REAL-TIME SUPPLY CHAIN SIMULATION: A BIG DATA-DRIVEN APPROACH

António A.C. Vieira
Luís M.S. Dias
Maribel Y. Santos
Guilherme A.B. Pereira
José A. Oliveira

ALGORITMI Research Center
University of Minho
Braga, 4710-057, PORTUGAL

ABSTRACT

Simulation of Supply Chains comprises huge amounts of data, resulting in numerous entities flowing in the model. These networks are highly dynamic systems, where entities' relationships and other elements evolve with time, paving the way for real-time Supply Chain decision-support tools capable of using real data. In light of this, a solution comprising of a Big Data Warehouse to store relevant data and a simulation model of an automotive plant, are being developed. The purpose of this paper is to address the modelling approach, which allowed the simulation model to automatically adapt to the data stored in a Big Data Warehouse and thus adapt to new scenarios without manual intervention. The main characteristics of the conceived solution were demonstrated, with emphasis to the real-time and the ability to allow the model to load the state of the system from the Big Data Warehouse.

1 INTRODUCTION

Supply Chains (SCs) are complex and dynamic networks, comprised by entities, such as suppliers and customers, wherein material and information exchanges occur, driven by demand and supply interactions (Simchi-Levi et al. 2008). Several activities take place in these networks, e.g., production and transport of raw materials. Ultimately, the goal of each entity in these networks may be seen as the fulfillment of their customers' orders, at a minimum cost, whilst improving their competitiveness. In other words, to efficiently manage raw materials receipt and timely schedule deliveries at the right time, place and quantities.

A proper assessment of the performance of a SC is not easy to determine. Thus, rather than being preemptive, companies often are limited to reactive approaches, due to the lack of proper tools. Simulation approaches can be used as proactive decision-making tools, for several purposes, e.g., test alternative scenarios, determine certain performance measures, or simply to animate logistics flows, allowing to discover new knowledge from raw data. However, the SC processes that originate said data are comprised of several activities, which, in their turn, generate huge amounts of data, nowadays referred to as Big Data. Thus, such decision-making tools require Big Data structures to provide quality data concerning several SC processes, to be used in the simulation model (Tiwari et al. 2018; Zhong et al. 2016). Such tool is currently being developed at a Bosch plant of the automotive electronics industry sector. This tool consists of a Big Data Warehouse (BDW), which supports the SC simulation model, by extracting raw data from selected data sources and transforming it in quality data that can be used by the simulation model.

The research that was conducted to develop the BDW was published (Vieira et al. 2019), as well as the prototype of the simulation model (Vieira et al. 2018). The previously presented model did not use real industrial data, but rather used the typically available approach in simulation, consisting of using random distributions, to validate the data model, i.e., the variables selected for the project. Thus, this paper extends

the previously published works, by presenting the modeling approach that was followed to provide the SC simulation model with real-time features, while also using the real industrial data stored in the BDW. To achieve this, data-driven concepts were adopted, which allowed the model to automatically retrieve such data and adapt to possible changes in the data source, without any manual interventions.

This paper is organized as follows. Next section discusses literature related with the use of simulation with structures to store and integrate data. Section 3 presents the SC under study and section 4 the framework adopted for this project. Next, section 5 presents the main development approach for this solution and section 6 discusses and illustrates how the real-time feature is achieved. Finally, conclusions and future research directions are provided in the last section.

2   RELATED WORK

The purpose of this literature review section is to analyze literature concerning the use of structures to store, integrate and provide data to simulation models. At the end of the section, the identified gaps are listed.

Bottani (2008) proposed a Simul8 model of the logistic movements in a warehouse. It allowed to observe the impact of RFID in a warehouse, including the data that would be generated by such identification tags e.g., expiry date and type of component. This data was collected and stored in a Data Warehouse (DW), which, was used to extract information from the simulation results. The authors emphasized the use of the DW structure to derive value-added information from the simulation results.

Gupta et al. (2009) proposed an ExtendSim model to simulate different query workloads in a DW and analyze its performance. The workload was managed by considering query scheduling, admitting and executing, while also allocating resources to meet the performance objectives. This way, it was possible to experiment workload schedulers without trying them on the real DW.

Advocating that typically used commercial tools present simulation results as averages that mask important aspects of the transient behavior of the system, Ehmke et al. (2011) developed a DW that stored and allowed analysis the obtained simulation results. Thus, it was possible to extract value-added information from the simulation results. The simulation model was developed in Plant Simulation and the authors showed the benefits of such approach in a real case study which considered transportation scheduling, operational procedures, among other aspects.

Postulating that effective data cleaning produces considerable savings, since bad decisions can be taken based on bad data quality, Li and Joshi (2012) proposed a model to help practitioners determine the best cleaning strategies. The authors used ProModel to test two different cleaning approaches. According to the authors, simulation allowed them to better understand the interactions among data cleaning approaches and the produced results, which could not have been effectively achieved with other methods.

Nageshwaraniyer et al. (2012) proposed an Arena model of a coal mine, in which a DW provided data to a simulation model. The simulation outputs were stored in the DW for further analysis and reporting. According to the authors, the approach allowed the best parameters of the system to be determined.

As postulated by Kugu et al. (2012), data mining methods allow information to be extracted from data sets. However, when a DW is new, there may not be enough data stored. Grounded on this, the authors developed a simulation model that generated data for the DW, hence allowing data mining to be applied.

