Amritpal Singh Bansal, Master's student in PSM

Development of a comprehensive data validation including multiple independent systems to increase quality in volume calibration

Master's thesis

To achieve the University Degree of Master in Production Science and Management (Dipl.-Ing.)
Submitted to
Graz University of Technology

Under the supervision of
Assoc.Prof. Dipl.-Ing. Dr.techn. Mario Hirz, Institute of Automotive Engineering, Graz University of Technology
and company mentor
Dipl.-Ing. Marhold, Maria-Elisabeth, Lead Engineer, Dataset Management, PTE / DEO, AVL Graz

Graz, May 2017

Affidavit

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZ online is identical to the present master's thesis.

17.05.2017
Date

Signature

Acknowledgment

Firstly, I would like to express my gratitude toward my thesis advisor at TU Graz, Professor Mario Hirz, from Institute of Automotive engineering for providing the mentorship for the thesis. His lectures, which extends from mechanical domains of automobiles to mechatronics, telematics, virtual development, and innovation, has broadened the ken of many students at TU Graz like me. Apart from this, I would like to thank Professor Siegfried Vössner, whose lectures about process optimization with programming based solutions enabled me to implement many innovative concepts in the research work.

I would like to thank from the bottom of my heart to my thesis advisor at AVL, Dipl.-Ing. Maria-Elisabeth Marhold, who gave me chance to work as a part of such a prestigious organization. The fact is, it was her vision that guided me toward different areas with her inspiring management skills, which stimulated my thinking capabilities toward the process, without which I would not be able to think for highly innovative solutions suggested in the thesis. Apart from this, I'm grateful to Markus Lindner-Mayerhold, who was the adviser for my thesis in early months but had to change the department. It was his idea that was generated with years of experience in calibration process that label values for vehicle variants are related to their attributes. I built upon his idea to develop the concept of data classification and data tags based on these attributes in various areas of calibration dataset management process, which possess major part in the thesis. With his encouraging attitude, I was able to prove using my programming skills that it was indeed possible and efficient to do so.

I would like to thank all my colleagues at AVL, especially the calibration pilots and calibration engineers in DEO department, working on calibration process at AVL, who provided their precious time to me with various interviews and teachings. It was them, working in the field, who shared the problems they face in the process, which was the root cause behind why I was able to innovate to generate such effective solutions. Special thanks to Kevin Oswald, who have been working with calibration dataset management since years and helped me countlessly on various occasions with very precise answers. Besides, I thank Anne Frey, a calibration engineer, who shared the office with me at AVL and helped me with random questions at random times for calibration process.

I sincerely thank my brother in India, who helped me a lot with affording to live in all these times of studies and research work. I would like to thank a retired language teacher named Christina Neppel living near Andritz, who helped me on many occasions during my studies and made me feel like a family, away from my home. Furthermore, I would like to thank my mother and my belated father, who nurtured me with their love which made me capable enough to reach my life goals.

Dedication

I dedicate this thesis to my belated father, Rajinder Singh, a well-known gears machinist in the industrial area of a small city in North India by the foothills of Himalayas. He always used to epitomize metal machines and mechanical designs from Central Europe. I believe his inspiration is one of the reasons that I find myself researching in the finest automotive research organization of Central Europe, with the help of knowledge and skills imparted by admirable Professors of TU Graz. Their teachings in every single lecture I find is composed with 'the words of knowledge', colossal as the beautiful Alps, if visualized, surrounding the university. Besides, I am highly thankful to all the people in Austria, who provided me with various supports in all these times in unthinkable manners.

Abstract

The automotive industry is going through the times which require high innovation and quality, elements which possess no limits but continues development. Continues development provides space to many complexities. Innovation and quality are being achieved by software shift in the automobiles. However, developing these software which are parts of automotive ECUs, is a complex task in each phase. Inefficient development or any defects in these software can result in serious issues, such as underperformance of automobiles, product recall, project delays etc. ECU software are composed of datasets, which are further composed of labels. Automobile performance can be highly enhanced by optimizing the values of these labels in datasets by calibrations on test-beds, which require collaborative efforts of teams in different organizations. Each label gives rise to data which is required to be managed carefully between these teams, carrying nearly 50% of typical project efforts. Thus, the quality of the calibration process can be increased by improving the data quality in the dataset management process. Currently, the datasets of ECU software in calibration process undergo validation and management by project specific ad hoc solutions developed with an atomistic view of the process. For many complex activities, there exist no solutions at all.

This master's thesis carries a holistic view for the calibration dataset management process to create solutions which can be applied to different projects. The goal is to find the dataset validation methods by optimizing their management to make sure the quality before they are delivered to customers. It does this by applying a renowned Six Sigma methodology called DMAIC together with BPMN (Business Process Modeling Notation). DMAIC defines the areas to improve, measure the current performances of methods, analyses six sub-process requiring improvement based on quality measurement, improves the methods with recommendation actions, and further suggests control plans for continuous improvement of the process. The calibration process was analyzed with different levels of details using BPMN diagrams which fit better than simple flowcharts to the complexities of the process with different elements and teams. Recommendation actions are composed of concepts of business process automation with a software tool, forecast analysis, and data visualization. The efficiency of these concepts for calibration process is proved by developing prototype software in C# using API's of calibration dataset management software called AVL CRETA, forecast calculations in Excel sheets, and case studies on data visualization with node-link diagram. Thus, the data quality of datasets is proved to be enhanced for calibration process at AVL with the potential of providing a competitive advantage in above domain and further creation of jobs.

Acronyms and Abbreviations

AVL	Anstalt für Verbrennungskraftmaschinen
API	Application Programming Interface
AS	Arbeitsseite (german)
AT	Automatic Transmission
BPA	Business Process Automation
BPMN	Business Process Modeling Notation
CAN	Controller Area Network
CMM	Capability Maturity Model
Det	Detection
DMAIC	Define, Measure, Analyze, Improve and Control
DoE	Design of Experiment
ECU	Electronic Control Unit
EU	Euro
FMEA	Failure Mode and Effect Analysis
FWD	Front Wheel Drive
GUI	Graphical User Interface
IEC	International Electrotechnical Commission
INCA-EIP	Experimental Target Integration Package
INCA-LIN	INCA Local Interconnect Network
INCA-MIP	INCA MATLAB Integration Package
LSL	Lower Specification Limit
MAD	Mean Absolute Deviation
MAPE	Mean Absolute Percentage Error
MCD	Measurement, Calibration, and Diagnostics
MIL-STD	Military Standard
MSE	Mean Squared Error
MT	Manuel Transmission
MVB	Master Validation Batch
Occ	Occurrence
ODX Flash	Open Diagnostic Data Exchange Flash
ODX LINK	Open Diagnostic Data Exchange Link
OEM	Original Equipment Manufacturer
P & D	Prevention and Detection
PPM	Parts Per Million
PPV	Pre-Production Vehicle
QG	Quality Gates
RPN	Risk Priority Number
RS	Referenzseite (german)
RWD	Rear Wheel Drive
SAE	Society of Automobile Engineers
Sev	Severity
SPICE	Software Process Improvement and Capability Determination
TQM	Total Quality Management
USL	Upper Specification Limit
VIP label	Very Important Person Label

Table of Contents

Chapter 1: Introduction

Over the recent years, the automotive world has encountered significant transformation. The new vehicle sales around the world have not demonstrated substantial growth, but instead, the automakers have been trying to gain market share by being distinct in the populous automotive market to gain a competitive advantage. The struggle of OEMs to differentiate their product in the market, with limited remained potential in the physical and mechanical domain, have overcome by the software shift in the Industry [1]. There has been significant growth in the use of electronics in the last decades, which has substituted various mechanical and hydraulic systems. Anti-lock braking system, adaptive cruise control, and electronic stability system, etc. are all examples of such systems. Advancements in the electronics systems in automobiles instigated the requirement of in-vehicle communication. In the beginning, ECUs involved being used as standalone components for a specific function. The strategy changed due to the need of integrated electronics system architecture, with the requirement of various functions to be scattered among different ECUs and information network among functions. These days, electronics and software encompasses 90% of all innovation, and software related activities comprise 50-70% of ECU development expenditure. Today, a premier car constitutes nearly 70 ECUs, linked by 5 system buses and transacting up to 2500 signals [2]. The figure below shows the evolution of software based embedded systems in the vehicles and growing ECU market size:
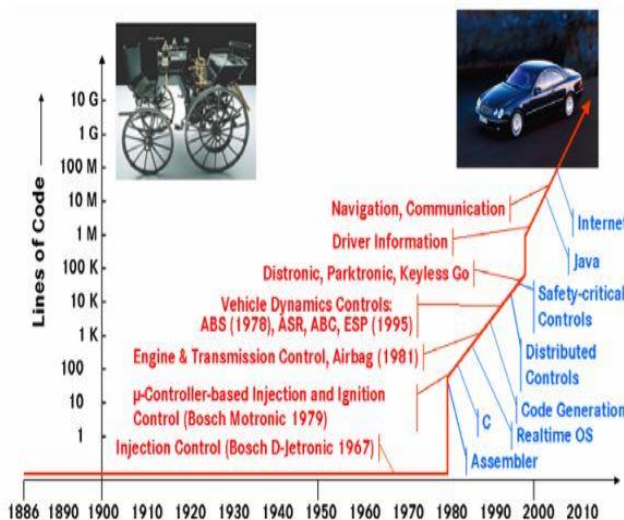
| Figure 1: History of use of software based embedded systems [3] | Figure 2: China Automotive ECU market size by application ($ Billion) [4] |
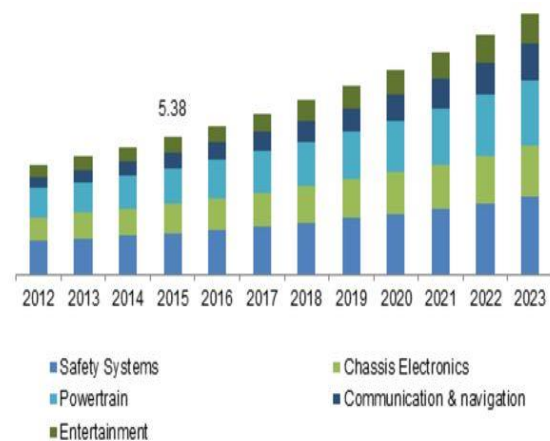
The growing trend of ECUs usage in automobiles consequently led to growing trend in software usage, making vehicles as software incentive IT systems as well. Consider the example of the Mercedes-Benz S-Class (W220) from 1998, which had more than 50 control units, running on more than 1 million lines of codes, with CAN-B, CAN-C, and D2B Bus systems. It was encompassed with sharing more than 100 bus signals, and more than 1000 single signals with about 100 MB embedded codes. While, the model W140 (1990) had about 1 MB of embedded code, and the model W221 (2005), had more than 500 MB of embedded code. In the year 2000, of 125 billion euro of automotive market volume, software share composed 20% of production costs, and the rest hardware. Meanwhile, in 2010, the software share of production costs increased to 37% [3]. The modern cars

incorporate greater hardware and computational capabilities as compared to Apollo spaceship, which led humans to the moon. As ECUs are not anymore a standalone component, ECU software compromise as an intensive element of ECUs, without which they are incapable of functioning. Increasing software functionalities has led the automotive industry to various complex challenges. Such challenges have caused various issues for organizations, such as leading to various project suspensions and expenditure overrun [2]. The ECU software currently finds a wide range of application in various automotive systems, and any defects or bugs in them may bring a major overhaul for an OEM. Various ECU software with their categories and attributes in the automobiles are listed below:

Sl. No	ECU software category	Attributes
1	Emission related control ECU	-Safety critical -Emission regulation compliance -Less user interface -Long product life cycle
2	Non Emission related control ECU	-vehicle usability concern -safety regulation(concern) -less user interface -long product life cycle
3	Telematic ECU	-service commitment -conditional user interface complexity -medium product life cycle
4	Infotainment ECU	-less safety related -fancy user interface -neat feature size -adapting consumer electronic features -shorter product life cycle
5	Off-board ECU	-interacts with on-board system -adaptive to multiple vehicle models

Figure 3: ECU software category [2]   Figure 4: Level of innovation in various aspects of vehicle by 2020 [5]
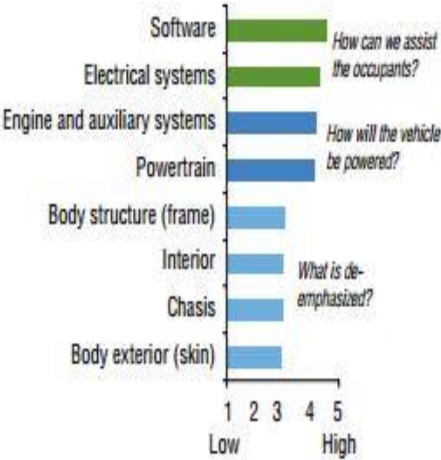
The number of defects in the automotive software is on the rise, following their growth in the automotive applications. OEMs are struggling with the fact that the current processes are establishing various issues with growth in software's content. Therefore, OEMs are facing issues for spending an enormous amount of costs due to warranty, of which about 30% costs accounts to defects of electronics and software, exceeding software and electronics warranty costs for 3 billion US$ per year [1]. Between the years 2005 to 2012, about 3.6 million vehicles were affected as a consequence of 32 automotive recalls. This encompassed electronics and software issues, which nearly doubled to the number of 63 between June 2015 to 2012, affecting about 6.4 million cars [6].

With respect to the volume of projects, 'complications, quality, time, expenditure, and accountability' are extensive issues of ECU software development projects. To conform to the requirement of high quality and reliable software, extensive works are being done for software verification and validation. Examples include development of software management tools, software standards, and defect prediction approaches etc. To provide the required quality and reliability for the software, testing is considered to be the essential segment of the development process, amounting up to 50% of total software development expenditure. Various defect prevention strategies aim at optimizing software development process, planning, and controlling for various testing elements for software development and analyzing the degree of completeness before the deadlines [7]. The defects can be reduced to a great extent with the use of automated tools, which also assists with saving

time, which further helps development teams to focus on other important issues in the process[8]. The defects timeline per vehicle along the automotive software development with various aspects is shown in the figure below:
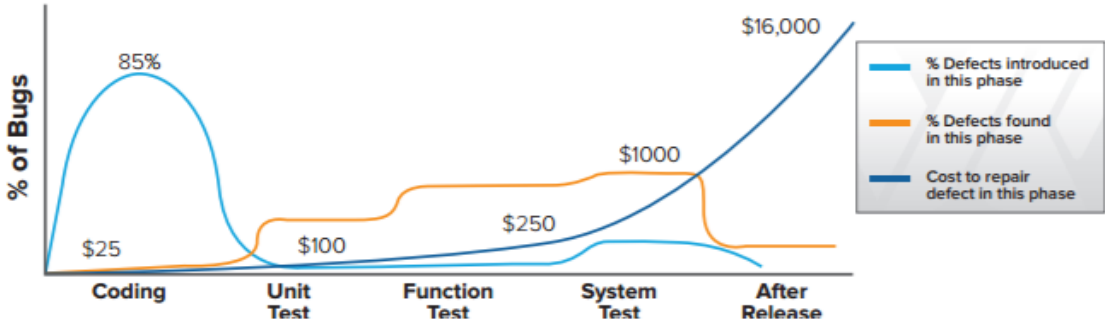
Figure 5: Defect detection with respect to repair costs [8]

1.1 Scope of the thesis

Even today, automotive software engineering finds itself in an immature state. Lifecycle management of automotive software is confined to the early stage. For example, numerous supplier, including various OEMs, are incapable of reaching Capability Maturity Model (CMM) level 2. Moreover, it's a major challenge to reuse same software for the different car models. For example, only moving from one car generation to another, with about 10% changes, demands 90% of software to be re-written. One of the major reason behind this, apart from hardware issue, is because of low automation in the development of automotive software. Automotive software engineering follows various process models like Capability Maturity Model Integration, Software Process Improvement and Capability Determination (SPICE), and V-model. Apart from these, it follows various standards such as IEC 61508, ISO 26262 and MISRA-C [9].To understand the development process of automotive ECU software, a general form of V-model can be considered, shown in the figure below:
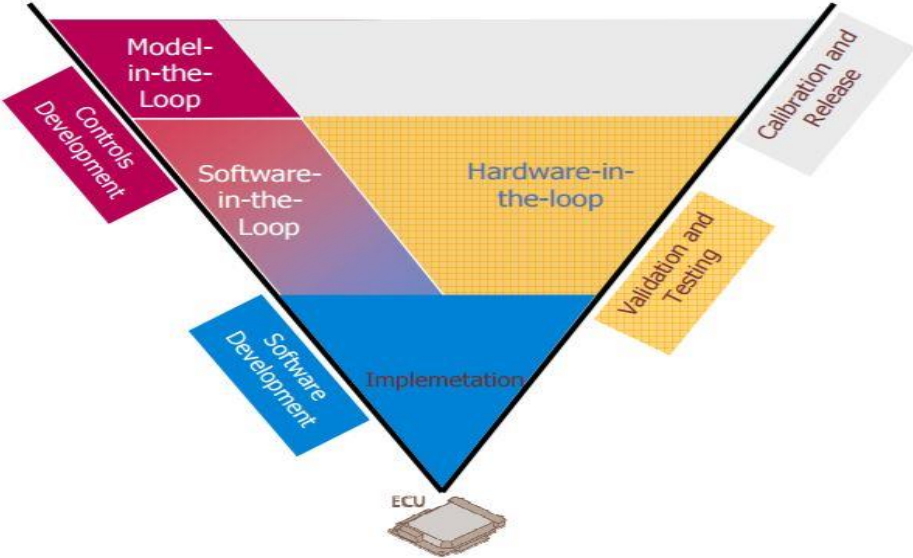
Figure 6: V-model for model based ECU development for ECU software [10]

The main goal of V-model is to provide fast and agile development, and test and validations for new controls strategies. Model-in-the-loop deals with optimizing the datasets for calibration, various release protocols for ECU sub-system models, analyzing communication for various ECU sub-systems, and cleansing plant models. The major section of software-in-loop is to optimize ECU software's architecture by functional validations, for example, for CAN bus composition and their loads. Hardware-in-the-loop deals with the validation of ECU software by validation with actual sensors and actuators signals, with respect to advance measurements of calibration [11].

This thesis deals with the calibration process and software release segment of the automotive ECU software development relating to model-in-the-loop process, as shown in the V-model in figure 6. Within the scope of the calibration process, which possesses a wide range of activities, it deals with the calibration dataset management processes. Software quality is a wide range of subject that has various aspects such as usability, efficiency, reliability, maintainability, and reusability. Among these, this thesis deals with the 'reliability' and 'maintainability' section of an automotive software quality, within the development process for different variants of cars. Reliability means that a software performs required functions without any failures. Maintainability means that an organization is able to make changes to a new revision of software with ease [11]. The required software quality is maintained in accordance with customer needs in creating a new revision for the software or making changes to software variant. The concepts of software development, which are related to unit testing, and software-in-the-loop are out of the range of this research work. This thesis compromises various issues associated with the calibration process and software release, which includes direct involvement of customer of a company for development purpose.

1.2 Problem definition and objectives of the thesis

ECU software calibration is an iterative process which consists of measurement and calibration during the runtime to optimize or efficiently tune the parameters/labels it encompasses. In brief, the process involves:
- Calibration engineers, who work on ECU algorithm's parameters/labels. Various teams of calibration engineers are formed based on the classification of the parameters/labels with respect to their function.
- Calibration pilots, who support the calibration engineers with dataset management and exchange of data and information with customers,
- Lead engineers, who lead various teams of calibration engineers, and manage the project, and
- Various other Individuals such as project managers, steering committees, and skill area managers etc.

