DATA-DRIVEN SIMULATION FOR PRODUCTION BALANCING AND OPTIMIZATION: A CASE STUDY IN THE FASHION LUXURY INDUSTRY

Andrea Nunziatini
Virginia Fani
Bianca Bindi
Romeo Bandinelli
Mario Tucci

Department of Industrial Engineering
University of Florence
Viale Morgagni 40/44, 50134, ITALY

ABSTRACT

The paper presents the definition of a data-driven simulation framework and its development using AnyLogic® software for production balancing and optimization in the leather luxury accessories industry. The model has been developed including an industry-oriented set of objects in order to easily replicate the resource configuration and layout mainly used within the analysed context. A case study is therefore conducted to validate the model and confirm its usability. The results of this work demonstrated that the proposed framework can be easily adapted to a shoe joiner using the pre-configured objects. Scenarios have been carried out based on what-if analyses regarding productivity, resources saturation and bottlenecks. This tool can therefore be used as a valuable support to production planning, scheduling optimization, workload and production cycle balancing.

1 INTRODUCTION

The objective of this paper is to propose a data-driven simulation model for production balancing and optimization in the leather luxury accessories industry. As widely reported in the literature (Fani et al. 2020), this sector is characterized by a high fragmented supply chain, with orders to be frequently rescheduled and high quality standards to be compliant. In this context, frequent changes of production mix have to be managed, often requiring the re-optimization or even re-design of production flows.

Despite this, standard production layouts and SC configurations could be found among different suppliers and brands (Fani et al. 2017). As a consequence, a common framework can be defined, which will be able to represent the majority of the industrial case studies belonging to this sector. In particular, the present work reports the preliminary results of a complex study aiming to propose a unique framework for modelling different actors of the luxury accessories SC (i.e. leather goods shops, maneuvers and shoe joiner) by using a data-driven object-oriented simulation approach. Main barriers to industrial applications of traditional discrete event simulation models, widely studied in the literature to solve production-related issues, refers to the fact that they do not allow real-time support to business decisions in dynamic contexts, due to the time-consuming activities needed to re-align parameters to changing environments (Lugaresi et al. 2018). Data-driven approach overcomes these limitations, giving the possibility to easily update input and quickly rebuild the model itself without any changes in the modelling. As first real-case scenario, the proposed model has been therefore validated within a shoe joiner.
In order to realize the proposed data-driven simulation model, AnyLogic® has been identified as the tool for modelling and analysing the entire SC, due to the flexibility of the simulation approach and to the Object Oriented (OO) architecture, able to replicate the different configurations of the suppliers belonging to the fashion SC.

The paper is structured as follows. In Section 2, a literature review on real-case applications of data-driven simulation has been presented. The proposed model has been detailed in Section 3, describing the formulation of this conceptual model. Its application in a case study has been shown in Section 4. Finally, in the last section the main conclusions of this work have been discussed.

2 DATA-DRIVEN SIMULATION APPLICATION

In this section, the literature review results on data-driven simulation will be presented. The review has been using the database "Scopus" with the keywords "data-driven" and "simulation" combined with "Manufacturing", "Production", "Planning" and "Scheduling" published in the last ten years. The result was 453 papers that were filtered using the abstracts: the inclusion criteria was the thematic focus on the data-driven simulation.

2.1 Data Driven Theoretical Approach and Industrial Plant Optimization

First, it is possible to adopt a dynamic way to automatically build a simulation model (Frick, R. 2011). Once the problem has been defined, data-driven simulation allows to generate the model and provide the necessary entities according to the database setting. Then, a simulation software can execute it, obtaining scenarios or experiments according to the needs. An example of application is to improve the flexibility of an automotive assembly line (Wang et al. 2011). In this case, one of the goals was to enable non-specialized users to use simulation models. This article shows that a generic and fully parameterized model was firstly designed in order to simulate a specific set of systems. Secondly, the data were prepared and a relational database was created using the IDEFIX® standard. The main blocks for material handling and assembly were therefore created with object-oriented programming developed on Arena®. More examples of optimized assembly lines can be found in the aerospace industry (Zhou et al. 2019), in order to reduce the very high costs through process optimization. However, product customization in this industry is extreme as well as line variability. A flexible layout was therefore the necessary compromise, supported by a simulation model that could quickly be updated.