Truong and his co-authors (Truong et al. 2013; Truong et al. 2014; Truong et al. 2016) proposed a framework which can be used for simulation models dealing with high volumes of data. The simulation model uses data stored in a DW, which combines real data from the system and also results from previous simulation experiments. Hence, the framework allows typical data aggregation and analysis of the simulation outputs. According to the authors, their framework would not be suitable for simulation models dealing with small amounts of data, due to the considerable time required to implement such framework, which deals with different applications, namely: the DW, the simulation models and the Business Intelligence applications. In (Truong et al. 2014), Truong et al. used this framework and proposed a way to automatically calibrate simulation models, i.e., the process of tuning model parameters, so that the reliability of the simulation model increases. According to the authors, despite the positive results, they did not succeed in implementing the calibration approach in their framework.

Rabe and Dross (2015) proposed a framework comprised of a DW and a simulation model to predict the outcomes produced by changes in the logistics network of a company. The DW provided data to the simulation model and analysis for certain business indicators, while a copy of such DW was created to store simulation results. These results were then used as reward criteria for reinforcement learning algorithms, which were employed in the simulation.

As the above reviewed studies suggest, the scope of most papers is reduced to a specific process, e.g. warehousing, not considering all the activities occurring in a SC. Furthermore, all the reviewed papers considered traditional databases or DW, not using the benefits of Big Data concepts. Furthermore, to the best of the authors' knowledge, this is the first study to use Big Data technologies to provide real industrial data, of an automotive electronics plant, to a real-time SC simulation model. The lack of studies focusing on SC was recently corroborated (Vieira et al. 2018; Kagermann et al. 2013; Tiwari et al. 2018).

3   SUPPLY CHAIN CHARACTERIZATION

This project is being developed at a plant of the Bosch organization, which is an automotive electronics manufacturer. Like traditional automotive SC, this is a dynamic, complex and big SC. Figure 1 shows the location of all suppliers (top), suppliers from Europe (bottom left) and from Portugal (bottom right).
The Braga plant received materials from around 500 suppliers located in roughly 30 different countries, especially from Europe and Asia, with Germany and Netherlands having more suppliers and shipments from Europe, and Malasya, Taiwan, China, Hong Kong and Singapore having more shipments from Asia. These suppliers shipped a total of 218 848 deliveries.
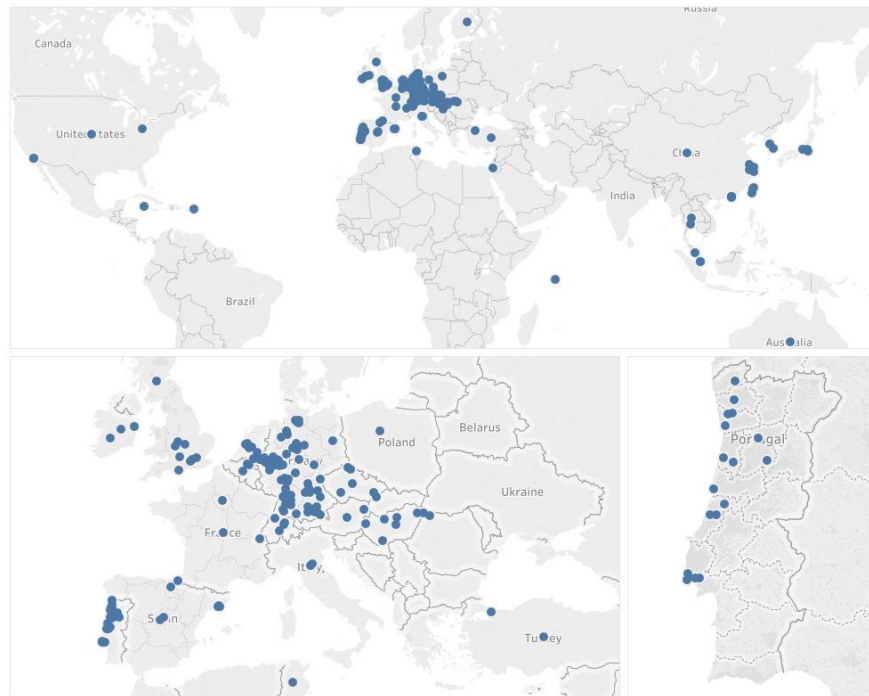

Figure 1: Location of all suppliers: world (top), Europe (bottom left) and Portugal (bottom right).

To comply with final customers' orders, the Braga plant places orders to its suppliers, which later culminate in arrivals. Most of these arrivals occur within the scheduled date. However, some suppliers provide the orders before the scheduled date. In these situations, these materials are stored in a special warehouse until the scheduled date is met. On the other hand, supplier deliveries may be delayed, potentially resulting in orders arriving after the scheduled data. To bypass this, when delays occur, special freights are

scheduled, which are usually faster but also have considerably higher associated costs. The internal movements are out of the scope of this paper, given its main research objective.

