

LSS4.0 - Creating Opportunities from existing Limitations

Tanawadee Pongboonchai-Empl

Doctoral candidate, Edinburgh Business School, Heriot-Watt University, Edinburgh, UK
(tp61@hw.ac.uk)

Dr. Dietmar Stemann

MTS Consulting Partner, Werl, Germany
(stemann@mts-consultingpartner.com)

Prof. Jiju Antony

Operations and Logistics Group, Edinburgh Business School, Heriot-Watt University, Edinburgh, UK
(J.Antony@hw.ac.uk)

Abstract

Purpose: This research explores Lean Six Sigma (LSS) limitations identified in previous studies and underpins the need to integrate Lean Six Sigma with Industry 4.0. Furthermore, this paper presents a use case for an opportunity offered by data-driven tools.

Method/Approach: Expert Interviews were conducted with senior Lean Six Sigma practitioners and academics in Europe. The data were analysed using qualitative content analysis and NVivo software.

Findings: This paper reveals limitations experienced by practitioners and their causal relationships. Additionally, new limitations have emerged, reinforcing the urgency to inject data-driven technologies into the traditional LSS Framework. Associated with this is the need for an LSS curriculum update.

Research limitations/implications: A new limitation found was the waterfall nature of the DMAIC framework. Therefore, an integration of Lean Six Sigma and Agile could be an emerging trend. Further research is planned by the primary author, i.e., designing an LSS4.0 framework including data-driven tools.

Originality/value: To the best of the authors' knowledge, this is the first qualitative study to examine causal relationships among LSS limitations. The authors present a use case as an initial remedial strategy to alleviate Lean Six Sigma limitations.

1. Introduction

Does Lean Six Sigma need a refresh? Lean Six Sigma (LSS) has been introduced approximately two decades ago and has been a successful improvement paradigm ever since (Antony *et al.*, 2017). Although LSS was applied by organisations in various sectors, e.g., manufacturing, services, financial, public and healthcare (Antony *et al.*, 2017; Singh and Rathi, 2019), the methodology is still based on traditional Lean and Six Sigma tools. Therefore, in the era of Industry 4.0 (I4.0) and digital technologies, it is valid to claim that LSS needs a makeover (Fogarty, 2015). In an increasingly digitised world with continuously growing data volumes, traditional quality management and process improvement methods reach their limits. As a data-driven improvement paradigm (Snee, 2010), LSS should embrace the opportunities provided by the data-driven tools of the 21st century. The integration of LSS with I4.0 and LSS with big data are emerging trends reported by Antony *et al.* (2019). Hence, the purpose of this study is to empirically underpin the demand for enhancing LSS by data-driven tools. As a starting point, the authors explore which LSS limitations justify a need for a step towards digitisation. This paper seeks to answer the following questions:

- 1) Which LSS limitations have LSS experts experienced in their own practice?
- 2) How are the LSS limitations related to each other?
- 3) Which LSS limitations underpin the need for integrating data-driven tools into the DMAIC roadmap?

Furthermore, this study aims to follow the emerging trend of integrating LSS with I4.0 while addressing LSS limitations. To that end, the authors present an easy-to-implement scenario in which data mining techniques are integrated into a DMAIC project. The use case is based on DMAIC because it is the most widely used framework in Six Sigma projects (Albliwi *et al.*, 2015; Martins Rosa and Broday, 2018; Panayiotou and Stergiou, 2020). As the Six Sigma DMAIC roadmap represents the core of LSS, the literature of both very similar disciplines, Six Sigma and LSS, are considered for this study and the terms are used interchangeably, depending on the respective context.

The following section provides background information about LSS, data mining as part I4.0 and the motivation for an integration. Section 3 describes the research method by explaining how the empirical part of this research was conducted. The findings are presented in section 4 as a response to the research questions. In section 5, a use case is presented as initial remediation, and we close the article with conclusions for research and practice as well as the limitations and further research directions.

2. Background Information

2.1 Lean Six Sigma

LSS is a hybrid improvement framework that combines Lean tools with those of Six Sigma. The Six Sigma DMAIC roadmap provides the missing structure to the Lean improvement philosophy and helps to put it into practice (George and George, 2002). Thus, LSS combines the virtues of Lean, i.e., speed and accuracy through waste elimination, and Six Sigma, i.e., precision and accuracy through minimised variation. A plethora of papers and books with LSS success stories have been published in academia since its introduction by Michael George (2002). As an enabler for improved performance and capability, increased profitability, and revenue growth, LSS has become the most widely used improvement strategy (Antony *et al.*, 2017; Sony *et al.*, 2020a). Experts referred to as ‘Belts’ apply statistical methods, graphical tools and creativity techniques to analyse a problem and elicit practical solutions to achieve improved bottom-line results. The belt system and the DMAIC framework were adopted from the Six Sigma methodology (Schroeder *et al.*, 2008; Pepper and Spedding, 2010). DMAIC consist of 5 phases that lead to process improvement: Define, Measure, Analyse, Improve and Control. Suitable tools are available for each phase to enable the project team around the Belt to successfully resolve the problem, improve the process and guide the operations team to sustain the improvement (George and George, 2002, pp. 35-36).