Data-driven simulation has been also applied in the field of construction sites (Akhavian et al. 2013), satisfying the need to quickly model reality by using real-time data. In fact, traditional modelling cannot detail the dynamic operations of construction sites, making essential to use a generator, Stroboscobe® in the article. Similarly to the construction environment, examples of data-driven simulation applications could be found in the mining field (Meng et al. 2013). The dynamism of the context was caused by the handling of materials, managed by a Material Handling System (MHS), making again the data-driven simulation essential for decision making. The structure of the MHS could also change rapidly, depending on the type of coal required. In addition, the transport tonnage must always be maximized. Remanufacturing is another industry that employs data-driven simulation (Goodall et al. 2019): this kind of manufacturing consists in several sequenced repairs made to regenerate the product. The uncertain nature of returns caused a highly variable environment, requiring an extreme need for support in evaluating different production scenarios in real time. Experiments with simulation models have been conducted in the past, but they often turned out to be an extreme simplification of reality (Akhavian et al. 2013). In the case study illustrated by the paper, a more structured approach has been described, in which RFIDs sensors directly communicated with the Delsi® simulator. This model worked successfully and another possible application was highlighted in the field of post-production services, such as maintenance.
2.2 Data Driven Approach and Production And Planning Optimization

It must be considered that the model is never a monolithic entity, but a composition of related parts (Huang et al. 2014). The definition of each part constitutes the constraints or domain of the model; the parameterization, initialization and composition of each part therefore represent the degrees of freedom of the corresponding domain. In general, the increasing demand for product customization drives companies to smart manufacturing (Fang et al. 2019). There are several examples of rapid changes in the production department that a traditional simulation cannot follow. The layout of production lines and the logistics system change continuously, as well as production cycles. All this information should therefore be data-driven in the model. The same requirements caused a radical change in the production mode of a digital workshop (Gao et al. 2019). A data-driven simulation was applied even to this context: a set of sensors was structured, then connected in real time to a database through a specific data transmission interface, KEPServerEX®. This information was processed by a simulator, Unity3D®, which monitored KPIs in real time and simulated future scenarios. The data-driven simulation has been successfully applied also to production planning, as integration with process planning change (Wang et. al. 2020). In this context, the phase of data collection has been crucial, since it was necessary to distinguish layout information from resources information (e.g. machines, staff). During the following phase, the model has been generated and it has been essential to define the logics and their verification. A classical approach could only perform a basic optimization of production planning: this method does not generate optimal schedules, capable of considering the state of the production system in real time (Kuck et. al. 2016). By exploiting the data-driven optimization, the effects of modified parameters can be easily estimated also in complex systems. Obviously, a platform of data exchange must be connected to the model, in order to automatically aggregate data, which refer to the modifications in progress. An ERP system has been integrated with a model using the FlexSim® software (Krenczyk et al. 2018). It has been therefore possible to follow both a semi-automated and an automated approach of model generation: avoiding the manual entering of the data has represented a great advantage in complex systems. Finally, data-driven simulation could also be employed to estimate the environmental and economic sustainability of a production system (Ojstersek et al. 2020). Production planning constitutes a multi-objective problem, whose areas of major interest are the consumption of energy, materials and resources. Moreover, the high degree of product customization implicates an increasing flexibility of manufacturing systems. The high variability of the system, regarding both mix and layout changes, has required the use of data-driven simulation. Finally, only one contribution (Fani et al. 2021) has been found in literature about data-driven simulation applied in the fashion industry. Even a wide set of parameter can be set by final users, the limitation of this article is that the proposed model was not data-driven in terms of resource and layout configurations. According to this, any change in layout could not be directly managed by the company but have to be modelled by the model developer at the simulation start-up.

