\, r \in I,\, L \in L_r,\, k \in K_r^p,\, t \in H \tag{13}$$

$$DW_{rd} = \{t | r \in P,\, d \in D_r,\, t \in H: T_{rd}^{dd} \ge t \ge T_{rd}^{ed}\}$$

$$R_{krlt}^{p},\, R_{krlt}^{c},\, R_{krlt} \in \mathbb{R}_+ \quad \forall\, r \in I \cup P,\, k \in K_r^p,\, l \in L_r,\, t \in H$$

$$\xi_{klt} \in \mathbb{R}_+ \quad \forall\, r \in E,\, k \in K_r,\, l \in L_k,\, t \in H$$

$$\Pi_{krlt} \in \mathbb{R}_- \quad \forall\, r \in I \cup P,\, l \in L_r,\, k \in K_r^p,\, t \in H \tag{14}$$

$$Y_l \in \{0, 1\} \quad \forall\, l \in L$$

$$N_{klt} \in \{0, 1\} \quad \forall\, k \in K,\, l \in L_k,\, t \in H$$

The objective functions considered in this work are the minimization of cost (see expression 15), that involves the storage, operational, backlog, and lot costs, and the maximization of profit (see expression 16), reflecting the economic value of the products.

Storage costs are associated with holding costs of intermediaries and products during the scheduling horizon 16a. Operational costs are related to the assignment of processing units to tasks are defined by expression 16b. Backlogged demand costs are given by expression 16c, and lot fix and variable costs are given by expression 16d. The economic value of the products is given by expression 16e.

$$\min Z1 = \text{storage costs (SC)} + \text{operational costs (OC)}$$
$$+ \text{backlog cos ts (BC)} + \text{lot cos ts (LC)} \tag{15}$$

$$\max Z2 = \text{value of the products (VP)}$$
$$- \text{storage cos ts (SC)}$$
$$- \text{operational cos ts (OC)}$$
$$- \text{backlog cos ts (BC)} - \text{lot cos ts (LC)} \tag{16}$$

$$SC = \sum_{r \in I \cup P} \sum_{l \in L_r} \sum_{k \in K_r} \sum_{t \in H} (c_r^{sto} R_{krlt}) \tag{16a}$$

$$OC = \sum_{k \in K_r} \sum_{l \in L_r} \sum_{t \in H} (c_k^{op} N_{klt}) \tag{16b}$$

$$BC = \sum_{r \in P} \sum_{t \in H} c_r^{bk} B_{rt} \tag{16c}$$

$$LC = \sum_{l \in L} (c_l^{fix} Y_l + c_l^{var} \beta_l) \tag{16d}$$

$$VP = \sum_{r \in P} \sum_{l \in L_r} \sum_{k \in K_r^p} \sum_{t \in H} (-\nu_r \Pi_{krlt}) \tag{16e}$$

4.3.3. Decision Making. The scheduling solutions obtained through this model are then evaluated by experienced planners and other participants involved in the scheduling process. In practice, it is desirable to produce more than one schedule, even for the same objective function, as the model may not represent the real problem due to simplifications considered. In some situations model constraints are linear approximations or aggregations designed to keep the problem computationally tractable. Thus, it could happen that solutions might not be preferred by planners or could be considered operationally infeasible. The assessment of the model will then be made by the quality of the solutions delivered and the computational time required to produce them.

To avoid this problem, several scheduling solutions are generated and compared during the decision-making process. For example, multiple scheduling solutions can be obtained by using the CPLEX solution pool feature.

Although the most common constraints of the scheduling problem are known and well-described in the literature, new types of constraints are often necessary when trying to solve real world scheduling problems. To address this issue, a knowledge base is kept with the purpose of describing new scheduling rules that are empirically followed by the planners. The scheduling rules are then evaluated and, if applicable, are converted into model constraints. Thus, the model is composed by a set of constraints that can be activated or deactivated in order to convey to the preferences of the planner. For example, case-specific extensions for dealing with layout, manpower, and maintenance constraints may be developed in the model.
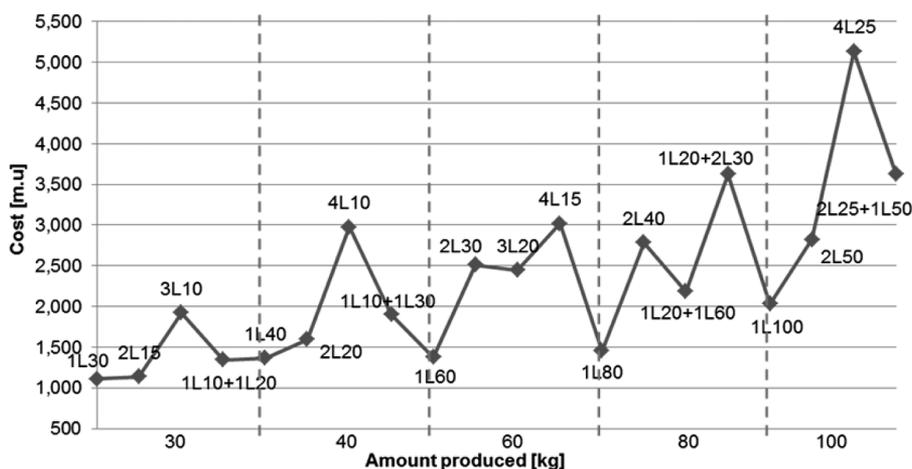
Figure 5. Production lots cost in monetary units (m.u.) for product P1 (1L30 = 1 lot of 30 kg; 2L15 = 2 lots of 15 kg; and 1L10 + 1L20 = 1 lot of 10 kg plus 1 lot of 20 kg).

