Lean Six-Sigma in Aviation Safety:
An implementation guide for measuring aviation system’s safety performance

Ilias Panagopoulos*, Chris Atkin, Ivan Sikora,
City University London,
Northampton Square, London EC1V 0HB, United Kingdom
*Tel: +44-79315-70582
*E-mail: ilias.panagopoulos.1@city.ac.uk

doi:10.5296/jss.v2i2.10438 URL: http://dx.doi.org/10.5296/jss.v2i2.10438

Abstract
The paper introduces a conceptual framework that could improve the safety performance measurement process and ultimately the aviation system safety performance. The framework provides an implementation guide on how organisations could design and develop a proactive, measurement tool for assessing and measuring the Acceptable Level of Safety Performance (ALoSP) at sigma (σ) level, a statistical measurement unit. In fact, the methodology adapts and combines quality management tools, a leading indicators programme and Lean-Six Sigma methodology to formally measure and continuously improve a stable and in-control safety management process by reducing safety defects and variability from core organisational processes and objectives. The implementation guide was empirically tested and validated with data collected and analysed within a period of nine months by the safety department of a complex aviation organisation operating a large transport aircraft fleet.

Keywords: Aviation Safety; Lean Six-Sigma; Measuring Performance
1. Introduction

According to the International Civil Aviation Organization (ICAO) Annex 19 (2013), ‘safety is the state in which risks associated with aviation activities are reduced and controlled to an acceptable level’. Indeed, safety is a system quality stemming from a legal and regulatory framework that stipulates strict and high performance targets as well as a number of activities that must be performed by air operators. According to ICAO Annex 19 each service provider shall, as a minimum:

- Provide continuing monitoring and regular assessment of safety performance,
- Ensure remedial action to maintain agreed performance,
- Aim at a continuous performance improvement.

At European level, the European Aviation Safety Agency (EASA, 2014), in parallel to management system requirements, outlined its new harmonised approach for establishing a Performance-Based Environment by introducing a clear set of indicators and targets against which the oversight performance of civil aviation authorities is assessed. However, ICAO and EASA do not provide a clear guidance on how stakeholders could measure safety performance. Besides, in aviation industry there is a level of uncertainty as to what extent existing methodologies for measuring performance are suitable for those operators who have achieved excellent safety records and in-control processes and as such look for further improvements.

1.1 Key Research Questions

Safety performance is the State’s (i.e. Civil Aviation Authorities) or the service provider’s (i.e. any organization providing aviation products and/or services) safety achievement as defined by its Safety Performance Indicators (SPIs) and Safety Performance Targets (ICAO, 2013). Consequently, this study further investigates the following key research questions:

1. What methodology could proactively measure system safety performance and improve the safety performance measurement process?
2. Could a conceptual framework assist the continuous improvement of the safety performance measuring process?

1.2 Aim

For addressing the key research questions, the paper aims to present a conceptual framework that will improve the safety performance measurement process and the aviation system safety performance. Figure 1 presents the framework, where the Safety-Performance Indicator Lean Sigma (Safety-PILS) model has been embedded within Define – Measure – Analyse - Improve and Control (DMAIC) continuous improvement process. This integration results in a continuous improvement framework that measures system safety performance and reduces the safety process variability. In addition, the study provides an implementation guide on how organisations could use this framework to design and develop a proactive, performance-based methodology for measuring Acceptable Levels of Safety Performance (ALoSP) at sigma (σ) level, a statistical measurement unit.
2. Measuring Safety Performance

An effective performance measurement system should monitor past performance and help to plan desired future performance (Gutierrez et al., 2014). According to Muller (2014), one of the main Safety Management System (SMS) objectives is to measure system effectiveness and improving safety performance. Since most accidents have multiple precursors that an accident is likely to happen, there is a common belief that even a small number of ‘leading indicators’ can identify increased risk of an imminent accident (Leveson, 2015). As a leading indicator example, instead of only measuring at the end of the year ‘number of runway excursions’ or ‘hard landings’ it would be more proactive and beneficial to monitor the ‘monthly rate of unstabilised approaches continued for landing’ that may lead over time to a runway excursion or a hard landing. Sikora (2013) defines a proactive monitoring system for airlines. Besides, Leveson (2015) discusses how to operationalise leading indicators as warning signals and through the Systems-Theoretic Accident Model (STAMP) proposes a proactive methodology for identifying leading indicators in an aviation system.