The calibration engineers work together under the supervision of lead engineers, on the ECU software parameters/labels. They optimize the labels using various software tools on test beds, chassis dynamometer, or durability runs etc. by making changes to them based on results obtained. These changes are made iteratively until the final optimized dataset of labels is formed. The entire process involves a large amount of activities for managing these datasets efficiently with the help of calibration pilot. Dataset management is carried out using a tool called AVL CRETA, which calibration pilots use to obtain and manage the datasets between calibration engineers and customers. AVL CRETA provides a central

repository which acts as a backbone for communication and optimal management of the vast amount of data for calibration datasets.
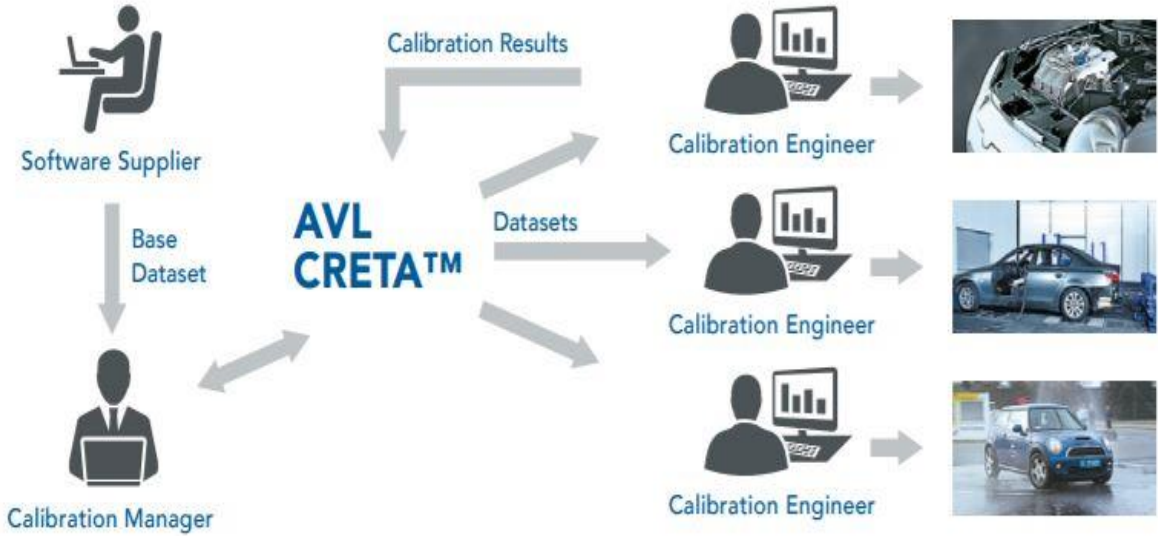
Figure 7: Brief depiction of dataset management for calibration process using AVL CRETA [11]

Currently, due to the growing complexity of various projects from various software suppliers or OEMs etc., various issues with dataset management have arisen. Calibration pilots are facing difficulties in certain sections of dataset management with AVL CRETA. There are issues with dataset validation methods, effective information flow between various teams or customer, and effective project management for the calibration process. All these problems emerge together to produce various quality issues for ECU software's dataset management, which is one of the major section of ECU software development. Various problems with dataset management include following aspects with respect to the functionalities that AVL CRETA provides in the calibration dataset management in the calibration process:

- Process workflow: Currently, the calibration process at AVL is facing various challenges with smooth workflow because of various manual activities to be performed in between all the phases of the calibration process. Each requires certain time to be executed and to be handled by project specific solutions, such as by the use of Excel files (ad hoc solution).
- Collaborations and traceability: Since the calibration process is divided into various parts due to rising needs of various activities for managing and validating datasets, there are problems occurring with collaborative work and traceability. This is because the data of each part of activities shares no common database, but are managed with ad hoc solutions with Microsoft Excel files saved in different places of project drive. For example, if an engineer makes changes to certain label/parameter of the dataset which influences many other labels/parameters, it may go undetected. Another example can be that engineers who perform durability run may provide wrong information to calibration engineers about dataset for its vehicle's variant on which the test was performed. Thus, there are problems with an efficient alert or notification systems and effective information flow in teams.

- Responsibility management: There have been huge issues of handling dataset elements like labels/parameters which belong to the responsibility of particular AVL calibration engineers. For example, various data associated with their labels such as values, dimensions, or comments etc. can be changed without their knowing, unintentionally or intentionally, by other team members or supplier or customer. As discussed before, sometimes labels that can influence labels in the responsibility of other engineers may be changed, resulting in inefficient responsibility management.

- Conflict free data merging: AVL CRETA provides merging of changes made to the dataset labels/parameters. The changes can produce certain issues for the same reasons discussed in the previous point. Conflict free data merging solutions are being made by project specific Microsoft Excel dataset validation solutions in AVL currently (ad hoc solutions).

- Reporting and documentation: Various reports and documentations for data management for individuals like AVL calibration project managers, customers or suppliers are being produced using the Excel macros with no connection with AVL CRETA database. This results in inefficient reporting and documentation for the calibration process management. The Excel reports or documentations are project specific (ad hoc solution)

- Project status control: Currently, the project status control tools provided by AVL CRETA for calibration milestone tracking, lacks various features required for efficient project management and cause various issues with customers. To provide more information to customers, project specific reports are made using the Excel macros, by manually downloading related data from AVL CRETA database (ad hoc solution)

- Using know-how: Various data involved from the know-how of calibration engineers is incapable of being used via AVL CRETA. For example, calibration engineers set values of some labels/parameters with their know-how. These values are saved and used by creating and programming Excels macros to map them to the dataset's labels values downloaded from AVL CRETA database manually. These solutions for dataset validation made by Microsoft Excel are project specific (ad hoc solution).

All these causes behind these issues merge together to reduce the software quality based on maintainability and reliability. Currently, individuals working with AVL calibration datasets management for projects are carrying an atomistic perspective toward the process. The sub-processes of calibration process like pre-calibration, testbed calibration, chassis dynamometer tests, durability run test, conflict free dataset management (dataset validation), and monitoring project etc. have been further divided into various segments. This occurred after analyzing the need for various new activities for dataset validation, arose from growing projects complexities like checking switch labels, VIP labels and analyzing the parameters of durability run etc. This causes a lot of time to engage with each of the separate segments associated with each sub-processes of the calibration process. This causes the team to lose precious time, creating barriers to focus on various other essential parts of the project. They manage data of datasets or prepare reports for customers or internal management, related to each segment of sub-processes, by programming Microsoft Excel macros for them. This causes team management various hassles with providing responsibilities to various employees for handling and creating Excel files. Moreover, calibration pilots are facing usability issues with AVL CRETA, as users in the field are unable to use efficiently various tools provided in it, due to the growing project complexities or the way calibration engineers works. These issues can be

visualized as in figure 8. Considering the points discussed, it can be concluded that, as a whole, the main reason behind the deterioration of ECU software quality with growth in the complexity of the projects, is the use of project specific ad hoc solutions based on Excel macros (with no connection to AVL CRETA database). These solutions demand various activities with various manual tasks. The other major issue being, the usability or incapable functionalities of AVL CRETA with respect to the growing project complexities, causing team management to perform various manual tasks, resulting in the generation of various defects in the ECU software's dataset management. This gives rise to need to develop solutions which can perform dataset validation and management for different projects in efficient manners.

Process sections with different teams   Each section with various set of activities   Process Islands

Process paritions, for example, conflict free data merging, durability run tests, Pre-calibrations etc.   Each process section segmented into various activities performed for them, such as, switch labels check, VIP labels check etc.   Management issues
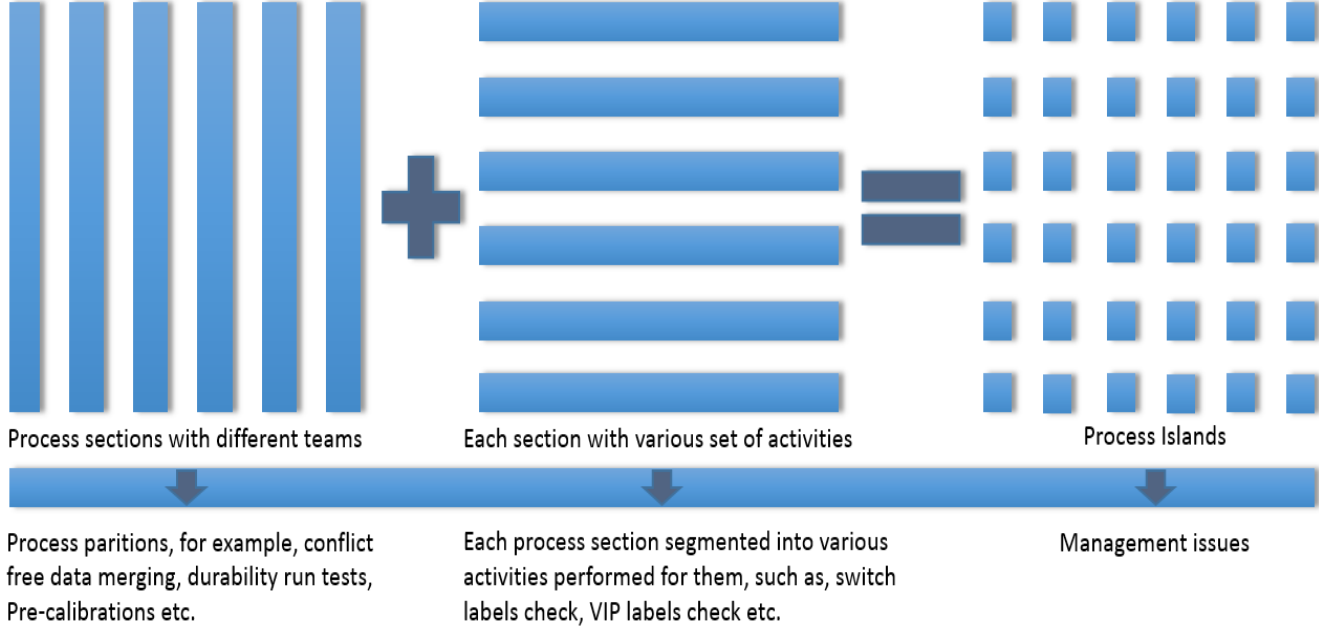
Figure 8: Segmentation of various parts of calibration process with various activities (visualization of overall problem in the process)

Considering these subjects, following objective of the thesis were determined:

- To study the current calibration dataset management process at AVL and various fundamental elements associated with it, such as AVL CRETA tool, its API and various project management methods like quality gates, used in the company,
- To use the Six Sigma 'DMAIC' methodology, and conduct a detailed evaluation of essential sub-processes and activities e.g. switch and VIP labels check for dataset validation, durability run tests, and monitoring of projects etc. A team for implementing various tools of Six Sigma has to be made at each required phase, with individuals related to them. The teams have to be made with respect to the sub-processes and activities to be evaluated, and interview various other engineers or team management affected by the process activities.
- To carry out the FMEA analysis as a part of DMAIC methodology to analyze the current solutions for each activity involved in various phases of dataset validation in calibration process to overcome various hassles in the workflow,
- To study the recommendation actions generated as an outcome of conducting FMEA analysis in the calibration dataset management process, and relate them in

order to generate a comprehensive dataset validation solution to increase the quality in the volume calibration,

- To prove the concepts behind the recommendation actions by programming in C#, a prototype linked to CRETA database, using its API. Collect opinions of various process stakeholders such as calibration engineers, pilots, lead engineers etc. for the results of the concepts behind recommendations actions, after they use the software prototype. For recommendation actions unrelated to the CRETA database, use other methods to prove the concept behind results.

Thus, the target of this research work is to develop dataset validation methods and other process optimization solutions to guarantee the dataset quality before delivering them to the customers. It has to make sure with the solutions that they replace difficulties encountered by current process solutions to optimize the calibration process for dataset management, which would in return enhance the ECU software quality provided by AVL List.

1.3 Report outline

The first chapter of the thesis discusses the significance of research work for current and future of automotive industry. It gives the brief description for the area of focus of the thesis in the automotive software development. It then explains the current situation in a concise manner in the form of case study, to state the problem statement and objectives of the research work carried out. The second chapter discusses the literature review, which provides a detailed knowledge of all the concepts that are related to the research work for ensuring the quality with dataset validation methods and optimal management for ECU software calibration process at AVL. It begins with the quality management concepts and discusses in detailed quality improvement methodology for Six Sigma called DMAIC. All the tools used in the various phases of DMAIC are discussed in detail. After DMAIC, ECU software calibration process is discussed in the new chapter, to enable the reader to acknowledge the complexities involved in the calibration process. This includes discussing various software solutions and files types used in the automotive industry for the calibration process. The thesis then leads the reader to a new chapter for project management and explains various concepts involved for an efficient data management in an organization. After this, in the same chapter, the concepts of business process management are discussed in detail. Various elements involved in business processes, used in the research work, are explained such as business process modeling notation (BPMN) and business process automation. Forecast analysis is then explained with the concepts of exponential smoothing, used later in the research work. In the end of the chapter, quality gates concept is discussed, as it is used extensively at AVL to handle the calibration projects. The fourth chapter deals with utilizing DMAIC methodology along with its various tools in each phase for current calibration process at AVL. It is divided just like the phases of DMAIC such as Define, Measure, Analyze, and Improve. Most of DMAIC phases involves using FMEA tool, as it was felt to be providing required information for the process in an efficient manner. The control phase is discussed in the second last chapter, as it is a result based phase. The second last chapter provides proofs to support the concepts behind various recommendation actions with a software prototype prepared by C# programming, Microsoft Excel calculations, and case studies. This chapter then includes discussing control phase of DMAIC to provide actions for continuous improvement for calibration process at AVL. In the end, the last chapter discusses the conclusion of the thesis and a final statement.

1.4 Assumptions and limitations

It is assumed, as a part of qualitative analysis involved in DMAIC methodology that all of the AVL employee involved in the research of the thesis, are well qualified and trained with their profession and possess extensive knowledge in their field of ECU software's dataset calibration process. They were considered to be supporting the research work with truthful and accurate feedbacks in various meetings or interviews involved with several employees. All of the employees involved had years of experience with their work domain in the calibration process and thus answered correctly with their firm skills and capabilities developed during these times. While making the research, above thesis finds it's certain limitations due to the following reasons:

- Certain employees were concerned about the confidentiality for their project.
- The confidential data of research work is replaced by dummy data, and certain information is written in broad general perspective for the concerns of maintaining the confidentiality of AVL projects.
- Another constraint occurred for the calculations of precise values for quantitative measurements. For example, in FMEA analysis, as there was no records or history created by the company for the number of defects found in the process, the experience of team members was used to judge the defect occurrence probability.
- Staff involved in utilizing the methodology of DMAIC had knowledge of it and Six Sigma as a part of their corporate training or education, but possess no certified belts of Six Sigma. However, to use the Six Sigma methodology, recommendations from a black belt owner engineer at AVL were taken. It is to be noted that the purpose of the thesis was to improve and ensure the quality up to certain level, but not up to the level of Six Sigma.
- Furthermore, the recommendation actions could not be applied in the calibration process of the company in the limited timeline of the thesis. Due to this, a prototype, dummy data, and case studies were used in order to prove the concepts behind the recommendation sections for their effectiveness.

As the reader would have a brief idea about the content of the entire thesis, next chapters for literature view explains fundamental concepts used in the studies and methodology of the research work.

Chapter 2: Quality optimization: DMAIC and its various tools

2.1 Introduction

This chapter introduces the essential factors necessary to gain the insight of the discipline of quality management, its various theories, methodologies, and tools. It describes the concepts related to process optimization and effective project management, such as business process modeling & optimization, quality gates, forecast analysis etc.

Quality has developed into a crucial mean of gaining competitive advantage in a market. There are several philosophies available in the subject of quality. Talking about quality, it would be unfair to not mention about Dr. W. Edwards Deming, who is considered to be the best quality expert in the world of quality. He transformed Japanese industry after second world war with his expertise in quality management [12]. His chain reaction theory describes that improved quality leads to cost decrease and less rework, which in turn, leads to improvement in the productivity, which would cause an organization to lead in the market by better prices, allowing them to stay in the business, which would finally provide more work, and hence create more jobs for an organization [13]. Based on products, goods, or services, quality has several distinct definitions available. One of the important definitions of quality for services is defined by Juran as [14]:

 "Quality means freedom from deficiencies –freedom from errors that require doing work over again (rework) or that result in field failures, customer dissatisfaction, customer claims, and so on. In this sense, the meaning of quality is oriented to costs, and higher quality usually costs less".

Juran describes an organized approach toward management for quality in an organization. This involves various quality management areas which are interrelated to each other and form a continuous improvement strategy like PDCA from Deming. This trilogy is related together and described using the following diagram:
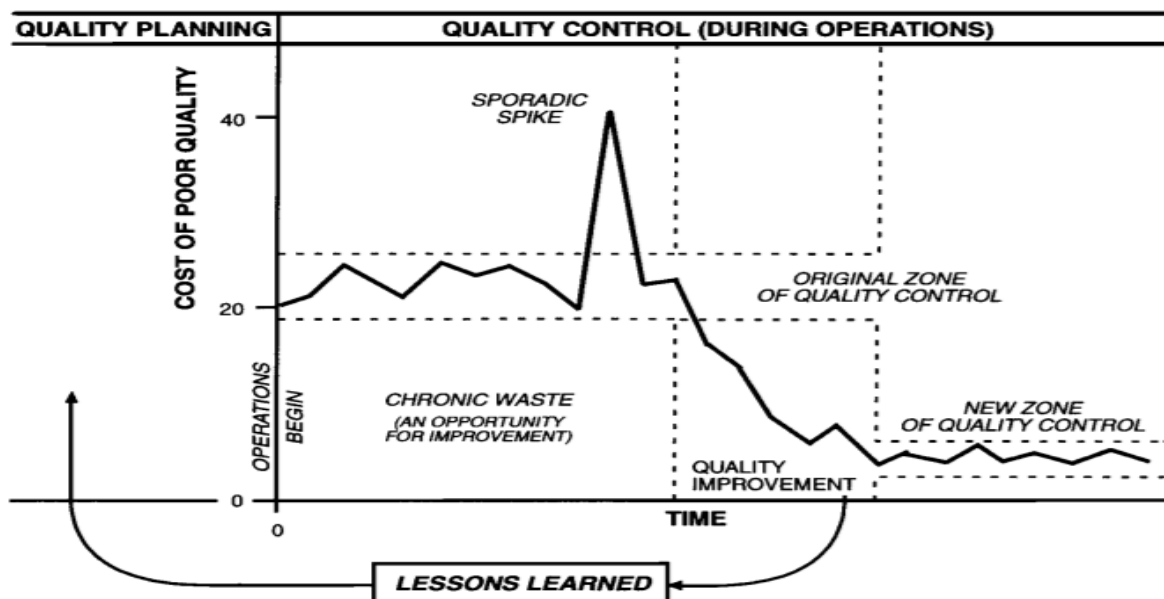
Figure 9: The Juran Trilogy, 1989 [15]

The three processes involved in this trilogy are discussed below [15]:

1. Quality planning: It involves developing a goal-oriented process or product. It determines customers and their needs. It then generates quality targets with respects to customer needs with minimum expenditure. The process is further developed to meet the targets, which are later proven along with its capabilities, under various conditions for an organization. According to Professor Christian Ramsauer from TU Graz, following methodologies or tools can be applied to the quality planning processes [14]:
   - Plan-Do-Check-Act
   - Quality function deployment
   - Robust design from Taguchi
   - Design of experiments (DoE)
   - Failure mode and effect analysis

2. Quality control: It is related to avoiding wastes growth and keeps an organization on the track of gains. It involves selecting required area for control, for which certain measures are developed. It then establishes a controlling process of measuring the deviation of actual results with respect to the standards generated by the process, which includes taking actions if inappropriate deviations are found in the process. According to Professor Christian Ramsauer from TU Graz, following methodologies or tools can be applied to the quality control processes [14]:
   - Seven quality control tools
   - 5-Why's investigation method
   - 8-D methodology
   - Statistical process control

3. Quality improvement: It is associated with lowering the cost for an ongoing process in an organization. Its most important aspect is to figure out recommendation actions as a part of learning and to pursue innovative means to acquire improved performances in the process [14]. It begins with realizing the requirement of improvement in a process to instigate the formation of improvement projects. It involves coordination of various individuals to evaluate the projects to determine causes and evaluate them to find various recommendation actions for the improvements. The actions then have to be proven to be efficient and effective under various conditions of a process. In the end, control strategies are determined in order to continuously gain benefits from improvement actions. Juan found out with his studies conducted on a number of top managers of various companies that most of them were dissatisfied with the quality improvement segment of quality in their companies. The reason was found out that they prefer the least priority to it, and thus possess the highest room of improvement [15]. According to Professor Christian Ramsauer from TU Graz, following methodologies or tools can be applied to the quality improvement processes [14]:
   - Six Sigma
   - Quality control circle
   - Suggestion system
   - Kaizen and the continuous improvement process
   - Lean production

2.2 Six Sigma

Sigma, donated as σ, is a statistical mathematics expression that measures the degree of inconsistency in a process. In order to attain Six Sigma for a process, the number of defects or errors per million opportunities should not exceed the number of 3.4. The main target of this measurement based approach is to concentrate on improvement of process and lowering the variations, by its projects. There exist two methodologies in achieving so [14]:
- DMAIC: Define, Measure, Analyze, Improve and Control
- DMADV: Define, Measure, Analyze, Define and Verify,