Finally, the user interface for the scheduling solutions plays also an important role, since the dynamic nature of the scheduling process requires the visualization of a considerable amount of information, as well as advanced interactive options. In our case, a prototype of a Gantt chart was built in Microsoft Excel to allow the test and assessment of the scheduling solutions. The evaluation of the produced pool of schedules is then supported by Gantt charts and additional indicators (see below), allowing the planners to choose the most adequate schedule for real implementation.

4.4. Case-Study Description and Results. As mentioned above, the case under analysis concerns the scheduling problem in a chemical−pharmaceutical industry where different products are to be produced. The decision-making tool is used to obtain optimized production plans and to perform an evaluation of alternative production processes. The scheduling decisions involve the determination of the processing units to be used by each process, the lot sizes, the total amount produced, and the delivery dates. Here, recipe design decisions (as described in Representation of the Scheduling Problem) are integrated with scheduling decisions in order to predict the impact of the process in the shop-floor, guaranteeing that units are used in an efficient way. For example, the selection of the processing units to execute a given process can be done taking into account the scalability, completion time, number of processing units used, material flows, and cost of the processes.

Since a discrete time model is being used, a critical modeling decision concerns the definition of the length for the time intervals. A thin discretization of time would theoretically result in better solutions but may conduct to models that are very difficult to solve. In our case, we assume a time interval (grid) of 8 h, since tasks processing times can be roughly approximated by multiples of 8. Moreover, schedules having time intervals of 8 h (1 working shift) work well in practice. Computational tests and discussions with the planners have shown that the computational time required to solve the short-term scheduling model is quite reasonable and acceptable in practice. The model was implemented using ILOG/CPLEX version 12.5.1, running on an Intel Xeon X5680 at 3.33 GHz with 24 GB of RAM.

In the following two subsections, we present some results that show the utilization of our scheduling model (defined in Optimization). Initially, we perform an analysis of the production processes involved and we discuss trade-offs related to the determination of the lot size, which are important to the scale-up

strategies followed by the chemical−pharmaceutical industry. Then, we derive short-term production schedules for time horizons of 1 and 2 weeks.

The network studied in this work considers four processes, responsible for the production of four products (P1 to P4) and that may share 9 processing units (7 reactors of different characteristics and 2 filters-dryers). The total number of tasks is 40, with processing times varying between 6 and 64 h. Processes data are available in the Supporting Information.

4.4.1. Evaluation of Alternative Processes. The impact of the lots definition on cost, capacity, and time is the first analysis to be performed. These are relevant indicators concerning the scale-up of the lot size. In practice, the company wants to determine what are the best units among all alternatives for producing a process in terms of cost, reactors capacity used, and completion time for different lot scale-up scenarios. Leading to different alternative processes. On the basis of this analysis and on the evaluation of full production schedules (see Scheduling Solutions), units are then assigned to the process, which is then submitted to validation. Figure 5 depicts the minimum cost of manufacturing 30, 40, 60, 80, and 100 kg of product P1, considering different lot sizes (e.g., 3 lots of 10 kg = 3L10). Table 1 shows some numerical values used in this analysis. These results were obtained by running the scheduling model with the cost minimization objective function (min Z1) and assuming the scheduling horizon of 1 week, discretized into 21 time intervals of 8 h.

When analyzing Figure 5 and Table 1, it can be seen that the production of 30 kg of product P1 has a cost always below 2000 m.u. The most costly production case is to consider 3 lots of 10 kg (3L10), with a cost of 1922.5 m.u., which leads also to the highest completion time. This indicates that the processing units are used inefficiently, as can be seen by the value of 27100 L of reactors capacity allocated to the process (see Table 1). Assuming a production of 40 kg, the costs increase by 23% considering just one lot; however, the completion time goes from 112 h to just 40 h. Again with the increase of the number of lots the process requires more reactors capacity and takes more time. An increase of the production to 60 kg (assuming a single lot) has almost no effect in the cost, and has no impact on the completion time and no impact on the reactors capacity allocated. Interestingly, the production of 3 lots of 20 kg results in a lower cost, in a lower reactors capacity, and in a lower completion time, when compared with the production of 2 lots of 30 kg. The production of 80 kg (assuming a single lot) increases the costs by

Table 1. Costs, Reactors Capacity Used, and Completion Time for Product P1[a]

amount produced (kg)	process/lots	cost (m.u.)	reactors capacity used (l)	completion time (hours)
30	P1/1L30	1108.6	15600	112
	P1/2L15	1134.6	15600	128
	P1/3L10	1922.5	27100	152
	P1/1L10 + 1L20	1341.5	19300	88
40	P1/1L40	1365.0	20300	40
	P1/2L20	1595.0	23000	64
	P1/4L10	2973.5	42300	136
	P1/1L10 + 1L30	1903.0	27100	144
60	P1/1L60	1375.0	20300	40
	P1/2L30	2511.8	35900	144
	P1/3L20	2446.0	34500	88
	P1/4L15	3021.6	42300	136
80	P1/1L80	1455.0	20800	40
	P1/2L40	2782.2	40600	72
	P1/1L20 + 1L60	2185.0	31800	80
	P1/1L20 + 2L30	3623.2	51100	160
100	P1/1L100	2030.0	28600	88
	P1/2L50	2823.7	40600	72
	P1/4L25	5131.2	71400	168
	P1/2L25 + 1L50	3631.2	51100	152

[a]1LOT30 = 1 lot of 30 kg; 2LOT15 = 2 lots of 15 kg; and 1LOT10 + 1LOT20 = 1 lot of 10 kg plus 1 lot of 20 kg, etc. Reactors capacity used = total reactors capacity allocated to the process $\Sigma_{k \in K_r} \Sigma_{l \in L_r} \Sigma_{t \in H} (V_{krl}^{max} N_{klt})$ calculated during the post processing phase of CPLEX.