Furthermore, Podgorski (2015) suggests that new methods are needed to ensure management system effectiveness and he proposes a method for prioritising performance indicators based on the utilisation of a certain set of Specific, Measurable, Attainable, Relevant and Timely (SMART) criteria. In addition, Andriulo (2014) proposes a lean framework for measuring the effectiveness of a management system in the automotive industry. Also, Ulfvengren (2014) argues that Lean methodology by integrating quality with existing management processes would achieve effectiveness and could demonstrate safety performance in compliance with new aviation safety regulations. Besides, Verstraeten et al (2014) introduced a framework of
SPIs for the aviation industry and concluded that unless a process for continuous safety monitoring is in place, the system of SPIs only provides a snapshot view of safety. In the aviation industry, EASA has recently established the Network of Analysts (NoA) SPI Sub Group for considering the subject of SPIs (EASp, 2014). Moreover, a Safety Management International Collaboration Group (SM-ICG) was created in February 2009 as a joint activity between key aviation authorities to encourage progress and harmonisation. Although various guidelines have been developed, measuring performance from SPIs will require some time for air operators to determine how SPIs represent safety performance (Roelen, 2012). Also, Karanikas (2016) argues that aviation authorities have not clearly defined the different meanings between ‘system effectiveness’ and ‘effective operation of a system’: the former regards the effects of the system on the organisation whereas the latter refers to how much satisfactory a system is operated.

To sum up, in aviation industry the measurement process of a set of pre-defined indicators for measuring system’s safety performance has not yet been introduced or standardised. In addition, the development and measurement of proper SPIs is not straightforward and the operational experience for measuring the effectiveness of SMS is very limited, since ‘there are many questions yet to be answered on measuring safety performance’ (Roelen, 2012). Consequently, the main challenge remains how to control and maintain performance within agreed safety specification limits and how to develop an objective methodology that will proactively investigate and measure system performance variability (±σ) from target.

3. Lean Six-Sigma for measuring safety performance

Lean-Six Sigma (L6S) has been applied in the manufacturing and healthcare industry since 1990 (Mason, 2015) and is considered as the integration of two management philosophies, Lean and Six Sigma (6σ), and has been seen as a robust improvement methodology (Tenera, 2014). According to Harmon (2014), Lean focuses on improving the flow of activities and 6σ focuses on improving the quality and consistency of process outputs. Ulfvengren (2014) in an effort to develop a Safe-Lean concept for an airline that demonstrates safety performance concluded that, ‘what can be really monitored is the normal variation of performance data.’ Nevertheless, this research study argues that Lean itself cannot bring any safety measurement process under statistical control. Therefore, organisations such as Motorola, Samsung, Sony and Honeywell are using 6σ for enhancing safety and as a continuous improvement management tool (Rehman, 2012). Six sigma is a statistical measure of excellence in process performance wherein process tolerance corresponds to 6σ with a maximum of 3.4 Defects per Million Opportunities (DPMO). Consequently, a process performing at 7σ corresponds to 0.019 DPMO, an outcome that may satisfy most aviation safety departments since the risk of an accident due to a defect is less than one accident per 10 million events (1×10⁻⁷). Besides, Harmon (2014) argues that 6σ is a good methodology for understanding the measuring process and the use of statistical techniques to analyse the outcomes.

4. The Conceptual framework

The conceptual framework for measuring safety performance composed of the Safety-PILS model embedded within the DMAIC continuous improvement process. The next sections of
this chapter are presenting this combined effort along with a practical implementation step guide for measuring system safety performance.