DMAIC is applied when an ongoing process is not operating efficiently or is unable to fulfill the customer requirements, while DMADV is applied when a completely new product or process need to be developed, due to the non-existence of any product or process for particular requirements. It is also used often when current product or process is unable to meet quality requirements of the customer, even after applying DMAIC [16]. ISO has standardized Six Sigma methodology with respect to DMAIC method, as ISO 13053-1:2011. It states that effective Six Sigma achievement using DMAIC method in a company occurs if certified individuals are assigned for Six Sigma implementation for a process. They can be certified as green, black, or master black belts based on their Six Sigma training or experience in using the methodology [17]. Sigma can be studied by the standard deviation in statistics and it can be related to normal distribution as [18]:

- The mean $\mu \pm 3\sigma$ encompasses 99.73% area,
- The mean $\mu \pm 2\sigma$ encompasses 95.46% area,
- The mean $\mu \pm \sigma$ encompasses 68.26% area.
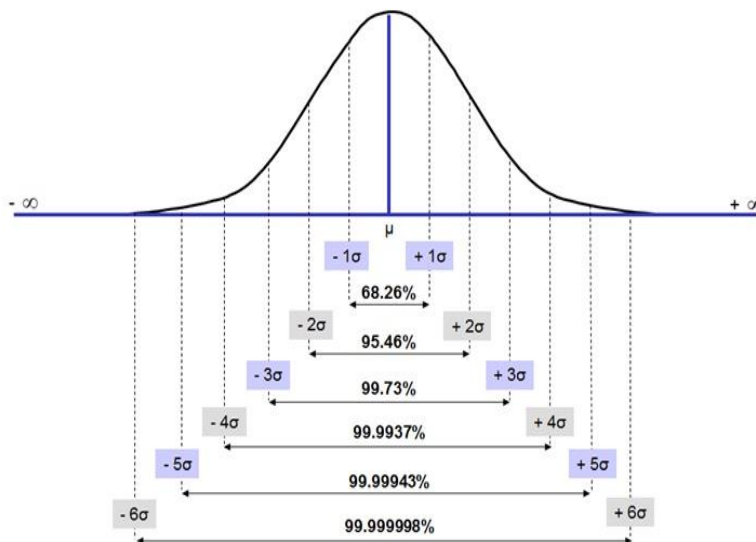
Figure 10: Six Sigma and its various levels in normal distribution [20]
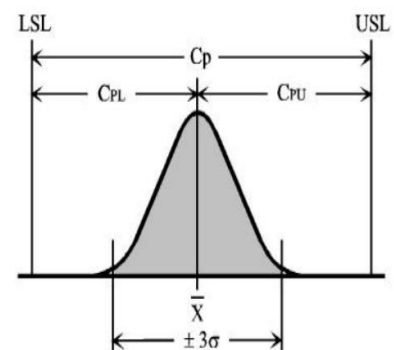
Figure 11: C_P, C_{PL}, C_{PU} in distribution [19]

It's a probability distribution, shown as N (μ, σ) i.e. it is characterized by mean and standard deviation. In general terms, a probability distribution is a theoretical frequency distribution. Around the mean, it possesses the highest value, which lessens as it moves away from the mean. Consider an example, where there are 30 assemblies to be made with 5 different

steps. This would result in 30x5 = 150 opportunities. If there exists one defect in each assembly, it would mean that there are 1 million defects in 1 million assemblies, resulting in 1000000/150= 6666 defects per opportunity, reaching to the sigma of level 3.97 [18], The formula to detect sigma level in the process is as follow [18]:

Level of Six Sigma $= 0.8406 + \sqrt{29.37 - 2.221 x ln(ppm)}$

Figure 11 shows lower and upper specification limits (LSL, USL). Considering C_P (a process's ability measure for meeting a specification), if the distribution is centered or not, and C_{Pk} (position of the total process variation with respect to mean of the specification), following are the levels achievable toward Six Sigma:

Distribution	Cp	Centered	Cpk	Expected Defects
±3.0 sigma	1.00	NO	0.50	66,810 ppm
±3.0 sigma	1.00	YES	1.00	2,700 ppm
±4.5 sigma	1.50	NO	1.00	2,700 ppm
±4.5 sigma	1.50	YES	1.50	3.4 ppm
±6.0 sigma	2.00	NO	1.50	3.4 ppm
±6.0 sigma	2.00	YES	2.00	0.002 ppm

Figure 12: Levels of Six Sigma and number of defects [19]

In 20 years with the use of Six Sigma, Fortune 500 companies of USA (Fortune magazine's largest 500 US corporations) were able to save about 427 billion US dollars [20]. Six Sigma Academy asserts that black belts can enable companies to complete up to four to six projects each year and save about 230,000 $ for each project [14]. These figures and above table are enough to reflect the potential that techniques like Six Sigma DMAIC possess for an organization. However, the goal of this research work is not to achieve any level of Six Sigma, but to use its methodology and its tools, to improve the current calibration process to ensure the dataset quality before it is delivered to the customers. Six Sigma, being the continuous process improvement strategy, is acknowledged as a methodology to execute total quality management (TQM) as well, which aims at quality optimization for processes [21]. DMAIC is the fundamental methodology to Six Sigma, used for optimizing and improving the business processes. However, DMAIC is not restrictive or limited to be used only for Six Sigma and can be utilized for the sake of improvement or optimization up to a certain level for processes as well. Hence, it can be utilized in the process optimization, which includes no target of achieving any level of Six Sigma, but rather just for improvement or optimizing the process [22].

2.2.1 DMAIC

Before DMAIC methodology is applied to a project, various attributes of the project shall be considered, which makes it adequate to be used for DMAIC. These elements for projects are given below [23]:
- Project shall present performance issues for an organization with unidentified solution,
- Project shall possess unambiguous numerical targets related with clearly defined group of metrics,
- Improvements for process can be identified with the help of metrics,

- Project should be able to follow correct procedure of DMAIC and utilize its various tools,
- Project should promote confidence for a team with the improvements,
- The benefits to a project shall be achievable in plausible time, generally between the periods of three to six months,
- Project should bring various business improvements, and
- Project should be considered with a perspective of improving the process, not just to acknowledge results based problem.

The project should then be inspected for efficient implementation of DMAIC as defined by Six Sigma philosophy. Defining the scope of the project is an important subject before beginning to work on it using DMAIC. A good scope involves a project that is able to complete in a plausible time of about six months. It should not make individuals involved frustrated with a wide variety of areas, data, or too much of required time [23]. DMAIC is a data-driven iterative process composed of five different phases, each involving utilizing tools among various tools belonging to the phases. These phases are applied step by step with the logical order and together form a backbone of any process optimization project. These phases allow an organization to effectively identify scope for the problems, evaluate the present performance, analyze the fundamental causes behind the troubles, and assess the recommendation actions provided for improvement and apply the changes made for long term improvements [23]. Various phases of DMAIC are discussed below with respect to the philosophies of Lean Six Sigma:

2.2.1.1 Define phase

The essential part of this phase is to define scope and limits of a project, analyze the needs of customers, and define the project targets. Although, before going through these components of define stage, a relevant team has to be formed [24]. It involves creating a project charter. A problem statement has to be defined, with confirmation from a team for the existence of problems in the current process. The problem statement should include severity, effect, and group involved. Based on the problem statement, a goal statement is derived. The team begins with a high-level view of a process, using often a tool called SIPOC. After carrying a high-level view, a detailed analysis of process is made by the team, generally by using simple flow-chart. Customer definition has to be derived with evaluation of their needs. Various other tools can be used as a part of Define phase in DMAIC, such as 'project charter', 'voice of the customer', 'value stream map' as well[25]. Some of the important tools used in this phase are discussed in detail later. Various questions shall be precisely answered in this phase such as [26], [27], [28]:
- What is the precise definition of the problem?
- What is the target of the organization?
- What is the boundary of a project? What is covered in the capacity of a team?
- What are the various resources available to a team, for utilization in order to achieve the target?
- Who are various customers associated with a project?
- What are the requirements of the customers with respect to quantitative terms?

2.2.1.1.1 SIPOC

SIPOC stands for suppliers, inputs, process, outputs, and customers. When dealing with process optimization or improvement, it provides a high-level overview of the scope of a process. Within a particular organization, a process then can be easily examined about various areas involved in it, before drawing detailed flow charts for interested parts. It involves defining the name of a project, after which outputs are identified. For each output, there exists a customer for a process, which is considered and defined. After it, inputs of a process are identified, and supplier to provide the inputs are listed. In the end, the sub-processes are written down, which together convert the input to output and give a brief overview to further create a detailed flow chart of interested process elements for optimization. An example of SIPOC is listed below [29]:

Process: Recruit Staff				
Suppliers	**Inputs**	**Process**	**Outputs**	**Customers**
Line Manager	Request to fill a vacancy	1. Specify needs 2. Authorise recruitment 3. Place adverts 4. Assess applicants 5. Offer appointment 6. Confirm start	New member of staff	Line Manager

Figure 13: Example of SIPOC for recruiting staff [29]

2.2.1.1.2 Flow Chart

As processes in an organization become more and more complex, changes made to them makes them less productive. This is because such changes generally lack efficient view and understanding of processes as a whole. Thus, at some point, they are required to be viewed from wider perspective in order to undergo improvement. The flowcharts are one such powerful tool. It picturizes various steps involved in the process in particular order and represents various activities and decisions performed for them. Common symbols used to design flowchart are shown as [30]:

Steps in a Process   Decisions or Alternatives   Direction of the flow

Figure 14: Common symbols used for creating flowchart [30]

Creating flowcharts enables individuals to study various steps involved in the process visually and more clearly with ease, which help to evaluate various problems associated with them, by enabling people for making discussion on a process in a group in an efficient manner. Thus, it can help to compose an ideal process, evaluate various unreasonable

steps in the process, determine various bottlenecks, identify duplication of activities, and helps determine various possibilities for improving the process. There exist various types of flowcharts, such as sequential flowchart, top-down flowchart, deployment chart, hybrid flowchart etc. [30]

2.2.1.2 Measure phase

The measure phase involves generating plausible metrics in order to monitor the advancement toward the targets of a project. The targets are basically reducing defects in the process. It consists of assigning numeric values to the elements under consideration according to a certain set of rules, which enables a reasonable evaluation of improvement [24]. After the project problems are defined in the first phase, this phase includes identifying various work processes associated with the problem. Various data about the performance of the processes are analyzed in this phase. It is often a situation in the organizations, that they do not collect the past data related to the performance of their processes. In such cases, metrics can be generated with the help of experience of individuals involved in the process. These performance figures are then used for obtaining best performance in the process [31]. Various tools which can be used in this phase, based on various situations, are listed below [26], [32]:

- To illustrate various attributes related to the process: Process mapping, lead time efficiency.
- To concentrate on require topics and arranging them in accordance with their importance: Pareto charts, cause and effect matrix, FMEA.
- To establish and structure the ideas: Brainstorming, multi-voting, nominal group technique.
- To compile data and make sure they are correct: check sheets, measurement accuracy.
- Acknowledge and remove variations: Run charts, process capability.

Various questions shall be precisely answered in this phase such as [27], [26], [28]:

- How good is an ongoing process operating for the organization?
- What are the various data we need to collect in order to anticipate the process capabilities?
- Is the process being statistically controlled?
- What are the various process inputs which cause serious impact on the output of the process?

FMEA can be used in the improvement phase as well, and thus is discussed later. Brainstorming is discussed below.

2.2.1.2.1 Brainstorming

There are various tools which are helpful to be used when utilizing FMEA tool for the process analysis, such as brainstorming, flow chart, root cause analysis, Ishikawa root cause analysis diagram etc. Ishikawa serves to be a good tool, to begin with, to analyze a particular process, for which brainstorming serves to be an effective technique to seek initial causes [33]. Topics of brainstorming are discussed by creating questions, which are the important factor for the brainstorming session. A question which is not too precise, nor

too vague, provide expansive thoughts and effective solutions to a team. Questions shall be formulated in a positive manner. This is because the negative ways would degrade the confidence of members and they can face blame games. The members participating should be chosen suitably, as too many members might lead to more ideas but can cause a chaos. Hence, a good balance shall be followed. There are various types of brainstorming techniques available such as silent brainstorming, brainwriting, reverse brainstorming, brainstorming large groups, autumn leaves, team idea mapping, question brainstorming etc. The choice of which type to choose is generally carried out based on various situations [34]. Brainstorming is a widely used tool to produce a large amount of ideas, options or possibilities related to a topic in a short period of time. An efficient brainstorming session involves no criticism and promotes creativity among team members with equalized involvement of each participant. It is a highly versatile tool which requires very little preparation and equipment. It is a very simple tool and involves following rules [35]:

1. Active Involvement of all participants. Silly ideas are welcomed, and no criticism, jokes, or inappropriate comments are allowed,

2. Build up on each other's idea,

3. A limited time is provided for the process, and ideas are shared and displayed to all of the participants,

4. Collected ideas are well clarified so that each member involved understand each point correctly. It is to be noted that inappropriate judgment about an idea is not appreciated.

5. If there are similar ideas, they should be combined and considered as a single entity.

Brainstorming possesses possibilities to come up with innovative ideas, and collects thoughts from past experience, in a relaxed atmosphere. It provides a simple and fast way to involve participant in generating ideas. The final ideas are established with the agreement of everyone. It is beneficial to organize the above technique with individuals of varying background, however, it should be taken care that it does not transform into just a chatting activity. Besides these points, one should make sure that participants do not start to become competitive with each other. Brainstorming is basically used for problem-solving, generating innovative ideas, and identifying a wide variety of solutions when the problem is wide [34].

2.2.1.3 Analyze phase

The main goal of this phase is to determine, validate, and choose essential root causes in the processes, in order to remove them to make improvements in them. A number of root causes can be found for each failure mode in the process by utilizing root cause analysis tools. These root causes can undergo metric evaluation and are evaluated until valid root causes are identified. Prioritizing the root causes is further carried out to recommend improvements in later phase [36]. Prioritizing is significant to be carried out, as often due to inefficient management, the top causes influencing the process, are buried under various root causes found for each cause, in this phase. If various root causes are not found, it simply signifies that team is unable to carry out a detailed analysis of the process. However, root causes should be found with respect to the available time for a project, in order to obtain efficient results for a team success [27]. Brainstorming can be an essential tool in evaluating root causes. These root causes have to be verified before moving on to the next phase, as they form the basis of whole process analysis. After a deep investigation

of a process, large data shall be visually arranged using various diagrams or charts. This assist team to easily handle large data in order to communicate issues related to process as a group or with various individuals involved [25].

Various questions shall be precisely answered in this phase such as [27], [26], [28]:

- What are various patterns exhibited by concerned outputs of the process?
- What is the information generated by evaluating inputs and outputs of the process?
- Using cause and effect diagram such an Ishikawa/Fishbone diagram, what are the various relationship between inputs and outputs?
- What are the detailed causes behind inputs which may have a huge impact on the process?
- What are the essential detailed inputs which shall be concentrated for further evaluation for process optimization?

Several tools can be used as a part of the analysis phase of DMAIC. However, not all of these tools have to be used. Following is the list of several common tools used in this phase [32], [26]:

- Pareto chart, where various causes are compared in a graph with percent weightage and bars.
- Fishbone diagram, also known as Ishikawa diagram for well-structured root cause analysis visualization.
- 5 Whys, for searching series of causes.
- Regression analysis, for visually analyzing the data with a graph.
- Time series plots, to check relation of various variables with time.
- Histograms, to visually see the significance of various variables.
- Tree diagram, to see the relationship and various root causes for a process.
- FMEA, to carry out a detailed analysis of root causes of the process.

2.2.1.3.1 Ishikawa/Fishbone diagram: cause and effect diagram

The European Organization of Quality defines the use of above tools in case of data acquisitions and data analysis [37]. As shown, histograms, control charts, and check sheets fits better for data acquisition, while scatter-plot, flowcharts, Pareto diagrams, and Ishikawa/Fishbone diagrams suit better to data analysis. Ishikawa diagram is a tool to evaluate various dispersions in the processes. It is used along with brainstorming technique to determine various root causes for the process problems and their various issues under consideration. It places the main causes found in a well-structured view in the outer area of the fish skeleton, with the fish head representing an overall problem. The various root causes evaluated with 5 Whys are summed up and related together in a fish skeleton, which assists teams in identifying causes, their relations, and their effects, which are written down over bones. These effects are spread into different branches as more causes are found by 5 Whys. The related causes are listed and related with further branches of skeleton bones. A typical form of the Fishbone diagram looks as shown in figure 16.
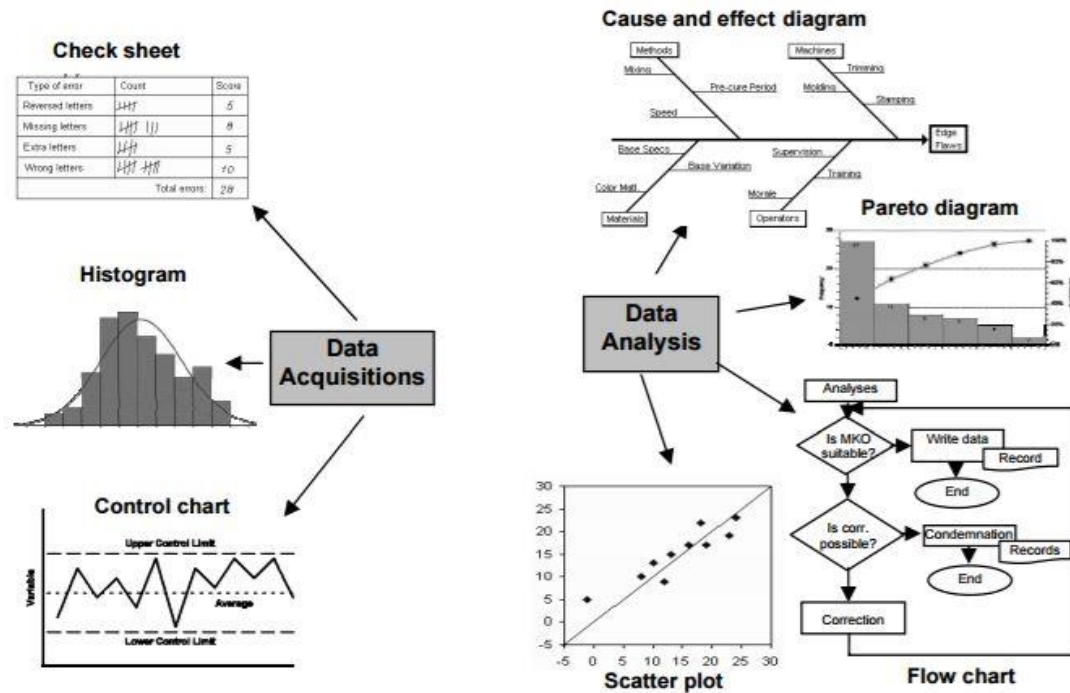
Figure 15: Cause and effect tool used in 7Q's methodology [37]
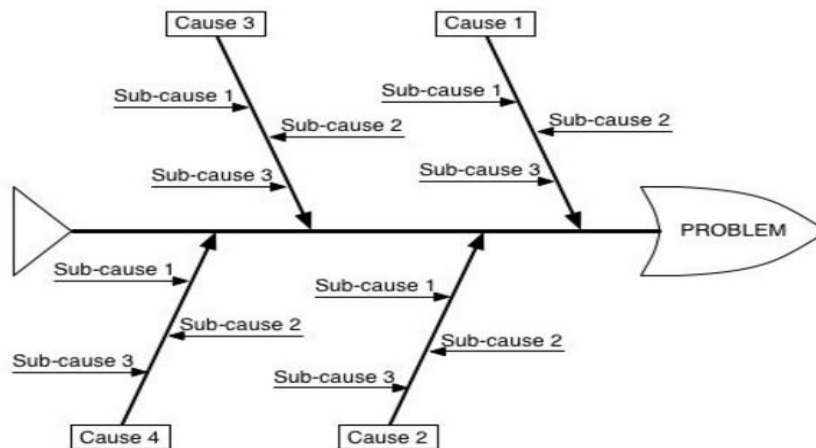
Figure 16: Fishbone/Ishikawa Diagram [38]

The answer to when and why to use Ishikawa root cause analysis diagram is given below [39]:

- To profoundly evaluate the process and prioritize its various elements to decide on which parts to concentrate on, in order to optimize the quality.
- To communicate with team, wide range of elements in a single pictorial view, and the various complexities of a process, as a whole.
- To generate, verify, and relate various issues associated with the problem, and later use the data for various other quality improvement tools like Pareto Chart, FMEA etc.