6% of the 60 kg production case but keeps the same completion time, while the production of 100 kg (assuming a single lot) increases the cost by 40% (assuming a single lot) and more than doubles the completion time, when compared with the 80 kg production case.

As conclusions, it can be said that the trade-offs associated with the scheduling decisions are related to the task-unit assignment and storage costs. The allocation of tasks to the smaller capacity units may result in lower operational costs but may lead to longer completion times, since tasks may need to occur multiple times or a higher number of changeovers may be required; both situations potentially leading to an increase in the storage costs.

4.4.2. Scheduling Solutions. Following the previous analysis, we now derive full schedules for the production of the four products, with different lot sizes and scheduling horizons of 1 and 2 weeks. Since all products are scheduled simultaneously, the trade-offs discussed above become more complex, resulting in the Gantt charts of Figure 6.

The first instance (INST1) is depicted in Figure 6a and considers the production of just one lot in a scheduling horizon of 1 week (21 time intervals of 8 h). Figure 6b shows the second instance (INST2) also based on a single lot, but now considering a scheduling horizon of 2 weeks (42 time intervals of 8 h). Figure 6c depicts the third instance (INST3) for the production of 2 lots of each product, in a scheduling horizon of 2 weeks. Table 2 summarizes the cost structure of each instance. The objective function utilized is the profit maximization max Z2 (see expression 16).

The most compact schedule is obtained with the instance INST1 (see Figure 6). Processing units need to accommodate the demand in just 1 week, and this leads to a high occupation rate of the processing units. By extending the scheduling horizon to 2 weeks (INST2), units can be used more efficiently, with a reduction of the costs and an associated profit increase (see Table 2).

For INST1, the profit is equal to 48438.8 m.u., while INST2 has a profit of 49441.4 m.u., which is at least 2% higher, since the solution of INST2 potentially has some margin for improvement because it is not an optimal solution (3.32% gap). Moreover, the schedule of INST1 is inherently more complex to execute in practice, since several production tasks are repeated in order to fulfill the demand, see Figure 6a). For example, in INST1, it can be seen that TASK1 of product P1 occurs 5 times, while in INST2, see Figure 6b), this task occurs only 2 times.

In instance INST3, we have defined 2 lots for each product in a scheduling horizon of 2 weeks. This scenario tends naturally to impose additional idle periods for the units, as a consequence of the changeover periods; nevertheless, the profit is comparable with the one obtained in INST1. Looking at the lot size decision variables of instance INST3, Product P1 has lots with 56 and 64 kg, resulting in a total amount of 120 kg. The demand of product P2 was 70 kg, resulting in one lot of 16.8 kg and another of 53.2 kg. The lots of product P3 have 23.5 and 26.5 kg for fulfilling a demand of 50 kg, and product P4 had a demand of 60 kg that was fulfilled through lots of 28 and 32 kg. As discussed in the above section, the lot size has impact on the task-unit assignments, thus lots of the same product may have been assigned to different processing units. Globally, INST3 has lower storage costs but has higher operational and lot setup costs.

Concerning the model performance, the computational time to obtain solutions is surely the main drawback when solving large instances. The numerical results demonstrate that CPLEX version 12.5.1 could not prove optimality for INST2 and INST3 during a computational time of 3600 s. However, it should be noted that, from a practical perspective, these solutions have been considered very satisfactory by the industrial practitioners.

5. IMPLEMENTATION

Although a huge progress has been recently achieved on the development of new mathematical formulations, conceptual frameworks, and ontologies, the implementation of optimization models in the industry is far from being trivial. The implementation challenges are due to a wide spectrum of issues that are related to (a) understanding of the model capabilities and limitations by the industrial practitioners; (b) definition of the model specifications and their impact on the decision-making process; (c) definition of the most relevant modeling trade-offs (model detail vs computational time versus quality of the solutions); (d) development of efficient models capable to be used across several functions inside a company; (e) models assessment; and (f) models scale-up to the development of robust software applications.