Although LSS enjoys great popularity, over 60% of LSS initiatives were reported to have failed (Glasgow *et al.*, 2010), and therefore, limitations, criticisms and failure factors have been discussed by numerous scholars. Limitations such as unavailability of data (Albliwi *et al.*, 2014), long implementation cycles (Sony *et al.*, 2019), wrong selection and prioritisation of projects (Snee, 2010) and unsustainable results Aboelmaged (2011) are valid motivational factors for an integration of LSS with data-driven tools and architectures. Additionally, the integration of LSS with big data is globally one of the top emerging trends investigated by Antony *et al.* (2020)

2.2 Integration with Industry 4.0 and big data: data mining as an opportunity

To date, there is no universally agreed definition of I4.0, but it is commonly associated with cyber physical systems, internet of things, big data analytics, cloud computing, additive manufacturing, augmented reality, autonomous robots, cyber security and horizontal & vertical integration (Rojko, 2017). The efficient analysis of huge volumes of inhomogeneous data, so-called big data, and the generation of knowledge from data has become a critical capability for companies and is therefore becoming increasingly important (Deuse *et al.*, 2014b). Companies started to focus on predictive and prescriptive analyses to determine the probability of future events and receive additional recommendations for action. Predictive analysis methods include data mining, statistics and artificial intelligence, which can be summarised as big data analytics (Russom, 2011). Data mining is also a sub-step of the KDD process (knowledge discovery in databases) developed by Fayyad, where specific algorithms are applied to extract patterns from the data (Fayyad *et al.*, 1996b). The origin of data mining resides in statistics and initially characterised the selective application of methods to confirm pre-formulated hypotheses (Fayyad *et al.*, 1996a; Berry and Linoff, 1997).

Meanwhile, further models have been developed to structure and guide data mining projects. One dominant model in the industry is CRISP-DM, Cross-Industry Standard Process for data mining (Chapman *et al.*, 2000). Unlike KDD, the CRISP-DM model was developed based on experience from practical data mining projects (Deuse *et al.*, 2014a). It consists of six iterative phases (see Figure 1). The process output is a prediction model based on algorithms, i.e., rules and functions that describe relationships between input and output. Depending on the intended use and the underlying database structure, the data mining expert can choose between numerous techniques, each again with a choice of algorithms. Various software-based tools, e.g., R, Python, RapidMiner and Knime, support the implementation. The second use case in this study uses the Knime standard software to create decision trees.

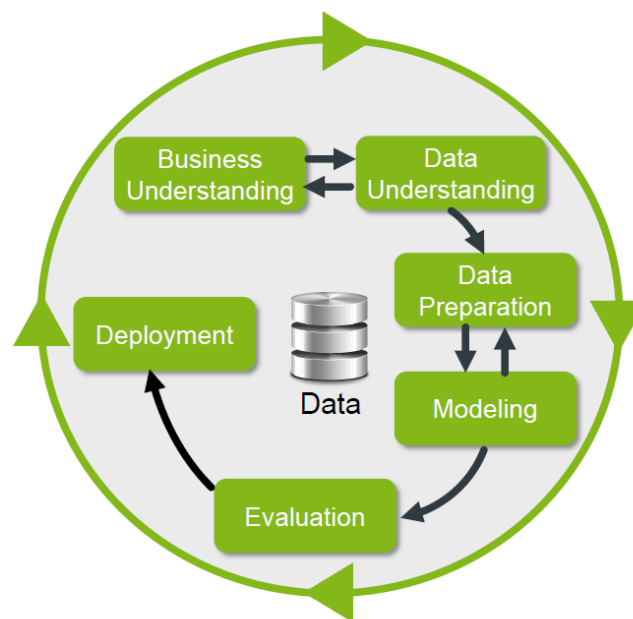


Figure 1: CRISP-DM Modell (Chapman *et al.*, 2000)

3. Methodology

To answer research questions 1 and 2, the first author conducted 13 Expert Interviews. These are more open than semi-structured interviews but provide more guidance and focus on the subject matter than unstructured or narrative interviews (Bogner and Menz, 2009, p. 46). An interview guide with four central topics was developed. The topics were: (1) Confirming the participant's expertise and getting to know the expert's context, (2) LSS Limitations, (3) I4.0 Opportunities, (4) LSS4.0 requirements. Candidates were selected purposively based on their diverse attributes to get a comprehensive view from a relatively small number of experts (Denscombe, 2014, p. 41).

Table 1 shows the candidates' attributes:

Table 1: Interview participants

Interviewee	Roles	Belt Level	LSS experience in years	Sector(s)
CO-I-1	Quality Consultant and Coach	Master Black Belt	20	Manufacturing/Services
CO-I-3	Operational Excellence lead and Coach	Master Black Belt	16	Manufacturing
OP-I-11	Operational Excellence lead and Coach	Master Black Belt	14	Services
AC-I-6	Academic and Quality Consultant	Master Black Belt	13	Finance/Manufacturing
OP-I-8	Operational Excellence lead and Coach	Master Black Belt	13	Manufacturing
OP-I-5	Operational Excellence lead and Coach	Master Black Belt	12	Manufacturing
OP-I-7	Operational Excellence lead and Coach	Master Black Belt	12	Manufacturing
OP-I-10	Operational Excellence lead and Coach	Master Black Belt	11	Manufacturing
AC-I-4	Academic and Quality Consultant	Master Black Belt	10	Finance/Public
OP-I-12	Head of Engineering	Black Belt	10	Manufacturing
AC-I-2	Academic and Quality Consultant	Black Belt	6	Manufacturing
OP-I-14	Operational Excellence lead and Coach	Black Belt	4	Manufacturing
AC-I-13	Academic and Quality Consultant	Green Belt	2	Manufacturing