4   PROPOSED FRAMEWORK

This section presents the framework defined for this project. Figure 2 depicts a summary of such framework.
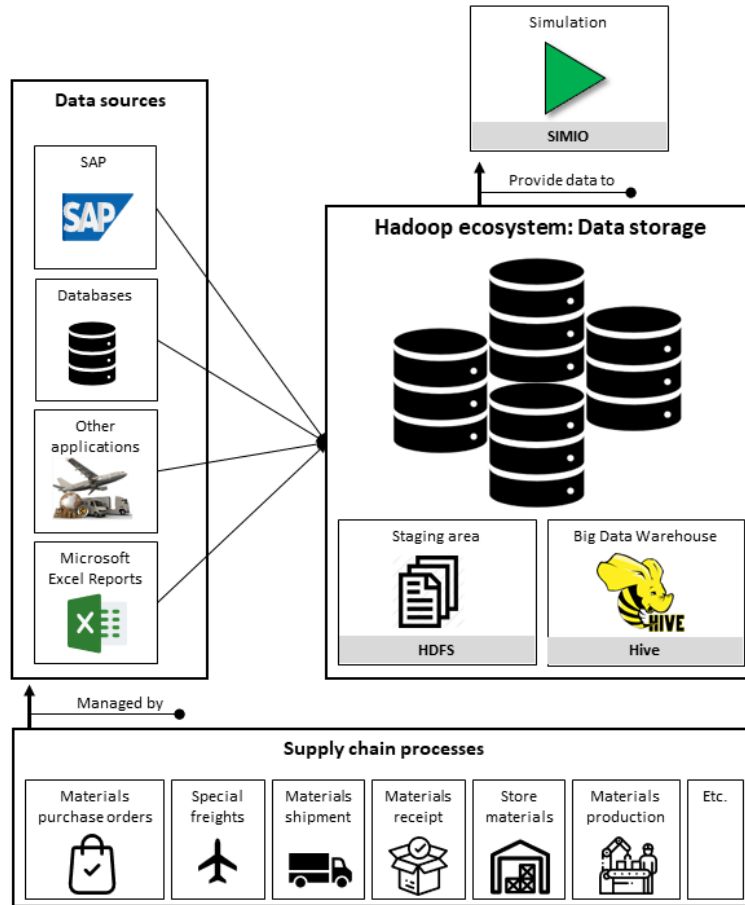

Figure 2: Framework defined for this project.

As can be seen, the tool consists of a BDW, which stores, integrates and provides the Big Data to a SC simulation model. Here, the Big Data term, as defined in literature, means that the data is too big, generated at rates too high, or too hard to process for existing tools, as corroborated by Madden (2012) and Costa et al. (2017). Therefore, for this purpose, the Big Data cluster of the organization is being used for this project.

As illustrated in the lower level of the above figure, the first step considered the identification of the relevant SC processes and their analysis. Raw materials purchase orders, material receipts, production consumptions and master data were some of the main SC processes that were considered.

Information systems, Enterprise Resource Planning (ERP), software and others are used to manage these processes, thus they generate data while users interact with them, comprising the data sources used for this project. The ERP used at Bosch is SAP, thus most of the data related with these processes is stored in this ERP. In it, users execute transaction codes, in order to visualize the intended data. To retrieve the intended data, it is possible to export it to different types of raw files, e.g. CSV. Some of the stored data is concerned with transaction events, e.g., logistic movements, purchase orders and others, for which SAP does not maintain historical data. To bypass this problem an additional database is used at Bosch, which

saves data from some of these SAP transactions, in different relational tables. This database stores not only logistics data, but also data of other departments, e.g. production, purchasing. Yet, not all transactions are currently being stored, as this is still a work in progress, meaning that data tables are regularly being added.

To collect data from these data sources, the traditional Extract-Transform-Load (ETL) process was followed, meaning that data is collected, the necessary data transformations are performed and, thereafter, files with this data are sent to the staging area, namely to the Hadoop Distributed File System (HDFS) system. The Talend software was used for this task. The same software was used to create the Hive tables, using the files that were stored in the HDFS.

Using the Hive tables, different purposes can be achieved: ad hoc querying, data analytics and visualization using software such as Tableau, and feed data to simulation models. In this project, the SIMIO simulation tool was used (Dias et al. 2016). Thus, the simulation software is connected to the Big Data cluster and the data tables are imported to the tool, therefore being possible to use them in the model. See (Vieira et al. 2019) for further details regarding the development of the Hive data tables. In addition, the simulation model may also generate files with simulation results, which can also be subject to ETL jobs to create additional Hive tables, enabling further data analysis and integration of these results in the analytical dashboards or even in the simulation model.

5 DEVELOPMENT APPROACH

The simulation model runs in a world-map view provided by Google Earth, where entities represent both internal and external logistic movements, i.e., orders to suppliers - and the respective material arrival - and material transfers that occur within the plant (e.g., store materials in the warehouse and send them to production). These movements comprise the types that occur in any SC. However, given the objective of this paper, only the external logistic movements are here addressed.

The developed model takes advantage of the data-driven approach, which consisted in modelling the data so that the SIMIO objects which create entities can create them based on the date attributes of the defined BDW tables. In its turn, these tables are the result of denormalizing data for each business process relevant for the simulation. Figure 3 shows an example of the tables that was imported to the simulation model, which specifies the attributes of orders that are placed to suppliers, as well as all attributes associated to the events related to that order, until it is shipped to the plant.

The figure shows fact tables associated to the business processes that are most relevant for the arrival of orders to the plant. As can be seen, the data model comprises the attributes related with placing orders to suppliers (Waterfall is the name of the SAP transaction used at the plant) and the later arrival (OTD stands for On-Time Delivery). These arrivals can, in their turn, originate early arrivals (Pending) or possible late arrivals, which, in their turn, originate onerous special freights to try to get the materials in time, in order to avoid possible production stoppages.

This Big Data approach here applied differs from other traditional approaches, which typically consist in storing data in multiple tables, since it fully denormalizes the data of the multiple processes in a single BDW table. While this exponentially increases the storage space occupied, it also ensures that no join operations are required throughout the simulation. In fact, given the SC characterization provided in section 3, i.e., the number of suppliers and different types of materials involved, if the thousands of entities that are constantly flowing in the system would need to perform search operations among different tables (12 tables in the example of Figure 3), the simulation model would be loaded with this computational effort, making the simulation run almost unbearable.