The above-discussed tools like 5 Whys, Brainstorming, helps to prepare Ishikawa diagram, which all together form root cause analysis. These are used together as shown in figure 17.
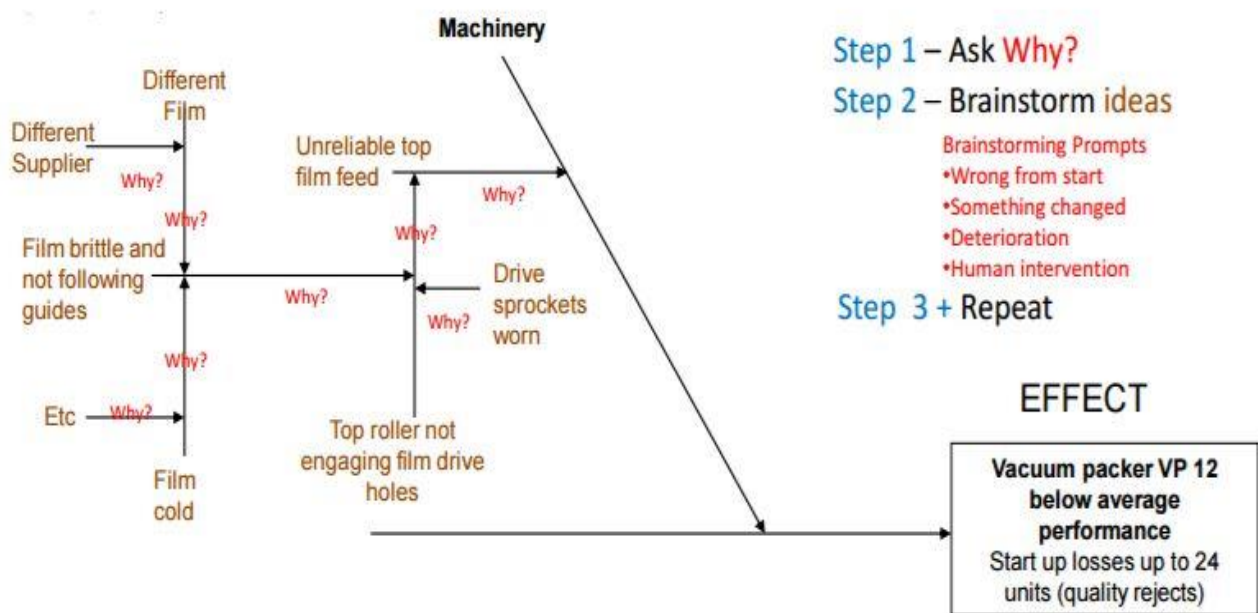
Figure 17: Example of 5 Whys analysis and Brainstorming used together for root cause analysis [40]

2.2.1.4 Improve phase

The fourth phase of DMAIC methodology deals with the determining and realizing the solution in some particular section of a process or in all as aggregate. The determined solution should be able to fix or prevent the defect, by eliminating the root causes of the problem. For example, solving problems from the various root causes listed in the Ishikawa diagram. Various simple approaches such as brainstorming, wearing six hats etc. can be used in this phase to seek plausible and satisfactory solutions. The choice using appropriate tool can be made depending on the situation. For example, in some cases, complicated techniques like Design of experiment is used. The solution shall be simple, innovative and reasonable. The solutions should assist to forecast for any possible unnecessary risks by making use of tools like FMEA [37]. After seeking the solutions based on their practicality, process maps shall be developed. These maps can assist individuals involved working with the process, to learn new methods of working. To verify that chosen solutions or decisions made are genuine, mini plan-do-check-act cycle (PDCA) can be implemented in order to evaluate the opinions of all the stakeholder of the process, which would help them to choose the finest solution [25]. Improve phase includes following aspects for process [28]:

1. Suitable documentation explaining how the problem solutions were collected.
2. Collected alternative improvement actions considered for recommendation.
3. Complete results obtained by pilot tests, i.e. data displays, analysis, experiments, and simulation analyses.
4. The explanation for how improvement plan would be executed on a full scale in company considering various issues such personal concerns, training, legal elements etc.
5. Evaluation of risks concerned with applying improvements in process, and how they can be handled.

There are many tools that can be used in this phase, for example:

2.2.1.4.1 Poka-Yoke

Poka-Yoke is a Japanese term coined by Toyota Motor Corporation Engineer Shigeo Shingo in 1961, for mistake (Poka) proofing (yoke). Any mechanism which eliminates root causes behind the mistakes causing errors or defects in the process to improve the quality is called Poka-Yoke. It enables users to avoid mistakes by correcting, preventing or notifying them. It purposes various recommendations which protect user to make mistakes in an environment they are present. This is achieved either by removing the process activities or tasks that cause mistake or by notifying a user of any error found. It can be used for any kind of process from production to service industry. It can be used for various types of situations, carrying any type of errors. Various errors such as following, considering the manufacturing process, can be overcome by using Poka-Yoke [41]:

- Individual did not follow or missed some elements for the correct procedure of process,
- Individual uses incorrect tools or machine settings,
- Individual mistakenly did not add all the parts or used incorrect parts for assembly or welding etc.

Following is an appropriate procedure, to utilize Poka-Yoke techniques for a process [41]:

- Evaluate the process,
- Evaluate the failure mode for the process and by using the techniques like root causes analysis,
- Determine appropriate Poka-Yoke strategy. These strategies can be complicated techniques as well,
- Try the strategy for the process to identify its effectiveness and train the individuals involved in the process.

There are basically two types of Poka-Yoke (software testing analysis and review conference paper) [42]:

- Prevention category: It is any mechanism in a process that makes it impossible to produce an error by an individual by avoiding the mistake.
- Detection category: It deals with notifying an individual for a mistake made. In this case, mistakes are made, notified by some mechanism, and correction is made later.

Although these techniques have been mostly used in the manufacturing world, It can be used in software development. It can be utilized as a part of producing error-free software, as it helps to avoid mistakes earlier in the process [42]. Various example of Poka-Yoke can be as follow [43]:

- USBs can only be used unidirectional. In order to properly insert USB in a right direction, USBs are designed with blockage stones on one side. This is a preventive type of Poka-Yoke.
- For providing road safety, cars are provided with engine software that lessens the power in unusual situations. For example, a driver may mistakenly press brake pedal along with accelerator or gives redundant power with respect to the road surface grip etc. This is a prevention type of Poka-Yoke.

- Motorcycles are provided with lights that turn on if motorcycle stand is still on while driving. This is a detection type of Poka-Yoke.
- A seatbelt sound system in cars produces sound to let car occupants know if they are not wearing the seat belts. This is another detection type of Poka-Yoke.

2.2.1.4.2 Failure mode and effect analysis (FMEA)

Murphy's Law states that "Anything that can fail, shall fail". The question arises that what would fail? What would be its impact? What could be done to overcome it? To overcome such issues in a process or design, FMEA methodology provides a well-structured approach toward quality improvement for a process, in a quantitative manner. FMEA shall be performed at the initial phase of design processes, and then should be carried along the other phases of a process. Just like DMAIC methodology, FMEA tool requires cooperation and agreement of various individuals involved in the different analysis of processes and should be carried out by specialists from each segment of the process. The value of FMEA is further optimized by using various other quality improvement tools along with it [44].

SAE J-1739 defines FMEA as: "An FMEA can be described as a systematic group of activities intended to [45], [46]:

a) Recognize and evaluate the potential failure of a product/process, and the effects of that failure,

b) Identify actions that could eliminate or reduce the chance of the potential failure occurrence, and

c) Document the process.

It is complementary to the process of defining what a design or process must do to satisfy the customer."

SAE developed a standardized practice of FMEA for the automotive industry. It was developed mutually with automotive giants: DaimlerChrysler Corporation, Ford Motor Company, and General Motor Corporation. The guidelines to apply FMEA in a company are provided as recommendations, which a company can mold to a suitable form to fit their specific user cases. In order to achieve the maximum amount of benefits from FMEA, it is recommended to be applied before a new process or when a design is implemented for an organization, as later changes would be costly [45], [46].

FMEA usage can be found in a wide range of activities. It was first used during the 1940s by the military, by Apollo mission in 1960s, and in the automotive industry in 1980s in the USA. There exist various standards of FMEA such as [47]:

- SAE J1739: Automotive industry based standard,
- IEC 60812: Electrotechnical commission based standard, and
- MIL-STD-1629A: Military based standard etc.

FMEA can be utilized on various occasions in an organization such as [47]:

- To find and counter the safety risks,
- To optimize the product performance by reducing defects,
- To optimize various verification methods and process controlling,
- While dealing with changes in current process or design,

- To evaluate important attributes for product or process, and
- To establish various maintenance plans etc.

The general overview including various important aspects and deliverables for carrying out FMEA in an effective manner is shown below [47]:
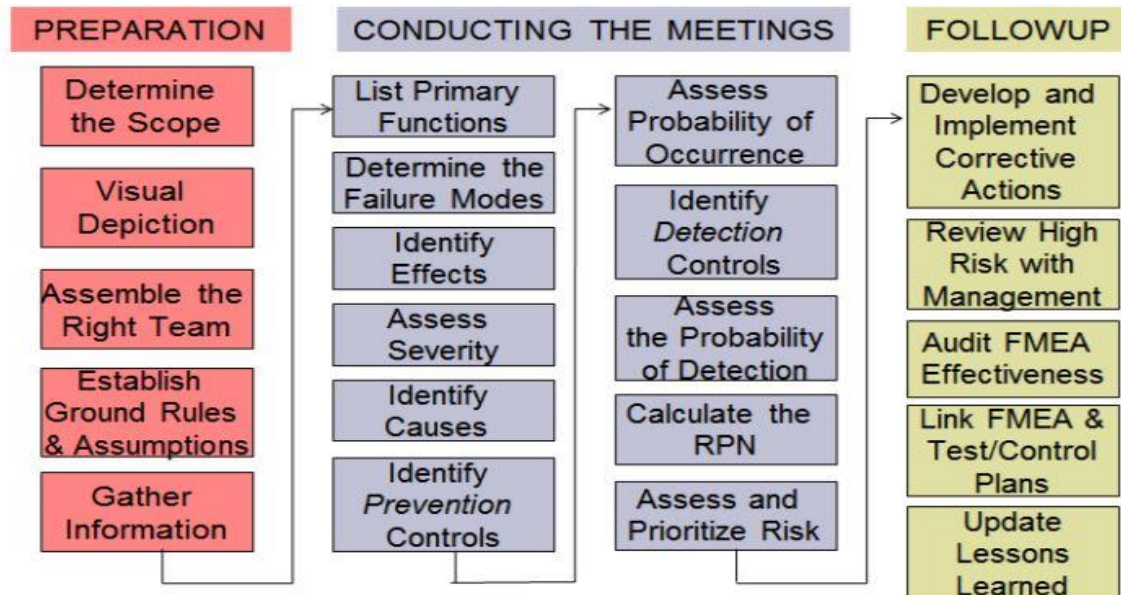
Figure 18: Roadmap to conduct FMEA analysis [47]

There is basically three type of FMEA according to SAE standard. The standard is divided into three sections based on the application of field, as follow [45], [46]

- Design: Focuses on various defect mode and effect analysis in product designs.
- Process: Focuses on various defect mode and effect analysis in processes.
- Machinery: Focuses on various defect mode and effect analysis in manufacturing machinery.

In an OEM, these FMEA types can be related to one another in a manner shown in figure 19, as mentioned by SAE standard. There are various steps defined by SAE standard for conducting effective and efficient FMEA analysis [45], [46]:

1. Various information regarding the project are filled inside an FMEA form such as FMEA number, who prepared the FMEA, team members, item (in which FMEA is carried out), various dates etc.
2. Process function: The process on which FMEA is carried out, is divided into various categories called process functions. A concise description of them is added to the columns.
3. Potential failure mode: A concise description is added to FMEA form, which describes how a process can go wrong, to meet the desired requirements of a process with respect to a process function. Various assumptions are made, for example, incoming material to this part of a process was correct. To identify failure modes, a process can be broken into sub-processes, or customer feedback can be used etc.
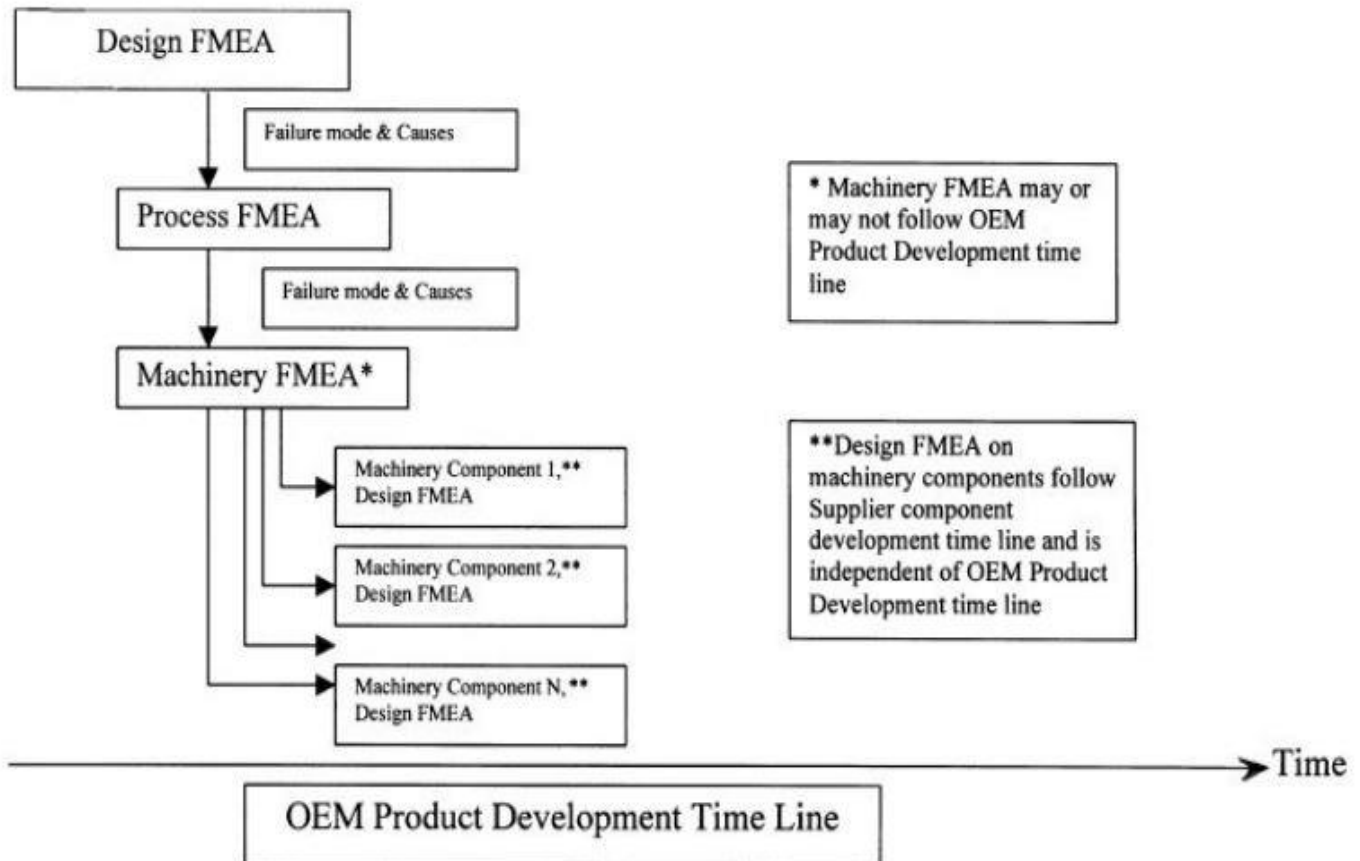
Figure 19: Relationship between different types of FMEAs in product development [45]

4. Potential effect of failure: Effects are defined, explaining how the failure mode identified could affect the customer. The customer can be an internal individual in the company or final end user of the product or process.

5. Severity: It is a rank related to the most rigor influence of a failure mode effect. It is, generally, a scale between 1 and 10. There is no standard ranking scale developed, and ranking scale for the effect can be derived by the team involved for a specific process based on various situations. If an analysis of severity cannot be assessed by the team involved, the concerned person shall be contacted.

6. Classification: Classification column in FMEA table can be used to mention any special process attribute that may require special process control. This column can be left empty if no such classifications are required.

7. Potential cause/mechanism of failure: This includes a description of various mechanisms or cause, through which failure occurs in a process. The description should be provided in the suitable terms, so that recommendation actions can be defined for them. Various causes associated can be listed, including root causes. The causes which solve the failure mode in terms of description of recommendation actions fits better the description for this column than others.
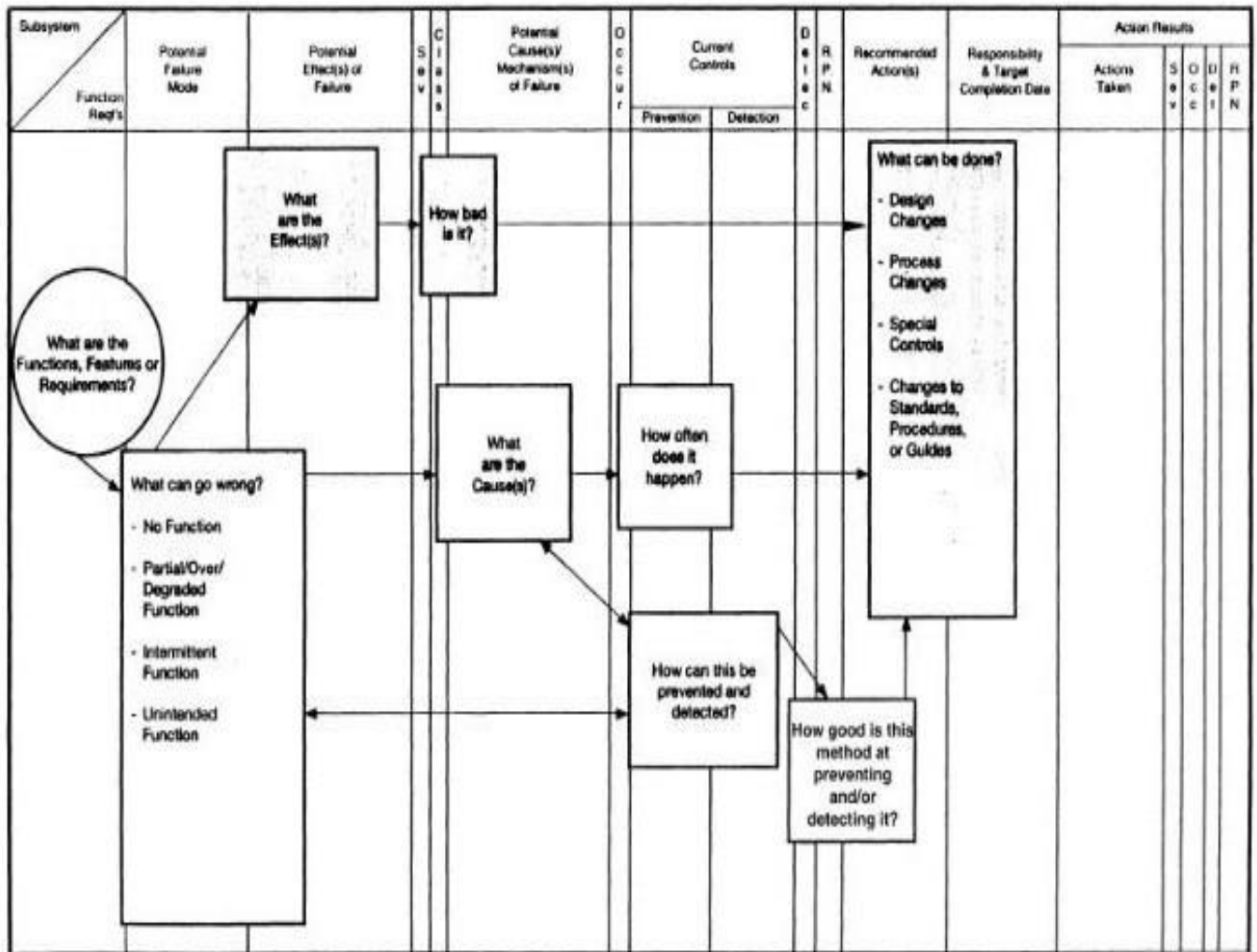
Figure 20: Preparation of FMEA table with respect to the FMEA procedure [45]

8. Occurrence: Occurrence describes the amount of probability of failure to occur in a process. It can be defined within a scale of 1 to 10, developed by the team specific to the process. The rating can be relative and may not reflect the actual occurrence of a failure. The goal of recommendation action is to reduce the occurrence ranking in the process. Statistical data can be used to assess the occurrence of a failure. If statistical data is not available, the team can evaluate rank by subjective assessment or historical data of similar process available.