The industrial context has motivated the present work, and based on the results obtained, some guidelines are now provided that may help academics and industry practitioners on the collaborative development and deployment of optimization tools. In fact, our experience suggests that the implementation of optimization models in real production environments can strongly benefit from a previous development of case-specific models, oriented to confined manufacturing systems. These first developments provide academics and industrial practitioners with the
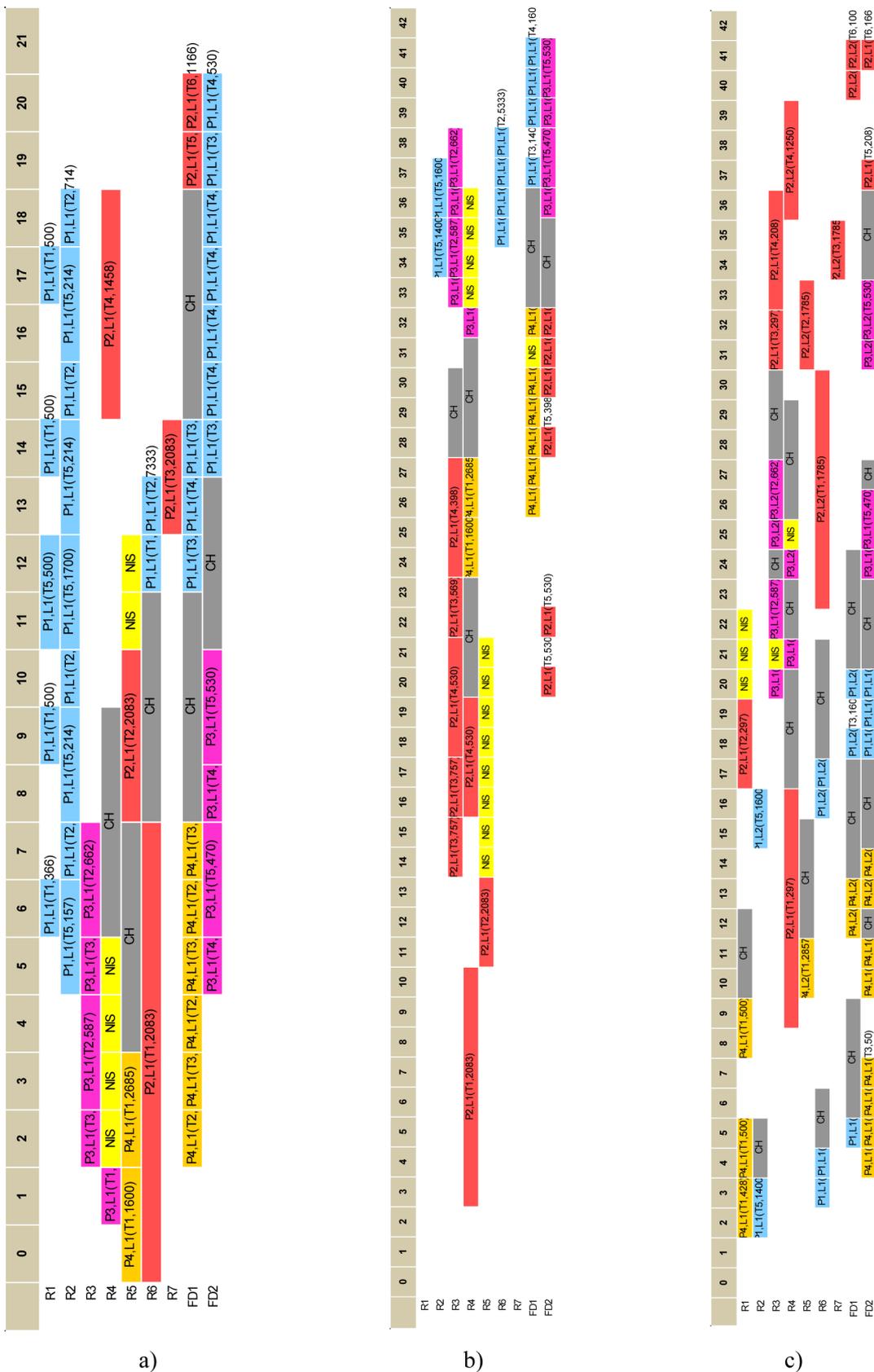
Figure 6. Production schedules: (a) INST1; (b) INST2; and (c) INST3.

necessary knowledge and confidence to address more complex problems. Thus, in general, the implementation approach should go from case-specific to more general models.

Following a model-based approach, as suggested in this paper, the definition of the scheduling problem should be done together with the definition of the model specifications. This will contribute

Table 2. Cost Structure of Each Instance

instance	profit (m.u.)	value of the products (m.u.)	storage costs (m.u.)	operational costs (m.u.)	backlog costs (m.u.)	lot costs (m.u.)
INST1	48438.8	56000.0	1646.2	5520.0	0.0	395.0
INST2[a]	49441.4	56000.0	1413.6	4750.0	0.0	395.0
INST3[b]	48754.7	56000.0	1450.3	5350.0	0.0	445.0

[a]Solution within 3.32% of the optimal solution. [b]Solution within 5.34% of the optimal solution.

Planning and scheduling frameworks
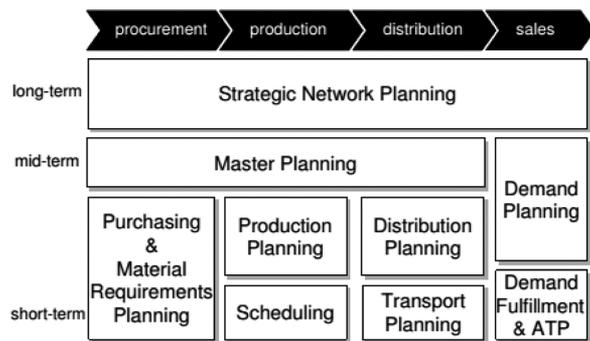
Figure 7. Supply chain planning matrix source.[6]
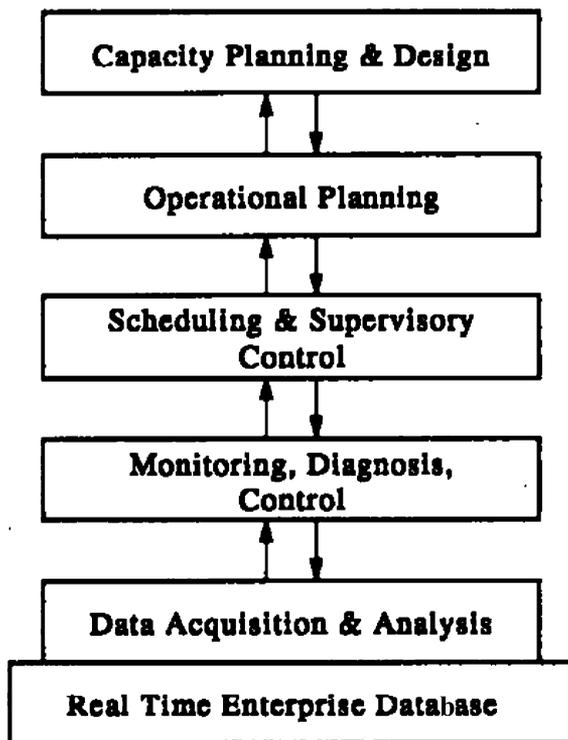
Figure 8. Process operations hierarchy source[7]

