Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Internal Supply Chain Digital Twin of a Pharmaceutical Company

João A. M. Santos [*,**] **Miguel R. Lopes** [*] **Joaquim L. Viegas** [**]
Susana M. Vieira [**] **João M. C. Sousa** [**]

[*] *Hovione Farmaciência S.A., Lisboa, Portugal*
[**] *IDMEC, Instituto Superior Técnico, Universidade de Lisboa, Portugal*

Abstract: A digital twin of a pharmaceutical company's internal supply chain is presented, along with a simulation-based rough cut capacity planning tool capable of giving estimates of the required monthly capacity for the different areas of the organization on the long-term. The work was a case study performed at a pharmaceutical company. The digital twin is delivered through a graphical user interface containing both its visualization and simulation tools. The proposed digital twin supplies accurate estimates of capacity needs to supply chain managers, giving the ability to easily visualize the resources required from different involved areas in the following 24 months.

Keywords: Digital Twin, Internal Supply Chain, Operations Research, Simulation-based Rough Cut Capacity Planning, Pharmaceutical Supply Chain

1. INTRODUCTION

1.1 Context

Current trends in logistics have placed the integrated supply chain (SC) as a process of unmatched capability in improving the overall performance of the organizations' SCs. The integrated SC considers all the phases of production and support as dependent, resulting in a complex network of people, machines, infrastructures and products. The planning and logistics behind these structures have undeniable impact but are extremely complicated. Strategies for optimization and improvement of these activities are being increasingly investigated.

An industry which has adopted the integrated SC is the pharmaceutical industry. This industry has faced a shift on its characteristics, *e.g.* the liberalization of the global markets (exposing companies to global competition) or the widespread appearance of generics [Shah, 2004]. Additionally, the number of drugs in the companies' portfolios has been steadily increasing, on account of the yearly number of approved drugs by the responsible entities being also rising [Mullard, 2019]. Both these reasons have prompted a shift in the industry: *big pharma* companies are nowadays more focused on research & development, leaving the manufacturing of the drugs to contract manufacturing organizations (CMOs). These organizations operate by manufacturing products by contract to other pharmaceutical companies. Their large, diverse and ever mutating product portfolios have created massive and complex SCs, which increasingly require optimization for a smooth operation.

⋆ This work was supported by Hovione FarmaCiência SA and FCT through IDMEC, under LAETA, project UID/EMS/50022/2019, to whom the authors express their gratitude.

The 21st century has seen the rapid development of modern technologies and computational power has become much cheaper. Concepts such as the Internet of Things (IoT), enabling the collection of data at distinct levels within an organization, have brought the required tools to optimize the integrated SC. Forecasting models, much more accurate than ever before, can now be developed, using collected data to obtain models based on the performance that has actually been previously achieved.

1.2 Related Work & Proposed Solution

The objective of this work is the creation of a Digital Twin (DT) of the internal SC of a pharmaceutical CMO, with the goal of increasing awareness to stakeholders and decision-makers by delivering information regarding past and present tasks and performance indicators. Additionally, the DT incorporates a forecasting tool, capable of performing simulation-based rough cut capacity planning (RCCP).

Very few applications of DTs on the SC or intangible assets were found. Kritzinger et al. [2018] provide a literature review of scientific articles on DTs, but these appear to be mostly DTs of tangible assets. Commercially available RCCP tools are frequent but no occurrence of simulation-based RCCP tools were found. Applied to the pharmaceutical industry, Costigliola et al. [2017] developed a discrete event simulation model of a QC laboratory, intended to be used as a decision-support tool for scheduling and planning (operational level); Lopes et al. [2018] presented a benchmarking platform to estimate the performance of new QC facilities (strategic level); Papavasileiou et al. [2007] developed a Monte Carlo simulation approach to task scheduling and cycle duration calculation.