9. Current process controls: Process controls describes the current approach toward controlling the errors or defects in the current process. The solutions can be either preventive or detection type, just like Poka-Yoke techniques described before. The current controls, thus, can be distinguished in FMEA table between these two error proofing types, in separate columns. If single columns are used in the table for process control, prefixes P & D can be assigned to each process control listed. The current process controls analysis can result in revision for occurrence ranking decided in the previous step. Thus, occurrence ranking and current process control for a process are related to each other.

10. Detection: Errors or defects ranking have to be defined based on the best process control available for the failure mode. It is to be noted that assumptions like

detection is low because the occurrence of the failure is low, are not valid. Consider the strength of current process to identify low-frequency modes of failure, or how it can prevent them to propagate further into the process. The ranking criteria can be defined by the team based on various situations. Lower the detection ranking, higher is the strength of current process to detect the error or defect for the process.

11. Risk Priority Number (RPN): This parameter is a result of multiplication of numerical values of occurrence of a defect, the severity of the defect, and detection that defect would go undetected. Similar to Pareto diagram, RPN helps the team to identify or order by significance, various concerned failure modes for a process. High RPN values signify greater concern for the team to work on it. However, low RPN value does not mean that the failure mode can be ignored. This is because, if the value of severity is high regardless of low RPN value, the team have to consider the process failure mode for the reliability of the process.

12. Recommendation action: Recommendation action shall first consider process failure modes with higher severity value, then higher RPN value, and then the other failure modes which members of the team decide to work on regardless of RPN value. High severity ranking deserves the highest attention of the team, after which, the priority shifts to other failure modes based on occurrence, and then on detection. An effective recommendation action requires process design or revision to reduce severity, occurrence, and detection of defects for a process. Defect detection type solutions can lead to increase in the cost of a process, as they first let a failure mode to happen. Due to which wherever possible, prevention type of solutions get more priority.

13. Responsibility: Name of a person responsible for carrying out recommendation action is added to this column with a date for completion of goals.

14. Actions taken: After the actions are applied to the process, a concise explanation is added for the action taken with the date of execution.

15. Revised ratings: After applying recommendation actions to the process, RPN ratings are re-evaluated for it. It is recommended to take improvement actions in order to determine the effectiveness of improvements. However, if no actions are taken, the column can be left blank. If in case, further actions are considered to be taken, the ranking can be re-evaluated in order to carry on the cycle of continuous improvement for the process.

2.2.1.5 Control

The function of this phase is to continue the benefits generated by carrying out various phases of DMAIC. The process is monitored continuously in order to attain sustainable and continual benefits, by creating a plan to control it. The records are carried out by a certain plan. A well-structured plan is prepared in order to monitor process performance for any deviation from the target by various control methods [48]. Following issues related to the process, shall be answered by this phase [31], [28]:

- Are recommendation actions bringing process improvements approved?
- How shall process be monitored in order to continue the effectiveness of recommendation actions?
- Who are the individuals responsible for the solutions for a process?
- What are the final results of carrying out DMAIC, around which future process deviations would be measured?

- What were other potential projects or areas found while carrying out DMAIC project for process improvement?

Basically, this phase involves sustaining the improvements by documenting new methods for a process, training employees with it, observe the new method, and figuring out the lessons learned. Many processes suffer insufficient effectiveness of DMAIC due to ignorance toward controls phase. In many cases, employees become too excited about improvement actions and ignore carrying out proper control activities toward a process. Processes can be easily influenced by certain environmental factors, inherited process variations, machines issues etc., and thus it is important to carry out control activities efficiently after improvements to a process. Common tools used in this phase are [26]:

- Control plan,
- FMEA,
- Control charts, and
- Error proofing.

2.2.1.6 DMAIC comparison to other similar methods

There exist various other methods which are similar to DMAIC, however, with different steps in execution. These are shown in figure 21.

It can be quite confusing to decide which methodologies to choose for a particular situation. This basically depends on the size of a problem, that is, if it is small, medium, or large scale, and if the purpose is a continuous improvement or a single problem. PDCA shown above is applied to medium sized problems, with a high concentration of plan phase. A3 method is short to medium based approach for problem-solving used in Lean management. DMAIC is used for large scale problems, exceeding more than 3 months, involving a large amount of data. It is basically a statistical approach but does not have to be always statistical. 8D is a customer oriented approach with fast reactions to customer complaints, in which first three steps shown in figure 21, are executed within three days in general [49].

PDCA	DMAIC	A3	8D/PSP
Plan	Define	Clarify the Problem	1. Create Team & collect Information
	Measure	Break down the Problem	2. Describe the Problem
		Set a Target	3. Define Containment Actions
	Analyse	Analyze the Root Cause	4. Analyze the Root Cause
		Develope Countermeasures	5. Define possible corrective Actions
Do	Improve	See Countermeasures	6. Implement corrective Actions
Check	Control	Evaluate Results & Processes	7. Define Actions to avoid Recurrence
Act		Standardize Success	8. Congratulate your Team

Figure 21: Other methods similar to DMAIC [49]

Chapter 3: ECU software calibration process and its various solutions/tools

3.1 Introduction

ECU software calibration is a complex process. It is important to understand how the process is executed in order to understand its complexities and management. The parameterization for vehicle's ECU software is executed in such a way that by manipulating or by calibrating various function algorithms' characteristic values, the various functions, belonging to control and diagnostics, can be conveniently used for different types of variants of system and various vehicle models, all without demanding any modifications for calculation routines. By utilizing various calibration tools, the calibration engineers are able to tune these characteristic values (also called labels) of various function algorithms. Calibration engineers carry out above process being on a fly mode, meanwhile concurrently, are able to collect ECU's signals including signals from various measuring devices and vehicle buses. Utilizing the laboratory model, one can only conclude on the basis of conditions, the various parameters (labels) belong to control algorithms. The acquisition and optimization on a test bench or driving trial for various labels/parameter values such as characteristic maps, curves or values belonging to various functional algorithms, composed securely in ECU software, is called ECU calibration [50]. It is a challenging task involving meeting various vehicle conflicting issues, such as minimizing NO2 vs CO2 emissions. Calibration processes become quite complex with calibration to be made in different vehicle models, and their further variants planned to be sold in different world markets with different requirements and standards. The complex process of calibration is generally divided among OEM, ECU supplier and R&D company. Calibration process involves dealing with thousands of ECU software parameters/labels, a number of interaction among various functions of ECU software, considering different ECUs. The whole process is iterative in nature, at various levels of details [51]. In an entire process, various tasks are performed as shown below:
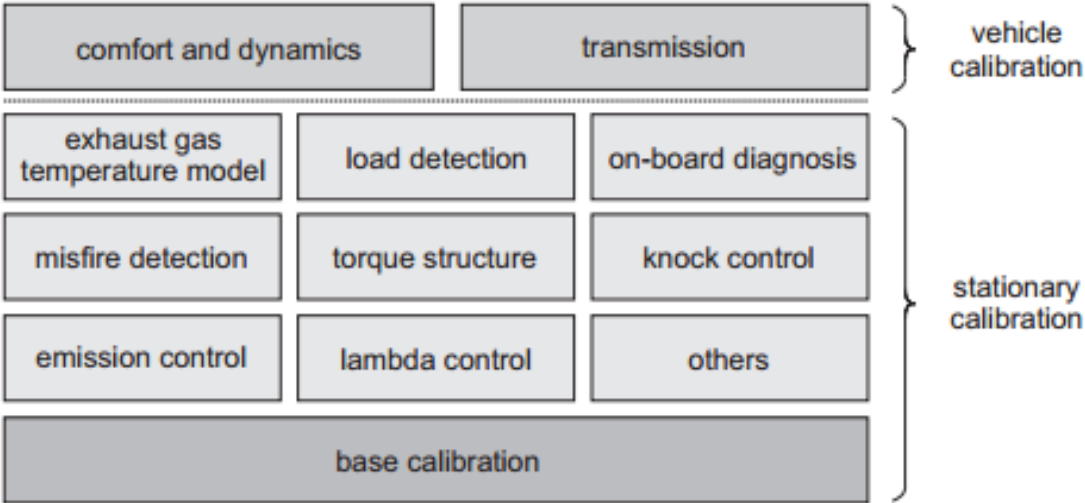
Figure 22: Example of various activities involved in calibration process [52]

ECU vehicle calibration involves various activities like calibrating different comfort and dynamics functions or calibrating for transmission. Stationary calibration is composed of a variety of activities as well. It involves handling various fundamental ECU functions such as torque structure and deals with the functions which are related to safety and clean performance of the engine. Necessary as a pre-condition, the base calibration is a most essential element for all calibration tasks. It involves specifying basic adjustment parameters. For example, defining and storing in the ECU maps, main injection time or injection pressure. Acquiring smooth engine maps for low dynamics parameters is another goal of calibration tasks [51].
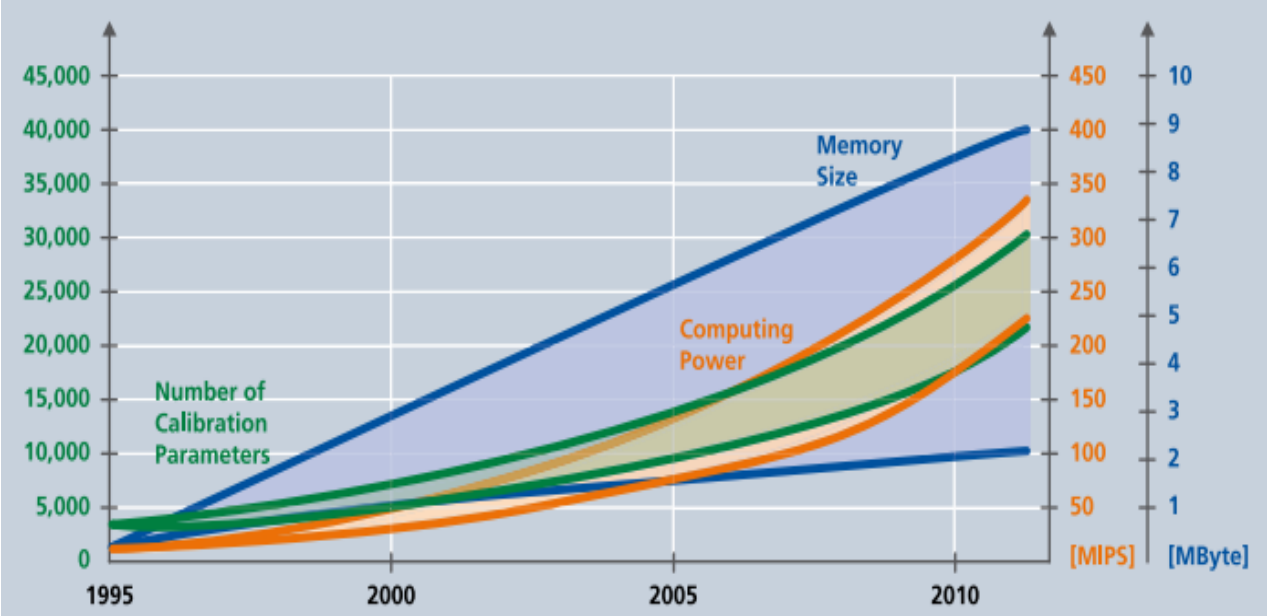
Figure 23: Advancement in memory size, calibration parameters (labels), and calculation power of control units of engine since 1995 [51]

3.2 ECU calibration approaches

For the ECU calibration process, two main approaches have been developed [53]:
1. Measurement based approach: This is simple and easy to use technique utilized during low-dimensional parameters space, i.e. when only few amount of parameters need to be optimized, such as 2 to 3 parameters. It requires no specialized engineers and results are conceived easily and rapidly. However, due to the concept of dimensionality, this approach flounder for optimizing the vast amount of parameters, which are common in vehicles these days. These days, this approach finds its application in some special situations only. This research work is related to model-based calibration process. However, for the sake of completeness of knowledge of calibration process, two different types of measurement based approaches can be briefly discussed.

There are two types of measurement-based approach. One is 'full factorial based optimization', which involves dense grid discretization of parameter/label space, and for each point, measurements are made. The second type of measurement based approach is, 'one factor at a time'. This involves working on one parameter/label at a time, meanwhile, other parameter values are kept constant. Figure 24, below shows an example of the measurement based approach of parameter/label x1 and x2. In this example, while working on parameter x1, x2 is kept constant, and while working for x2, x1 is made constant, after it is optimized. Thus, this is one-dimensional based optimization process.

In large ECU software with a large number of parameters with relationships with each other, this one-dimensional based approach become inappropriate to be used.

(a) Full Factorial Optimization   (b) One Factor at a Time
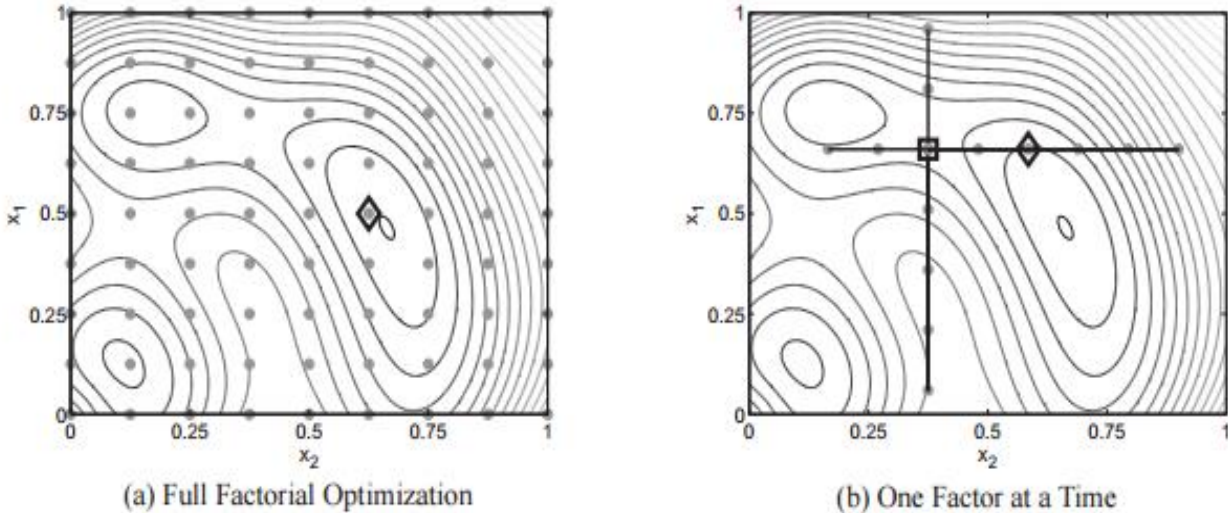
Figure 24: Types of measurement based approach [53]

2. Model-based approach: In this approach, a black or gray box models are formed by utilizing test benches for measurements. There are mainly two types of model-based approaches for calibration. Model-based offline optimization is one such approach. It possesses strong distinction between test bench measurements and modeling on PC. Following diagram shows the concept of model-based offline optimization [53]:
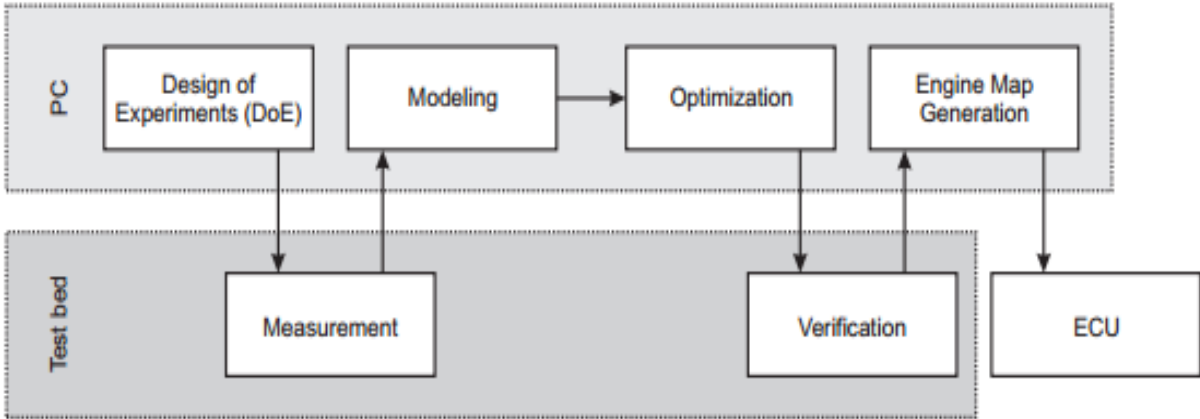
Figure 25: Model-based offline calibration optimization process [53]

Firstly, by using techniques like DoE, planning of experimental design is carried out. After the test bench measurements are acquired, parameter modeling is carried out by calibration engineers. After achieving optimized values for parameters/labels, the values undergo verification in the testbed. After ensuring correct verification of values generated, engine operation maps are released and saved in ECU. The main disadvantage of this techniques is that calibration engineers are unable to check model quality while carrying out measurements in the test bench. This can be overcome by another type of approach called 'model-based online optimization'. Contrary to offline optimization, it involves no strong distinction in measurement on test beds, modeling, and optimization. This leads to calibration process with the constant interaction of modeling and optimization algorithms with a test bench. This approach give calibration engineers an advantage to know when the optimal amount of measurements are made on the test bench, and which segment of

measurement need to be concentrated, for maximizing the information acquisition, providing efficient use of test bench by saving time for use. The following figure shows how this approach is carried out:
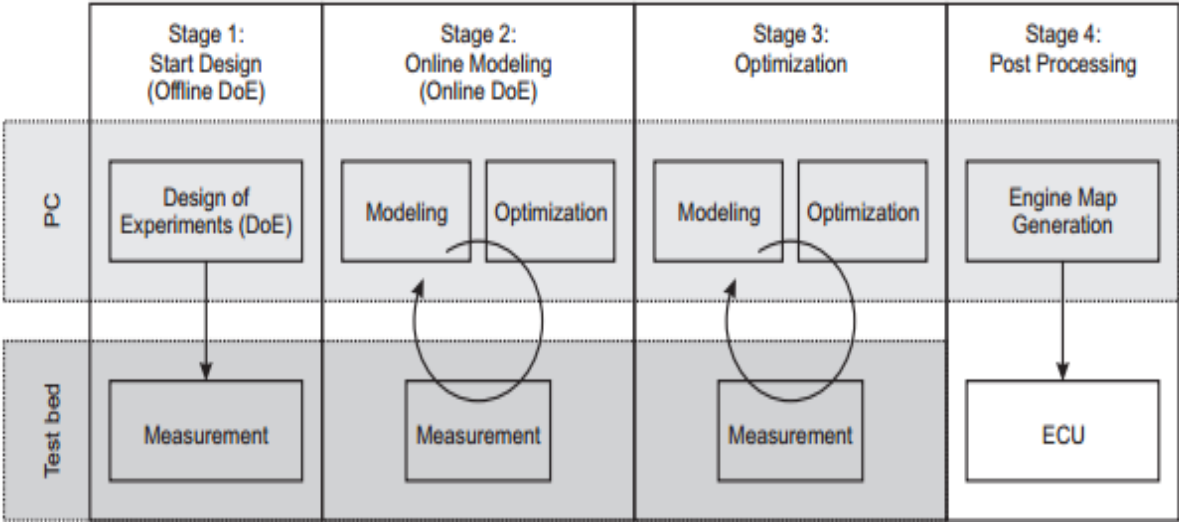
Figure 26: Model bases online approach [53]

This approach involves various phases of calibration. It begins with offline kind of optimization, where starting design is prepared and measurements are made on the test bench. These are further optimized in later stages or phases, with complete connection with a test bench, in distinct operation points, and errors are reduced further. In the last stage, engine maps are released and updated to ECU [53].

3. Static and dynamic modeling based calibration: There exist two form of calibration called static and dynamic modeling. Most of the calibration functions are related to static calibration, as there is no possibility to calibrate them in a transient state of a vehicle, while dynamic calibration provides possibilities for dynamic measurements, and various emission functions which are dependent on the transient state of a vehicle can be better analyzed [53].