Thus, the attribute that triggers the creation of entities is the idReleaseTransmissionDate, which stores the dates of the orders placed to suppliers. When these entities are created, the process depicted in Figure 4 is executed. It should be noted that these processes are executed by tokens, which is a SIMIO concept that executes the processes, in place of the entities; it is also possible to have several tokens executing the same process associated to the same entity, in parallel. Lastly, note that the processes here presented were rather simplified, with the purpose of explaining the approach - the real ones are considerably larger.
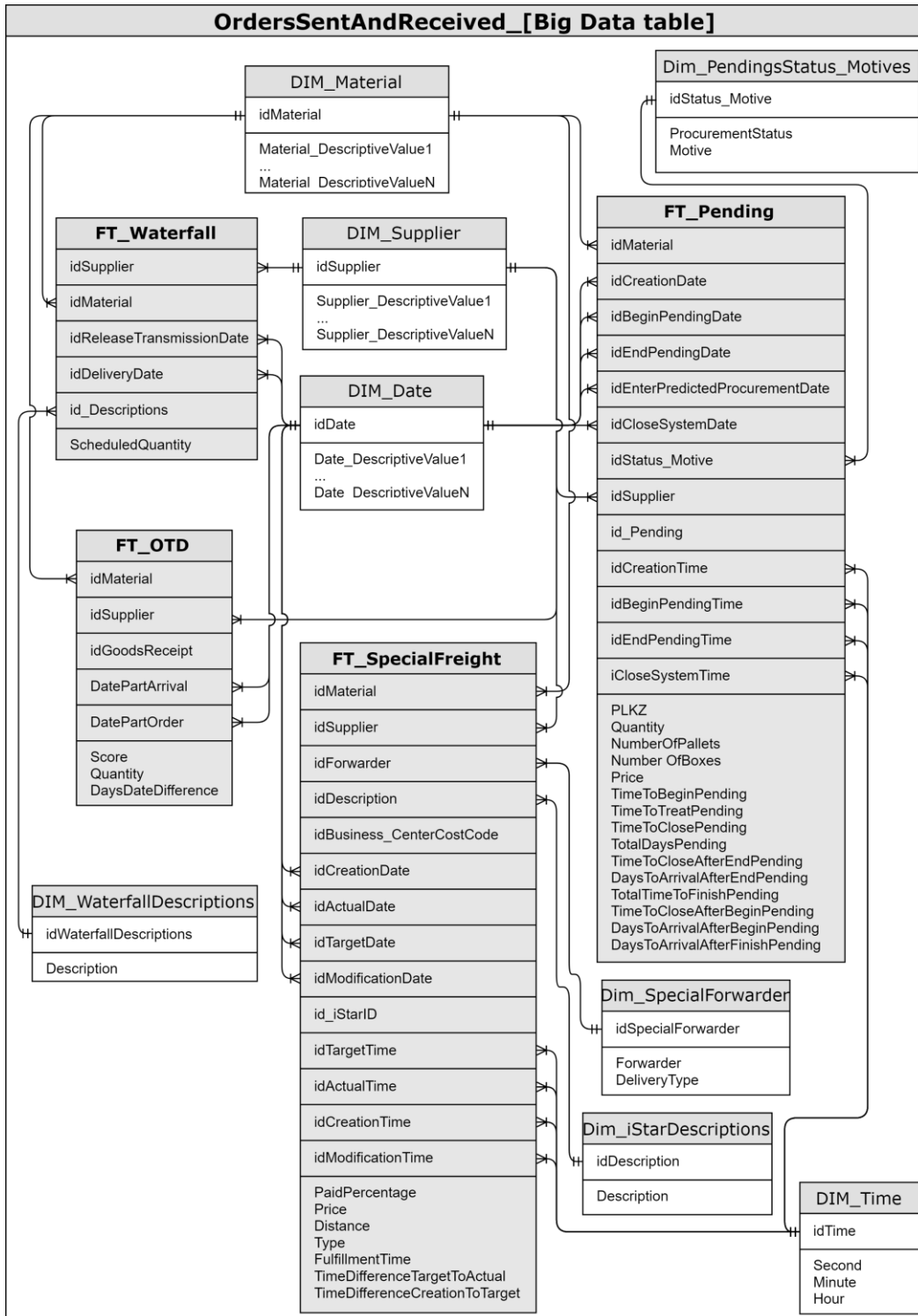
Figure 3: Illustration of the Big Data approach followed in this project.

As Figure 4 shows, the token starts by making some assignments to the created entity, related to the associated data row. Thereafter, the entity is transferred to Free Space, using the "Go to Free Space" Transfer step, and travels to the location of the supplier. This movement represents the activity of placing

an order to a supplier and was modelled to occur at a very high speed. This allows users to visualize orders being placed and the duration is considered when setting the total order lead time. When arriving at the location of the supplier, the token executes a series of Decide steps, responsible to model alternative SC scenarios. Finally, the token verifies if the entity has a valid transit time (higher than 0) and executes a new process, which is responsible for setting the order's lead and transit time durations for the cases in which the values of the BDW cannot be used. This was necessary in order to face some of the data issues that were verified, which are out of the scope of this paper, nonetheless.
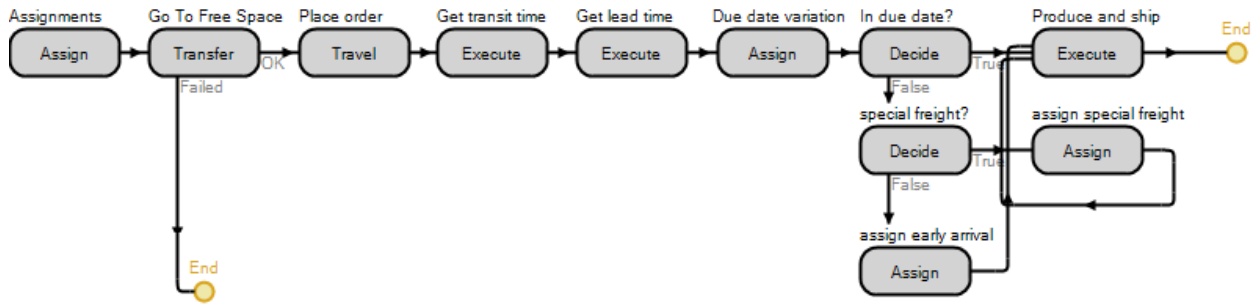