All candidates were Six Sigma or LSS belts; 10 out of 13 had more than 10 years of experience, while 9 out of 13 participants had experience applying data-driven tools in conjunction with quality management. The remainder (4/13) were at least aware of the potential offered by I4.0 technologies. The interviews were conducted one-on-one in virtual meetings and took between 30 and 60 minutes. All interviews, except one, were recorded and transcribed. Participants were assigned pseudonyms, and transcripts were anonymised. Qualitative Content Analysis applying a deductive category system was selected as the most appropriate analysis method for Expert Interviews (Gläser and Laudel, 2010, pp. 46-47). A category system was deduced from theory so information could be structured (pp. 197-198). This research's category system was built based on previous research about LSS limitations, Key criticisms and Critical Failure Factors reported by Antony *et al.* (2020), Sony *et al.* (2020b) and S. Albliwi *et al.* (2014). After consolidating similar elements, 48 limitations were included in the first dimension of the category system. Similarly, 20 categories for LSS Opportunities were defined and included in the second dimension.

As the interviews with German native speakers were conducted in German and with non-German speakers in English, translation work was necessary. The interview transcripts were initially analysed in the original interview language to preserve the narratives' meaning. Therefore, the first author had created a category system in English and German. For more rigour, the categories were back-translated into English and again into German by a professional proof-reader, an English native speaker with proficient German language skills. NVivo coding functionality was used to assign text passages to categories for extraction. New categories were added to the category system for text passages that did not fit into any existing category (Gläser and Laudel, 2010, p. 217). The first author provided participants with interview transcripts for review, including corresponding categories and requested confirmation that the assigned categories reflected the intended meaning to avoid researcher bias. This procedure, along with the purposive sampling and the review of analysis results by the second author, enhanced the research quality by ensuring trustworthiness (Krefting, 1991). The relevant information was extracted into an NVivo framework matrix and then paraphrased. During this process, extracts in German were paraphrased directly in English to unify the

language. The timing of translation was relatively late, but it was a legitimate choice (Santos Jr *et al.*, 2015).

4. Findings

Based on the 13 interview transcripts, 181 text references were extracted and assigned to 69 categories. 40 relationships were found. Most of the 46 LSS limitations synthesised from the literature could be confirmed, except for 9, e.g. “Overemphasis on variance reduction in processes”. Some top limitations, key criticisms and critical failure factors from the literature could be explored in terms of causal relationships. The interviews revealed causal chains of the limitations/failure factors/criticisms. The causal relationships were discussed and confirmed between the first and the second author. The top limitation, “High failure rate” in LSS projects (Antony *et al.*, 2020), is a concern raised by experts queried. Further, “Lack of data and poor data” (Albliwi *et al.*, 2014) was confirmed to be a major cause for project failures. The causal relationships are revealed in the following graph:

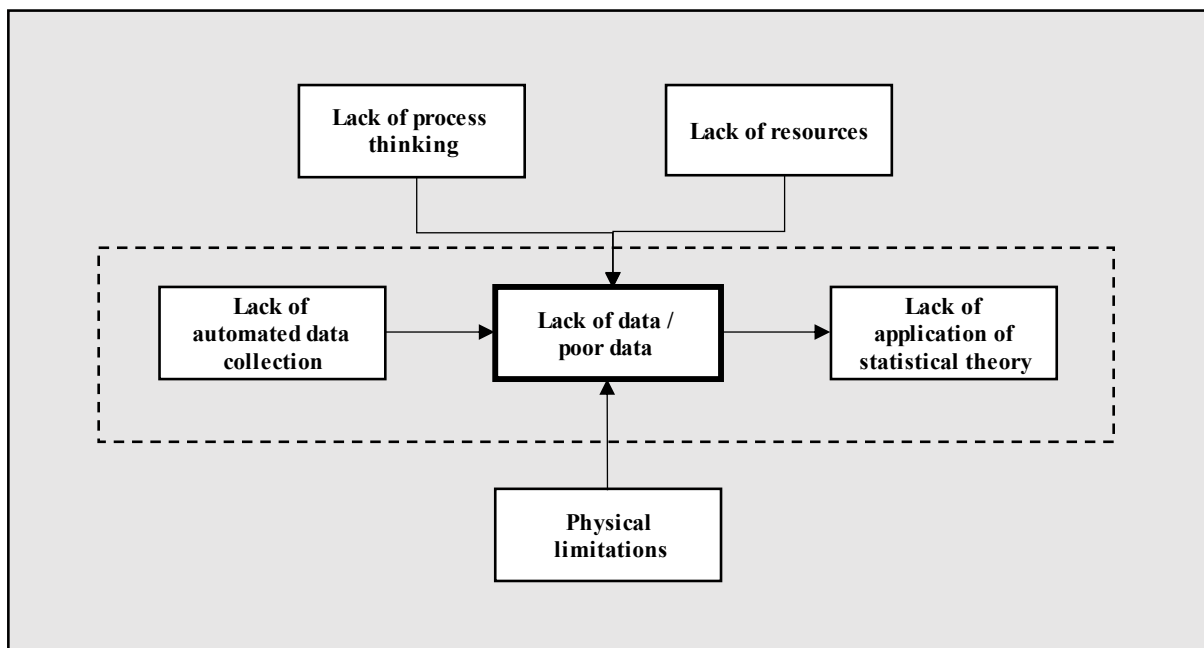


Figure 2: Causal chain for Lack of data / poor data

Companies that still collect their data manually typically suffer from a lack of data or poor data quality. The main reasons for this are that humans cannot collect as many datasets as sensors and IT systems, and manual processes are error-prone by nature. Other reasons discovered during this study are lack of process thinking and lack of automated data collection.