The presented solution is a DT to support the internal SC of the pharmaceutical organization in study. The DT largely follows the basic concept of DT provided by Ivanov et al. [2019], where the authors describe DTs as a *combination of simulation, optimization and data analytics*. Unlike the solutions proposed by Costigliola et al. [2017] and Lopes et al. [2018], the proposed tool is intended to act as a decision support tool to the overall internal supply chain and not only to the QC laboratories. Additionally, the proposed solution aims at giving insights for a long-term timeframe. It provides a framework for visualizing operational information regarding the plant's activities, schedule operations and historical data and provide accessibility to the simulation tool. The simulation-based RCCP tool calculates the approximate monthly utilized capacity for the different areas, based on the performance that has actually been previously achieved. Furthermore, an approach similar to the one used by Papavasileiou et al. [2007] was employed, in order to verify interferences between tasks using assets considered to be bottleneck, in the sense that they are constrained and can only be used by one project at a time.

With the main goal of improving visibility into the internal SC and predicting capacity utilization on the long-term, a DT with a simulation-based RCCP tool was selected due to its clear advantages over current capacity estimators, which do not include any measure of variability. This tool can aid the SC decision-makers in planning on the long-term. Using the tool's current and past information, along with its capacity of generating capacity utilization predictions, they can make decisions in a more data-driven manner.

1.3 Digital Twin

A DT can be defined as a *dynamic virtual representation of a physical object or system, using real-time data to enable understanding, learning and reasoning* [Bolton et al., 2018]. Although its definition varies from source to source, the basic idea consists on a digital representation of an asset (being it tangible [entity] or intangible [system]), which uses IoT to receive meaningful real-time data and reaches conclusions based on both the developed model, how it has performed in the past and how it is performing at the present. The basic schematic of a DT is shown in Figure 1.

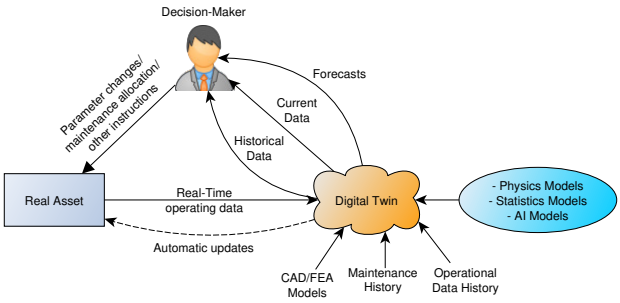

Fig. 1. Basic schematic of how a DT works. Note that the *automatic updates* transition may be optional.

DTs are frequently confused with monitoring tools and simulation models. In reality, DTs bring together both concepts, effectively delivering a visualization tool with improved simulation models [Madni et al., 2019]. DTs differ from simulation models in the sense that they receive real-time data to generate better predictions. Regularly, simulation models have complete descriptions of the object or system in study, but often lack its historical performance and almost always lack their current state. By having both, the generated simulations can be verified and improved by simulating on past data and the predictions that the model can create are based on current states, which will deliver more data-driven and accurate responses. DTs also supersede monitoring tools in the sense that all the data that these tools possess and display is also available by DTs. Additionally, DTs have access to forecast data, created by its simulation models.

1.4 Supply Chain

The internal SC of a CMO can be briefly summarised by the schematic of Figure 2. It is comprised of 4 areas: manufacturing (M), quality control (QC), quality assurance (QA) and warehouse (WH). The production of a component can be divided into two stages, the manufacturing itself and the quality release (QR) which considers all the stages after production and before shipment to the client. During the manufacturing stage, the M area is producing the components, while the WH area dispenses the raw materials (once at the beginning and often a series of times throughout the stage). Additionally, one or more in-process control (IPC) tests are performed by the QC area. The QR stage is characterized by two parallel workstreams (besides the WH final product storage, which happens once at the beginning): batch production record (BPR) and QC. The first workstream starts with the BPR review by the M area and ends with the BPR release by the QA area. The QC workstream starts with the QC release (QC R), followed by the QC release review (QC RV). There can be multiple QC R-RV workstreams, but their QC R phase always starts at the beginning of the QR stage and their QC RV phase always starts after the corresponding QC R finishes. Lastly, the QA area releases the QC analysis; this step starts after the last QC RV phase finishes.

1.5 Rough Cut Capacity Planning

On a manufacturing organization, available capacity is measured for a given production plant, area or workcenter and for a specific range of time. It corresponds to the total available time (in the considered period) multiplied by the number of resources related to the selected scope. For this work, capacity planning is done at the long-term time horizon (up to 2 years).