3.3 Calibration process in industry: AVL List

Due to the increase in strict emission legislature laws, large variances in vehicles created a challenge for calibration validation with conventional methods, which would cause a shortage of manpower, and test equipment in these situations. The use of approaches like V-model for the development process had put forward a requirement of a model-based approach for AVL. This is because in such development process, a lot of test activities take place along with development process, many in the end of the project, and with frontloading in earlier stages. The model-based approach involves firm, a real time-based semi-physical model of the engine along with powertrain model and virtual testbed combo, with various automated solutions or tools [54].

The model-based approach is divided into various phases, such as concept phase, which involves utilization of model-in-the-loop environment for simulating performance and emission behaviors with real driving conditions. This enables calibration project to acquire emission or performance pre-calibration to be set more precisely to requirements before first engine start, by setting systems such as emission after treatment. The test bed calibration and validation for hardware-in-the-loop calibration are made in the application phase. The whole process is very well monitored constantly using dataset management

systems, with efficient exchange of ECU software datasets and various other data, providing required dataset quality for the calibration process. For the first engine start, pre-calibration can be obtained by utilizing calibration data for earlier development phases including using semi-physical modeling approach, based on components features and performance targets, which is later developed to empirical model using DoE data [54].
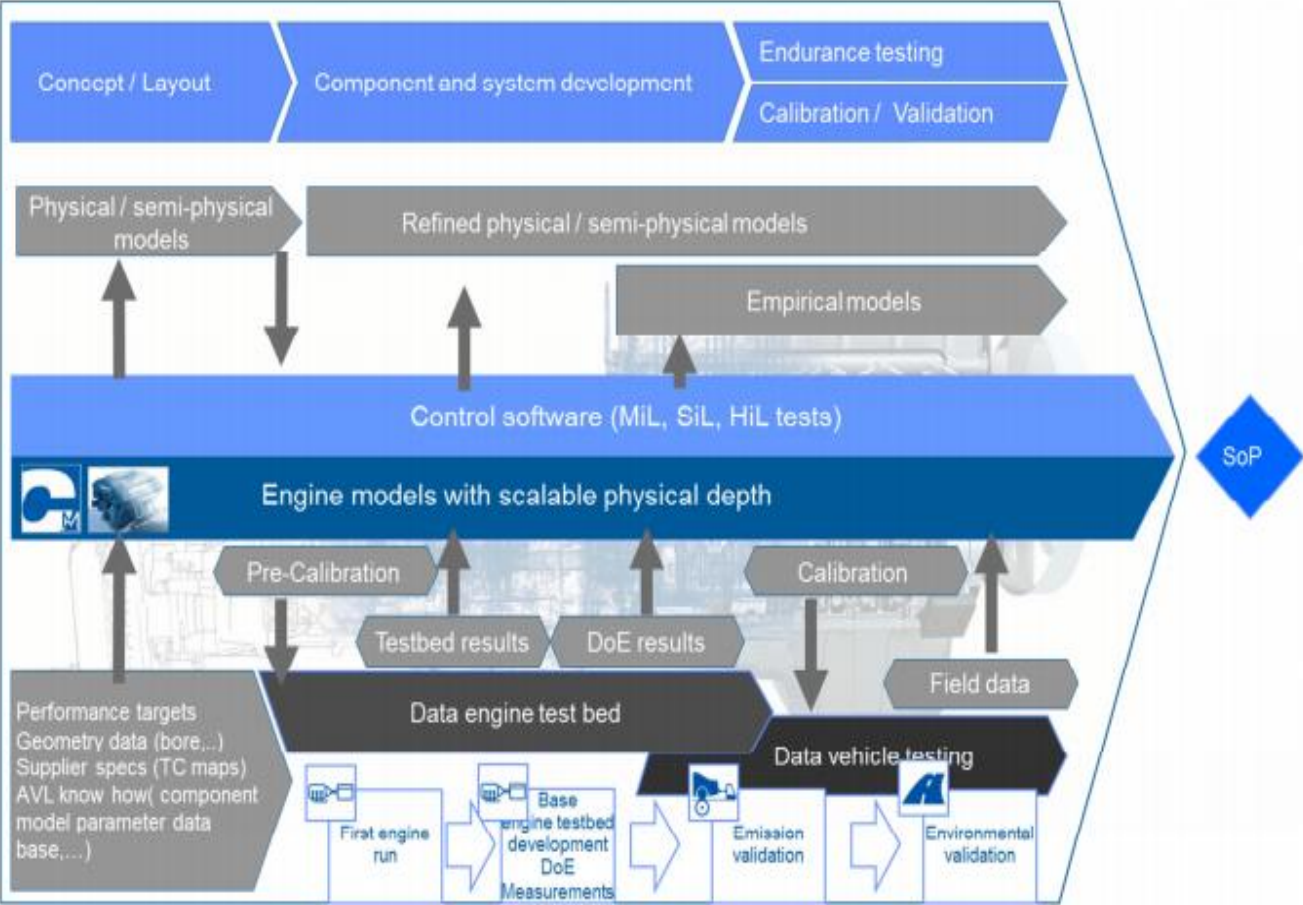
Figure 27: Broad overview of model-based Calibration process [54]

3.3.1 Dataset Validation

Validation has become an important aspect for calibrations after new requirements emerged for the automotive industry. Dataset validation has to be performed over all variants of vehicles under all distinct conditions for driving. Calibration can undergo various errors due to complexities emerged with new approaches [54]:

- Various errors emerged in the process, even though, no technical support is incorrect.
- Various errors emerged due to complexity in the systems. For example, a calibration belonging to a particular component is correct, but is incorrect as a whole, due to ignorance of various components relations to each other.
- Various errors emerged as a result of a deficiency of time and resources to calibration engineers.

These errors can be effectively avoided by various techniques such as by utilizing efficient data management tools, automated hardware-in-the-loop validation for dataset considering various component ranges before releasing any dataset, and simulated supported fleet validation with vast spectrum. This enables engineers to prevent these

errors even before detection. Various methods of model-based calibration can be implemented efficiently using series of tools as shown below [54]:
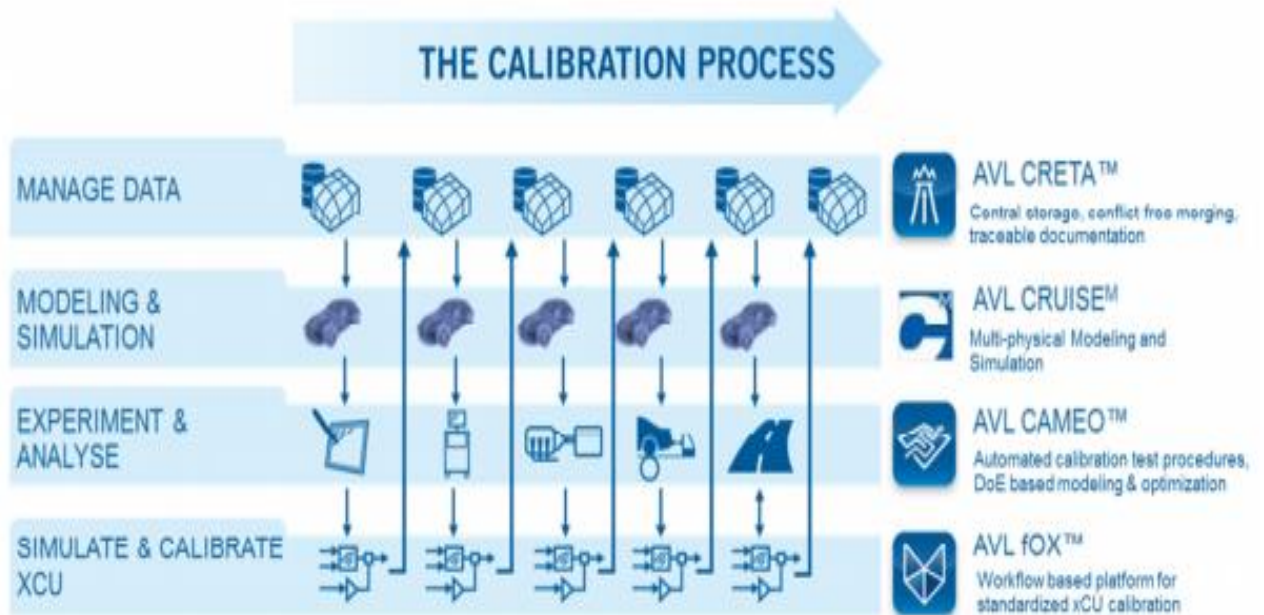
Figure 28: Calibration process elements used in AVL [54]

3.4 Calibration tools/solutions

The calibration process is composed of a number of activities or task carried out in a certain order. Moreover, the development stage of engine and vehicle involves various changes to hardware and software. This results in basic calibration tasks to be repeated for each change in configurations, resulting in large amount of work for calibration engineers to carry out the calibration sequences, and to collect and analyze related measurement data, causing loss of time and money for an organization. Such issues have resulted in great demand for automation of the process, to replace manually carried out tasks to improve calibration quality [55]. Figure 29 reflects the need of automation in ECU calibration processes due to various factors.

3.4.1 INCA tools for measurement, calibration, and diagnostic (MCD)

Actual script generation for an ECU software provides a challenge in calibration process for automation. Scripts are generated in a coding language, and calibration engineer would have to be trained for programming, which is often not their specialty. INCA flow provides the possibility of graphical coding in the form of flow charts. This tool facilitates the calibration with following issues [57]:

- Allow automation of procedures,
- Graphical algorithms (visual solutions) require fewer efforts for calibration engineers to understand and provide for calibration procedures and ease of maintenance,
- Efficient designs by collections of ready to use measurement, calibration, and diagnostic methods contained in various libraries.
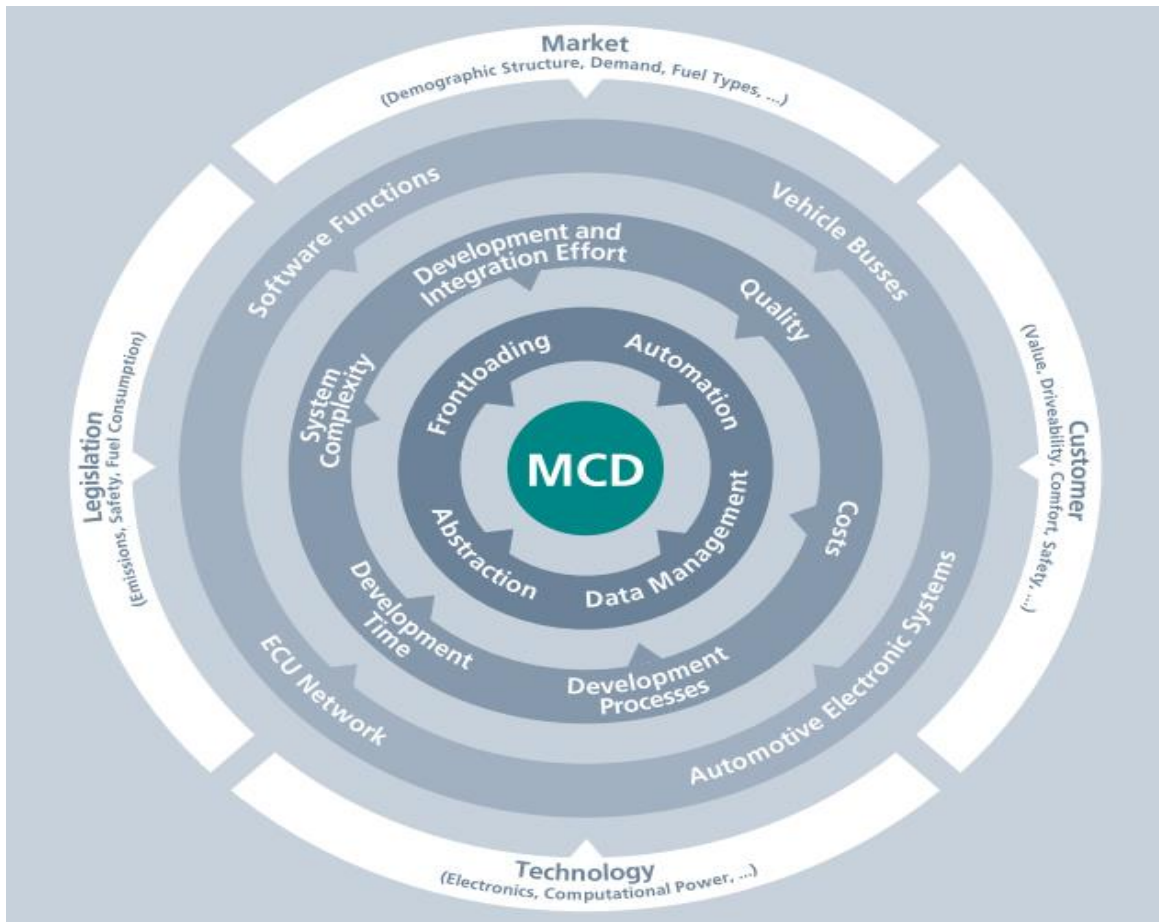
Figure 29: Relationship of measurement, calibration and diagnostics automation requirement with various elements [51]

The following figure depicts how INCA flow is utilized in calibration process:
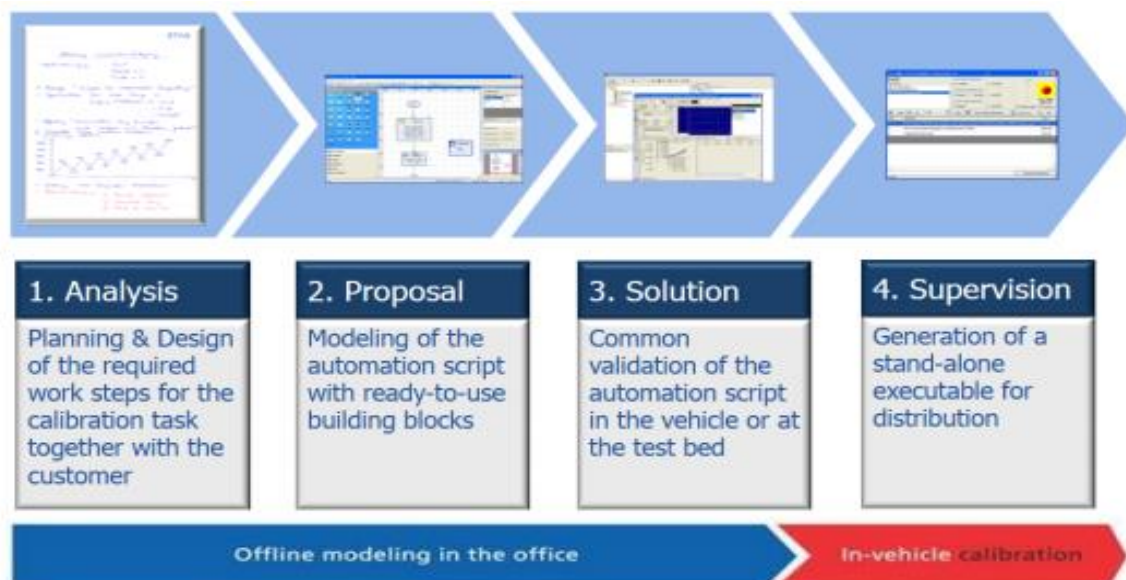
Figure 30: Calibration using INCA flow [55]

Calibration parameters (labels) can be modified online by utilizing test automation provided by INCA flow. Data recording are automated and can be evaluated straightly by the tool. Hence, the requirement of automation in the calibration process is quite obvious after reading above issues [55].

ETAS INCA provides series of tools for measurement, calibration, and diagnostic in the calibration process, of which INCA flow is just one part. Some of these tools are listed below [56]:

- INCA-MIP: Provide interfaces for ease of use of MATLAB to analyze measurement signal and parameters of calibration from control units.
- INCA-LIN: Provide display for monitoring of LIN Bus signals acquired.
- INCA FLEXRAY: Utilizing XCP Protocol, supports measuring and calibration of parameters of Flex ray ECUs.
- INCA-EIP: Facilitates rapid prototyping systems functions calibrations.
- INCA Base Product: Provides basic measurement and calibration, with calibration data management.

Other tools in INCA product family are measure data analyzer, ODX-LINK, ODX-Flash etc. Apart from calibration, all these tools can be used for various other purposes such as on and off-board diagnostics, vehicle testing on real roads, frontloading by providing the possibility of calibration on testbeds, virtual environments with prototyping and simulation, and process integration etc. With all these product family elements, INCA is used all over the ECU development process, starting from function models initial calibrations to test bench pre-calibration for ECUs, to final vehicle optimization. INCA can be utilized from PC simulations to vehicle instrumentations, including calibration data offline management and analysis of measurements. Apart from interfaces for automation systems, various tools in INCA product family facilitate calibration engineers with a vast variety of open and standard interfaces for the purpose of data exchange, configuration, documentations, and automation. They provide possibilities to combine customer applications. Various family products of INCA and their relationships can be shown with following diagram [51].
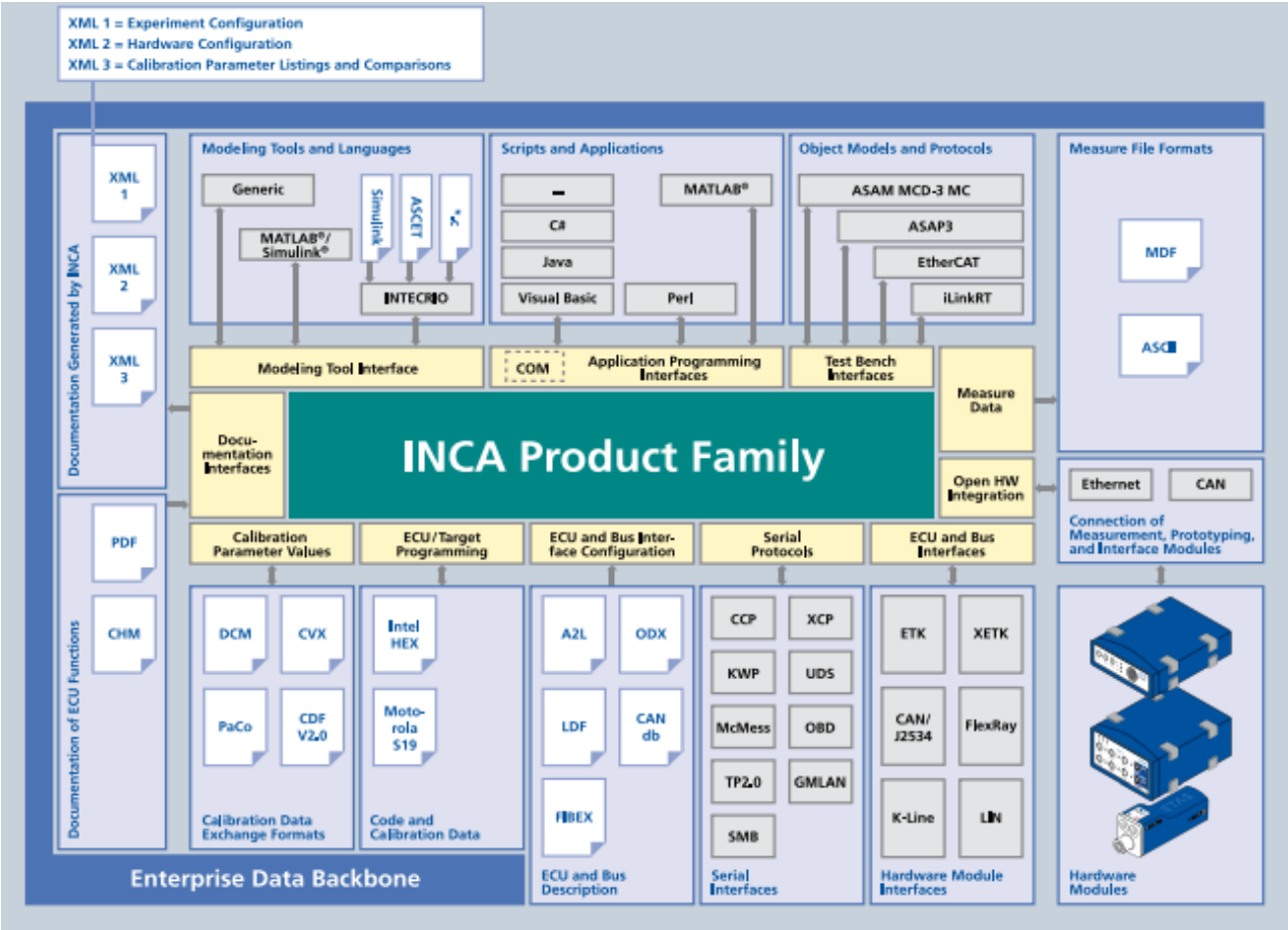
Figure 31: INCA product family with various deliverables [51]

3.4.2 Calibration lifecycle management solution: AVL CRETA

Due to the increasing complexities of the calibration process, there exist many challenges. For example, due to increase in variants for vehicles and their complicated structure, it becomes quite problematic to manage data with respect to a project with various variants. Complex data management or data exchange become extremely difficult to be handled by emails or project drives for hundreds of engineers working at various locations on various variants. With a number of tools being used for the calibration process as discussed before, the monitoring of project become a significant issue [57]. Same calibration parameters/labels are often used repetitively for different calibration projects, such as calibration for throttle's fine adjustment or setting of the exhaust flap, remains same for different projects. But it may happen that their calibration is done redundantly for various cases due to lack of history in central data storage, causing wastage of time and costs, and task redundancy in the process [58]

The above issues can be resolved by utilizing AVL CRETA tool in the calibration process. It provides central storage repository, which optimizes calibration data management process and dataset exchange, with efficient knowledge management between teams. It helps to clearly divide complex projects tasks and workloads with a large number of variant, into various segments called work-packages, belonging to a particular component of the project. CRETA provides a component library to help calibration engineers to release their work directly to their components, without the need of understanding complex structuring of the entire project as a whole. Various features are provided by CRETA, like assigning attributes to variants for correct data management and comparisons. Using work-package, responsibilities and roles are well defined and grouped for the teams of calibration engineers. The component library provides functionalities of creating calibration history which help to re-use calibration work and knowledge, which further avoid redundancy in work and provide correct data management. With segmented work into work-packages, CRETA provides ease of monitoring project status with its tools [59].The simplified implementation of CRETA in calibration process can be shown in figure 32.