In conclusion, the data-driven approach is already present in other industrial sectors but it has still not been tested in a low tech sector like the fashion industry. There is only one application of a data-driven simulation model in the fashion industry and it is not data-driven in terms of resource and layout configurations. Therefore, compared with literature just described, the added value of this work is to propose a new approach for production balancing and optimization in this sector.

3 THE PROPOSED FRAMEWORK

This section will detail the data-driven approach used to build the proposed framework, developed to cover the peculiarities of the leather luxury accessories industry and to be easily readapted to different companies belonged to the analysed sector.

First of all, one of the main goal is to overcome the classic approach, in which the static twin model is composed of two sets of data: static setup information and live data. For instance, setup information refers to layout configuration such as the number of production areas and workers, while live data constitute the input data that will be processed during the simulation. They also include the information regarding the
production cycle that every entity will follow. In the proposed approach, the dynamic twin model is composed of live setup information and live data stored together into a single database. The live setup information allows the software to generate the model at simulation start and, subsequently, to process live data as a traditional static twin.

According to the literature review, dynamic contexts such as the fashion industry require a quick-changing use of simulation, based on automatic models creation: in the classical approach, the model is developed through the design phase, while in the data-driven one it is generated during every simulation run. In detail, the data of the model are stored in a database and read at the model start-up in order to create a wide range of objects, such as workstations and production areas, according to the parameters previously defined. A process flowchart description has been defined for each object to establish the path and/or the process that the entity entering the object must follow. If changes are needed to represent the process flowchart moving from one to another resource, even similar (e.g. two workstations of an assembly line), two different objects have to be created.

Summarizing, the proposed framework based on data-driven simulation is constituted of objects and their respective process charts and parameters, that have to be defined a priori to cover the wide range of objects to be represented in order to analyse production processes in the leather luxury accessories industry.

Once the single elements of the framework are defined, these objects have to be connected, in order to define a path for each entity that will be processed by the simulation model. To obtain this result, it is necessary to implement customized functions, requiring a basic knowledge of the programming language of the software in use (e.g. Java® in AnyLogic®). This function will be able to define the correct destination of entities to move them along the model, object by object, from the “source” block where they are generated to the “sink” block where they exit the model. The design of the data-driven model ends once the objects, their process charts and their parameters have been defined. After that, the objects have to be stored in a database to be generated at the model start up.

Once the simulation model has been set according to the described data-driven approach, the model run traditionally but with some clear advantages. In fact, traditional model require to specify the path for every entity in the process chart, making each model only able to simulate one specific scenario at a time. Despite classic models, the parametric approach used in the proposed data-driven simulation framework allows to include the information to define specific paths on a database, as well as to adjust the route of entities by simply varying the input data. However, only parametric models are not able to modify their own process chart, for example by adding or eliminating workstations: if an object has not been created during the modelling phase, the entity will never be able to cross it. The advantage of combining data-driven and parametric approaches is that, once the objects have been defined, it will be sufficient to modify their number in the database in order to increase or decrease the modelled objects; and it will be sufficient to modify the production cycle in order to change the path, therefore the sequence of objects, that the entity will have to cross.

Starting from the described evidences that demonstrate how dynamic contexts require data-driven simulation model, the purpose of this paper is to define an industry-oriented framework able to model the production process of leather luxury goods, such as shoes, bags or other accessories. First, common characteristics between the analysed production processes have been investigated, such as the use of specific containers for moving raw materials and semi-finished products. These containers can be handled by conveyor belts, AGVs (Automated Guided Vehicles) or directly workers. Production areas, each representing a sub-set of production processes (e.g. cutting, assembly and packing), are connected by loading and unloading points, typically called "bays". Production can be managed as job-shops or lines, according to the specific processes included in each production area. Production phases usually alternate with transport phases: at the exit of a production area, articles are placed in the loading bay to be picked up and taken to the unloading bay of the next production area, as well as internal material flows have to be managed picking up material from one workstation and taken to the next one within the same production area. To simulate the described flows, it is therefore necessary to model the "bay" and "production area" objects. An noticeable limit of this approach is that, if it will be necessary change the logic of process, it
will have to be changed the relative object and only the developer will have the skills to do it. Furthermore, buffer of raw materials, semi-finished and/or finished products have to be modelled across the production process, as well as at the beginning and end of production lines. Warehouses are quite frequent elements to be modelled in such contexts, since they represent the physical connection between the various actors of the supply chain. For this reason, an adequate object must be created and interfaced with transporters. Regardless of the transporter employed, it is essential to use a "fleet" block that manages all tasks of internal transport.