OP-I-10: “...in the company that I now work we really have the problem that we don’t have many data sources. ... Then we manually make an Excel or Minitab sheet where we combine those parameters and the project Y that we want to measure. And if we’re lucky, we find the p-

value that proves that this is indeed the case, and then we can start problem-solving. But it's still very opinion-based."

AC-I-6: *"In practice you are sometimes content to get 35 reliable data points to perform an analysis with."*

Without sufficient reliable data, it is not possible to conduct statistical analyses and draw informed conclusions.

From the experiences reported by the interviewees, numerous LSS limitations call for an integration of LSS with I4.0/big data technologies. This emerging trend was also reported by Antony et al. (2020). Figure 3 depicts which limitations justify the need for an integration of digital tools into DMAIC:

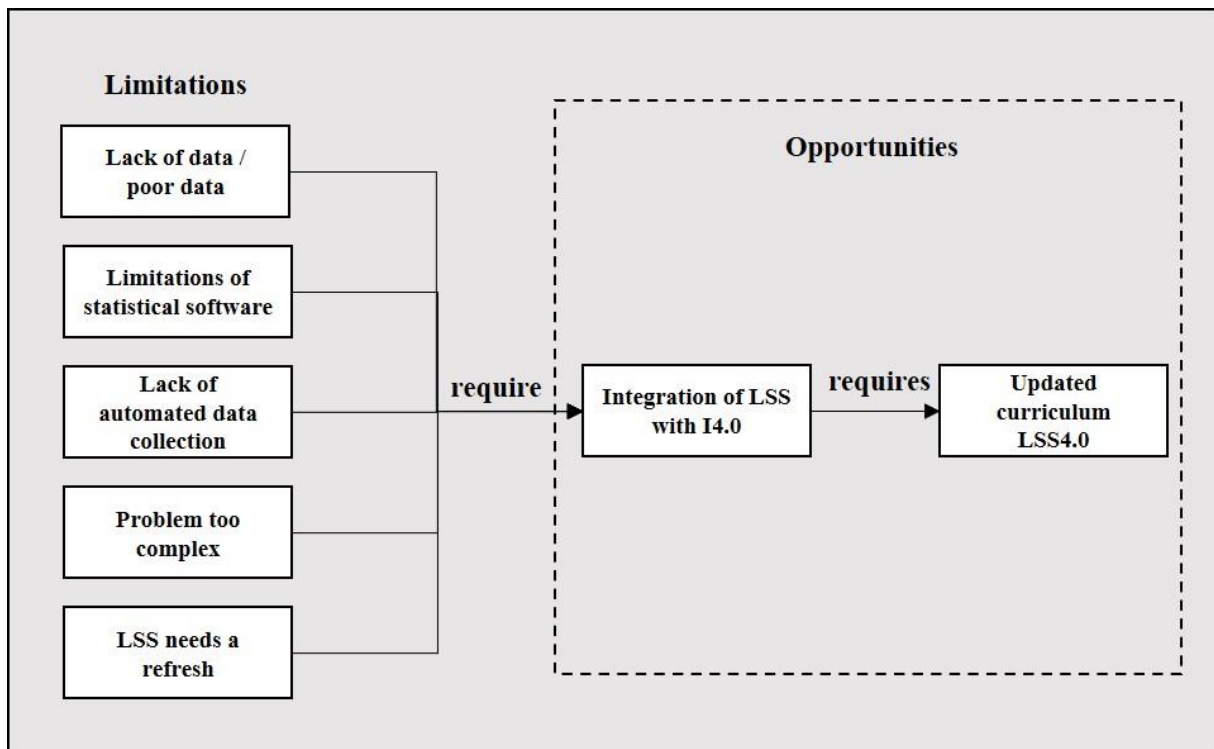


Figure 3: Relationships between LSS Limitations and Opportunities

As described above, a lack of data/poor data quality is still a common limitation:

AC-I-6: *"In practice, you are sometimes content to get 35 reliable data points to perform an analysis with."*

OP-I-10: *"It is difficult to query data that is stored in different systems. This makes data analysis very complicated... I would love to see that automated."*

Without sufficient data, statistical theory cannot be applied:

OP-I-10: *“Lack of data is the biggest problem. Due to this, we have to rely on opinion-based tools instead of using statistical tools.”*

If there are not enough data points or the data quality is doubtful, statistical tools are ruled out. On the other hand, with too many data sets, traditional statistical tools are completely overwhelmed, or the processing is extremely slow. Another downside is that updates are not performed dynamically, and it takes significant time and effort to gather the data from different sources and to enter them into tables. Consequently, DMAIC projects would profit from technologies that automate data collection from different systems and provide processed, good-quality data to a single data source.

A limitation that emerged from the data is that traditional tools used in DMAIC, e.g., XY-Matrix or Ishikawa-diagram, rely on the subjectivity of experts. Moreover, statistical tools cannot handle the complexity of causes and effects when a process has 60-70 input variables. As a result, the increased complexity of products and processes requires Lean Six Sigma to integrate data-driven techniques, such as process mining, data mining, and big data analytics. OP-I-5: *“But what you don’t see is if you look at that systematically via DMAIC, then there might be interactions between if we change something in the recipe which is then causing that we need to adjust the process settings... So, there might be influences which we don’t see, and that’s where the big data comes into play...”*

Big data analytics can reveal dependencies between influencing factors in complex processes that would not be visible when using conventional LSS tools.

