

## **A METHODOLOGY FOR CONTINUOUS QUALITY ASSURANCE OF PRODUCTION DATA**

Jon Bokrantz  
Anders Skoogh  
Jon Andersson

Jacob Ruda  
Dan Lämku

Department of Product and Production  
Development  
Division of Production Systems  
Chalmers University of Technology  
Hörsalsvägen 7  
Gothenburg, SE-41296, SWEDEN

Department of Virtual Manufacturing Methods and  
IT Tools  
Volvo Car Group  
Karossvägen 2  
Gothenburg, SE-41531, SWEDEN

### **ABSTRACT**

High quality input data is a necessity for successful Discrete Event Simulation (DES) applications, and there are available methodologies for data collection in DES projects. However, in contrast to standalone projects, using DES as a day-to-day engineering tool requires high quality production data to be constantly available. Unfortunately, there are no detailed guidelines that describes how to achieve this. Therefore, this paper presents such a methodology, based on three concurrent engineering projects within the automotive industry. The methodology explains the necessary roles, responsibilities, meetings, and documents to achieve a continuous quality assurance of production data. It also specifies an approach to input data management for DES using the Generic Data Management Tool (GDM-Tool). The expected effects are increased availability of high quality production data and reduced lead time of input data management, especially valuable in manufacturing companies having advanced automated data collection methods and using DES on a daily basis.

### **1 INTRODUCTION**

There is no doubt that today's business environment within the automotive industry puts huge pressure on car manufacturers. In order to gain a competitive advantage, they need to meet high and rapidly changing customer demands, which in turn requires fast development of flexible, high performance, and cost-effective production systems. Interestingly however, the average Overall Equipment Effectiveness (OEE) in Swedish manufacturing industry have for more than two decades been consistently reported to be around 50-60% (Bokrantz et al. 2015). Clearly, reduced lead times in product realization projects and increased efficiency in both existing and future production are immense challenges that the automotive industry is facing. To meet these challenges, companies have focused on utilizing virtual tools and methods for product and production development. For example, Volvo Car Group (VCG) have included Discrete Event Simulation (DES) as an important tool in their concurrent engineering processes (Jägstam and Klingstam 2002). In order to reap the benefits of DES, numerous authors have stressed the need to adopt a systematic simulation methodology. There is in fact a general agreement on the appropriate structure for such methodologies (Robinson, 2004), where e.g. Banks et al. (2004) and Law (2007) are some of the most informative sources. Within simulation projects, the data collection phase have been argued to be particularly time-consuming (Ülgen et al. 1994, Robertson and Perera 2002), and empirical studies have shown that it constitutes on average 31% of the total project time (Skoogh and Johansson, 2007). Poor data availability is a major reason for long data collection time (Perera and Lyanage 2000),

and authors have therefore proposed both data collection methodologies (Lehtonen and Seppälä 1997; Perera and Lyanage 2000; Skoogh and Johansson 2008) as well as software solutions to reduce the lead time for input data management, e.g. the Generic Data Management Tool (GDM-Tool) (Skoogh, Johansson, and Stahre 2012). However, a common denominator of all the proposed methodologies, both regarding overall simulation methodologies and data collection methodologies, is that they stem from a project-based approach to simulation.

Standalone simulation projects are often run over the course of several months (Robinson 2004), in which lengthy data collection phases are often acceptable. In contrast, there has been a shift towards increased use of DES as day-to-day engineering tool in manufacturing companies (Randell 2002), and there is an emerging need for purely data-driven production analyses on a close to real-time basis, e.g. tracking system losses and identifying bottlenecks without requiring simulation. Clearly, this situation is vastly different from the typical standalone simulation project. In particular, it completely changes the demands on data collection procedures since high quality production data ideally should be available and ready to use at any given time. This is impossible to achieve by purely relying on project-based data collection methodologies, meaning that there is an obvious gap in available research.

Therefore, the aim of this paper is to propose a structured methodology for continuous quality assurance of production data. This is achieved through a case study within VCG. The proposed methodology will contribute towards increased availability of high quality production data, which will improve input data management to DES, reduce overall lead time in development projects, and aid in improving both existing and future production systems.

### **1.1 Case description, VCG Torslanda**

This case study is performed at VCG in Torslanda (VCT), an automotive industry manufacturing plant in Sweden. By 2020, VCG aims to sell 800,000 cars worldwide, a 45% increase in annual sales in 10 years. Until then, they also strive to achieve a development lead time of 20 months from program start to production of the first saleable vehicle. In parallel, VCG aspires to reach at least 85% OEE in running production. In regard to simulation, VCG have integrated DES in their concurrent engineering projects through continuously improved long life-cycle models (Jägstam and Klingstam 2002). Within these development projects, rapid evaluations of design concepts are expected, and detailed simulation experiments related to production improvements are often requested to be answered within 24 hours. Hence, a key factor for reducing development lead time and increasing production efficiency is to enable a better use of production data. Therefore, a new production disturbance follow-up system has recently been implemented for all new production lines. In addition, a recent study mapping of the availability of input data for DES within VCT identified many issues with data quality (Carter and Ludvigsson 2014). In this study, the current state of input data management for DES and general data collection procedures are mapped based on three previously completed cases in concurrent engineering projects, referred to as E1, E2, E3; see further details in section 3.1.