Figure 9. Planning information flow in a manufacturing system source.[4]
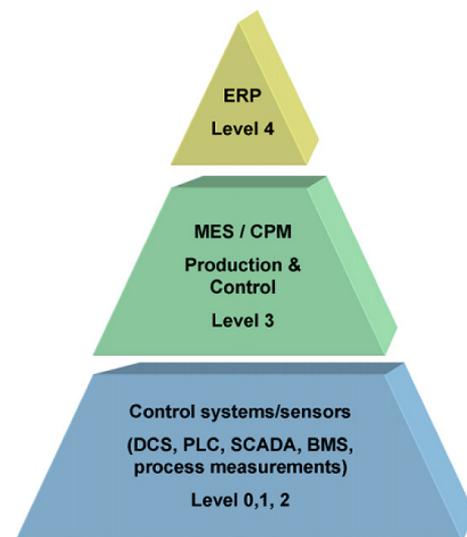
Figure 10. Automation pyramid source.[8]

to the alignment of the problem requirements with the modeling capabilities.

Test and evaluation of models are also critical activities, since they provide valuable information for the identification of the problem constraints and the development of solution approaches, targeting the improvement of the model performance. In this direction, generic and flexible ways of delivering model solutions are required. Powerful prototype visualization tools to enable a fast analysis of solutions can in general be built with a reasonable effort.

The solution methodology proposed takes into account some of these implementation challenges and aims at contributing to the development and deployment of optimization models in industry. A major advantage of the methodology comes from the fact that it is based on components that build the
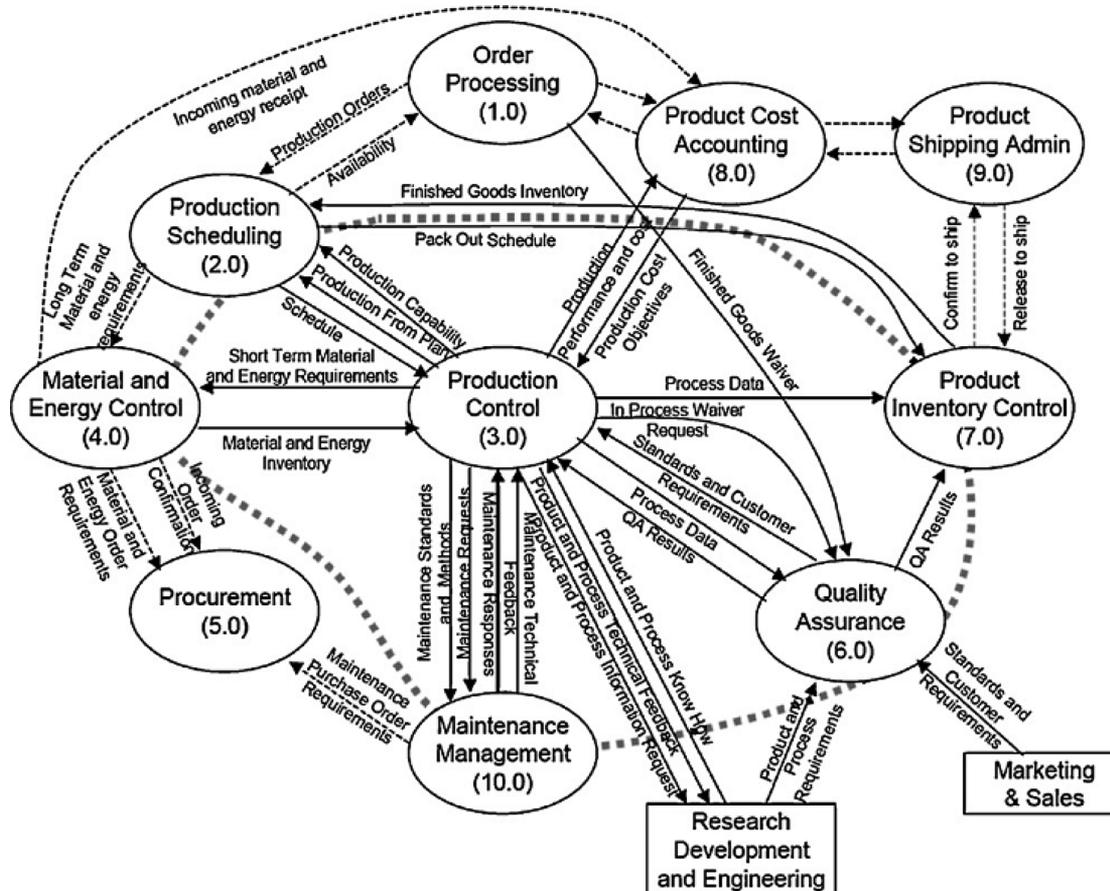
Figure 11. Purdue reference model.[11]

adequate information context (problem representation and decision-making process) and thus support the design of consistent models. We believe that methodologies as the one proposed in this work can clearly lead to further improvements in the area.

6. CONCLUSIONS

The existing planning and scheduling frameworks and ontologies provide a view on general requirements, information flows, decisions, and integration needs of the scheduling decision-making processes. They present building blocks for structuring complex and highly integrated systems. Nevertheless, these frameworks do not yet give an answer to the question of how planning or scheduling should be done in a specific company or industry. This happens because planning and scheduling decision-making processes are usually case-specific, and the available building blocks are not enough to define and integrate optimization tools in companies.