Interview results have underpinned that “...it’s time for Lean Six Sigma to have a refresh.” (Fogarty, 2015, p. 17), and the results assert that the opportunities of artificial intelligence should be exploited (George and George, 2002, p. 102). The experts interviewed also confirmed a need to complement the curriculum with data science tools to be prepared for integration:

OP-I-7: *“We must start with the usual Green Belt education and offer an additional qualification for Green Belt 4.0, so the Green Belts can advocate for digitalisation.”*

Engineers and operators are no IT experts. Therefore, they need to be trained to manage the information flood generated by modern machines and IT systems and interpret them correctly. This long journey must start as soon as an organisation thinks about digitalisation.

Some experts claimed that the DMAIC waterfall approach should allow iteration cycles as in the Agile methodology. In practice, iterations are inevitable.

OP-I-11: *“If the cause is unknown, it does not make sense to set a timeline...Projects need an agile approach, DMAIC as a waterfall approach is not helpful! ...you cannot wait for weeks to start with the Improve phase. Spending weeks in Analyse; it is too static... Things change faster in a dynamic market.”*

Consequently, ‘agile DMAIC’ could become another emerging trend. Agile DMAIC could mean that improvement goals are split into smaller, incremental but independent objectives, and project teams may iterate between the DMAIC phases if needed.

5. Discussions

Findings from this study have elicited a need to integrate LSS with I4.0 as a remedial strategy for pressing LSS limitations, such as poor data quality or increased complexity. They are pressing because if they are not addressed, companies will put their competitiveness at risk (George *et al.*, 2019, p. 102). Therefore, the authors present a use case that offers companies the opportunity to implement an integration scenario relatively quickly.

Use Case: Integrating data mining into a Green Belt DMAIC project to reduce errors

In a Green Belt project aimed to reduce chip formation occurring during the drilling of edge connectors for printed circuit boards, a tool for data mining, Knime, was used to enhance the DMAIC phases. Knime is a standard software tool that uses configurable workflows (Figure 4, 5 and 7) to perform calculations at each step. Chip formation during the drilling process caused 7% of defects. The goal was to reduce defects to 3% or less which represented a benefit of €150k p.a. For the Measure phase, 100k data sets were pulled from the system with 100 process parameters and raw material characteristics. The output of the drilling process was chip formation 1/0 (1 = yes/0 =no) and was measured by a camera system (automatic inspection system with 100% check). A workflow was customised to perform data engineering, i.e., data cleansing and dimension reduction. The team used the algorithms “correlation” and “low variance procedure” to reduce the dimensions of the “big data” set.