There exists, numerous functionalities provided by AVL CRETA, some of which are listed below [11]:

- Assigns responsibilities for calibration parameters, also called labels, among calibration engineers, and work-packages among various teams. Well defined labels responsibilities, provide conflict free merging of labels from various sources.
- Label values from different datasets can be compared for changes. Simplified HTML and Excel reports for the complete history of software and calibration can be easily generated.
- Different variants of the project can be well structured and examined by explorer provided in the tool. There exist the possibility to create delta variants to reduce the burden of variants management.
- Drop-boxes and delivery templates are provided for effective delivery of results of calibration belonging to a correct variant. Metadata like label maturity information can be stored along with any attachments.
- Using powerful search functions, calibrations from components library or old projects can be used for pre-calibration and promote effective calibrations in everyday work.
- Unified activities for merging datasets, carrying out changes to ECU software, and exchange of data provide efficient project management.

- Advance data mining algorithms, help detects errors in delivered or received datasets.
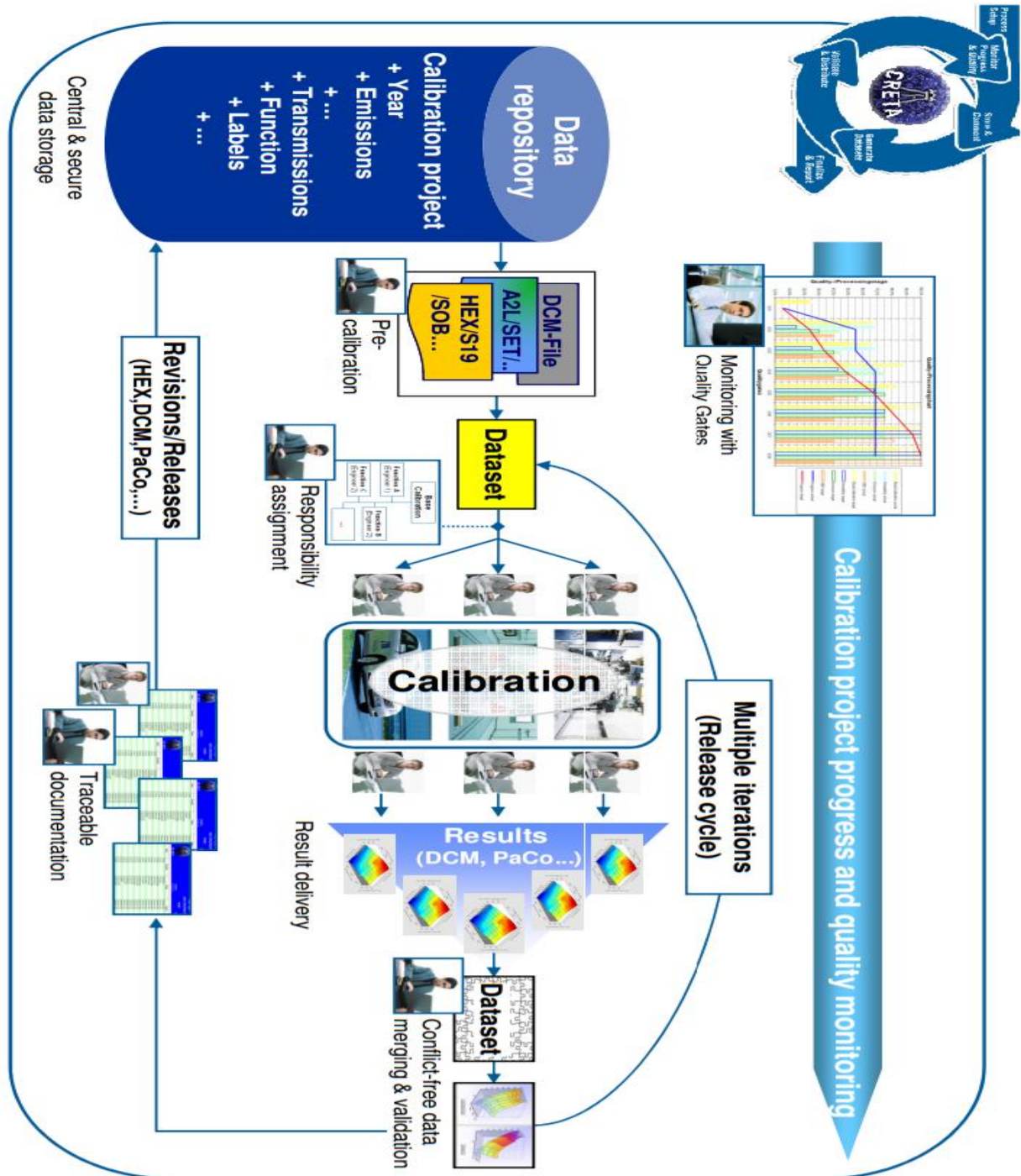
Figure 32: AVL CRETA in the calibration process [60]

AVL CRETA provides help for the calibration process in handling a large number of variants of calibration parameters. Calibration engineers are able to make sure that correct calibration dataset is provided for right vehicle variant. It is easily connectable to various other development software like INCA, CANape, Matlab etc. and provide a reliable solution for calibration management. Figure 33 shows the amount of time for activities involved in typical calibration projects, and thus, depicts the importance of data management tool like AVL CRETA in the calibration process, with data handling time requirements amounting up to 51% of whole calibration projects in general.
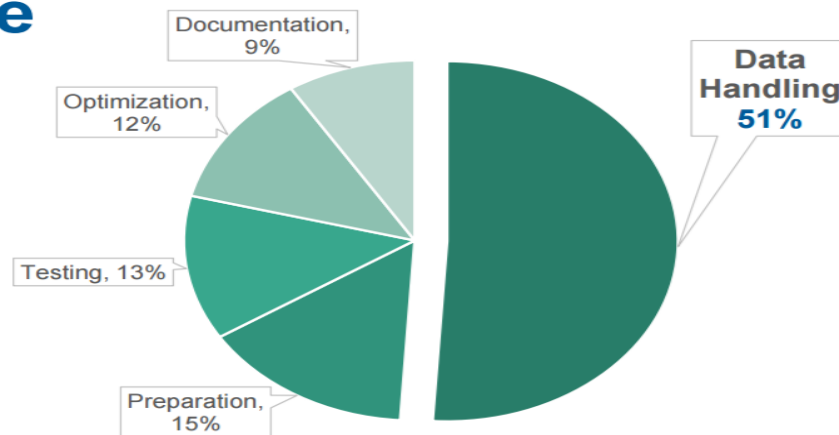
Figure 33: Typical activities time in a calibration project [61]

3.4.3 Accelerated calibration data post-processing solution: AVL CONCERTO

AVL CONCERTO provides graphical calibration data visualization and analysis solution for the distinct type of data obtained from different sources like test beds, test beds environment, chassis dynamometers, or from durability runs of vehicles. It basically provides browsing data efficiently, data presentation and its management. Apart from this, it enables processing data obtained in batches, various calculations, and then generating reports. It can obtain various format data and squeeze them to obtain effective results easily with its wizard supported design.
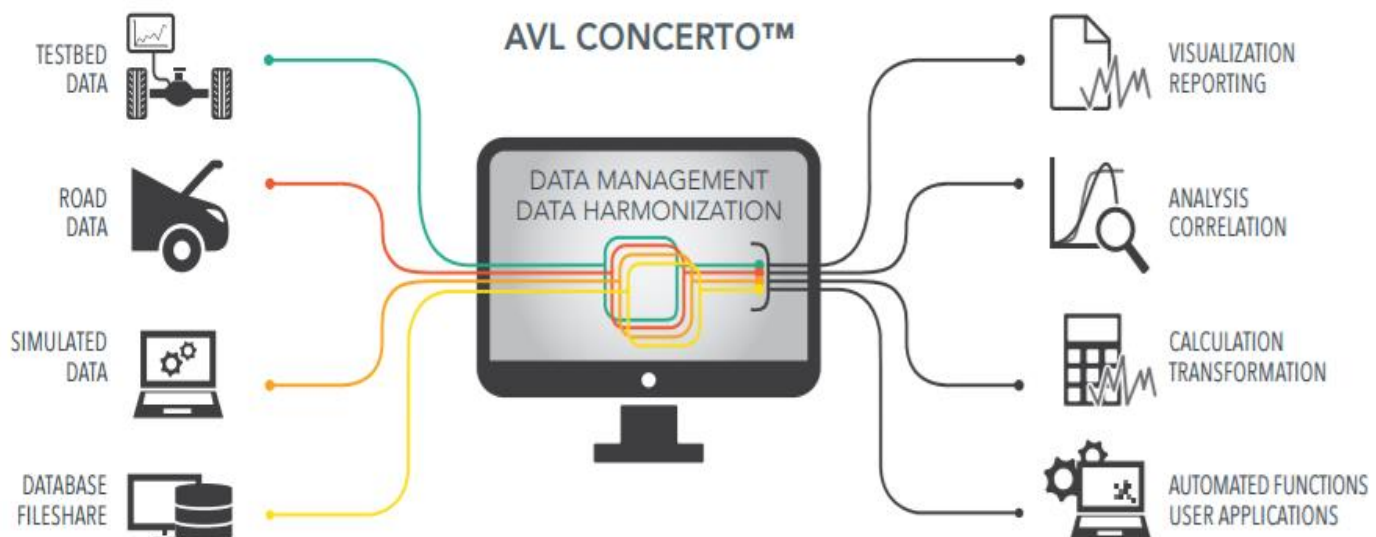
Figure 34: AVL CONCERTO in calibration process [62]

AVL CONCERTO, which has made calibrations comparison as well as their results comparison possible, is utilized under following situations [62]:

- While managing vast amount of files composed of distinct contents,
- To organize data at various places in a reasonable manner,
- To analyze same repeated cycles of calibration tests as single (like average),
- To merge data with distinct format, resolution, and length of file,
- To process data in batches in an automated manner,

- To execute calculations using standardized or customer specific algorithms,
- To prepare high-quality reports and then print them etc.

3.4.4 ECU software calibration process files

Set of various parameters/labels, at a particular time, contained in edit history for ECU calibration is called calibration dataset. Datasets encompass all ECU parameters and are represented by various files such as database file (A2L file), memory image content (HEX file) etc. Dataset act as the main object used in the calibration process to carry out versioning and configuring parameters/labels of ECU software [57]. As shown in figure 31, following files are significant in calibration process [63], [57]:

- A2L file: It is basically an ECU software file format for the calibration process. In earlier days, calibration parameters had to be modified with source code and data logging was performed using variables in the source code. As ECU software became more complex, the development was divided into various domains such as function development, software development, and calibrations etc. By creating A2L file format, ASAM MCD developed a standard format (ASAM MCD-2 MC) to separated calibration relevant information from ECU software with a detailed description of calibration and measurement variables excluding information like code details. For the goal of reading and writing, it contains information for device interface for ECU. It helps to provide extensive support for lookup tables with various dimensions, static or calculated axis points or calculated calibration parameters. A hierarchical grouping can be carried out for calibration and measurement variables to perform function-oriented calibration. Thus, data description is written in an easy to parse and importable well-structured ASCII format (*. A2L). A2L file has modular structure, divided into four main parts as- i) AML: containing interface specific parameters, ii) General ECU Implementation: containing general description and standardized structure of ECU, iii) IF_DATA: containing configuration interface keywords, and iv) A2L objects: encompasses CHARACTERISTIC keywords to define calibration parameters and MEASUREMENT keyword for measurement parameters. A2L file is provided by ECU supplier to access data inside ECU software for ECU.

- HEX File: It is basically a dataset file for calibration process for an ECU software (A2L file.) Once ECU is connected, the changes to ECU software are made using HEX file. It is a binary file format and works by gaining some part of the information from A2L file to know about functions of various part of ECU memory. HEX files are used in various ways due to various purposes. It reduces need for accessing ECU for reading and writing, with the principle of memory cache. During calibration, the main memory contains mirror memory, the content of which is saved into hard disk when the program is excited for next calibration. Besides, Intel HEX file provide the possibility to calibrate in the offline model.

- DCM File: These files contains changed labels of a dataset (HEX file) and their changed values by calibration engineers during their tests. Several DCM files are merged together after various tests performed to a dataset, to create a new revision of ECU software dataset (.HEX file).
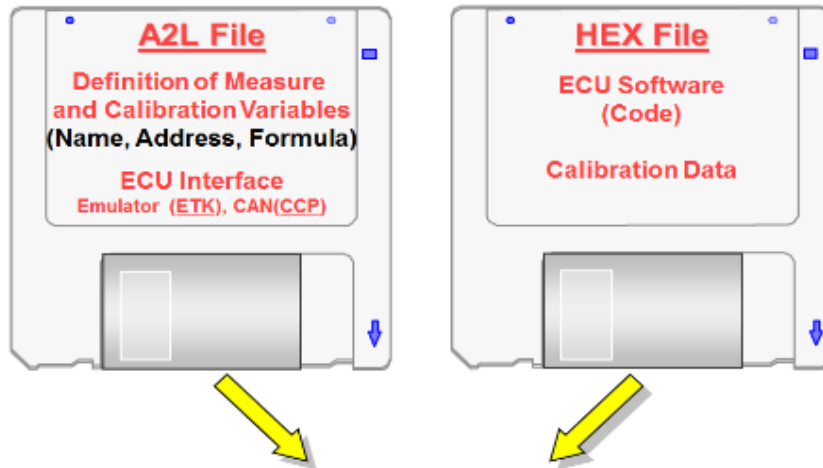
Figure 35: HEX file and A2L file relation [64]

The figure above depicts the relationship between HEX file and A2L file. A HEX file belonging to particular ECU software cannot be used for other ECU software.

Chapter 4: Various elements of project management

4.1 Data Management

Data management is an essential element for an organization's ability for efficient project management [65]. Data management act as a crucial aspect to optimize revenue generation, handle costs, reduce various risks, and raise confidence for organizations facing ever increasing volume of data, to gain competitive advantage. Data management is defined as "Development, execution, and supervision of plans, policies, programs and practices that control, protect, deliver and enhance the value of data and information assets" [66]. According to DAMA international, data management encompasses various areas of knowledge for effective and efficient data management, mentioned in a wheel, as shown below:
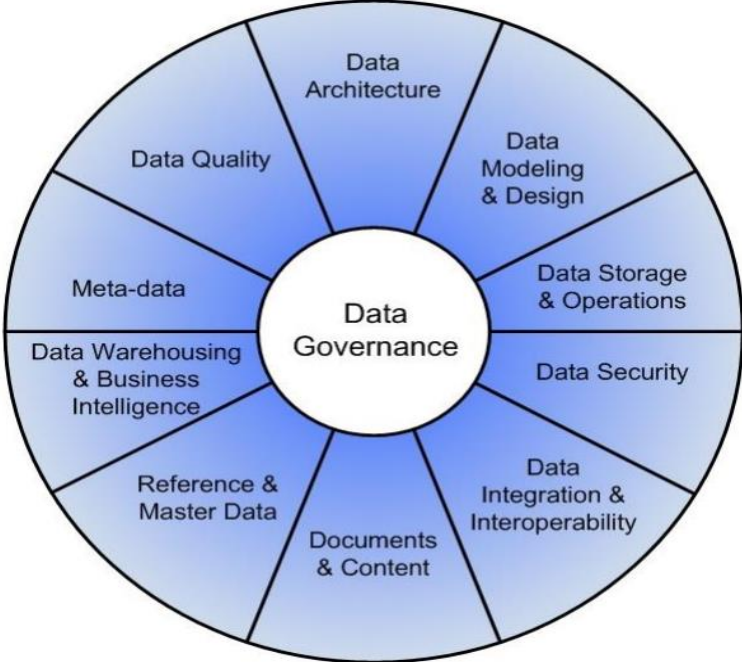
Figure 36: Various knowledge areas of data management [67]

4.1.1 Data quality

Data quality is often considered in organizations, to overcome various problems related to data. These problems can be effectively evaluated by root cause analysis approach to eliminate the source of flawed data by building business cases, by carrying out a holistic view over the data quality management approach to benefit the organization. Data quality is influenced not by unique one-time events or changes but consistently by new and insidious ways. These ways involve data quality teams to consider not just various failure points but also the ways to monitor, measure or track the status of current data quality. For example, measuring data quality with some metrics, which would further let the organization know, in a quantifiable manner, if they have achieved desired goals by improving quality in their data. Moreoever, they can communicate with their clients with

which level of quality assurance they can provide service with their results. The data quality can be broken down into various aspects or dimensions like data consistency, completeness, uniqueness etc., which can be related to various business case elements to know which data aspects are to be concentrated the most. The various dimensions of data quality are listed below [68], [69]:

- Uniqueness: It provides a measure for no redundancy or duplication in recording the data.
- Accuracy: It reflects the degree to which data represent right value. It defines the extent to which data is correct, reliable and certified. Accuracy is a measure of proximity between correct value and data value.
- Consistency: It is the degree to which data is present in a compatible form so that the semantic rules determined for the dataset are not offended. For example, the sum of the number of employees at each site must be equal to the total employee of an organization.
- Completeness: It is a degree to which data is plentiful for the concerned task. It is to ensure that all the information required to fulfill the task is present efficiently.
- Timeliness: It is the degree to which, the activity is related to the age of the data. It is a delay among change in actual correct value and change in value in the database. Apart from age, it has another element called volatility, which means how frequently data changes its correct value and reflects instability.
- Currency: It is the degree to which, how up-to-date information is. Data is up-to-date, if it is correct, even though it is influenced by changes with time.
- Conformance: It determines if data instances are either stored, exchanged, or presented in an acceptable format for the domain of values, and is consistent with similar attribute values.
- Referential integrity: Designating an identifier which is unique for data objects within user's environment, provide ease of data management.

Various technical methods can then be used to check if data quality is in required state. If not, the process of data validation and cleansing could be carried out [68].