Figure 4: Process executed to model orders being placed to the respective suppliers.

Afterwards, the token verifies if the associated order was delayed, calculates the respective arrival date deviation and verifies if it is now a special freight. Finally, the token executes the Execute "Produce and ship" step, which models the production and shipment of the order back to the plant.

6   REAL-TIME SUPPLY CHAIN SIMULATION

In what concerns simulation, as mentioned by Barlas and Heavey (Barlas and Heavey 2016), a real-time simulation is capable of initializing its run with data from its sources. For Fowler and Rose (Fowler and Rose 2004), one of the main pillars of a real-time simulator is its ability to automatically build models without users' interactions. Alluded to the above, the real-time feature of this solution concerns its ability to automatically retrieve data from the BDW (which is the only data source of the simulation model) and adapt the model (e.g., different suppliers or plant location) to meet what is stored in the data.

Apart from the automatically adjusting to the data it receives, the authors argue that an additional feature is required for a real-time SC simulation model, which consists in allowing users to load the state of the system, when analyzing specific time slots. The main reasons for this are now discussed. First, in typical simulation models of SC systems, the simulation time is considerably high, to capture the lasting activities that take place, e.g.: shipment, production. Second, the considerable amount of entities generated by such systems and the Big Data usually involved, contributes to increase the elapsed time required to run a simulation. Finally, on one hand, decision-makers may be interested in analyzing specific time slots, albeit they still need to wait a considerably time to reach the intended simulation time. On the other hand, if their focus of analysis is in the beginning of the simulation, it is possible that the simulation has not reached a steady-state, which does not allow a proper analysis. The next two subsections cover the features above described and implemented in the proposed solution.

6.1   Real-Time

The purpose of this subsection is to show how the real-time feature could be achieved. In light of this, the automatic connection to the BDW is automatically achieved by setting up the necessary connections between the simulation software and the Big Data cluster of the plant. Thereafter, the data retrieved from this data source is managed in the software as if it was a normal data table. In its turn, to ensure that the simulation model can automatically adjust to changes in the data, some modeling approaches were crucial, which are listed below:

- **Avoiding physical objects**

Eliminating the number of physical objects that are placed in the model ensures that when the data changes or new data is added there is no need to update the model, since the only physical object that comprises the model is the one which sets the Braga plant. This way, the system logic is stored in the data, and the processes (which were described in section 5) allow each entity to retrieve its own characteristics and activities from the data and react accordingly. Hence, the only physical object that comprises the simulation model is the one which sets the location of the Braga plant. Therefore, dragging it to any location sets the new plant's location and the remaining physical locations (e.g., supplier's locations) are set by the data and do not require any object to be placed in the model.

- **Using Free Space Movement**

In most simulation tools, entities travel between objects or locations using some type of connector. However, SIMIO offers the Free Space Movement feature, which allows entities to travel between objects without requiring any link or connection between them or even move freely in an orthogonal tridimensional space. To achieve this, the movement of entities needs to be modeled using processes, which adds complexity to the development, as it is not possible to connect two distinct locations. As all activities of an entity are saved in said entity as its characteristics, it knows the locations to where it needs to travel. Thus, in the processes it is possible to model these movements using the appropriate building blocks, which establish the entities' activities, trajectories to follow and other aspects, such as possible transportation delays, disruptions and others.

To demonstrate that the model is capable of automatically updating to new data, the data was changed, in order to exclude all suppliers from Europe and concentrate them in countries from America, Asia and Africa (to be easier to verify the differences from the real data); most of these fictional suppliers' locations considered new countries that are not listed in the original data. Moreover, the object referring to the Braga plant was also moved to Luanda, Angola. Figure 5 shows the model during run time, which considers the above described changes to the data.

Figure 5: New plant location and new suppliers' locations.

As the figure shows, the model could automatically adjust to the new data, without any intervention. Furthermore, the figure also shows entities traveling to the new plant's location, to deliver orders that had been previously placed to the fictional suppliers.