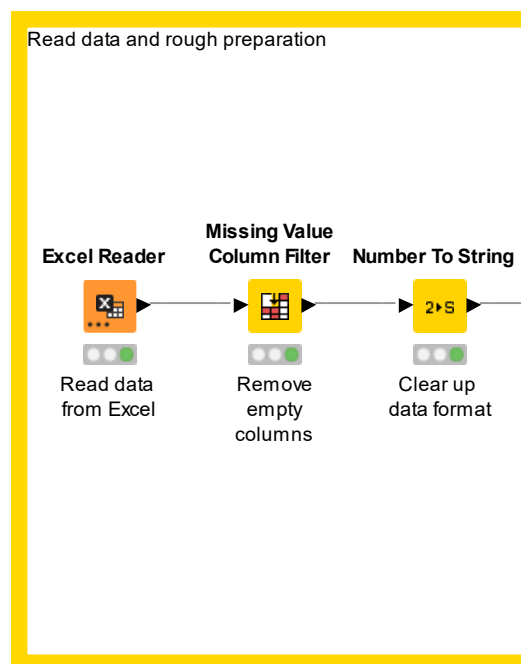


Figure 4: Workflow in Knime for reading and preparing data

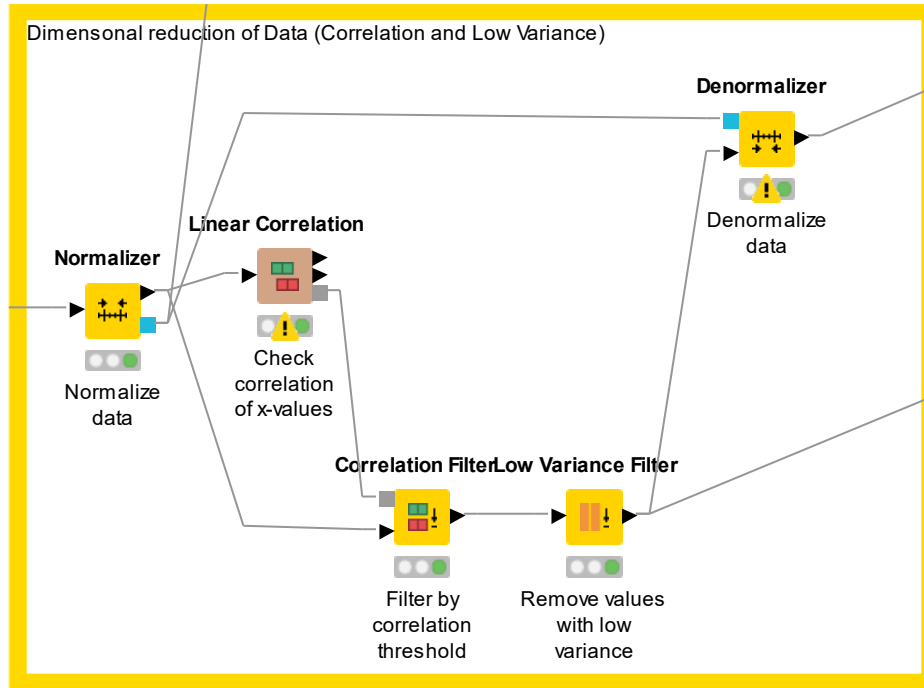


Figure 5: Workflow in Knime for Dimensions Reduction

The correlation algorithms helped eliminate associated input variables (Figure 6). Subsequently, the team performed an attribute agreement analysis for the camera system. The standard was defined by the quality department. The agreement percentage (camera versus standard) were above 80%, which meant that the camera system delivered a solid data quality. After checking the data quality, the team conducted a binomial capability analysis (using the attribute p-chart). The analysis returned a proportion on average of 0.068 chips for the product with the highest volume per year.

	MK01 - 17/18/19	MK30 - 17/18/19	MK28 - 17/18/19	MK20 - 17/18/19	MK22 - links - ...	MK22 - rechts - ...	MK 83 - Ebenh...	MK138 - Säbel...	MK31 - Recht...	MK31 - Recht...	Spanbildung - ...	Presse
MK01 - 17/18/19	corr = +1	corr = +1	corr = +1	corr = +1	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a
MK30 - 17/18/19	corr = +1	corr = +1	corr = +1	corr = +1	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a
MK28 - 17/18/19	corr = +1	corr = +1	corr = +1	corr = +1	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a
MK20 - 17/18/19	corr = +1	corr = +1	corr = +1	corr = +1	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a
MK22 - links - 17/...	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a
MK22 - rechts - 17...	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a
MK 83 - Ebenheit ...	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a
MK138 - Säbel - 1...	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a
MK31 - Rechtwink...	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a
MK31 - Rechtwink...	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a
Spanbildung - 17/...	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a
Presse	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a	corr = n/a

Figure 6: Correlation Matrix

For the root-cause analysis, a decision tree was created. Figure 7 shows the corresponding workflow.

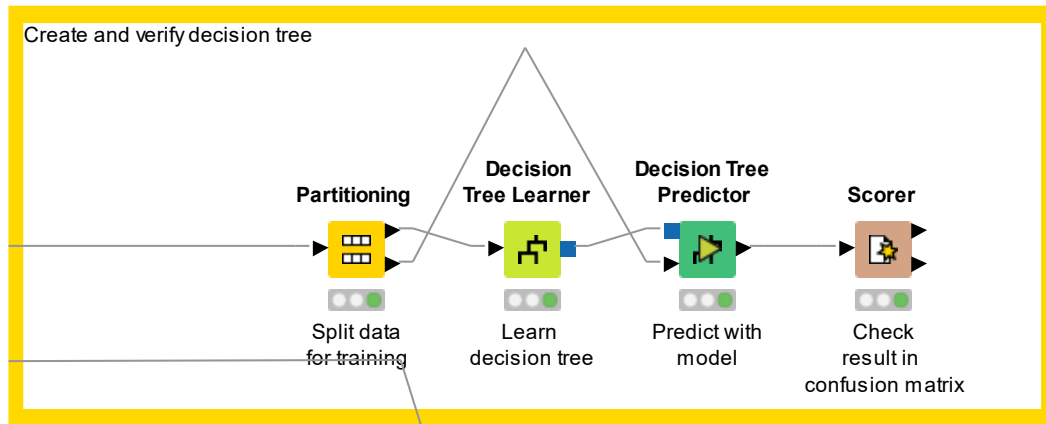


Figure 7: Knime workflow (decision tree part)

The variables' effect on the output was analysed using a decision tree. The higher a variable is located in the decision tree, the stronger its influence on the classification for chip formation (Han *et al.*, 2012, p. 331). In the second level of Figure 8, we can see five different machines with the strongest influence on chip formation. The probability for chip formation differs among the machines. Three of them form nearly zero chips, while machine 1 and 4 cause the vast majority. The two levels that follow show which combinations of variable settings led to more chip formation. This information helped the project team decide which machines to use and chose the ideal variable settings to avoid chip formation and, therefore, defective parts. The team used creativity techniques to generate improvement solutions and assessed the criteria (effectiveness, easy to implement, cost) for reducing chip formation. Hence, the team implemented the solution with all corresponding steps; documentation, training and communication. After the implementation, the chip formation was 2.7 % (compared to 6.8 % before improvement). In the Control phase, the team introduced monitoring for chip formation with a control chart (p-chart) to sustain the improvement results. Overall, the project delivered a benefit of €75k (no additional investment was necessary). After successful project completion, the Green Belt candidate acquired a Six Sigma certification.