This paper proposes a solution methodology for doing the production scheduling of chemical batch plants, supported by a MILP model. Our methodology has integrated some characteristics of the existing frameworks and was applied to a real case in the chemical−pharmaceutical industry. We have developed a MILP discrete-time model based on the one proposed by Moniz et al.;[56] however, other models could also be used. The data used by the MILP is automatically taken from the process representation tool developed in this work. The model is then run for different scenarios. Scheduling solutions and key performance

indicators are represented in Gantt charts and tables to support the decision-making process.

On the addressed case-study, two types of analyses were done. First, an evaluation of the processes alternatives and their associated costs was performed. Second, production schedules for scheduling horizons of 1 and 2 weeks were produced. Numerical results show that the model performed well in all instances. However, CPLEX could not prove optimality for the larger problem instances. In general, the developed framework proved to be very useful for the company and provided a solid base for structuring the scheduling data.

The proposed methodology has a set of advantages, which are related to the general representation of the scheduling problem that can be used by several departments in the company and to the integration of the decision-making process with the optimization model. In our view, this has been a missing unifying point that could promote the adoption of planning and scheduling optimization tools in the industry. Methodologies should clearly define how tools should be applied and used in the company decision-making processes. In this field, research work is still required to map current planning and scheduling practices into coherent methodologies capable of efficiently using tools, to systematically deliver planning and scheduling solutions.

The experience presented in this work clearly shows the need for new innovative approaches and further levels of cooperation between academia and industry, to address the still open challenges in the adoption of advanced optimization approaches for industrial companies.

■ APPENDIX

Planning and Scheduling Frameworks

Various frameworks have been proposed, including work by Myer and Rohde[6] (Figure 7), Bassett et al.[7] (Figure 8), Pinedo[4] (Figure 9), Harjunkoski et al.[8] (Figure 10), and ANSI/ISA-95[11] (Figure 11).

Notation

Indices

d delivery period
l lot
k task
p product
r resource (processing unit, intermediary or final product)
t time interval

Sets

D_{rt} demand of product r at time interval t
DW_{rld} delivery window of lot l and resource r (final product) at delivery period d
E processing units
H scheduling horizon
I intermediaries
I^{NIS} intermediaries subject to a nonintermediate storage policy
I_k^{NIS} intermediaries produced by task k and subject to a nonintermediate storage policy
L lots
L_r lots associated with resource r
L_k lots associated with task k
K_r tasks that require resource r (processing unit, intermediary, or final product)
K_r^c tasks that consume resource r (intermediary or final product)
K_r^p tasks that produce resource r (intermediary or final product)
P products
R production resources

Parameters

τ_k processing time of task k
$\nu_{kr\theta}, \nu_{kr\theta}^p, \nu_{kr\theta}^c$ production/consumption proportion of resource (intermediary or final product) r in task k at time Θ relative to the start of task
c_r^{sto} cost of storage of products and intermediaries r
c_k^{op} operational costs of task k
$c_l^{fix}, c_l^{var.}$ lot fix and variable costs
$c_{rll'}$ changeover time in processing unit r from lot l to lot l'
f_l^r tasks k associated with processing unit r and lot l
q_l^{min}, q_l^{max} minimum and maximum lot l size
R_{rt}^{max} maximum resource availability of resource r (intermediary or final product) at time interval t
R_{krl}^{init} resource r (intermediary or final product) availability of lot l at task k in the beginning of the planning horizon
T length of the scheduling horizon
T_{ld}^{ed} earliest time interval of lot l at delivery d
T_{ld}^{dd} latest time interval of lot l at delivery d
v_r value of product r
$V_{krl}^{min}, V_{krl}^{max}$ minimum and maximum capacity of resource r (processing unit) for task k of lot l

Variables

β_l amount of product manufactured by lot l — lot size (continuous)
ξ_{klt} batch size of task k and lot l at time interval t (continuous)
$\prod_{krlt}$ delivery of resource (final products) r of lot l at time interval t available from task k (continuous)
N_{klt} binary variables that are equal to 1 if task k starts lot l at time interval t
R_{krlt} resource r (intermediaries or final products) availability, produced by task k of lot l at time interval t (continuous)
R_{krlt}^c amount of resource r (intermediaries or final products) consumed from task k of lot l at time interval t (continuous)
R_{krlt}^p amount of resource r (intermediaries or final products) produced by task k of lot l at time interval t (continuous)
Y_l binary variables that are equal to 1 if l lot has been produced

■ ASSOCIATED CONTENT

⑤ Supporting Information

Networks of processes P1, P2, P3, and P4 and data tables of product demand, delivery dates, and backlog costs; minimum and maximum lot sizes, economic value and lot setup costs; materials initial, minimum, and maximum availability, inventory costs, and storage policy; processing units characteristics; changeovers times of each unit and between products; numerical results of the processes evaluation; and numerical results of the production schedules. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: samuel.moniz@inesctec.pt.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Klatt, K.-U.; Marquardt, W. Perspectives for process systems engineering: Personal views from academia and industry. *Comput. Chem. Eng.* **2009**, *33* (3), 536−550.

(2) Mendez, C. A.; Cerda, J.; Grossmann, I. E.; Harjunkoski, I.; Fahl, M. State-of-the-art review of optimization methods for short-term scheduling of batch processes. *Comput. Chem. Eng.* **2006**, *30* (6−7), 913−946.

(3) Grossmann, I. Enterprise-wide optimization: A new frontier in process systems engineering. *AIChE J.* **2005**, *51* (7), 1846−1857.

(4) Pinedo, M. *Scheduling: Theory, Algorithms, And Systems*; Prentice Hall: Upper Saddle, N.J., 2002.

(5) Stephanopoulos, G.; Reklaitis, G. V. Process systems engineering: From Solvay to modern bio-and nanotechnology.: A history of development, successes and prospects for the future. *Chem. Eng. Sci.* **2011**, *66* (19), 4272−4306.