4.1 The Safety–PILS model

The Safety-PILS model provides guidance on how organisations could design, implement and use a proactive, performance-based measurement tool for assessing and measuring ALoSP. For the purposes of this study, safety performance is measured at sigma (σ) level.

Figure 2. Illustration of the Safety-PILS model, inspired from Ishikawa (1968)

Nevertheless, the Safety-PILS model shown in Figure 2 provides a holistic view on how organisations could set leading performance indicators and monitor metrics\(^1\) on the top of identified root-causes that affect safety performance or how to set lagging indicators and feedback metrics on the top of safety outcomes (e.g. number of accidents). In fact, the above model adapts and combines quality management tools, a leading indicators programme and L6S methodology to continuously improve a stable and in-control safety management process. In particular, an Ishikawa-fishbone (Ishikawa, 1968) root-cause and effect diagram is used for establishing leading and lagging performance indicators in an aviation system aiming at within ±1.5 sigma (σ) tolerable safety limits.

---

\(^1\) The Organisation for Economic Co-operation and Development define metric ‘as a system of measurement used to quantify SPIs or how the SPI is being measured’ (OECD, 2008)
In addition, the Safety-PILS model aims to control and maintain safety performance within agreed Upper or Lower Specification Limits (i.e. USL, LSL) and to develop an objective methodology that will proactively investigate and measure system performance variability within ±1.5 sigma (σ) from an ALoSP target. Furthermore, Figure 3 shows how the overall performance of an integrated safety management system implemented within an aviation organisation could be affected by its components. The intricate relationship during operation of all these components eventually results in the overall safety performance for the organisation’s management system, usually captured by indicators relating to safety occurrences. In fact, the core advantage of the Safety-PILS model is that applies the Central Limit Theorem (Chern et al, 2015); since it repeatable uses a large size of data and means, the distribution of the sample means will finally approach a normal distribution.

4.2. The DMAIC process

Safety-PILS model assists operators to comprehend and design their safety system in accordance with the agreed safety performance targets and specification limits. Accordingly, the next step for the operator is to follow the DMAIC for continuously improving the overall system’s safety performance measurement process. Through DMAIC process, the operator could apply L6S methodology for measuring both the performance of each established indicator and system safety performance variability at sigma level from core safety objectives. However, the sequence of DMAIC steps and the times could vary widely, according to the size, the type and the complexity of the project. Since safety measurement is a ‘data-driven process’, operators should mainly examine whether the data are normalised, the process is in statistical control and capable to achieve the desired outcome. As the results of this research study indicate, DMAIC is a useful methodology that could assist operators to accomplish this task.

In general, the ‘Define’ phase rolls out the tools such as the Voice Of the Business (VOB) which identifies pre-actions in the measurement process. Also, multiple regression identifies
the condition of optimality on root causes and effects in the pre-action process. The ‘Measure’ phase reveals the continuous assessment of measurement process. In ‘Analyse’ phase, the vital root causes that impact the responses are identified. The ‘Improve’ phase concentrates on optimising the vital root causes by implementing potential solutions and in ‘Control’ phase, the confirmation run is conducted, the results are obtained at σ level and process variation is controlled or eliminated within acceptable levels (Srinivasana, 2014).


Table 1 shows the Phase-I implementation guide for measuring system’s safety performance.