## **2 QUALITY, VALIDITY, AND CREDIBILITY IN PRODUCTION DATA**

Important factors to achieve a collection of high quality production data are presented in this section, with a particular focus on theory covering data validation and data quality indicators to achieve credibility in input data. Clearly, a key factor for any simulation application is to work in an organized manner. However, even the most well-performed simulation studies could be rejected if one fails to achieve acceptability of the results. Eriksson (2005) studied this topic, and argues that decision makers' ignorance of simulation results relates to that that simulation holds no special place amongst other decision support methods. Balci (1990) however highlights that acceptability can be influenced by striving to achieve high credibility of the simulation results, and therefore presents a hierarchy of credibility assessment stages, in which data validation is included.

Sargent (2005) defines data validity from a simulation perspective as 'ensuring that the data necessary for model building, model evaluation and testing, and conducting the model experiments to solve the problem, are adequate and correct.' Therefore, any validation activity requires a structured methodology (Balci et al. 2000), and several authors have suggested ways to perform organized data validation (Balci 1990, Rabe, Spieckermann, and Wenzel 2009). A practical example is e.g. to utilize face validation by collaboration with process experts (Skoogh and Johansson 2008). Validation and verification of input data can also be partly automated (Perera and Lyianage 2000), e.g. by using the GDM-Tool (Skoogh, Johansson, and Stahre 2012). However, Sargent (2005) argues that unfortunately, there is not much that can be done to determine whether the data are correct. Nonetheless, Balci et al. (2000) explains that in order to achieve credibility of data, it is critical to assess the following *Data Quality Indicators*:

**Data Quality Indicators** *Accessibility, Accuracy, Clarity, Completeness, Consistency, Currency, Precision, Relevance, Resolution, Reputation, and Traceability.*

These indicators can be seen as guiding criteria to ensure high quality production data. However, not all data require high accuracy and validity (Skoogh and Johansson 2008), meaning that quality, validity, and credibility always need to be assessed in relation to the objectives (Balci, 1990, Williams and Ülgen 2012). However, in the case of long-life cycle models, data integrity needs to be checked continuously (Ülgen et al. 1994), and Rabe, Spieckermann, and Wenzel (2009) suggests that if regular updates of data are needed, a suitable process needs to be in place or prepared.

### **2.1.1 Roles, responsibilities, and data requirements within an organization**

Knowledge and experience is essential in simulation, and every simulation project team should include designated roles and responsibilities to avoid project failure, e.g. leadership, client, modeling, system experts, data providers etc. (Balci, 1990; Randell 2002; Robinson 2004). However, in contrast to literature regarding overall simulation projects, none of the articles proposing data collection methodologies deliberately suggests any specific roles or responsibilities. This is very unfortunate since it is evident that data requirements can largely vary between organizational functions. For example, using data to calculate energy consumption of machines (Dettman et al. 2013) differ from using it for predictive maintenance (Tretten and Karim 2014), and minor stoppages should be calculated as a lump sum to be manageable for operators (Ljungberg 1998). Furthermore, simulation are often ignored in the specification of the collection system and databases, resulting in that learning and understanding all data sources can be very time-consuming (Skoogh and Johansson, 2007). To tackle this issue, Randell (2002) argues that data should be collected in such a way that it is useful for a variety of activities, and he therefore proposes a generic framework that describes the appropriate data structure.

The essence of collecting data for various purposes is to be able to provide the right information to the right user with the right quality at the right time. This idea of a pull flow of information have been utilized in information and knowledge management systems for product development, where e.g. Yang and Cai (2009) explains a lean supermarket thinking aimed towards fulfilling 4 criteria: (1) information and knowledge are always fresh and up to date, (2) information and knowledge are sufficient to serve all the needs, (3) we know where each information and knowledge are stored, and (4) the information is ready to be pulled at the right time, the right kind, and right amount.

### **2.1.2 Success factors and pitfalls in data collection for simulation**

There are plenty of available literature highlighting success factors and pitfalls in simulation. In table 1, descriptions of success factors and pitfalls related to the data collection phase are presented, based on Law and McComas (1986), Sadowski (1989), Law and McComas (1991), Musselman (1994), Ülgen et

al. (1994), and Williams and Ülgen (2012). Note however that Ülgen et al. (1994) already compiled previous pitfall papers, and table 1 therefore serves as a general overview of potential difficulties.

Table 1: Pitfalls and success factors in the data collection phase in simulation projects (summary from references above).