6.2 Loading the State of the SC

This subsection discusses how the model is able to load the state of the system from the BDW. In a normal run, the simulation would only create the orders that are placed after such beginning date. However, there may be orders that are in transit, or under supplier lead time, for instance, which would be ignored. Thus, it is necessary to search among the tables, for rows that represent orders in these situations, and consider them to the model. Figure 6 shows the process that was created to achieve this.
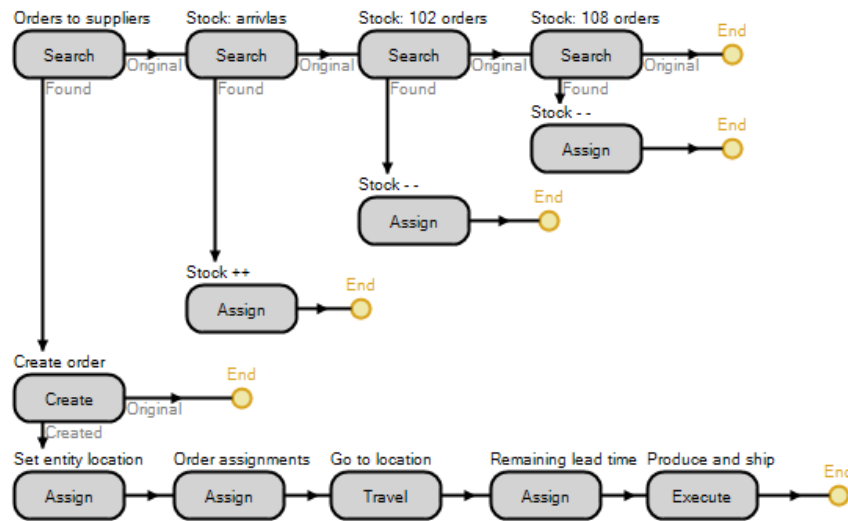

Figure 6: Process executed to load the state of the model from the BDW, at the start of simulation run.

The process starts by executing the Search "Orders to suppliers" step, which searches the table with data from the orders and arrivals, looking for rows referring to orders that should arrive to the plant after the simulation start date, but whose orders were placed before the simulation start date. For each data row that is found meeting the specified search condition, the process continues by executing the Create "Create order" step, which creates one entity for each found data row and associates the created entity to the respective row. Thus, it is possible to make assignments to the order according to the table row associated to it. Next, the Travel "Go to location" step sends the created entity to the location of the supplier and calculates the remaining lead time in the next Assign step. Finally, the Execute "Produce and ship" step makes the created entity execute the process which allows them to continue the normal process flow described in section 5. This process was also considerably simplified, like the ones discussed in section 5.

Apart from creating the orders that were placed before the simulation start date, but should arrive after the simulation start date, it is also necessary to calculate the stock at a given date. The steps to do so are depicted in Figure 6, however, as noted earlier, the internal movements are out of the scope of this paper. Figure 7 shows the state of the system on the 1st of April being loaded.

As can be seen, when the model starts, there are already entities in the respective suppliers' location, in contrast to running the model without loading the state of the system, where the simulation starts without orders on the suppliers' end; the stock is also loaded through the calculations above described. Furthermore, the transit and lead times of each order are also updated, thus it is ensured that, when the simulation continues, the orders arrive at the plant at the arrival date specified in the BDW; naturally, the orders that have their lead time reduced with this approach will travel at considerably higher speeds, in order to reach the plant in the intended date.
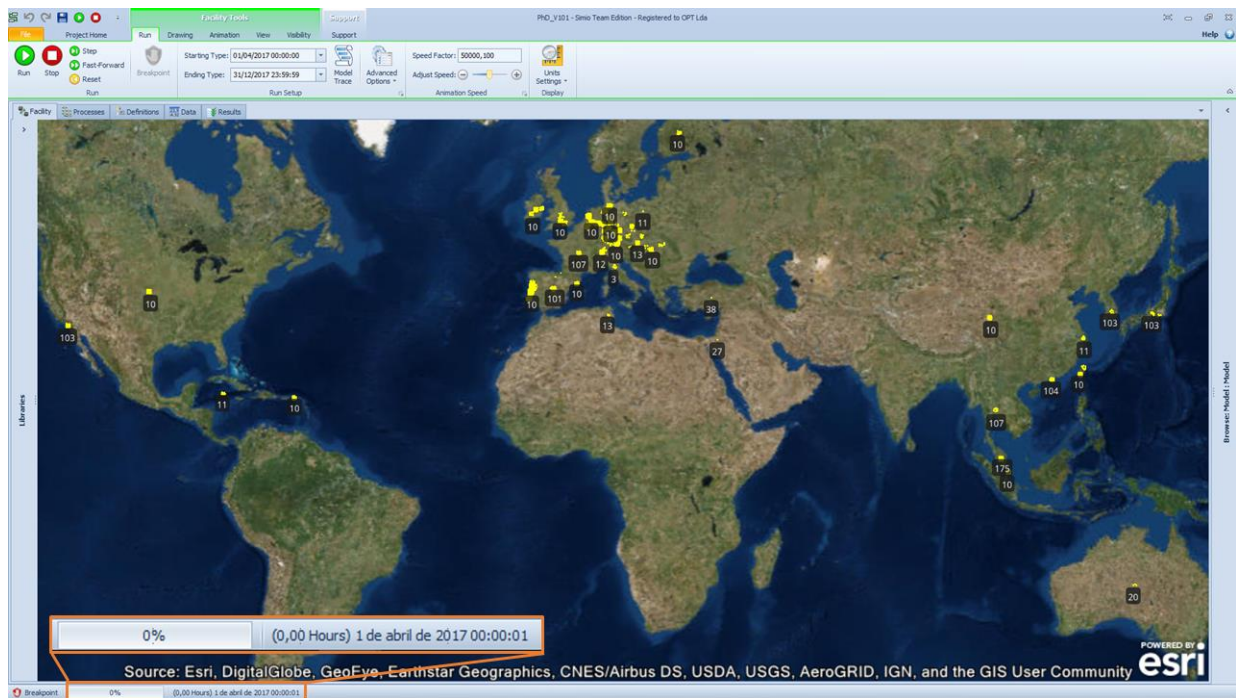
Figure 7: Loading the state of the system from the BDW.