Table 1. Implementation guide– Phase I

<table>
<thead>
<tr>
<th>Phase-I. Safety-PILS model utilisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Design Safety-PILS model for VOB</td>
</tr>
<tr>
<td>- Define the driven KPI for the VOB</td>
</tr>
<tr>
<td>- Set the VOB Targets and the LSL-USL based on industry standards</td>
</tr>
<tr>
<td>- Set SPIs on the VOB- Critical to Safety (CTS) characteristics</td>
</tr>
<tr>
<td>- Set metrics on each VOB SPI and the associated LSL-USL</td>
</tr>
<tr>
<td>2. Correlation and Multiple Regression Analysis</td>
</tr>
<tr>
<td>- Identify correlation between cause/effect</td>
</tr>
<tr>
<td>- Define the driven KPI for the VOB</td>
</tr>
</tbody>
</table>

Phase I: Design Safety-PILS model for VOB

Key terms and concepts in Lean are the Voice of the Customer (VOC) and the VOB. The VOC reflects the customer needs and the customer perceptions of operator’s products and services, whereas VOB usually refers to what an enterprise strives to achieve, such as key organisational targets and core objectives. Since VOC and VOB identify improvement opportunities and consume outputs from a process, need to be in harmony (Arafah, 2015).

Figure 4 shows how the safety department of a complex air operator utilised the Safety-PILS model for designing the VOB related to flight safety. The driven KSPI has a clear organisational target and USL, meaning to achieve a Safety Occurrence rate/1000 Flying Hours (FH) of 0.6 and not exceeding the rate of 1.1 by the end of 2016.
Moreover, all leading and lagging SPIs are drawn with their associated specification limits and a clear target defined by the operator. Finally, each SPI consists of a set of monitor and feedback metrics and, each metric should perform within specification limits for achieving its target. Table 2 shows two indicative indicators, the SPI_02: Runway Excursions (RE) and SPI_06: Loss Of Control (LOC) with their associated metrics. In particular, SPI_02 and SPI_06 with the metrics ‘Unstable-De-stabilised approaches (all)’ and ‘Take-off configuration warnings events’ have been selected as a research sample for further explaining and validating the conceptual framework implementation guide. The VOB indicators have been selected based on the Pareto Analysis (Brook, 2014) results, past experience and literature review. Also, the study examined the correlation exist among VOB indicators and revealed a moderate to strong correlation, since all examined Pearson’s coefficient (R) values were ranging from 0.6-0.8. Finally, the regression analysis revealed that ‘Unstable - De-stabilised approaches’ is the metric that accounts the most for the variation in the VOB process output.

Table 2. VOB: Indicative Safety Performance Indicators (SPI) & associated Metrics

<table>
<thead>
<tr>
<th>SPI_02: Runway excursions (RE)</th>
<th>SPI_06: Loss of Control (LOC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deep Landing events</td>
<td>Stick-shake and alpha floor events</td>
</tr>
<tr>
<td>Unstable-De-stabilized approaches (all)</td>
<td>Take off Configuration warnings</td>
</tr>
<tr>
<td>Unstable-De-stabilized approaches (all) continued for landing</td>
<td>Low speed during cruise events</td>
</tr>
<tr>
<td>High speed touchdown events</td>
<td>Low speed during approach events</td>
</tr>
<tr>
<td>High speed rejected take-off events</td>
<td>Percentage of pilot’s readiness rate for proficiency</td>
</tr>
<tr>
<td>Take-off landing events involving loss of</td>
<td>Pilot’s utilisation effectiveness</td>
</tr>
</tbody>
</table>
Phase II: Apply Six Sigma DMAIC methodology

Phase II shown in Table 3, starts with the Data Collection Planning (DCP), the hypothesis testing and control chart selection. During DCP the Safety Office had to decide on what type of data is most appropriate to collect for measuring the VOB SPIs and metrics, what resolution is needed, what statistical tool should be used to interpret the data and what the sample size and frequency should be.