- 
- Poor data availability is a major problem, especially due to inaccurate information and lack of formalized operation procedures. Strive to collect all possible data from various sources, but be cautious for optimistic data, and make a sensitivity analysis if needed.
  - Make assumptions about data, but document them. Alert the client about data weaknesses, but review and validate assumptions before presenting them, and make sure to reach agreement.
  - Variation, e.g. machine breakdowns, have the most effect on the system, so never replace a distribution by its mean. Instead, use a distribution fitting software to represent any variability with an appropriate statistical distribution. In fact, fitted distributions of bad data are as bad as unreliable estimations, and wrong conclusions can be drawn from data that appears to be good.
  - Data collection should start early, and the data structure should be defined before model building starts. However, data collection should not start if no macro data are available, if experts disagree on estimations, or if the data are not crucial. Ask questions about the system, collaborate with data collection people to make data collection plans, and strive to exploit concurrency in data collection. Examine all data and learn when it was collected, detect early if data are unusable, and question all data. For long life-cycle models, continuously check the integrity of input data, regularly test time-sensitive input data, and set up long term data collection procedures for revalidating input data.
- 

### 3 METHODOLOGY OF THE STUDY

The aim of this study was to propose a methodology for continuous quality assurance of production data. To achieve this, an embedded representative single-case study approach was adopted (Yin 2003). Based on three cases from completed concurrent engineering projects (E1, E2, E3), the current state of input data management for DES and general data collection procedures were mapped within VCT. The 20 largest problems in the current state were identified and used as input to the proposed methodology.

#### 3.1 Gathering of empirical data

Three sources of empirical evidence were used: semi-structured interviews, direct observations, and reviews of archival records. Direct observations consisted of walkthroughs with the simulation analyst, focusing on explaining the use of DES and current input data management procedures. Reviews of archival records were conducted by studying production data in the follow-up systems and data collection guideline documents (e.g. cycle-time definitions) (Yin 2003). First, direct observations and an interview was performed with a simulation analyst, focusing on mapping the current approach to input data management. The interview template was constructed based on literature review and observations, and covered a first general part regarding simulation methodology and data collection methodology, and a second part aimed regarding the three cases E1, E2, and E3. The cases revolved around assignments that the simulation analyst had faced in previous concurrent engineering projects, namely:

**E1** *Determine the standalone throughput of a specific production line by verifying its availability.*

**E2** *Determine the performance of a new laser station by estimating technical availability on all components that affect throughput.*

**E3** *Determine the buffer size between the new laser station and the subsequent arc weld station to ensure sufficient throughput.*

Thereafter, three additional interviews were conducted, for which the interview templates were based on the initial interview and observations with the simulation analyst. These included one maintenance engineer responsible for implementing the new disturbance follow-up system, one maintenance engineer working with disturbance data collection, and one laser equipment expert who was largely involved in the E2 and E3 case. The equipment expert is the process owner for all laser equipment in the body shop, and connected in a world-wide collaborative network of Volvo plants, universities, and equipment manufacturers. The first interview lasted for 2 hours, the subsequent interviews for 45-75 minutes, and all were audio recorded and transcribed within 24 hours. Although a sample size of 4 interviewees is rather limited, the participants are in fact the key persons involved in the collection of production data at VCT.

### 3.2 Analysis and presentation of findings

The results from the three sources were triangulated by developing converging lines of enquiry, which is suitable for single-case studies (Yin 2003). A strategy of relying on theoretical propositions was adopted, which is a way of letting available theory guide the analysis in order to focus attention on certain, relevant aspects and ignore other, irrelevant, aspects. This enables a focus on the most significant parts of the study (Yin 2003). The 20 largest problems with current input data management to DES and general data collection procedures are presented first, followed by a description of the current state at VCT. Thereafter, the proposed methodology is presented, aimed towards solving the problems in the current state.

## 4 CURRENT STATE AT VCT

Based on the interviews, direct observations, and reviews of archival records in regard to E1, E2 and E3, the current state of data input data management for DES and general data collection procedures at VCT are presented in this section. In particular, the 20 largest problems in the current state are highlighted.

Table 2: The 10 largest problems with current procedures for input data management for DES.

- 
- Lack of formal methodologies for simulation in general, and data collection in particular.
  - Lack of time to perform data collection.
  - Poor data availability (particularly disturbance data, where the complete raw data log has full accessibility for only the past 30 days).
  - Poor quality of *category A* data (particularly disturbance data, where minor stoppages < 1 minute and specific stop type categories are automatically removed in data logs with full accessibility).
  - Large time requirement for manual collection and transformation of breakdown input data
  - Transformation of input data requires large involvement of equipment experts, who are difficult to find and time-consuming to contact.
  - Correction and calculation phases are aggravated by poor data quality and inconsistent disturbance classification, where it is particularly difficult to determine what stoppages that affects technical availability, and distinguish between stopping vs. non-stopping disturbances.
  - Input data transformation is questioned by project leadership due to assumptions, estimations, and judgments in correction and calculation phases, a process that also lacks documentation.
  - Input breakdown data (MDT/MTTR with fitted statistical distributions) are rejected by project leadership due to diverging views of expected disturbance patterns in equipment.
  - The rejection of simulation results causes uncertainties in projects, with possibly large financial consequences.
- 

It is obvious that the current approach is not sufficient to reach the high set goals for fast response time of DES experiments (table 2). Particularly problematic is the poor data reputation that has caused

uncertainties in projects. A benefit of concurrent engineering within VCT is that the simulation analyst is integrated in the project team, in which DES experiments are conducted continuously and evaluated collaboratively. However, the simulation analysts summarizes the situation by saying ‘If you pick raw data, you are questioned. If you present filtered data, you are questioned. Yes, it is a lot of judgment, it is unreliable data, and in this particular case (E3), I am not really sure that the data that I have used are correct. I mean, I don’t trust it either.’. Figure 1 shows the current approach to input data management for DES at VCT.