By analysing the common characteristics of the different contexts above-mentioned, a generic framework for simulation of leather luxury accessories industries has been defined, as reported in Figure 1.

The boundaries of the proposed framework can be summarized as follow:

- the production layout is a job-shop or a line with small batch flow;
- items are processed into workstations or production areas;
- the logic of the queues between different workstations, or production areas, is First In First Out (FIFO);
- a transport system move items between production areas, placed inside specific containers before being handled by workers, conveyor belts or AGVs.

![Figure 1: Example of objects configuration in the proposed model.](image)

The transport system can be modelled or not within the framework, according to the environment physical characteristics: if the time to move items from one production area to the next one is not included in the so-called "shadow times", the populations of "bays" and "conveyors" will be created; otherwise, the semi-finished products will be instantaneously moved from one "production area" to another. When the transport system is modelled, the following hypotheses must be verified:

- the transporters are AGVs, conveyor belts or workers;
- the transport takes place through "bays";
- the "transporters" are managed by "fleet" block.

This conceptual framework can be applied not only within a single plant of a company, but also in order to represent the entire fashion supply chain headed by the brand owner. Changing perspective and focusing on the multiple levels of the supply chain, it is possible to understand and model how the involved actors interact, describing in the database at least the production areas of each suppliers.
This approach allows to model all the supply chain actors, making possible to rank suppliers, test the bullwhip effect on the entire supply chain and improve the coordination among actors. Summarizing, this conceptual framework is capable of:

- modelling the production areas in which the single supplier organise its processes;
- modelling an entire supply chain.

The application of data-driven modelling derives from the need of using a single tool to support different realities. For example, a shoe joiner could be modelled as composed of production areas served by workers, while a smart assembly plant may need modelling AGVs: as declared, the proposed framework can be applied for both contexts through the correct setting of the database, where only the populations (i.e. workers’ and AGVs’ respectively) needed have to be populated and consequently automatically generated at the model start-up.

Each object must be both adherent and flexible, in order to adapt to different realities. However, customization of the process charts could be necessary to accurately model reality. Independently of that, a well-structured framework drastically reduces the model realization time.

After theorizing the conceptual framework, it is possible to detail the elements of the model. First, every object must be able to interface with the others. According to this, the enter and exit function will be modelled in a parametric way: this has been done using "wireless" connections in the palettes of AnyLogic® simulation software. Secondly, every process chart has to include incoming queues and processing times. Furthermore, the objects "production area" and "bay" will be modelled in a different way. The “production area” object will include the drop of items from the container received as input, the movements of items between different nodes, the jobs directly processed by workers and the pick of items into a different container to be sent to the next step. On the other hand, the modelling of “bay” will include only the transport phase, since the content and its container will never be separated. For this reason, the "transporter" object will also be modelled as an AGV, a worker or a conveyor belt. As an example a bay object with AGV is shown in Figure 2.

![Image of bay object with AGV](image)

Figure 2: Example of a bay object with AGV.

In this case the object has the parameters in the upper part of the image and the flow chart divided into two lines: the first dedicated to sending items and the second dedicated to receiving items.

The last part of the model configuration is related to the measurement of KPIs. A log file is generated by the model according to a template defined according to the model layout. In detail, start and end of every event are recorded and clustered by item code, usually called Stock Keeping Unit (SKU), and number of order or batch. Even a quite common set of KPIs have been defined, specific measurement could be
Nunziatini, Fani, Bindi, Bandinelli, and Tucci

requested to better fit user needs. Since the data-driven approach can enable non-experts to easily access to simulation results, the definition of a clear output and an intuitive dashboard is crucial for the use of the tool in real scenarios.