8th INTERNATIONAL CONFERENCE ON LEAN SIX SIGMA

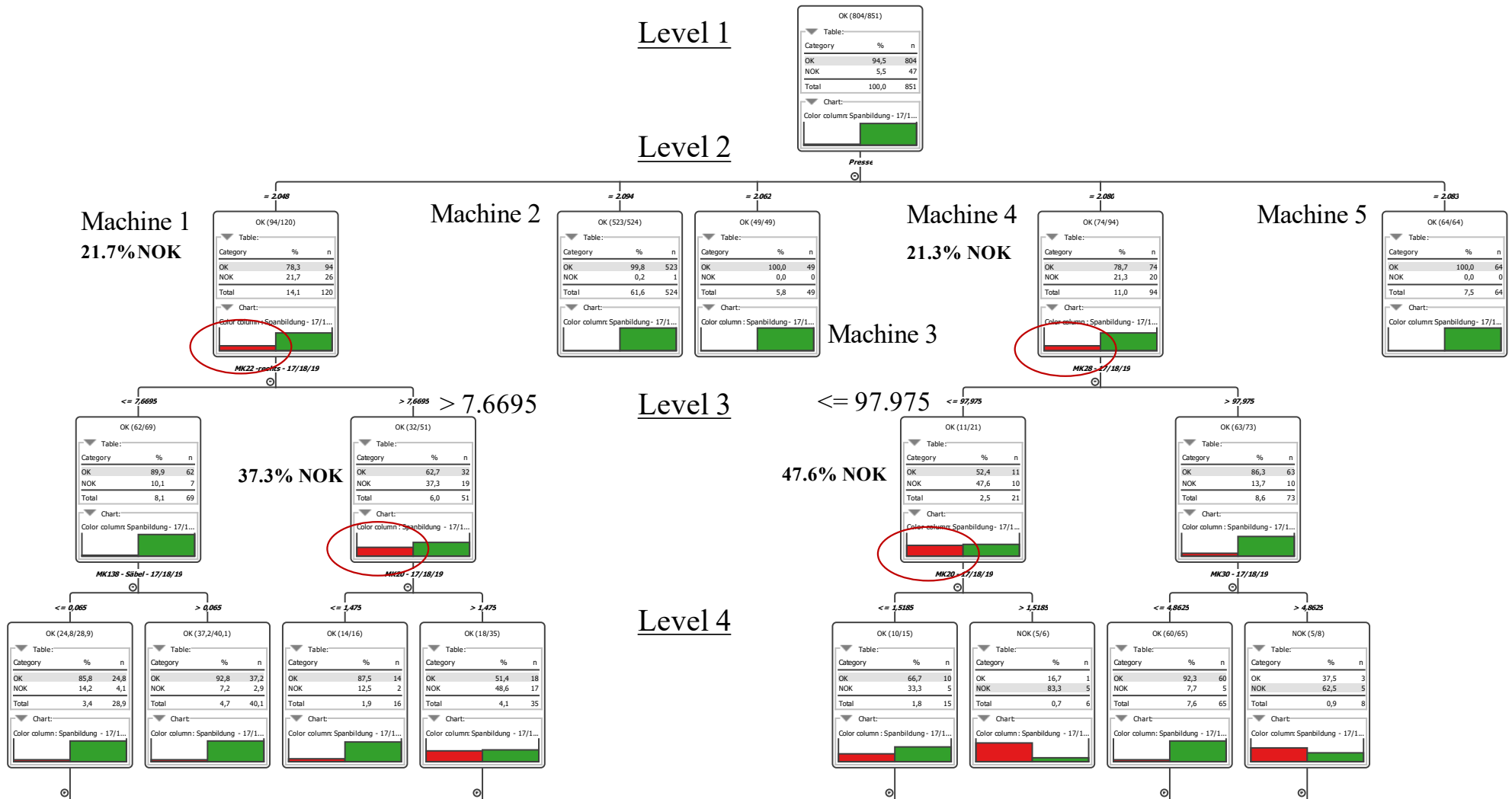


Figure 8: Decision tree

6. Conclusion

This research sought to explore which LSS limitations practitioners have encountered in their improvement projects and how the limitations link to each other. The interviews revealed causal relationships of LSS limitations which had not yet been studied in the academic literature. Additionally, new limitations have emerged, such as ‘lack of automated data collection’, ‘limitations of statistical software’ and ‘problem is too complex for traditional LSS’. These new limitations, along with previously reported limitations related to data quality and availability, call for an integration of I4.0 technologies into the traditional LSS framework.

As an initial response to this urgent need, the authors presented a proven use case integrating data mining into a DMAIC project. The scenario demonstrated that an integration with data-driven tools does not have to be costly and can start with small steps. However, Belts must be trained so they can apply new LSS4.0 techniques. Consequently, the LSS curriculum also needs to be updated to include data science skills.

7. Limitations and future research directions

This study has been intended as a first step for confirming the need for an LSS 4.0 framework. The continuation of this research is planned by the first author, i.e., to design a comprehensive LSS4.0 framework considering more data-driven concepts, e.g., process mining, robotics process automation and machine learning.

The authors additionally call for further research into Agile DMAIC, LSS project selection based on big data analytics, or integrating Design for Six Sigma (DFSS) with Industry 4.0.