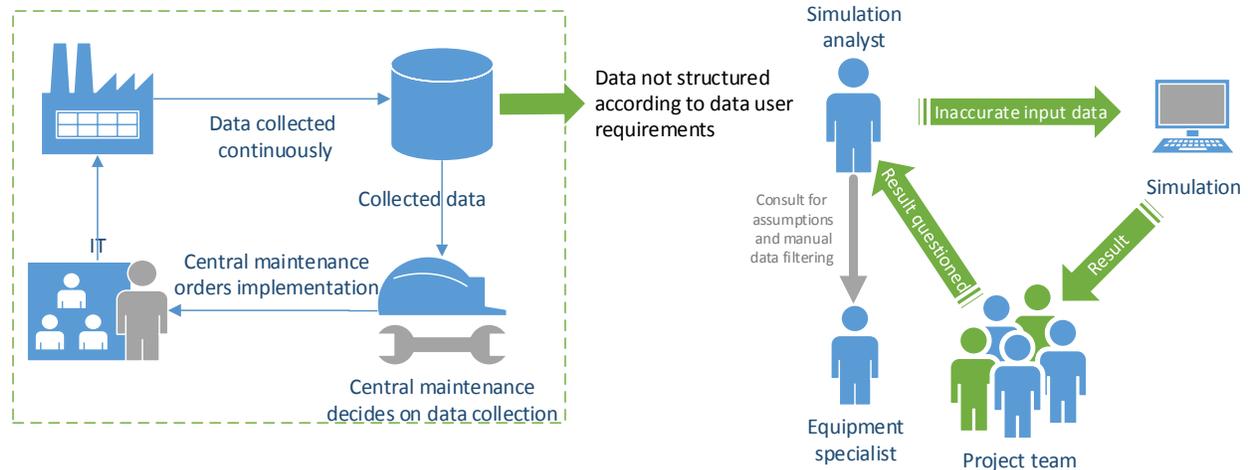


Figure 1: Current state of input data management for DES at VCT.

Evidently, a major challenge for the simulation analyst is to collect and transform disturbance data to input breakdown data with fitted statistical distributions, and these data have particularly poor credibility. Therefore, the subsequent interviews largely revolved around disturbance data. Based on these interviews and complementary observations and reviews of archival records, the 10 largest problems with general data collection procedures are presented in table 3.

Table 3: The 10 largest problems with current procedures for general collection of production data.

- Lack of communication between central maintenance department and data users, resulting in low involvement from data users and few collected and implemented data requirements.
- Lack of roles, responsibilities, and formal work procedures to collect and implement data requirements, aggravated by lack of time and resources for developing data collection systems.
- Large variety in data requirements between organizational functions, resulting in a big challenge to fulfill all the different needs, especially due to the lack of formal data requirement specifications.
- Large problems in development and implementation of disturbance classifications and cause codes for operator commenting, aggravated by conflicting views on the level of detail of the comments.
- Operators and maintenance technicians, responsible for commenting disturbances, are the key roles in identifying failure root causes, but they lack education and influence in how data are collected.
- Lack of documentation of changes in equipment, processes or data structure that affects the representativeness of production data (i.e. data currency)
- Very few people are using production data, resulting in low value in collecting high quality data, aggravated by that some data files are not designed to be easily interpretable.
- Minor stoppages are automatically excluded in the reports used by production and maintenance personnel, resulting in that up to 60% of all disturbances remain unnoticed and unresolved.

- 
- Equipment experts perceive the collected disturbance data as rather unusable, wish to have better support and higher data quality, want to be involved in how data are collected, and have the knowledge to decide how data should be collected for their specific equipment.
  - Large opportunities in collecting valuable data are wasted due to lack of roles and responsibilities for implementing data requirements, resulting in that e.g. equipment with advanced condition monitoring sensors have never been connected to the overall systems.
- 

The problems in table 3 clearly illustrate the need for a more structured approach to data collection. In particular, all interviewees mention the lack of communication and important roles and responsibilities. One of the maintenance engineers talks about implementing the new disturbance follow-up system, and explain: ‘The central maintenance department is responsible for collecting requests from different sources within the organization. We had some discussions about data requirements, but not really that concrete. We took some things into account, but since the new system is so open and allowing for changes, it was not worth going to the bottom with this.’. The other maintenance engineer also talks about data requirements, and says: ‘Actually, our only responsibility is to ensure that the system can receive data, not that anyone sends it to us. But there have been no requirements sent to us, not from production or anywhere else.’. Interestingly, the equipment expert mention: ‘We don’t have any possibility to set up the systems so that we can get the data that we want and need. Maybe the central maintenance department can, but they have not asked us’. Remarkably, when all three are asked what people should be involved in developing data collection procedures, they all refer to each other and explain that a much better communication and collaboration would be very useful.