(6) Meyr, H.; Wagner, M.; Rohde, J. Structure of advanced planning systems. In *Supply Chain Management and Advanced Planning*; Springer: New York, 2005; pp 109−115.

(7) Bassett, M.; Dave, P.; Doyle, F.; Kudva, G.; Pekny, J.; Reklaitis, G.; Subrahmanyam, S.; Miller, D.; Zentner, M. Perspectives on model based integration of process operations. *Comput. Chem. Eng.* **1996**, *20* (6), 821−844.

(8) Harjunkoski, I.; Nyström, R.; Horch, A. Integration of scheduling and control: Theory or practice? *Comput. Chem. Eng.* **2009**, *33* (12), 1909−1918.

(9) Engell, S.; Harjunkoski, I. Optimal operation: Scheduling, advanced control and their integration. *Comput. Chem. Eng.* **2012**, DOI: 10.1016/j.compchemeng.2012.06.039.

(10) ANSI/ISA-88. *Batch Control Part 1, Models and Terminology, (see also IEC 61512-01)*. ISA: Research Triangle Park, NC, 1995.

(11) ANSI/ISA-95. *Enterprise-Control System Integration. Part 1: Models and Terminology*. ISA: Research Triangle Park, NC, 2000.

(12) Venkatasubramanian, V.; Zhao, C.; Joglekar, G.; Jain, A.; Hailemariam, L.; Suresh, P.; Akkisetty, P.; Morris, K.; Reklaitis, G. Ontological informatics infrastructure for pharmaceutical product development and manufacturing. *Comput. Chem. Eng.* **2006**, *30* (10), 1482−1496.

(13) Muñoz, E.; Espuña, A.; Puigjaner, L. Towards an ontological infrastructure for chemical batch process management. *Comput. Chem. Eng.* **2010**, *34* (5), 668−682.

(14) Henning, G. P. Production scheduling in the process industries: current trends, emerging challenges and opportunities. *Comput.-Aided Chem. Eng.* **2009**, *27*, 23−28.

(15) Barbosa-Póvoa, A. P. A critical review on the design and retrofit of batch plants. *Comput. Chem. Eng.* **2007**, *31* (7), 833−855.

(16) Li, Z.; Ierapetritou, M. Process scheduling under uncertainty: Review and challenges. *Comput. Chem. Eng.* **2008**, *32* (4), 715−727.

(17) Maravelias, C. T.; Sung, C. Integration of production planning and scheduling: Overview, challenges and opportunities. *Comput. Chem. Eng.* **2009**, *33* (12), 1919−1930.

(18) Verderame, P. M.; Elia, J. A.; Li, J.; Floudas, C. A. Planning and Scheduling under Uncertainty: A Review Across Multiple Sectors. *Ind. Eng. Chem. Res.* **2010**, *49* (9), 3993−4017.

(19) Kallrath, J. Solving planning and design problems in the process industry using mixed integer and global optimization. *Annals of Operations Research* **2005**, *140* (1), 339−373.

(20) Ierapetritou, M.; Floudas, C. Effective continuous-time formulation for short-term scheduling. 1. Multipurpose batch processes. *Ind. Eng. Chem. Res.* **1998**, *37* (11), 4341−4359.

(21) Janak, S. L.; Lin, X.; Floudas, C. A. Enhanced continuous-time unit-specific event-based formulation for short-term scheduling of multipurpose batch processes: Resource constraints and mixed storage policies. *Ind. Eng. Chem. Res.* **2004**, *43* (10), 2516−2533.

(22) Shaik, M. A.; Floudas, C. A. Improved unit-specific event-based continuous-time model for short-term scheduling of continuous processes: Rigorous treatment of storage requirements. *Ind. Eng. Chem. Res.* **2007**, *46* (6), 1764−1779.

(23) Vooradi, R.; Shaik, M. A. Improved Three-Index Unit-Specific Event-Based Model for Short-Term Scheduling of Batch Plants. *Comput. Chem. Eng.* **2012**, DOI: 10.1016/j.compchemeng.2012.03.014.

(24) Schilling, G.; Pantelides, C. A simple continuous-time process scheduling formulation and a novel solution algorithm. *Comput. Chem. Eng.* **1996**, *20*, S1221−S1226.

(25) Maravelias, C. T.; Grossmann, I. E. New general continuous-time state-task network formulation for short-term scheduling of multi-purpose batch plants. *Ind. Eng. Chem. Res.* **2003**, *42* (13), 3056−3074.

(26) Sundaramoorthy, A.; Karimi, I. A simpler better slot-based continuous-time formulation for short-term scheduling in multipurpose batch plants. *Chem. Eng. Sci.* **2005**, *60* (10), 2679−2702.

(27) Castro, P.; Barbosa-Póvoa, A.; Matos, H. An improved RTN continuous-time formulation for the short-term scheduling of multi-purpose batch plants. *Ind. Eng. Chem. Res.* **2001**, *40* (9), 2059−2068.

(28) Kondili, E.; Pantelides, C.; Sargent, R. A general algorithm for short-term scheduling of batch operations. I. MILP formulation. *Comput. Chem. Eng.* **1993**, *17* (2), 211−227.

(29) Shah, N.; Pantelides, C.; Sargent, R. A general algorithm for short-term scheduling of batch operations. II. Computational issues. *Comput. Chem. Eng.* **1993**, *17* (2), 229−244.

(30) Pantelides, C. C. In *Unified Frameworks for Optimal Process Planning and Scheduling*; Cache Publications: New York, 1994; pp 253−274.

(31) Barbosa-Póvoa, A. P.; Macchietto, S. Detailed design of multipurpose batch plants. *Comput. Chem. Eng.* **1994**, *18* (11−12), 1013−1042.