Capacity planning can be defined as the process of determining and evaluating the amount of capacity required for future manufacturing operations. This capacity can often be in terms of labor, machinery, warehouse space or supplier capabilities. The RCCP step comes as the capacity plan at the tactical level, which regards the master production schedule (the plan made by the company at the long-term, regarding production, inventory and workforce). RCCP is a capacity planning tool made at the long-term, used to adjust the required and available capacity and to change the master schedule or available

capacity, if necessary. Using the results from an RCCP, the master schedule can be modified in order to solve capacity inconsistencies by moving scheduled dates, increasing/decreasing scheduled production quantities or subcontracting additional workforce, for example.

2. IMPLEMENTATION

2.1 Data Extraction

Data was extracted from the company's enterprise resource planner. This data was divided into 4 categories:

- **Campaigns durations:** for each campaign manufactured in the production plant, information regarding the project, manufacturing duration and quality release duration were extracted. The extracted data and its relation to the reality can be graphically observed in the timeline from figure 2.

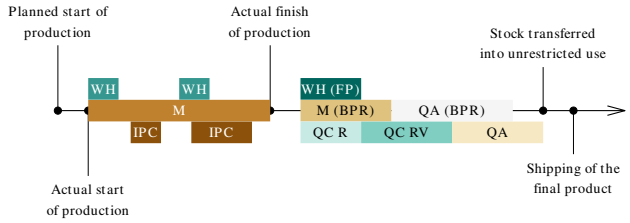

Fig. 2. Extraction data points and their relation to the real processes. The bars represent the real starts and ends of the different stages, while the vertical lines represent the dates that could be extracted.

- **Planned Orders:** orders that have been confirmed and are within the long-term timeframe. The extracted information regarded the project, planned start date, planned deadline and batch size.
- **Required Resources:** *worker-hours* and *equipment-hours* required by each project. The extracted data contemplated for each project all the tasks in the project's sequence of operations, which included information regarding the tasks area, effort and duration. The efforts were then aggregated by area and project.
- **Available Capacities:** available capacity per area and month.

2.2 Distributions Fitting

To model the variability inherent to the processes, the collected data regarding manufacturing and QR durations was studied, with the objective of fitting a theoretical probability distribution function (PDF) to the observed durations, in such a way that allowed their values to be randomly sampled by the simulation tool. Note that for confidentiality reasons, the time units are not particularized, expressed instead in $[TU]$. The basic steps for choosing theoretical PDFs and fitting them to the observed data were:

(1) **Selecting the set of contender PDFs:** since the extracted data did not have sufficient granularity for continuous distributions, only discrete PDFs were considered. Law [2015] proposes the use of the poisson, binomial and negative binomial distributions. Note that these PDFs belong to the exponential distributions family, which do not take negative values, having a lower bound at 0. Since the observed processes duration frequently also had a lower bound (inherent to chemical processes), an offset was applied to the theoretical PDFs, in such a way that x became $x - x_0$.

(2) **Preprocessing the data:** given the similar nature of the processes, it can be assumed that all the projects duration follow approximately the same behaviour; to this end, only projects with more than 30 observations (representative) were used to discover the most representative theoretical PDF.

(3) **Removing outliers:** due to the relatively small datasets, with projects having no more than 500 observations, *Tukey's fences* were used to remove outliers, setting the value $k = 1.5$. This outlier removal algorithm showed good results, with outliers correctly (and conservatively) removed and manifestations of variability correctly kept.

(4) **Statistical study:** an empirical method for determining the PDF based on the distributions' statistical properties is through the Cullen and Frey graph. This method considers the kurtosis and the squared of the skewness to make a prediction about the most appropriate PDF, in the manner depicted in figure 3. As can be seen from the figure, the negative binomial PDF appears to be the most appropriate one, according to the existing data.