In this current state map, a number of vital roles and responsibilities can be identified: (1) the simulation analyst, (2) the equipment expert, (3) the central maintenance department, (4) the PLC expert, (5) the project leadership, and (6) the shop floor personnel.

## **5 PROPOSED METHODOLOGY**

Through the use of interviews, direct observations, and reviews of archival records, the 20 largest problems with the current procedures for input data management for DES and general data collection are identified from three cases: E1, E2 and E3. Based on the current state mapping, the proposed methodology for continuous quality assurance of production data is presented in figure 2.

It is noteworthy how many of the problems in regard to input data management for DES (table 2) directly relate to the data quality indicators (Balci et al. 2000), e.g. poor clarity, completeness, reputation etc. Moreover, poor data availability, lack of documented assumptions, issues with distribution fitting, and lack of procedures for validating input data, directly relates to the pitfalls and success factors in table 1. In particular, validation of the laser breakdown data (E2, E3) was extremely difficult due to poor data quality, despite applying face validation by collaborating with an equipment expert (Skoogh and Johansson 2008). Moreover, as many of the identified problems can be attributed to poor communication and a lack of necessary expertise, it is not strange that the establishment of roles and responsibilities in simulation projects are emphasized (Randell 2002, Robinson 2004). Also, the current data structures are not designed for simulation (Skoogh and Johansson, 2007), and the lack of roles and responsibilities for updating production data and documenting changes stresses the need to prepare a suitable process for this (Rabe, Spieckermann, and Wenzel 2009).

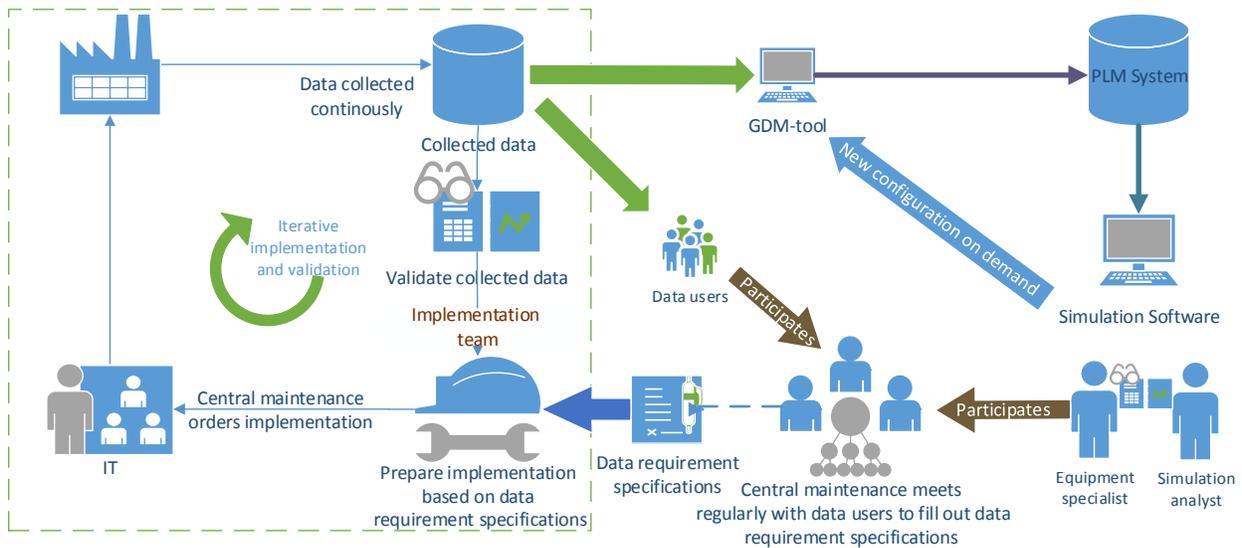


Figure 2: Proposed methodology for continuous quality assurance of production data.

The proposed methodology describes work procedures for continuous quality assurance of production data by collecting and implementing data requirements from various data users, as well as a structured approach to input data management for DES. The methodology enables a much better communication between the central maintenance department and data users. The central maintenance department should be responsible for calling regular data requirement meetings, in which data requirement specifications are filled out by interested data users, e.g. simulation, production and maintenance personnel etc. These documents specifies the users' data requirements based on the data quality indicators proposed by Balci et al. (2000). Thereafter, the data collection implementation team is responsible for reviewing and implementing the data requirements specifications. The roles within this team should consist of representatives from the (1) central maintenance department, (2) equipment expertise (e.g. laser, robot), (3) PLC programming, (4) manufacturing engineering (e.g. robot and layout design), and (5) shop floor operators. Within this team, the operators should play a leading role since they are the only ones who can truly ensure that representative data are collected (e.g. commenting disturbances). In fact, it would be ideal if data requirements are not implemented unless it brings value to the operators, and the primary focus should be on implementing their requirements. The implementation team should meet on a regular basis, where the focus is on translating data requirement specifications into data implementation plans, specified in data requirement implementation checklists. Thereafter, the central maintenance department is responsible for ordering the changes from the IT department. Examples of potential changes could be redesigning the raw data disturbance log to be applicable for simulation, or integrating available condition monitoring data. The implementation team holds the responsibility for data validation, consisting of initial face validation of implemented data requirements, and continuous face validation of the collected data in the case of changes in production equipment, processes, data structure etc.