7   CONCLUSIONS

Being complex and dynamic networks, SCs are prone to events that affect its performance. In light of this, simulation may be used as an effective tool to assess such uncertainties. This paper presented a real-time decision-support tool, comprised of a BDW and a data-driven SC simulation model. Albeit some studies exist which reported the use of simulation and structures to store and integrate data, no study was found coupling the benefits of Big Data with simulation, which this paper achieved to do. Thus, this application of both areas in a SC, as well as the development approach that allowed this solution to have real-time features, can be seen as the major contribution of this paper.

Typical SC simulation studies target specific problems, resulting in models too tailored for such problems. In fact, the literature concerning these cases is well established. However, especially in the advent of Big Data, the authors argue that this approach should shift towards using the real industrial data stored in these structures, as these provide significative level of detail to allow complex and detailed simulations of several scenarios, hence not being limited to specific problems.

Apart from the mentioned wider level of detail, since the simulation model has access to real industrial data, it should also be able to automatically adapt to possible data changes, hence matching the main characteristics of a real-time simulation. In fact, in the solution presented in this paper, elements such as suppliers' locations, entities' movements and others were not manually placed, since the model automatically adjusts its elements to the data it gets from the BDW; i.e., the model is drawn by the data and not by the users. This way, it is expected that the interest of managers in these type of tools increases, as it is not necessary to update the model, which typically would require additional technical knowledge and often limits the applicability of simulation models. This is particularly relevant in the case of SCs, since these are dynamic networks, subject to constant changes.

This paper also contributed to literature by providing the experience of working on a SC simulation model in a Big Data context, namely by describing the development approach that allowed to confer real-time features to the simulation solution. In particular, the data-driven concepts that were adopted, as well as other additional modelling approaches, which allowed the simulation to automatically adapt to the data

it receives from the BDW. In fact, when starting to develop an artifact, if certain features are mandatory (e.g., real-time) it is crucial to ponder on the modelling approach which will allow such feature to be achieved. Therefore, the authors hope that other researchers also find this work useful, when seeking approaches to allow SC simulation models to work in real-time.

Finally, the solution's animation should also be emphasized, as it has a direct integration with Google Maps and runs in a 3D world map view, which enriches its visualization. Moreover, the solution is also able to load the system's state from the BDW. This is important, because it allows users to analyze different time slots without having to wait for the simulation to reach the point in question.

The real-time part solution discussed in this paper applies solely to the SC simulation model, which can automatically adapt to changes, or to new data from the BDW. However, this is still dependent on manually refreshing the BDW, which considers a manual step, related with the extraction of data. Thus, future research should include efforts to make this BDW also being able to operate in real-time. Finally, with the proposed tool, the next research steps would naturally be to use it to test different disruption scenarios and to make predictions, thus achieving the main benefit for which simulation can be used in this type of problems. In fact, this is already being performed, but was not included in this paper as it falls out of the scope of the research objective established.

ACKNOWLEDGMENTS

REFERENCES

Barlas, P. and C. Heavey. 2016. "Automation of Input Data to Discrete Event Simulation for Manufacturing: A Review". *International Journal of Modeling, Simulation, and Scientific Computing* 7(1): 1630001.

Bottani, E. 2008. "Reengineering, Simulation and Data Analysis of an RFID System". *Journal of Theoretical and Applied Electronic Commerce Research* 3(1): 13–29.

Costa, E., Costa, C., and Santos, Maribel Y. 2017. "Efficient Big Data Modelling and Organization for Hadoop Hive-Based Data Warehouses". In *European, Mediterranean, and Middle Eastern Conference on Information Systems, EMCIS 2017,* 7th - 8th, Coimbra, Portugal, 299:3–16.

Dias, Luis M. S., Vieira, Antonio A. C., Pereira, Guilherme A. B., and Oliveira, Jose A. 2016. "Discrete Simulation Software Ranking - A Top List of the Worldwide Most Popular and Used Tools". In *Proceedings of the 2016 Winter Simulation Conference*, edited by T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, 1060–71. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Ehmke, J. F., Großhans D., Mattfeld, D. C., and Smith, L. D. 2011. "Interactive Analysis of Discrete-Event Logistics Systems with Support of a Data Warehouse". *Computers in Industry* 62(6): 578–86.

Fowler, J. W., and Rose, O. 2004. "Grand Challenges in Modeling and Simulation of Complex Manufacturing Systems". *Simulation* 80(9): 469–76.

Gupta, C., Mehta, A., Wang, S., and Dayal, U. 2009. "Fair, Effective, Efficient and Differentiated Scheduling in an Enterprise Data Warehouse". In *Proceedings of the International Conference on Extending Database Technology: Advances in Database Technology, EDBT*, March 24th -26th, Saint Petersburg, Russia, 696–707.

Kagermann, H., Helbig, J., Hellinger, A., and Wahlster, W. 2013. *Recommendations for Implementing the Strategic Initiative INDUSTRIE 4.0: Securing the Future of German Manufacturing Industry*. Frankfurt, Germany: Platform Industrie 4.0.

Kugu, E., Altay, L., and Sahingoz, O. K. 2012. "Using Agent Based Modeling and Simulation for Data Mining". In *Proceedings of the International Conference on Neural Information Processing, ICONIP,* November 12th - 15th, Doha, Qatar, 7664 LNCS:258–65.

Li, Y., and Joshi, K. D. 2012. "Data Cleansing Decisions: Insights from Discrete-Event Simulations of Firm Resources and Data Quality". *Journal of Organizational Computing and Electronic Commerce* 22(4): 361–93.

Madden, S. 2012. "From Databases to Big Data". *IEEE Internet Computing* 16(3): 4–6.