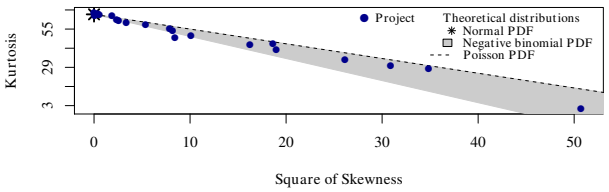

Fig. 3. Cullen and Frey graph for the projects in study

(5) **Fitting the theoretical PDFs to the data:** to fit the theoretical PDFs to the existing data, an optimization process is performed, with the objective of minimizing a *goodness-of-fit* (GoF) test. The GoF chosen was the Chi-Squared (CS), due to its widespread use. The mathematical formulation of the optimization is as shown in equation 1. Note that it assumes p_{i_x} with x as being one of the 3 PDFs and the optimization parameters $\{s, t, \lambda, p\}$ according to the chosen PDF.

$$\begin{aligned} \min_{t,s,\lambda,p} \quad & \chi^2 = \sum_{i=1}^{n} \frac{(O_i - Np_{i_x})^2}{Np_{i_x}} \\ \text{s.t.} \quad & t, s, \lambda > 0 \\ & p \in [0, 1] \end{aligned} \tag{1}$$

The average CS values of the optimization were 324.2 for the Poisson PDF, 476.3 for the binomial PDF and **46.3** for the **negative binomial** PDF. Given the results of the fitting process and the conclusions from the statistical analysis, the Negative Binomial PDF can be seen as the most appropriate PDF, for the processes being modeled.

2.3 Simulation-based RCCP Implementation

The first consideration made for the construction of the RCCP algorithm was the adoption of an upper bound for the obtained PDFs, point at which these are truncated. Since a lower bound already exists, the adoption of an upper bound removes the possibility of obtaining far-out values, which could badly influence the results of the simulation.

The simulation-based RCCP tool was based on Monte Carlo simulation. This class of computational algorithms relies on random sampling of values in order to find a pattern or tendency, and theoretically, is able to solve any problem with probabilistic interpretation. In the case at hand, it was seen that the manufacturing and QR durations had probabilistic characteristics that could be measured, which are propagated to the area's efforts. The mathematical formulation of the Monte Carlo method is done through the problem's probability space (sample space and function that assigns probabilities to the events), which can be seen in expression 2 ($N \equiv$ *number of orders*).

$$\begin{cases} \boldsymbol{\Omega} = \{D_{M_1}, D_{QR_1}, D_{M_2}, D_{QR_2}, \cdots \\ \qquad \cdots, D_{M_N}, D_{QR_N}\}, \Omega \in \mathbb{N} \\ \boldsymbol{P(x)} = \prod_{i=1}^{N} P(D_{M_i}, s_{M_i}, p_{M_i}, x0_{M_i}) \cdot \\ \qquad P(D_{QR_i}, s_{QR_i}, p_{QR_i}, x0_{QR_i})_j \end{cases} \tag{2}$$

with $P(x, s, p, x0) \equiv$ *Negative Binomial PDF*

Regarding the probability space of each PDF, each value of D_{Mi} or D_{QRi} corresponds to a duration in terms of manufacturing or QR, corresponding to project i, taking a value from $\{D_{min_i}, \cdots, D_{max_i}\}|_{M,QR}$. Note that the maximum value derives from the truncation that was previously mentioned.

For a simulation as the one whose results are shown in section 3.3, a total of 547 orders were considered, resulting in a sample space with around 10^{1392} scenarios. Calculating the monthly capacities for every scenario and observing the most probable utilized capacity for every month and every area would be computationally unfeasible. The use of Monte Carlo simulation is justifiable in such a situation.

The implemented algorithms are based in the earliest due date (EDD). The EDD method settles the start dates and then adds the manufacturing duration plus the QR duration to arrive at the *earliest due date*. The algorithms were ran using parallel computation, to improve their efficiency. These can be subdivided into 5 sections:

- **Sampling:** considering the set of planned orders, each campaign's manufacturing and QR durations are sampled from their respective PDFs. The values are re-sampled until being contained within the confidence level set.
- **Monthly utilized capacity:** for each planned order, the corresponding daily required capacity for each area is calculated (based on the sampled values and on the efforts from the recipes). After all the planned orders are considered, the capacities are aggregated by month.
- **Bottleneck asset (BA) utilization:** given a set of assets considered to be BAs, the tasks on each BA regarding every project are obtained, by being scaled according to their duration on the recipes and on the sampled manufacturing duration.
- **Aggregation of simulations:** both the monthly utilized capacity and the BA utilization have to be aggregated, since a series of i independent results exist, with i being the number of Monte Carlo iterations chosen. For the monthly utilized capacities, the aggregated results are the median of the monthly capacities, with a deviation of plus or minus 1 IQR. Regarding the BA utilization, the start date and duration of each process are used, with the start being its median plus or minus its IQR, and the end being the median plus or minus IQR of the sum of start and duration.

  One last step is performed, which is detecting the interferences between BAs utilization. The process is done after the aggregation and is done BA-wise, with different BAs being independent. The algorithm detects if there are any interference between two tasks of the same BA. If so, the tasks are considered to have *interferences*. Additionally, tasks that have interferences but only on their extended length (considering the minimum IQR of the start and the maximum IQR of the end), are considered to have *possible interferences*. All the other tasks have *no interference*.
- *(Optional)* **Optimization:** the optimization process is optional and has the objective of removing any interference between the tasks. It can be performed by the median, correcting the interferences between BAs themselves or by a factor of the IQR, either 0.25, 0.5, 1 or 2. This way the method basically operates by considering the tasks as an extended version of each, by the chosen factor. After the optimization, the resulting monthly utilized capacities are also calculated. Consider $\mathbf{BA} = \{BA_1, \cdots, BA_m\} = \{BA_i\}$, $i \in [1, m]$ as the set of BAs, chosen by the user. i corresponds to the index of the BA, and m is the total number of BAs. Each BA features a series of activities from different projects; considering the BA with index i (BA_i), $\mathbf{P_i} = \{P_{1_i}, P_{2_i}, \cdots, P_{o_i}\} = \{P_{j_i}\}$, $j_i \in [1, o_i]$. Here, j_i corresponds to the index of each activity (for BA_i), with the number of tasks being o_i.

$$\begin{aligned} \min_{n} \quad & n = \sum [P_{j_i} \cap P_{k_i} \neq \emptyset] \\ \text{s.t.} \quad & i \in [1, m] \\ & j_i, k_i \in [1, o_i] \\ & j_i \neq k_i \\ & n \geq 0 \end{aligned} \tag{3}$$

3. RESULTS

3.1 Convergence

Evaluating the convergence of the results is a fundamental step, since it can give clear insights into whether or not the simulation is in fact trying to reach a representative

solution. In theory, the more simulations ran, the greater the confidence of an average value being the most representative. For a specific scenario, the simulation was ran for a series of different number of iterations, ranging from 10 to 50000. Considering that the medians of utilized capacities for each area at 50000 iterations are certainly the best estimates of the monthly utilized capacities (named as the *50000-result*), the relative error between each iteration's median and the *50000-result* was calculated, and the area-wise total is presented in table 1. The relative error becomes neglectable from 500 iterations on.

Iterations	Total Relative Error
50	6.0%
500	0.7%
5000	0.5%

Table 1. Total relative error between the areas at each number of iterations and the *50000-result*.

Regarding the BAs, the convergence was studied in a slightly different manner: the number of interferences, possible interferences and no interference was recorded per simulation and its evolution with the number of iterations was obtained. The results showed that initially, the number of interferences decreases while the number of no interferences remains constant and the number of possible interferences increases. At around 500 iterations, this convergence stops and there is no more fluctuation in the number of interferences. This behavior can be explained by asset utilizations that become better defined, reducing the number of interferences; these former interferences become possible interferences which explain the rise in such category. After stagnating, increasing the number of iterations will reduce the variance which leads to a reduction in possible interferences, resulting in an increase in the number of assets without interferences.

3.2 Validation

The process of validating the model is important to verify whether or not the simulations obtained are trustworthy, and how they perform when compared with the baseline estimations, obtained without simulation, using the most common durations of the processes. The data used for the validation regarded a period of 9 months at the pharmaceutical company in study. Note that these capacities were not explicitly available and were therefore calculated according to the rules described in section 2.3 from the real processes' durations. Additionally, the first 3 months are not considered since they are comprised of orders within the short and medium terms. To benchmark the results of the simulation were compared against the baseline estimations. The results for the Manufacturing area are shown in figure 4.