The second part (right part of figure 2) of the proposed methodology describes an input data management procedure for DES, consisting of extracting data, performing face validation with equipment expert in relation to the implemented data requirement specification, and creating a configuration file for data transformation in the GDM-Tool. Formal documentation of all assumptions should be made during the GDM-Tool configuration, which serves to reduce the potential for rejection of input data by project leadership (table 2). The GDM-Tool generates the desired input data, e.g. MDT/MTTR with fitted statistical distribution, which is stored in a PLM system and pulled to be used in simulation on short notice. This configuration is only required once for a specific data type, and the GDM-Tool can be set to continuously perform updates to always ensure that the information is fresh and up to date.

The lean supermarket thinking (Yan and Cai 2009) have been adopted in the methodology in the following ways: (1) Keeping the data fresh and up to date using GDM-Tool updates. (2) Ensuring that the information is sufficient to serve all the needs by implementing specific data requirements. (3) Know where each information is stored by keeping transformed data from GDM-Tool in the PLM system. (4) Information ready to be pulled at the right time, the right kind, and right amount by collecting data in accordance to the users' data requirements. This thinking is also adopted in the implementation team constellation since it holds all the required expertise, and enables a pull flow of information and knowledge from within the organization.

## 6 DISCUSSION

This paper presents a methodology for continuous quality assurance of production data based on a case study at VCT, in which 20 major problems with the current state of input data management for DES and general data collection procedures are identified. The proposed methodology suggests the use of data requirement specifications based on data quality indicators (Balci et al. 2000). In the context of breakdown input data for simulation, this could e.g. include the following demands: Data need to have an *accessibility* so that a 12-month history is directly retrievable. Data need to have a *clarity* that enables a distinction between stopping and non-stopping disturbances. Data need to have a *completeness* so that all stoppages are logged, including minor stoppages. Data should have a *resolution* so downtime is traceable to a component level, e.g. a specific robot gripper. Note that the intention is to only specify the relevant indicators, and unimportant indicators are decided appropriately by the implementation team. The indicators were chosen to support the fundamental idea of specifying data quality requirements for specific users, and thus enable data to be collected in accordance to specific objectives (Balci, 1990; Skoogh and Johansson 2008; Williams and Ülgen 2012). Therefore, it is suggested to utilize previous research for the basic data structure (Randell 2002), and then adapt it to the submitted data requirements specifications.

Articles highlighting pitfalls and success factors provide general, but valuable, practical tips for how to avoid common simulation problems (table 1). Although the proposed methodology addresses data quality issues on a more detailed level, it will still help to avoid pitfalls and grasp the opportunities in success factors. For example, it improves overall data availability, and ensures appropriate representations of variation thanks to high quality data and use of the GDM-Tool to fit distributions. It also mitigates general data collection problems by establishing a structured approach to data collection throughout the organization, as well as enable continuous control of data-integrity and tests of time-sensitive data through regular GDM-Tool updates.

A critical aspect that has been identified in this study is the ability to track changes that influences data currency (table 2), e.g. replacements of equipment. This is both relevant for simulation and operations since data analytics can support follow-up of improvements. It is therefore vital that procedures, roles and responsibilities for documenting such changes are in place. However, this does not hold a specific solution at the moment, but will be tackled in future research.

Another issue to address is the problem with low acceptability of simulation results due to poor data credibility (table 2). Although the proposed methodology does not specifically address the requirements for project leadership to trust the input data, an awareness of a formal approach to ensure high data quality, in combination with documentation of assumptions, will most likely contribute towards improved credibility and reduced risk of project uncertainties. Of course, development projects can directly influence the data collection by submitting data requirement specifications.

Numerous previous studies have proposed detailed methodologies for data collection and rapid input data management for simulation projects (Lehtonen and Seppälä 1997; Perera and Lyanage 2000; Skoogh and Johansson 2008). Hence, this study does not intend to specifically support standalone simulation projects, and instead focus on continuous data collection. Also, the 20 identified problems primarily revolve around *category A* data, and this study therefore focuses on improving the management of such

data. However, the intention is to include *category B* and *C* data in the methodology through further studies. Of course, this work is still primarily on a conceptual level, and much effort is required to operationalize the methodology before implementation.

Moreover, without a better use of information and knowledge, most manufacturing companies will struggle to achieve sufficiently high production efficiency. Within VCT, the current use of production data to support production improvements will likely be insufficient to reach 85% OEE. Looking ahead, the emerging trends within the manufacturing industry, e.g. industry 4.0, place even higher demands on collecting, storing, and analyzing production data. Nonetheless, storing transformed and updated input data from GDM-Tool in the PLM system enables a rapid use of data in simulation, which contributes towards the possibility to present detailed simulation results within 24 hours. Note however that if a new GDM-Tool configuration is required for the needed data, 24 hours will be insufficient.