References

- Aboelmaged, M. G. (2011) 'Reconstructing Six Sigma barriers in manufacturing and service organisations', *International Journal of Quality & Reliability Management*.
- Albliwi, S., Antony, J., Abdul Halim Lim, S. and van der Wiele, T. (2014) 'Critical failure factors of Lean Six Sigma: a systematic literature review', *International Journal of Quality & Reliability Management*, 31(9), pp. 1012-1030.
- Albliwi, S., Antony, J. and halim Lim, S. (2015) 'A systematic review of Lean Six Sigma for the manufacturing industry', *Business Process Management Journal*.
- Antony, J., Snee, R. and Hoerl, R. (2017) 'Lean Six Sigma: yesterday, today and tomorrow', *International Journal of Quality & Reliability Management*.
- Antony, J., Sony, M., Dempsey, M., Brennan, A., Farrington, T. and Cudney, E. A. (2019) 'An evaluation into the limitations and emerging trends of Six Sigma: an empirical study', *The TQM Journal*, 31(2), pp. 205-221.
- Antony, J., Sony, M. and Gutierrez, L. (2020) 'An Empirical Study Into the Limitations and Emerging Trends of Six Sigma: Findings From a Global Survey', *IEEE Transactions on Engineering Management*, 10.1109/tem.2020.2995168, pp. 1-14.
- Berry, M. J. and Linoff, G. (1997) 'Data mining techniques: for marketing, sales, and customer support 1997', *John Willey & Sons*.
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C. and Wirth, R. (2000) 'CRISP-DM 1.0: Step-by-step data mining guide', *SPSS inc*, 9, p. 13.
- Denscombe, M. (2014) *The Good Research Guide: For Small Scale Social Research Projects*. 5 edn. Open University Press.
- Deuse, J., Erohin, O. and Lieber, D. (2014a) 'Wissensentdeckung in vernetzten, industriellen Datenbeständen', *Industrie*, 4, pp. 373-395.
- Deuse, J., Wiegand, M., Erohin, O., Lieber, D. and Klinkenberg, R. (2014b) 'Big Data Analytics in Produktion und Instandhaltung', *Instandhaltung im Wandel. Industrie*, 4, pp. 33-48.
- Fayyad, U., Piatetsky-Shapiro, G. and Smyth, P. (1996a) 'From data mining to knowledge discovery in databases', *AI magazine*, 17(3), pp. 37-37.
- Fayyad, U., Piatetsky-Shapiro, G. and Smyth, P. (1996b) 'The KDD process for extracting useful knowledge from volumes of data', *Communications of the ACM*, 39(11), pp. 27-34.
- Fogarty, D. J. (2015) 'Lean six sigma and big data: continuing to innovate and optimise business processes', *Journal of Management and Innovation*, 1(2), pp. 2-20.
- George, M. L., Blackwell, D. K., George Jr, M. and Rajan, D. (2019) *Lean Six Sigma in the age of artificial intelligence: Harnessing the power of the fourth industrial revolution*. McGraw-Hill Education.
- George, M. L. and George, M. L. S. (2002) *Lean Six Sigma : Combining Six Sigma Quality with Lean Production Speed*. McGraw-Hill Education.
- Gläser, J. and Laudel, G. (2010) *Experteninterviews und qualitative Inhaltsanalyse* (4 vols). VS Verlag für Sozialwissenschaften.
- Glasgow, J. M., Scott-Caziewell, J. R. and Kaboli, P. J. (2010) 'Guiding Inpatient Quality Improvement: A Systematic Review of Lean and Six Sigma', *The Joint Commission Journal on Quality and Patient Safety*, 36(12), pp. 533-AP5.
- Han, J., Kamber, M. and Pei, J. (2012) '8 - Classification: Basic Concepts', in Han, J., Kamber, M. and Pei, J. (eds.) *Data Mining (Third Edition)*. Boston: Morgan Kaufmann, pp. 327-391.
- Krefting, L. (1991) 'Rigor in qualitative research: The assessment of trustworthiness', *American journal of occupational therapy*, 45(3), pp. 214-222.

- Martins Rosa, A. and Broday, E. (2018) 'Comparative Analysis between the Industrial and Service Sectors: A Literature Review of the Improvements through the Application of Lean Six Sigma', *International Journal for Quality Research*, 12(1).
- Panayiotou, N. A. and Stergiou, K. E. (2020) 'A systematic literature review of lean six sigma adoption in European organisations', *International Journal of Lean Six Sigma*.
- Pepper, M. P. and Spedding, T. A. (2010) 'The evolution of lean Six Sigma', *The International Journal of Quality & Reliability Management*, 27(2), p. 138.
- Rojko, A. (2017) 'Industry 4.0 concept: Background and overview', *International Journal of Interactive Mobile Technologies*, 11(5).
- Russom, P. (2011) 'Big data analytics', *TDWI best practices report, fourth quarter*, 19(4), pp. 1-34.
- Santos Jr, H. P., Black, A. M. and Sandelowski, M. (2015) 'Timing of translation in cross-language qualitative research', *Qualitative health research*, 25(1), pp. 134-144.
- Schroeder, R. G., Linderman, K., Liedtke, C. and Choo, A. S. (2008) 'Six Sigma: Definition and underlying theory', 26, pp. 536–554.
- Singh, M. and Rathi, R. (2019) 'A structured review of Lean Six Sigma in various industrial sectors', *International Journal of Lean Six Sigma*.
- Snee, R. (2010) 'Lean Six Sigma – getting better all the time', *International Journal of Lean Six Sigma*, 1(1), pp. 9-29.
- Sony, M., Antony, J. and Naik, S. (2020a) 'How do organisations implement an effective LSS initiative? A qualitative study', *Benchmarking: An International Journal*, 27(5), pp. 1657-1681.
- Sony, M., Antony, J., Park, S. and Mutingi, M. (2020b) 'Key Criticisms of Six Sigma: A Systematic Literature Review', *IEEE Transactions on Engineering Management*, 67(3), pp. 950-962.
- Sony, M., Naik, S. and Therisa, K. K. (2019) 'Why do organisations discontinue Lean Six Sigma initiatives?', *International Journal of Quality & Reliability Management*, 36(3), pp. 420-436.