The results show that the simulated capacities are a better approximation than the baseline estimations and that the simulated values are within an acceptable distance from the real utilized capacities. Numerically, the overall relative error between the simulated capacities and the validation data is 6.75%, while the baseline estimation's error is 14.64%. This shows a significant reduction in

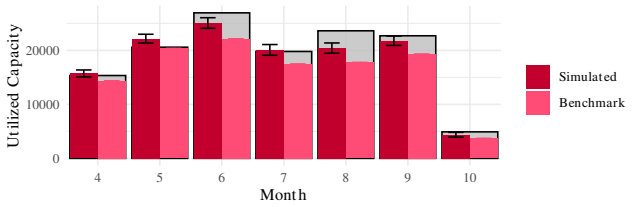

Fig. 4. Capacities validation and benchmarking, for the manufacturing area. Note that the grey regions indicate the actual utilized capacity.

the error using the simulation and that the simulated results are actually within reasonable distance to the actual results. Considering all the areas, the simulation error is 13.10%, while the baseline error is 17.42%. While this shows less significant improvements, it is worth noting that the manufacturing results are the most important due to the fact that its absolute values are much higher than the remaining areas. Additionally, it was possible to observe that the errors from simulation tended to be frequently by excess, while the errors from the baseline estimate tended to be by deficit. Errors by excess tend to be preferable than by deficit, which is another advantage of the simulation-based RCCP.

3.3 Forecasting

To test the forecasting capabilities of the simulation-based rough cut capacity planning tool, a specific scenario was ran, corresponding to a series of planned orders and orders on short/medium term (current orders), with their respective project and planned start date. While the current orders cannot be modified in any way, it is important to consider them in the simulation since they affect the utilized capacity in the following months. A total of 299 planned orders and 248 current orders were considered, for 500 iterations. The resulting monthly capacity utilization graphs for manufacturing is shown in figure 5. Regarding the BAs utilization, the Gantt chart represented in figure 6 shows a set of assets considered unique and unchangeable, with the clashes between tasks depicted.

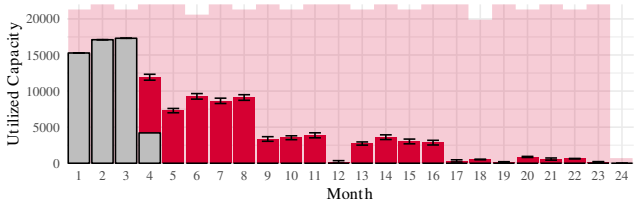

Fig. 5. Forecasted capacity evolution per month on the manufacturing area. The full color bars correspond to the actual simulated capacity for the month and area in question, with an error bar indicating ± 1 IQR. The grey bars correspond to the capacities from the current orders. The shaded background area corresponds to the limit capacity of each area.

An optimization step is then possible. The user can select the desired scope of the optimization, whether by the median or by 0.25, 0.5, 1 or 2 IQRs. The objective of this process is detecting and removing clashes between tasks

Fig. 6. Gantt chart of the BA's utilization. Tasks color-coded as grey signify that they are current orders, which can no longer be modified.

by their median value or by a factor of their IQR and calculating the resulting utilized capacity. Considering the example previously described and an optimization by 0.5 IQRs, the resulting Gantt chart is as shown in figure 7.

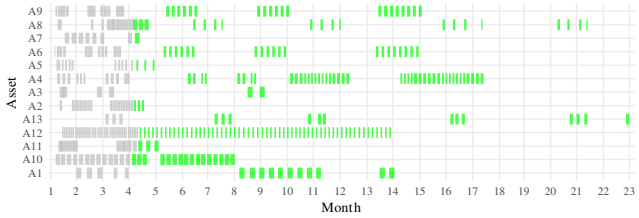

Fig. 7. Gantt chart of the BA's utilization after optimization. Grey tasks ≡ *(un-optimized)* current orders; green tasks ≡ planned orders *(optimized or not)*.