## **7 CONCLUSIONS**

The aim of this study is to propose a structured methodology for continuous quality assurance of production data, with the intention of increasing the availability of high quality data for various data users. Based on a case study at VCT, poor communication and a clear lack of roles, responsibilities, and formal work procedures for data collection were identified, in combination with inefficient manual DES input data management procedures. By identifying 20 major problems in the current state, it is evident that a more structured approach could generate drastic improvements.

The proposed methodology describes a team-based approach to implementation of data collection procedures based on submitted data requirement specifications from various data users, and it specifies the necessary roles, responsibilities, meetings, and formal documents. This intends to enable data to be collected according to specific, but diverging, objectives within an organization. In addition, an approach for input data management for DES based on a collaboration between simulation analysts and equipment experts is also presented, including extraction, validation, transformation, and storage of DES input data using the GDM-Tool. This work primarily supports the use of production data on a continuous basis, especially valuable to manufacturing companies who have reached far in the implementation of advanced computer applications for collection and storage of production data, and use DES as a day-to-day engineering tool. In addition, the proposed methodology is suitable for enabling purely data-driven production analyses to be used as a daily tool for production improvements.

## **8 FURTHER RESEARCH**

A limitation of single-case studies is the lack of generalizability of the results (Yin 2003). Therefore, the results from this study will serve as the basis for an extended multiple-case study within the StreaMod project. This will be aimed at evaluating the proposed methodology, ensuring a generic applicability within the manufacturing industry, and quantifying the effects of adopting it in its entirety.

## **ACKNOWLEDGMENTS**

The research project “Streamlined Modeling and Decision Support for Fact-based Production Development (StreaMod)” is funded by VINNOVA, Swedish Agency for Innovation Systems. The authors would like to thank all of the participating interviewees at VCT for their valuable contribution. This work has been performed within Sustainable Production Initiative and the Production Area of Advance at Chalmers. The support is greatly appreciated.

## **REFERENCES**

- Balci, O. 1990. “Guidelines for successful simulation studies”. *Proceedings of the 1990 Winter Simulation Conference*, ed. O. Balci, R.P. Sadowski, and R.E. Nance.

- Balci, O., W.F. Ormsby, J.T. Carr III, and S.D. Saadi. 2000. "Planning for verification, validation, and accreditation of modeling and simulation applications". *Proceedings of the 2000 Winter Simulation Conference*, ed. J.A. Joines, R.R. Barton, K. Kang, and P.A. Fishwick.
- Banks J, J. Carson II., B. Nelson, and D. Nicol 2005. *Discrete-event system simulation*. 4<sup>th</sup> ed. Upper Saddle River: Prentice- Hall.
- Bokrantz, J., A. Skoogh, T. Ylipää, and J. Stahre. 2015. "Handling of Production Disturbances – classification and registration in the manufacturing industry". *Submitted to the International Journal of Production and Operations Management*, April 2015.
- Carter, A. and M. Ludvigsson. 2014. "Data Acquisition Methods for Discrete Event Simulation – A Case Study at Volvo Cars Torslanda". Master Thesis, Chalmers University of Technology, Sweden.
- Dettman, T., C. Andersson, J. Andersson, A. Skoogh, B. Johansson, and P Forsbom. 2013. "Startup methodology for production flow simulation projects assessing environmental sustainability". *Proceedings of the 2013 Winter Simulation Conference*, ed. R. Pasupathy, S.H. Kim, A. Tolk, R. Hill, and M.E. Kuhl.
- Eriksson, U. 2005. "Diffusion of Discrete Event Simulation in Swedish Industry – On the Way to an Increased Understanding". PhD thesis, Chalmers University of Technology, Sweden.
- Jägstam, M., and P. Klingstam. 2002. "A handbook for integrating discrete event simulation as an aid in conceptual design of manufacturing systems". *Proceedings of the 2002 Winter Simulation Conference*, ed. E. Yücesan, C.-H. Chen, J. L. Snowdon, and J. M. Charnes.
- Law, A.M., and M.G. McComas. 1986. "Pitfalls in the simulation of manufacturing systems". *Proceedings of the 1986 Winter Simulation Conference*, ed. J. Wilson, J. Henriksen, S. Roberts.
- Law, A.M., and M.G. McComas. 1991. "Secrets of successful simulation studies". *Proceedings of the 1991 Winter Simulation Conference*, ed. B.L. Nelson, W.D. Kelton, G.M. Clark.
- Law, A.M. 2007. *Simulation modeling & analysis*. 4<sup>th</sup> ed. McGraw-Hill, New York.
- Lehtonen, J.-M., and U. Seppälä. 1997. "A methodology for data gathering and analysis in a logistics simulation project". *Integrated Manufacturing Systems* 8:351-358.
- Ljungberg, Ö. 1998. "Measurements of overall equipment effectiveness as a basis for TPM activities". *International Journal of Production and Operations Management*, 18(5):495-507.
- Musselman, K.J. 1994. "Guidelines for simulation project success". *Proceedings of the 1994 Winter Simulation Conference*, ed. J.D. Ted, S. Manivannan, D.A. Sadowski, and A.F. Seila.
- Perera, T., and K. Liyanage. 2000. "Methodology for rapid identification of input data in the simulation of manufacturing systems". *Simulation Practice and Theory* 7:645-656.
- Rabe M., S. Spieckermann, and S. Wenzel. 2009. "Verification and validation activities within a new procedure model for V&V in production and logistics simulation". *Proceedings of the 2009 Winter Simulation Conference*, ed. M.D. Rossetti, R.R. Hill, B. Johansson, A. Dunkin, and R.G. Ingalls.
- Randell, L. 2002. "On Discrete-Event Simulation and Integration in the Manufacturing System Development Process". PhD thesis, Lund University, Sweden.
- Robertson, N., and T. Perera. 2002. "Automated data collection for simulation?". *Simulation Practice and Theory*, 9:349-364.
- Robinson, S. 2004. *The Practice of Model Development and Use*. Wiley, Chichester.
- Sadowski, R.P. 1989. "The simulation process: avoiding the problems and pitfalls". *Proceedings of the 1989 Winter Simulation Conference*, ed. E.A. MacNair, K.J. Musselman, P. Heidelberger.
- Sargent, R. G. 2005. "Verification and validation of simulation models". *Proceedings of the 2005 Winter Simulation Conference*, ed. M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, 130-143.
- Skoogh, A., and B. Johansson. 2007. "Time-consumption analysis of input data activities in discrete event simulation projects". *Proceedings of the 2007 Swedish Production Symposium*. Gothenburg, Sweden.
- Skoogh A., and B. Johansson. 2008. "A methodology for input data management in discrete event simulation projects". *Proceedings of the 2008 Winter Simulation Conference*, ed. S.J. Mason, R.R. Hill, L. Mönch, O. Rose, T. Jefferson, J.W. Fowler.