Table 3. Implementation guide for measuring system’s safety performance – Phase II

<table>
<thead>
<tr>
<th>Phase-II. Safety-Apply Six-Sigma DMAIC methodology</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Data Collection Planning (DCP) for Hypotheses tests</strong></td>
<td></td>
</tr>
<tr>
<td>- Hypothesis Testing</td>
<td></td>
</tr>
<tr>
<td><strong>2. Control Chart selection - road map</strong></td>
<td></td>
</tr>
<tr>
<td>- Control Chart selection for each VOB SPI and Metric</td>
<td></td>
</tr>
<tr>
<td>- Identify special causes: If none the process is In-Control</td>
<td></td>
</tr>
<tr>
<td><strong>3. Measurement System Analysis (MSA)</strong></td>
<td></td>
</tr>
<tr>
<td>- Where does the variation of data comes from?</td>
<td></td>
</tr>
<tr>
<td>- Is the process Accurate and Precise?</td>
<td></td>
</tr>
<tr>
<td><strong>4. Process Capability</strong></td>
<td></td>
</tr>
<tr>
<td>- Is the process capable (i.e. efficient)? At what sigma level?</td>
<td></td>
</tr>
<tr>
<td><strong>5. Analyse the data</strong></td>
<td></td>
</tr>
<tr>
<td>- Identify root cause and attractive areas for improvement</td>
<td></td>
</tr>
<tr>
<td>- Identify best and feasible solutions</td>
<td></td>
</tr>
<tr>
<td><strong>6. Pilot solutions</strong></td>
<td></td>
</tr>
<tr>
<td>- Demonstrate that piloted solution provides a Return of Investment (ROI)</td>
<td></td>
</tr>
<tr>
<td><strong>7. Define Control Plan and Roll-out improvement</strong></td>
<td></td>
</tr>
<tr>
<td>- Monitor the Control Plan to sustain the change</td>
<td></td>
</tr>
<tr>
<td><strong>8. Measure total system safety performance</strong></td>
<td></td>
</tr>
<tr>
<td>- Voice of the safety Process (VOP)= VOB</td>
<td></td>
</tr>
</tbody>
</table>

For the purpose of this study the VOB sample size was 2.0 and the sampling frequency for collecting data was twice a month. The subgroup size for the VOB was identified as 1.0. The type of the safety data was quantitative-discrete, collected within a period of 6 months, meaning January– June 2016, and analysed with Minitab 17 statistical software and tools (Brook, 2014) during July-September 2016. In addition, the hypothesis summary report of Paired t-tests, STDV and Anderson Darling normality tests revealed that the data initially
followed the normal distribution curve and there was no special cause for any detected variation that needed further investigation (i.e. value p>0.05).

Next, the appropriate Statistical Process Control Charts for understanding the performance of the examined indicators have been selected. Since, a single chart may not allows the operator to model a process in a way that gives the needed understanding, the safety office implemented multiple charts drawing from the same set of data. The Xbar chart plotted the mean and examined no special causes of variation (i.e. no points outside control limits), meaning that the process was ‘in control’ and the Moving Range (MR) chart indicated that the average was moving downwards. Finally, the control charts and in particular MR charts were considered as effective since the data were normally distributed. During Measurement System Analysis (MSA) the Analysis of Variance (ANOVA) reports revealed that 7.32% of the total variation reported was caused by the gauge. Since this number was lower than 10% was considered acceptable. Also, Part-to-Part variation was 0.38651 which was big enough and therefore, good for the measurement system. The ‘Data by operator’ graphs revealed the variance recorded for each metric. Although some variation was still apparent, the metrics were measured consistently with each other. Finally, the ‘Part operator interaction’ graph revealed that the line for each metric followed about same pattern and the averages varied enough that differences were clear, which was also reasonable. To this end, it appeared that the measurement system was adequate for the operator’s needs.

One of the next critical steps was the Process Capability analysis. Process capability could assist the operator to understand if the measurement process of a particular SPI or metric is capable or efficient (i.e. the process fits within USL/LSL). Besides, to determine the σ level the particular indicator performs. The analysis revealed that $C_p$ (i.e. potential capability) was greater than $C_{pk}$ (i.e. actual capability), meaning that the potential process was centred. In addition, the $C_{pk}$ was less than 1, meaning that not only a special cause but also a common cause of variation was going to produce unacceptable variation (i.e. defects). $C_p$=0.17 means that only 17% of the process fit within USL/LSL and $C_{pk}$=0.08 means the process was 80% over one specification limit. The control charts revealed that there was no special cause of variation, meaning that this process was in-control. However, neither the actual process nor the potential was capable. At this point the operator had to apply solutions or to take mitigation measures such as, improving oversight through Flight Data Monitoring, improving the Standard Operating Procedures, increasing Flight Crew awareness and enhancing Flight Simulator training requirements.