Validation is the final phase of the modelling process. Data-driven approach requires a more complex validation phase, due to the necessity to validate every component of the whole simulation model, such as objects and paths, as well as each functionality implemented by coding.

4 CASE STUDY

The proposed framework has been applied to a shoe joiner working in the fashion luxury market, in order to test its adaptability to real context.

Footwear sector is characterized by a quite rigorous production process, divided into sequential phases, each of them linked to a different production area, classified as:

- **cutting**, performed by a cutting die or by hand, in case of valuable articles;
- **preparation**, where the absence of possible defects is verified, as well as the correct association between parts. Marking, fleshing and edging of the pieces are also checked as required during this phase;
- **shoe joiner**, where all the operations to realize the finished upper are performed;
- **assembly**, where the bottom is assembled through various systems, depending on the desired result;
- **cleaning and finishing**, where the eventual colouring and final retouching are performed.

In detail, the involved company produces fashion luxury shoes, moving from sandals to boots. Every single phase of the production process can slightly differ, according to the products type and line. Since the case study has been developed based on real data, generic codes and process times will be reported in the paper due to privacy issues, without losing generality of results and accuracy in the validation phase.

The developed model has been created starting from the following assumptions:

- the production layout can be represented as a production line with small batch flows;
- the processing logic of every queue of the model is the FIFO;
- semi-finished goods are processed only within dedicated areas, called "workstations";
- semi-finished products are carried out inside a specific container from one "workstation" to the next one;
- within every "workstation", each item is processed by a single operator at a time. Workstations can be replicated to increase production capacity (i.e. each first item of the FIFO queue on enter can be processed by two parallel workstations served each one by a dedicated operator, making quicker the production flow);
- from each "workstation" the item pass through the "bays" that have two line, the first dedicated to send to the next "workstation" and the second dedicated to receive from the previous "workstation".

In order to better explain the case study, the plant layout is shown in Figure 3.

The KPI needed for the model validation were already available, as company measures them on the field. Such measurements have indeed been carried out by the staff dedicated to performance monitoring, that has taken care of both the execution and the statistical analysis. In conclusion, reference data were provided, expressed as average values, ready for the validation phase.

In order to validate the model, the simulated data needed to calculate the required KPIs are already traced and logged in the proposed framework.
The proposed process chart has been confirmed and the database updated with real data. The first table to be loaded was the one related to the populations of the “workstation” objects. Table 1 is an example and contains transcoded information, in order to protect the real data.

![Plant layout](image)

**Figure 3: Plant layout.**

**Table 1: Workstation database.**

<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>Capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Workstation_1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>Workstation_2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>Workstation_3</td>
<td>9</td>
</tr>
</tbody>
</table>

The second table describes the production plan as list of SKU to be produced, as well as the related quantity, shoe size, and requested date. An example filled with transcoded data is reported in Table 2.

**Table 2: Production plan database.**

<table>
<thead>
<tr>
<th>ID</th>
<th>SKU</th>
<th>Size</th>
<th>Quantity</th>
<th>Requested date</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Shoe_1</td>
<td>5</td>
<td>20</td>
<td>01/09/2020</td>
</tr>
<tr>
<td>2</td>
<td>Shoe_2</td>
<td>6</td>
<td>20</td>
<td>01/09/2020</td>
</tr>
<tr>
<td>3</td>
<td>Shoe_3</td>
<td>7</td>
<td>20</td>
<td>01/09/2020</td>
</tr>
</tbody>
</table>

The third table of the database lists the workers to be included in the simulation model, and the fourth table is the production cycle and contains the most sensitive data, because it shows the processing time requested for each activity referred to a specific SKU.

During the validation phase of the model, KPIs related to productivity, takt time, resources saturation and queues size were examined. The output data needed for the simulated KPI measurement have been
extracted from the log file created by coding on AnyLogic®. In order to make this information available and usable for the company, it is essential to prepare programmed spreadsheets that effectively show the KPIs of interest. Starting from the log table, in fact, data related to productivity, queues size and resources saturation have been highlighted.