Nageshwaraniyer, S. S., Meng, C., Maghsoudi, A., Son, Y.-J., and Dessureault, S. 2012. "Simulation-Based Decision Support System for Sustainable Coalmining Operations". In *Annual Conference and Expo of the Institute of Industrial Engineers,* May19th-23th, Orlando, Florida, USA, 1574–83.

Rabe, M., and Dross, F. 2015. "A Reinforcement Learning Approach for a Decision Support System for Logistics Networks". In *Proceedings of the 2015 Winter Simulation Conference*, edited by L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti, 2020–32. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Simchi-Levi, D., Kaminsky, P., Simchi-Levi, E., and Shankar, R. 2008. *Designing and Managing the Supply Chain: Concepts, Strategies and Case Studies*. Tata McGraw-Hill Education.

Tiwari, S., Wee, H. M., and Daryanto, Y. 2018. "Big Data Analytics in Supply Chain Management between 2010 and 2016: Insights to Industries". *Computers & Industrial Engineering* 115: 319–30.

Truong, T. M., Amblard, F., Gaudou, B., and, Sibertin-Blanc, C. 2014. "To Calibrate & Validate an Agent-Based Simulation Model : An Application of the Combination Framework of BI Solution & Multi-Agent Platform". In *Proceedings of the International Conference on Agents and Artificial Intelligence*, March 6th -8th, Angers, France, 2:172–83.

Truong, T. M., Amblard, F., Gaudou, B., and Sibertin-Blanc, C. 2016. "CFBM - A Framework for Data Driven Approach in Agent-Based Modeling and Simulation". In *International Conference on Nature of Computation and Communication, ICTCC,* March 17th -18th, Rach Gia, Vietnam, 168:264–75.

Truong, T. M., Amblard, F., Gaudou, B., Sibertin-Blanc, C., Truong, V. X., Drogoul, A., Huynh, H. X., and Le, M. N. 2013. "An Implementation of Framework of Business Intelligence for Agent-Based Simulation". In *Proceedings of the Symposium on Information and Communication Technology*, December 5th -6th, Danang, Vietnam, 35–44.

Vieira, A. A. C., Dias, L. M. S., Santos, Maribel Y., Pereira, G. A. B., and Oliveira, José A. 2018. "Simulation of an Automotive Supply Chain in Simio: Data Model Validation". In *European Modeling and Simulation Symposium, EMSS 2018*, September 17th -19th, Budapest, Hungary, 294–301.

Vieira, António A. C., Pedro, L., Santos, Maribel Y., Fernandes, J. M., and Dias, Luís S. 2018. "Data Requirements Elicitation in Big Data Warehousing". In *European, Mediterranean, and Middle Eastern Conference on Information Systems, EMCIS,* October 4th - 5th, Limassol, Cyprus, 341:106–13.

Vieira, António A. C., Dias, Luís M. S., Santos, Maribel Y., Pereira, Guilherme A. B., and Oliveira, José A. 2018. "Setting an Industry 4.0 Research and Development Agenda for Simulation – A Literature Review". *International Journal of Simulation Modelling* 17(3): 377–90.

Zhong, R. Y., Newman, S. T., Huang, G. Q., and Lan, S. 2016. "Big Data for Supply Chain Management in the Service and Manufacturing Sectors: Challenges, Opportunities, and Future Perspectives". *Computers and Industrial Engineering* 101: 572–91.

AUTHOR BIOGRAPHIES

ANTÓNIO A. C. VIEIRA is an Invited Assistant Professor and Researcher of the ALGORITMI Research Centre, University of Minho, Portugal. He holds a MSc degree in Systems Engineering at the University of Minho. He is making his doctoral studies in Advanced Engineering Systems for Industry at the University of Minho, in partnership with Bosch Car Multimedia Braga. His research interests include Modelling and Simulation, Logistics and Transportation, and Big Data. His email address is antonio.vieira@dps.uminho.pt.

LUÍS M. S. DIAS is an Auxiliar Professor and Researcher of the ALGORITMI Research Centre, University of Minho, Portugal. He holds a PhD degree in Production and Systems Engineering from the University of Minho, Portugal. His research interests include Simulation, Systems Performance, Operational Research and Systems Visual Modeling. His email address is lsd@dps.uminho.pt.

MARIBEL YASMINA SANTOS is an Associate Professor at the Department of Information Systems and Researcher of the ALGORITMI Research Centre, University of Minho (UMinho), in Portugal. She received the Aggregated title (Habilitation) in Information Systems and Technologies (IST) from UMinho in 2012 and a PhD in Information Systems and Technologies from UMinho in 2001. She was Secretary-General of the Association of Geographic Information Laboratories for Europe (AGILE) from May 2013 to June 2015. Her research interests include Business Intelligence and Analytics, Big Data and (Big) Data Warehousing. Her email address is maribel@dsi.uminho.pt.

GUILHERME A. B. PEREIRA is an Associate Professor and Researcher of the ALGORITMI Research Centre, University of Minho (UMinho), in Portugal.. He holds a PhD degree in Manufacturing and Mechanical Engineering from the University of Birmingham, UK. His research interests include Operational Research and Simulation. Currently he is Pro-Rector of the University of Minho. His email address is gui@dps.uminho.pt.

JOSÉ A. OLIVEIRA is an Auxiliar Professor and Researcher of the ALGORITMI Research Centre, University of Minho (UMinho), in Portugal. He holds a PhD degree in Production and Systems Engineering from the University of Minho, Portugal. His research interests include Optimization with Heuristic Methods in Systems Engineering. His email address is zan@dps.uminho.pt.