- Skoogh, A., B. Johansson, and J. Stahre. 2012. "Automated input data management: evaluation of a concept for reduced time consumption in discrete event simulation". *Simulation: Transactions of the Society for Modeling and Simulation International*, 88(11):1279-1293.
- Tretten, P., and R. Karim. 2014. "Enhancing the usability of maintenance data management systems". *International Journal of Quality in Maintenance Engineering*, 20(3):290-303.
- Ülgen, O.M, J.J. Black, B. Johnsonbaugh, and R. Klunge. 1994. "Simulation methodology – a practitioner's perspective". *International Journal of Industrial Engineering, Applications and Practice I*.
- Williams, E.J., and O.M. Ülgen. 2012. "Pitfalls in managing a simulation project". *Proceedings of the 2012 Winter Simulation Conference*, ed. C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A.M. Uhrmacher.
- Yang, K. and X. Cai. 2009. "The integration of DFSS, lean product development and lean knowledge management". *International Journal of Six Sigma and Competitive Advantage*, 5(1):75-99.
- Yin, R.K. 2003. *Case Study Research: Design and Methods*. 3<sup>rd</sup> ed. Sage Publications Inc, Thousand Oaks, CA.

## **AUTHOR BIOGRAPHIES**

**JON BOKRANTZ** is a PhD student within the area of Production Service Systems & Maintenance at the Department of Product and Production Development, Chalmers University of Technology. Jon has a background in Production Engineering, and his research focuses on improvements in maintenance organizations. His email address is [jon.bokrantz@chalmers.se](mailto:jon.bokrantz@chalmers.se).

**ANDERS SKOOGH**. PhD, is an Assistant Professor at the Department of Product and Production Development at Chalmers University of Technology. He is a research group leader for Production Service Systems & Maintenance. Anders is also the director of Chalmers' Master's program in Production Engineering and a board member of Sustainability Circle ([www.sustainabilitycircle.se](http://www.sustainabilitycircle.se)) with responsibilities for research collaboration. Before starting his research career, he accumulated industrial experience from being a logistics developer at Volvo Cars. His e-mail address is [anders.skoogh@chalmers.se](mailto:anders.skoogh@chalmers.se).

**JON ANDERSSON** is a researcher in the field of virtual production system at the Department of Product and Production Development, Chalmers University of Technology. He is currently active in the field of input data management for simulation of production flows. Besides that, Jon is involved in projects targeting sustainability for manufacturing and visualization of production flows. His e-mail address is [jon.andersson@chalmers.se](mailto:jon.andersson@chalmers.se).

**JACOB RUDA** is active at the Virtual Manufacturing Engineering department at Volvo Cars. He has global responsibility for Discrete Event Simulation studies in Body Shop, Paint Shop and Final Assembly. His e-mail address is [jacob.ruda@volvocars.com](mailto:jacob.ruda@volvocars.com).

**DAN LÄMKULL**. PhD, is a Research Leader and Method Developer within the Department of Virtual Manufacturing Methods and IT at Volvo Cars. He has been involved in Virtual Manufacturing research for more than twentytwo years. His current research includes Digital Human Modeling, Assembly and Disassembly Simulation, Laser Scanning, Virtual Operator Training, Lean Plant Design/Layout and Discrete Event Simulation. He is currently Project Manager for the VINNOVA-funded StreaMod project. His e-mail address is [dan.lamkull@volvocars.com](mailto:dan.lamkull@volvocars.com).