During study period, several intermediate process performance reports of the ‘VOB before’ implementing solutions took place for re-evaluating the actual and potential capability of the process. In one report, it appeared that the actual capability was still poor with 21.74% of the process being out of specification limits with DPMO=217391. Nevertheless, the process remained centered since $C_p>C_{pk}$, thus, the most cost effective way to reduce variation was to discover and then eliminate (or reduce) any special causes might appear in the follow-up measurement phase. In fact, the overall actual process appeared to perform between 2-3σ level performance. In addition, by implementing the proposed solutions the potential of the process was to improve since only 1.34% of the process could probably remain outside the
specification limits with an expected number of DPMO=13391. Therefore, the operator proceeded by implementing the proposed solutions. Also, the diagnostic report of the ‘VOB before’ implementing solutions verified that data were continuing to follow a normal distribution (p=0.448). Therefore the operator concluded that the abovementioned Cpk results were valid.

Besides, in one case the Xbar-R chart showed four special causes of variation, as shown in Figure 5. In fact, the existence of a special cause is not necessary a problem since it can add to or detract from the total process variation. The process performance report of the VOB that was received right ‘after’ implementing solutions, re-evaluated the actual and potential capability of the process. It appeared that the actual capability was initially improving, the process was remaining centred and the data were continuing to follow a normal distribution. Besides, the Xbar-R chart showed none special causes for the existing variation.

Figure 5. Xbar-R chart showed for VOB before- Special causes of variation

To conclude with, the final report provided the capability comparison for VOB ‘before’ and ‘after’ the implementation of solutions, as follows:
- The percentage out of the specification limits has been significantly reduced by 51%, meaning from 16.43% before to 7.97%.
- The STDV (i.e. variation) was significantly reduced by 7%.
- The actual process performance has been increased by 0.52%.
- The σ level has been increased by 43%.
- The DPMO have been reduced by -84577.
- The actual σ performance was 1.41+1.5σ shift which equals to 2.91σ level performance.

5. Findings and Recommendations

By the end of Sep 2016 the VOB (i.e. occurrences rate) continued to improve meaning that the percent defective was 0.68% and the percentage yield or acceptance rate was 99.32%. These results indicated that the total system’s safety effectiveness was approaching 4σ level performance with the potential to have DPMO=6210 (i.e. safety occurrences per one million flying hours). Nevertheless, the study had few limitations. Consequently, the paper recommends the application of the conceptual framework to different settings, different sample and type of SPIs/metrics and the qualitative validation of the results by interviewing Subject Matter Experts (SMEs). As further research the study recommends the application of Genetics Algorithms and Simulation to the implementation guide. In this case, a metaheuristic procedure could sample a set of solutions and select or generate an algorithm that provides a sufficiently good solution to this safety performance optimization problem.

6. Conclusions

In the aviation industry, the process for developing SPIs, measuring safety data and analysing safety performance is neither clear nor yet standardised. The study introduced an integrated, empirical tested conceptual framework that may satisfy Authorities’ requirements for establishing a performance-based approach in aviation safety. Furthermore, the study identified and filled the gap existing in the literature and proposed an implementation guide for measuring aviation system safety performance. In particular, the paper introduced the Safety-PILS model that integrates with the DMAIC continuous improvement process. This integration develops a meaningful methodology for measuring performance, though with limitations. The study revealed that the application of L6S methodology can enhance the measuring process. To this end, the proposed guide is a new way of thinking for designing a safety case aims to achieve desired outcomes within agreed specifications limits.

References


**Copyright Disclaimer**

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).