(32) Pinto, T.; Barbosa-Póvoa, A. P. F. D.; Novais, A. Q. Optimal design and retrofit of batch plants with a periodic mode of operation. *Comput. Chem. Eng.* **2005**, *29* (6), 1293−1303.

(33) Wassick, J. M.; Ferrio, J. Extending the Resource Task Network for Industrial Applications. *Comput. Chem. Eng.* **2011**, DOI: 10.1016/j.compchemeng.2011.01.010.

(34) Sundaramoorthy, A.; Maravelias, C. T. A general framework for process scheduling. *AIChE J.* **2011**, *57* (3), 695−710.

(35) Floudas, C. A.; Lin, X. Continuous-time versus discrete-time approaches for scheduling of chemical processes: a review. *Comput. Chem. Eng.* **2004**, *28* (11), 2109−2129.

(36) Sundaramoorthy, A.; Maravelias, C. T. Computational Study of Network-Based Mixed-Integer Programming Approaches for Chemical Production Scheduling. *Ind. Eng. Chem. Res.* **2011**, *50*, 5023−5040.

(37) Castro, P. M.; Grossmann, I. E. New continuous-time MILP model for the short-term scheduling of multistage batch plants. *Ind. Eng. Chem. Res.* **2005**, *44* (24), 9175−9190.

(38) Amaro, A.; Barbosa-Póvoa, A. Planning and scheduling of industrial supply chains with reverse flows: A real pharmaceutical case study. *Comput. Chem. Eng.* **2008**, *32* (11), 2606−2625.

(39) Kopanos, G. M.; Méndez, C. A.; Puigjaner, L. MIP-based decomposition strategies for large-scale scheduling problems in multiproduct multistage batch plants: A benchmark scheduling problem of the pharmaceutical industry. *European Journal of Operational Research* **2010**, *207* (2), 644−655.

(40) Stefansson, H.; Sigmarsdottir, S.; Jensson, P.; Shah, N. Discrete and continuous time representations and mathematical models for large production scheduling problems: A case study from the pharmaceutical industry. *European Journal of Operational Research* **2011**, DOI: 10.1016/j.ejor.2011.06.021.

(41) Castro, P. M.; Harjunkoski, I.; Grossmann, I. E. Optimal short-term scheduling of large-scale multistage batch plants. *Ind. Eng. Chem. Res.* **2009**, *48* (24), 11002−11016.

(42) Méndez, C.; Henning, G.; Cerda, J. An MILP continuous-time approach to short-term scheduling of resource-constrained multistage flowshop batch facilities. *Comput. Chem. Eng.* **2001**, *25* (4−6), 701−711.

(43) Susarla, N.; Karimi, I. Integrated Campaign Planning and Resource Allocation in Batch Plants. *Comput.-Aided Chem. Eng.* **2010**, *28*, 1183−1188.

(44) Harjunkoski, I.; Grossmann, I. E. A decomposition approach for the scheduling of a steel plant production. *Comput. Chem. Eng.* **2001**, *25* (11), 1647−1660.

(45) Harjunkoski, I.; Saliba, S.; Biondi, M. Production Optimization and Scheduling across a Steel Plant. In *21st European Symposium on Computer Aided Process Engineering*, Pistikopoulos, E. N.; Georgiadis, M. C.; Kokossis, A., Eds.; Elsevier: Amsterdam, 2011; Vol. 29, pp 920−924.

(46) Schulz, C.; Engell, S.; Rudolf, R. *Scheduling of a Multi-Product Polymer Batch Plant*; Citeseer: Pittsburgh, 1998.

(47) Wang, K.; Löhl, T.; Stobbe, M.; Engell, S. A genetic algorithm for online-scheduling of a multiproduct polymer batch plant. *Comput. Chem. Eng.* **2000**, *24* (2), 393−400.

(48) Till, J.; Sand, G.; Urselmann, M.; Engell, S. A hybrid evolutionary algorithm for solving two-stage stochastic integer programs in chemical batch scheduling. *Comput. Chem. Eng.* **2007**, *31* (5), 630−647.

(49) Castro, P.; Barbosa-Póvoa, A. P.; Matos, H. A. Optimal periodic scheduling of batch plants using RTN-based discrete and continuous-time formulations: A case study approach. *Ind. Eng. Chem. Res.* **2003**, *42* (14), 3346−3360.

(50) Wassick, J. M. Enterprise-wide optimization in an integrated chemical complex. *Comput. Chem. Eng.* **2009**, *33* (12), 1950−1963.

(51) Applequist, G.; Samikoglu, O.; Pekny, J.; Reklaitis, G. Issues in the use, design and evolution of process scheduling and planning systems. *ISA Trans.* **1997**, *36* (2), 81−121.

(52) Grossmann, I. E.; Caballero, J. A.; Yeomans, H. Mathematical programming approaches to the synthesis of chemical process systems. *Korean J. Chem. Eng.* **1999**, *16* (4), 407−426.

(53) Amaro, A.; Barbosa-Póvoa, A. Supply chain management with optimal scheduling. *Ind. Eng. Chem. Res.* **2008**, *47* (1), 116−132.

(54) Maravelias, C. T. General framework and modeling approach classification for chemical production scheduling. *AIChE J.* **2012**, DOI: 10.1002/aic.13801.

(55) Reklaitis, G. Scheduling approaches for the batch process industries. *ISA Trans.* **1995**, *34* (4), 349−358.

(56) Moniz, S.; Barbosa Póvoa, A. P.; Pinho de Sousa, J. New general discrete-time scheduling model for multipurpose batch plants. *Ind. Eng. Chem. Res.* **2013**, *52*, 17206−17220.