In terms of monthly utilized capacity, it could be seen that while before being optimized, the results tended to be more concentrated on the initial months, greatly decreasing to near zero values afterwards, after being optimized the utilized capacities tended to be more distributed. This makes complete sense since the tasks are also more distributed, leading a greater distribution of the projects and consequently to capacities less concentrated on the initial months.

3.4 User Interface

Besides being a platform for delivering the simulation-based RCCP tool, a graphical user interface (UI) was built with the objective of delivering clear, intuitive and customizable ways of showing information on a global view, regarding operations per building or project, the key performance indicators and their evolution, and the schedule of activities, both that happened in the past, or that will happen in the near future. Additionally, the UI features a Project database containing a series of information regarding each project. All of these allow the users to observe past tendencies and patterns, enabling stakeholders to have decisions more data-driven. The user interface was developed using the *R* programming language and its capabilities for front-end application development, through the *shiny* package.

4. CONCLUSIONS

The developed tool was successful in delivering its two components: visualization and simulation. Its visualization capabilities can improve the users' awareness into the supply chain status, intuitively observe the scheduled tasks on the near future and deduce patterns of *e.g.* seasonality, given the historical behaviour. Regarding simulation, the results obtained were promising, delivering better estimates than non-simulation RCCP tools. This can be beneficial in a company since it allows decision-makers to better negotiate changes to the master schedule at the long-term, reducing the necessity for solving capacity deficiencies or surpluses at the short/medium-term, which is often riskier and more expensive. The methodologies here employed can be applied to other industry domains with little need for extensive modifications. Only the mapping of the internal SC (such as the one shown in figure 2) is industry, and often company-dependent.

The main limitation of this method is its inability to correctly deal with certain situations which are common but frequently cannot be reflected in the numerical data. Additionally, the simulation tool does not account for differences in the planned versus actual start date.

REFERENCES

Bolton, R.N., McColl-Kennedy, J.R., Cheung, L., Gallan, A., Orsingher, C., Witell, L., and Zaki, M. (2018). Customer experience challenges: bringing together digital, physical and social realms. *Journal of Service Management*, 29(5), 776–808. Doi:10.1108/JOSM-04-2018-0113.

Costigliola, A., Ataíde, F.A., Vieira, S.M., and Sousa, J.M. (2017). Simulation model of a quality control laboratory in pharmaceutical industry. *IFAC-PapersOnLine*, 50(1), 9014–9019.

Ivanov, D., Dolgui, A., Das, A., and Sokolov, B. (2019). Digital supply chain twins: Managing the ripple effect, resilience, and disruption risks by data-driven optimization, simulation, and visibility. In *Handbook of Ripple Effects in the Supply Chain*, 309–332. Springer.

Kritzinger, W., Karner, M., Traar, G., Henjes, J., and Sihn, W. (2018). Digital twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine*, 51(11), 1016–1022. Doi:10.1016/j.ifacol.2018.08.474.

Law, A.M. (2015). *Simulation Modeling and Analysis*. McGraw Hill Education, international fifth edition.

Lopes, M.R., Costigliola, A., Pinto, R.M., Vieira, S.M., and Sousa, J.M. (2018). Novel governance model for planning in pharmaceutical quality control laboratories. *IFAC-PapersOnLine*, 51(11), 484–489.

Madni, A.M., Madni, C.C., and Lucero, S.D. (2019). Leveraging digital twin technology in model-based systems engineering. *Systems*, 7(1), 7.

Mullard, A. (2019). 2018 FDA drug approvals. Doi:10.1038/d41573-019-00014-x.

Papavasileiou, V., Koulouris, A., Siletti, C., and Petrides, D. (2007). Optimize manufacturing of pharmaceutical products with process simulation and production scheduling tools. *Chemical Engineering Research and Design*, 85(7), 1086–1097. Doi:10.1205/cherd06240.

Shah, N. (2004). Pharmaceutical supply chains: key issues and strategies for optimisation. *Computers & chemical engineering*, 28(6-7), 929–941. Doi:10.1016/j.compchemeng.2003.09.022.