Concerning simulation setting, ten runs of one-month length have been launched, as this timing corresponds to the horizon of the company’s real production plan. By analysing the simulated productivity, a warm-up period of three days has been considered and therefore excluded from the data used for validation in each simulation run. In Figure 4 it is visible where the productivity trend of each simulation run is shown. Values on the vertical axis are only indicative as the real values are protected by privacy.

The remaining days was analysed in terms of resources saturation, productivity and lead time. Finally, these data were compared with those coming from the real measurements, thanks to the software Minitab® and, more specifically, the Paired T-test. Again, for privacy issues, it is not possible to report the numerical values obtained from the validation. This is a very important phase as it certifies that the output of the model coincides with the real one, therefore a potential future development could be a case study not protected by privacy where it is possible to describe this phase in detail. Especially, during the validation it is important to test every single object so as to be able to use the model with every potential combination of them.

After the validation of the model as-is, the company wanted to understand how to use simulation to increase its productivity. According to this, to-be scenarios based on what-if analysis have been simulated, making possible to identify a new configuration to meet the required productivity target.

In order to achieve this result, the described procedure has been followed. The first step involved the simulation of a scenario in which every resource in the company, like workstations and workers, was doubled. Starting from the saturation data, it has then been possible to identify inefficiencies like bottlenecks, iteratively solve them. This try and error process, in fact, attempted to solve the bottlenecks highlighted in the previous phase and balance the workload on each resource. The process ends when the target production is reached. This procedure does not require changes to the simulation model, but only a quick update of the database: therefore, even non-specialist users are demonstrated able to use this tool.

5 CONCLUSION
This paper presents the preliminary results of a data-driven simulation model, developed in order to be used in the leather fashion accessories industry. The results of this work demonstrated as this framework can be adapted to a shoe factory, in particular to a shoe joiner. The generated output is perfectly manageable by any professional figure, able to use the Microsoft Excel® software. With these premises, it is clear that this
tool can be used as a valuable support to the production planning, as it is able to update quickly as a function of both the workload and the production cycle. It can provide scenario analyses as well as what-if analyses regarding productivity, saturation and bottlenecks. Furthermore, from an industrial point of view, the application of this tool can reduce significantly the time for line balancing optimization in comparison with traditional methods. The disadvantage is that for the first use it must be validated as a model standard and perhaps adapted to the production context by modifying the objects appropriately. After this phase it no longer requires the intervention of the developer. Next steps of this research will cover the application of this model in other real contexts belonging to the leather luxury accessories industry, such as bags producer, as well as its scalability will be evaluated.

REFERENCES


Nunziatini, Fani, Bindi, Bandinelli, and Tucci


AUTHOR BIOGRAPHIES

**ANDREA NUNZIATINI** is a Research Fellow at the Industrial Engineering Department of the University of Florence. The issues he is dealing with are related to production scheduling and optimization. His e-mail address is andrea.nunziatini@unifi.it

**VIRGINIA FANI** is a Post Doc Research Fellow at the Industrial Engineering Department of the University of Florence. The issues she is dealing with are related to production processes optimization along the supply chain, with particular focus on the peculiarities that the fashion companies have to face. Her e-mail address is virginia.fani@unifi.it

**BIANCA BINDI** is a Post Doc Research Fellow at the Industrial Engineering Department of the University of Florence. The issues she is dealing with are related to supply chain strategy and performance measurement, with particular focus on the peculiarities that the fashion companies have to face. Her e-mail address is bianca.bindi@unifi.it

**ROMEO BANDINELLI** is an associate professor at the University of Florence, Department of Industrial Engineering. He is member of the IFIP 5.1 "Global Product development for the whole lifecycle". His e-mail address is romeo.bandinelli@unifi.it

**MARIO TUCCI** is a full professor at the University of Florence, Department of Industrial Engineering. President of Industrial Engineering Programmes, former Coordinator of Doctoral Programmes on Industrial and Reliability Engineering, former Director of Project Management Master Programme, Convenor of UNI CT523 Industrial Automation (mirror of ISO TC184). His e-mail address is mario.tucci@unifi.it