4.1.2 Data cleansing

Data cleansing or data cleaning or scrubbing involves detection and removal of errors and inconsistencies from data for the purpose of optimizing data quality. Due to the various reasons like misspelling, while entering data, user forgot to put in information etc., there can be various data quality issues present in a file or database. Due to various segments involved in a process with various data sources, the need of data cleansing increases significantly, as each data resource may cause various issues like data redundancy, inconsistency etc. A major requirement of data cleansing is perceived with data warehouses, which are an essential part of a decision making for an organization. Data warehouses involve loading and consistently refreshing large amounts of data from various sources, which creates a high probability for various sources to compromise bad data. A data cleansing strategy should perform various functions, such as detection and removal of every significant errors and inconsistency for individual or multiple integrated sources. It is appropriate to perform data cleaning by an automated tool, for example by data mapping, which replaces requirements of manual checks or programming, and should cover many areas for application [70]. A typical data cleaning method consists of three sub-processes: error detection, data validation, and error correction. Data cleansing

principles are similar to data quality principles. The main principles for data cleansing process are listed below [71]:

1. An excellent planning provides effective data management for an organization. For satisfactory data quality policy, the information management chain shown in figure 37, considers data cleaning as a central element of the policy. Applying data cleaning and validation as a part of the culture of an organization, improves its relationship with customers or suppliers, by optimizing data quality.
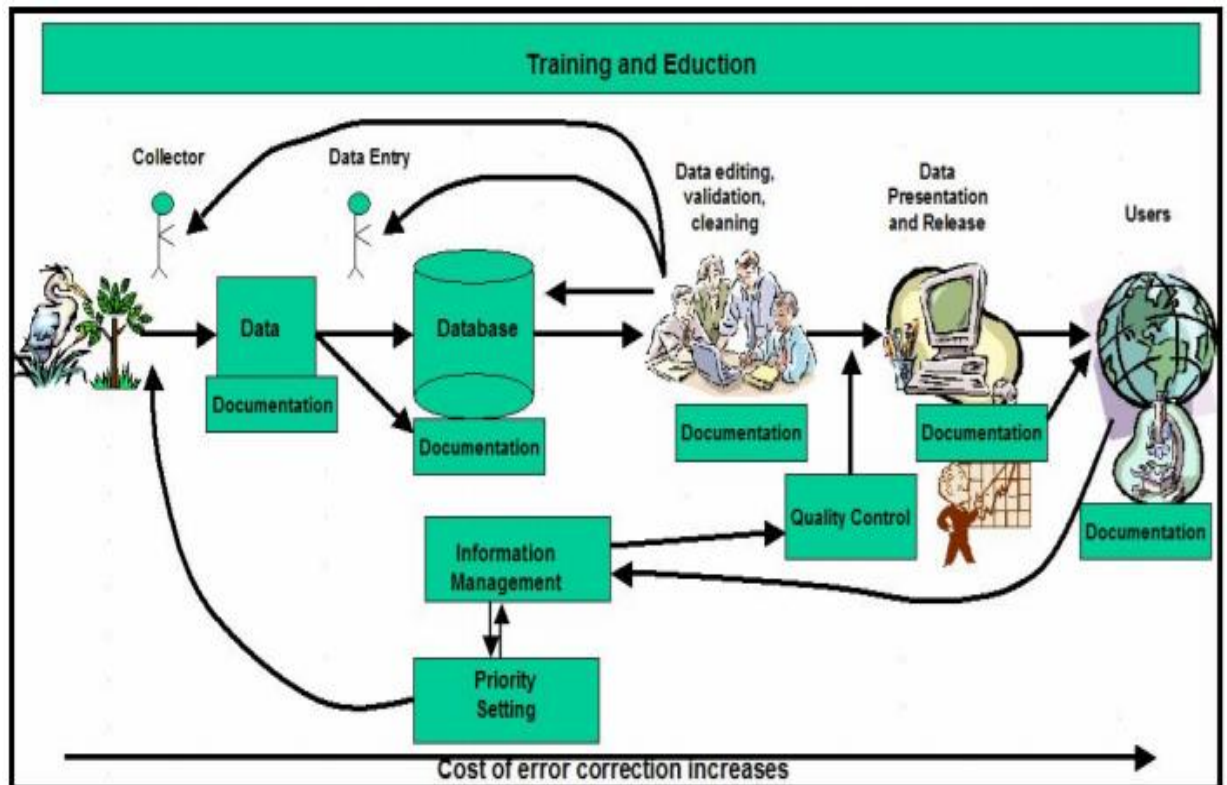
Figure 37: Shows cost of error decreases with earlier data cleansing for which training, documentation, and education is major part [71]

2. Before performing data checking, validation or correction, organizing data in a well-established manner would ease the data cleaning process. For example, organizing data on location or with attributes like the first alphabet, dates etc., would present efficient manner for handling the data.

3. Prevention is considered to be better than fixing errors. Wherever possible, it is recommended to prevent errors, for example, by utilizing drop down menu options during data entry instead of typing data. If preventive actions cannot be implemented, detection of error shall be carried out in such a way, that error in future are reduced.

4. Data cleansing responsibility belongs to all. The data custodian may hold the primary responsibility, but users or collectors, who possess the best place for error detection, should be able to respond to the custodian queries for errors in their data. Thus, it is recommended to build a partnership relation between each other for efficient data quality. The process should, thus, follow a two-way street feedback, including a clearly defined accountability for data cleaning.

5. Data shall be prioritized along with sorting the data. Prioritizing should concentrate on crucial data of high value which can be cleaned with lowest efforts, such as with

automation or batch processing. Data should be handled with appropriate measure and quality targets to monitor the performance of a process.

6. Data cleaning process should be well transparent and well documented to prevent reoccurrence of errors. It is important to note down which data has been cleaned and by whom. Metadata shall be available with a record of information at various levels of the dataset.

For the purpose of measuring data cleaning performance, scores have to be defined for various data quality dimensions, in order to measure the data quality for collection. By estimating these scores for various dimensions of data quality, an individual can be addressed in a quantifiable manner for the necessity of data cleansing, and also to the effectiveness of performed data cleansing for an organization. This also allows specifying which dimensions of data quality to concentrate more for the purpose of optimizing the data qualit. There exist many approaches for efficient and effective data validation and cleansing [72].

4.1.2.1 Organizing data: Data classification

As data cleaning involves sorting the data, it is quite important to understand the concepts which provide data sorting. Employees spend a significant amount of time just for searching the required data when dealing with unstructured raw data. This results in leaving them interacting with many inappropriate items and unrequired information. In many organizations, sensitive information is handled in inefficient manners leading to the poor identification, monitoring and ineffective controls for information. This often occurs due to lack of common face of reference like classification or grouping information which can, based on user case, differentiate sensitivity of information for efficient management along the lifecycle. Organizing or classifying data often, possess major challenges due to lack of harmony or accord between the organizations and their stakeholders. Classification is a simple concept that separates data into refined groups, demanding no comprehensive enterprise schema for implementation. It enables clarity in identifying, tracking and managing data without affecting operations [73]. Data classification is defined as a process which categorizes data by grouping it, with respect to identical attributes. Data classification enables significant benefits to various organizations such as the ones utilizing multiple, large scale applications. Working with a number of applications in a process can cause problems for smooth information flow. This can lead to requirements for organizations for integration of processes. For example, processes which require data securities for various important data, or companies facing challenges with data storage management due to a large amount of data generation etc. [74].

Data classification finds applications in various fields of applications, such as analytics, business value, data mining, integration, security and data securities/restrictions etc. It encompasses implementing metadata to data to manage it. Basically, there exist two different approaches to handle data: Data tagging, and Classification labels. Data tagging deal with providing basic information about a file, for example, if data file type is Excel or Word etc. Whereas, Classification labels are further development of tagging process, and deal with providing the user with information of context of data consisted in file or document. There are basically many methods of data classification, some of which are (in the context of management) listed below [75]:

1. User-driven classification – It involves classifying data based on know-how of user with complete knowledge about the concerned data. Generally, solutions employing this

concept enables classification by user clicking and selecting the appropriate label or attributes for it. It can be complicated to implement as it may require users to be trained, with pre-requisite as possessing good knowledge about data, considering access and permission options etc.

2. Automated classification – Based on file content or by manipulating the content which is required to be classified for change etc., various tool can automatically classify data without manual efforts. It can be of two types: 'real-time', which utilize some type of agent or 'batch type', which can be done by a local agent or remotely.

3. Data-at-rest scans – It is applied while using common servers or cloud environments like Dropbox or OneDrive etc. It can be used for data which is not structured, for example, some XML or PDF file. Semi-structured data are stored and shared to search tagged documents and files. Classification is done in semi-structured data once a scan is done. They need to be classified with respect to specific content.

PLAN DO CHECK ACT model provide an adequate manner for classifying data. 'Plan' can include determining properties of data, responsibility for carrying out classification and create a profile. 'Do' involves implementing classification after agreement from concerned individuals. 'Check' step involves ensuring if tool employing classification is efficiently conforming requirement. The last step 'Act', may involve reviewing data and provide reclassification if revision is required to adopt to changes made in data. The above process can be visualized as following [76]:
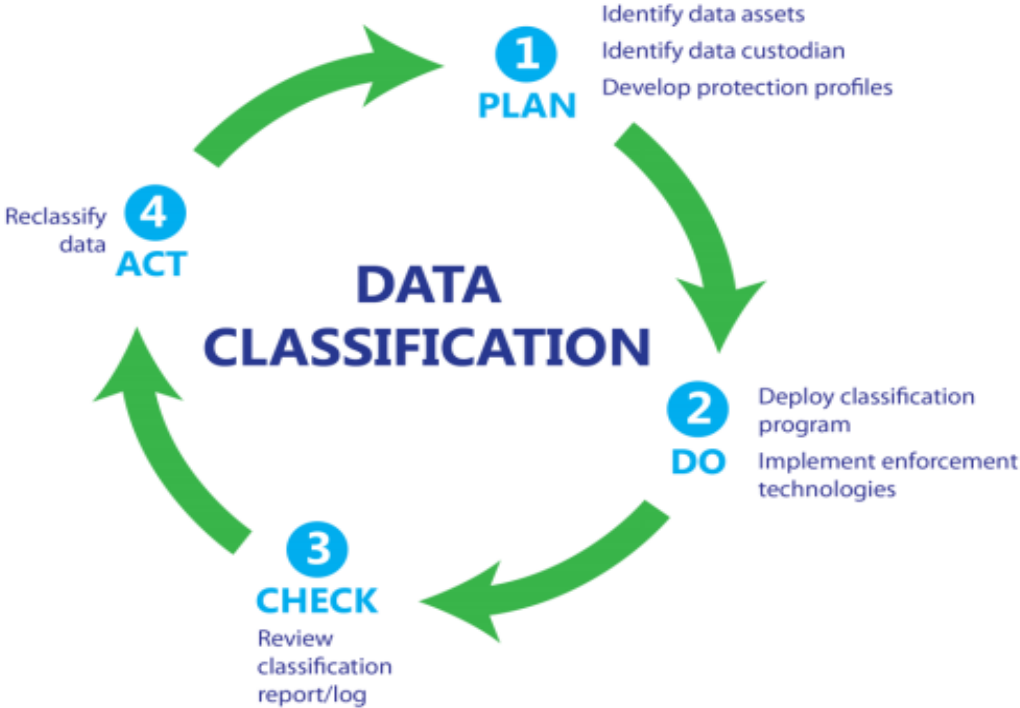
Figure 38: PDCA for data classification process [76]

For effective data classification, various important points can be considered such as [76]:

- Meta-data tags – A temper resistant support characterizing a permanent or temporary label adhering to the sensitive data for analyzing or tracking the data.
- Automated and user determined classification – It is beneficial if the significance of data can be automatically be assessed and manually as well, to add various information to data to make it up-to-date with changes.

- Inheritance – Data tagging should be able to transfer to related components. For example, from parent or source file to a child or derivative file for maintaining consistency in managing data.
- Multi-layered classification – The data should be able to classify with many individuals attributes/labels at the same time, which would make it possible to filter the required data with selected attributes/labels for various kind of analysis.
- Rule prioritization & conflict resolution - The ability to resolve the hierarchy of multiple meta-tags that describe the same file with competing or overlapping governance rules. An information protection solution which supports multi-layered data classification must allow logic gates, to be predefined with default prioritization, in order to make autonomous decisions for rules.

4.1.3 Data visualization

"A picture is worth a thousand words – especially when you are trying to understand and gain insights from data". Visualizing data is more efficient than written data, such as in the forms of tables with columns and rows, or compared to text lines. There are various ways to visualize data such as by graphs like scatter plot, line graphs, box plots, pie charts etc. It may occur that visual creator makes wrong choice to choose the right tool to visualize data. A good choice can be made by clearly understanding the data and its composition. For example, what information to depict, and how user process the information from the data [77]. In the times of information age, data is increasing exponentially, where we have been growing faster with storing and collecting the data than analyzing it. Analyzing these messy, inconsistent data, has become crucial for making decisions. Following diagram depicts the scope of visual analytics [78]:
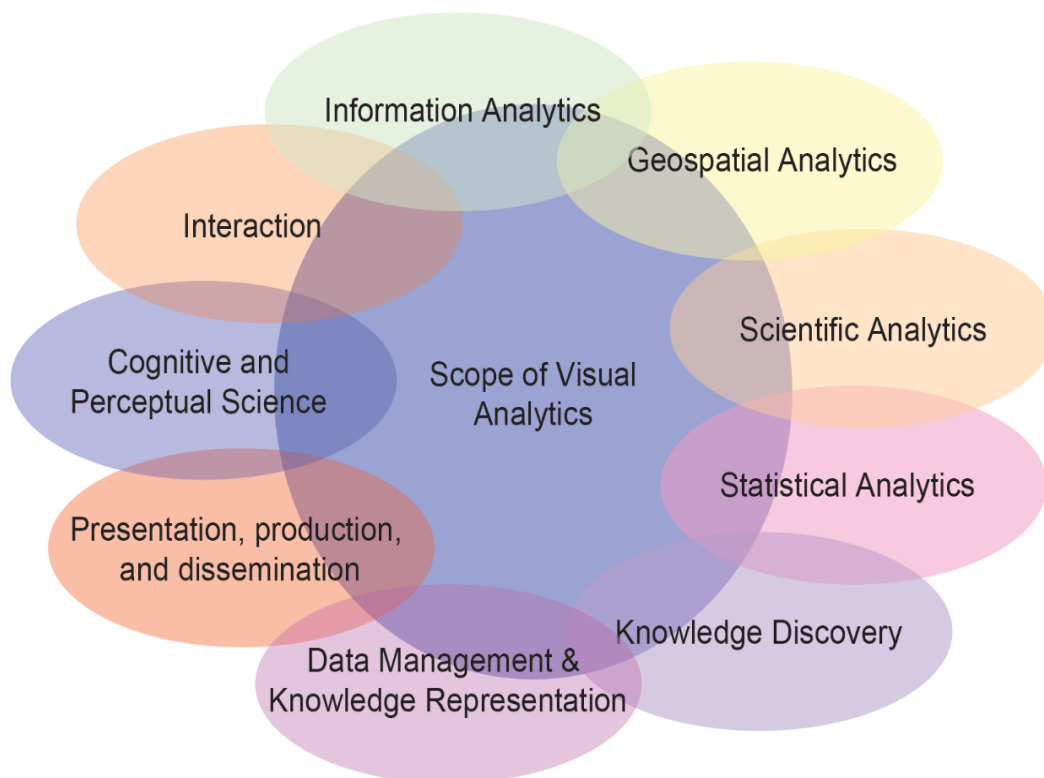
Figure 39: Scope of visual analytics [78]

In the business processes, data visualization carries two goals basically [79]:

- Explanatory: Here a user has a question, which is answered by viewing the data. Various business dashboards fall under this category.
- Exploratory: Here a user view data first, then certain questions arise in their mind, which are answered by viewing the data. For example, while comparing various datasets. These are cyclical in nature without any end point.

The two main reasons to visualize data can be as follow [80]:

- Better analyzing the data structure and its relationships,
- To provide information of data to a large audience with ease.

One prominent graphical way to represent data is node-link diagrams, also called network diagrams. These are extremely useful when data has structure embedded in them, there's no obvious link between elements under consideration, and they possess certain common properties/relation between them. If such requirements occur for data, it is useful to replace data structured in matrix tables with a nodes-link diagram. They are basically made of nodes, which represent elements of data, and their relationships are represented by arrows or edges. The direction of arrows depends on relationship content. They are of two types [80]:

- Directed - This represent a directional relationship, for example, a person 'owns' car, but vice versa is not true.
- Undirected – These are non-directions relationships, for example, one person is 'sibling' to other. This is true with vice versa as well.

An example of network or node-link diagram is shown below:
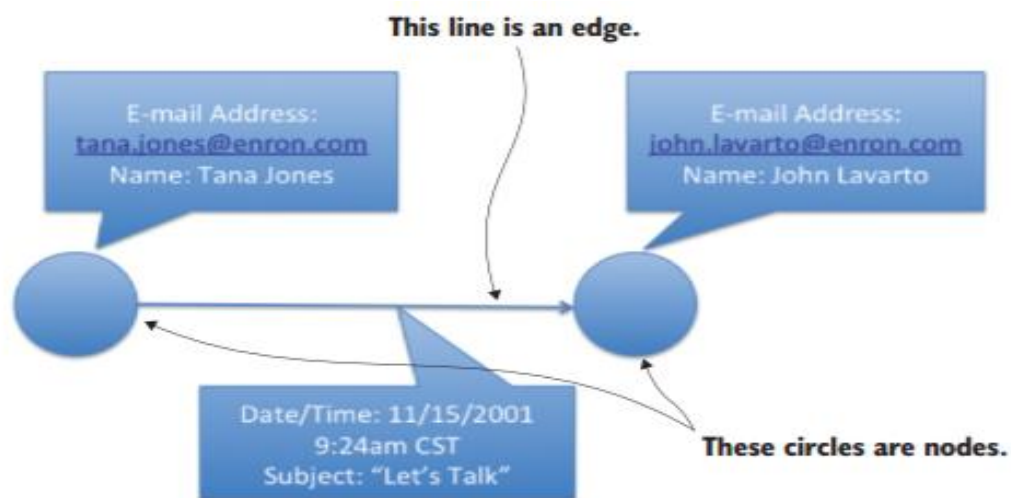
Figure 40: E-mail sent between two individuals [80]

4.2 Business process, modeling, and automation

An authentic process consists of every activity performed to accomplish requirements for individual or organization with certain expectations. A process encompasses various elements such as input and output, flow units which proceed along all process, networks of various tasks and buffers associated with them, resources which can be capital or labor type, and information structure to make decisions. "A business process is a network of

connected activities and buffers with well-defined boundaries and precedence relationships, which utilize resources to transform inputs into outputs for the purpose of satisfying customer requirements" [81]. Business process encompasses various elements such as sub-processes, which are further composed of activities, which are further made from a group of tasks. Thus, it can be decomposed in a manner shown in figure below:

Business Process	Subprocess	Activity	Task
Procurement	Vendor selection	Vendor survey	Documentation of outside vendor
Development engineering	Hardware design	Engineering change	Convening the Change Board
Office administration	Providing secretarial services	Calendar management	Making a change to existing calendar

Figure 41: Decomposition of a business process [15]

Business processes are made in most of the companies on the basis of functions, with each functional areas as different departments, performing specific work process for a sub-process. This is done in order to utilize specialized skills of different teams by assigning domain specific individuals [81].

Business processes can be analyzed to make improvement in an organization, such as in process quality management methodologies. After analyses, processes can be redesigned to make them more effective, efficient, and adaptable [15]. Due to increasing complexities, because of adoption of a large number of individuals and tools in a process, has made integration of processes a significant topic. There exist many such forms for integration such as application-to-application (A2A) within an enterprise, business-to - business (B2B) between various enterprises, and e-commerce integration etc. Apart from such technical integration, there exists a technique which involves integration and support for processes encompassing various business logics and information requirements called Business process integration. Business process integration provides an organization with communication, automated process eliminating human efforts, clarity for visible monitoring, and support for decisions. Business process integration is realized by tools, which by utilizing some mechanism, arranges and coordinates the data flow for a business process. Such tools also contain templates categorizing work into various sub-work-packages, which together accomplish a business process goal [82]. It involves reducing the effort required for communication and coordination. The level of integration of process can be evaluated by time require for the flow of information, optimized by reduced number of steps and handoffs in the process.

4.2.1 Business process modeling notations (BPMN)

BPMN are graphical notations which are internationally accepted standard, for process modeling. They are used to illustrate various steps and their logics involved in a process. They help to lower the gap between business process and their employment by constructing a standardized bridge. BPMN is based on flowcharting technique and includes all the information and activities of the process, allowing efficient communication for a process in a clear and efficient manner [83].
BPMN does not have to be constructed completely correct for a particular process but can be constructed to depict the process in general or simplified form. It contains various

categories of elements, each represented in BPMN as shown in figure 42. Various elements in BPMN are explained below [84], [85].

- Flow objects: Flow objects are represented in BPMN as shown in figure 42. Various flow objects are 'activities': work executed for a process, 'events': actions that occur in a process which effects activities such as their logic or timing in a process, 'gateways': decisions points which control divergence or convergence of a sequence.
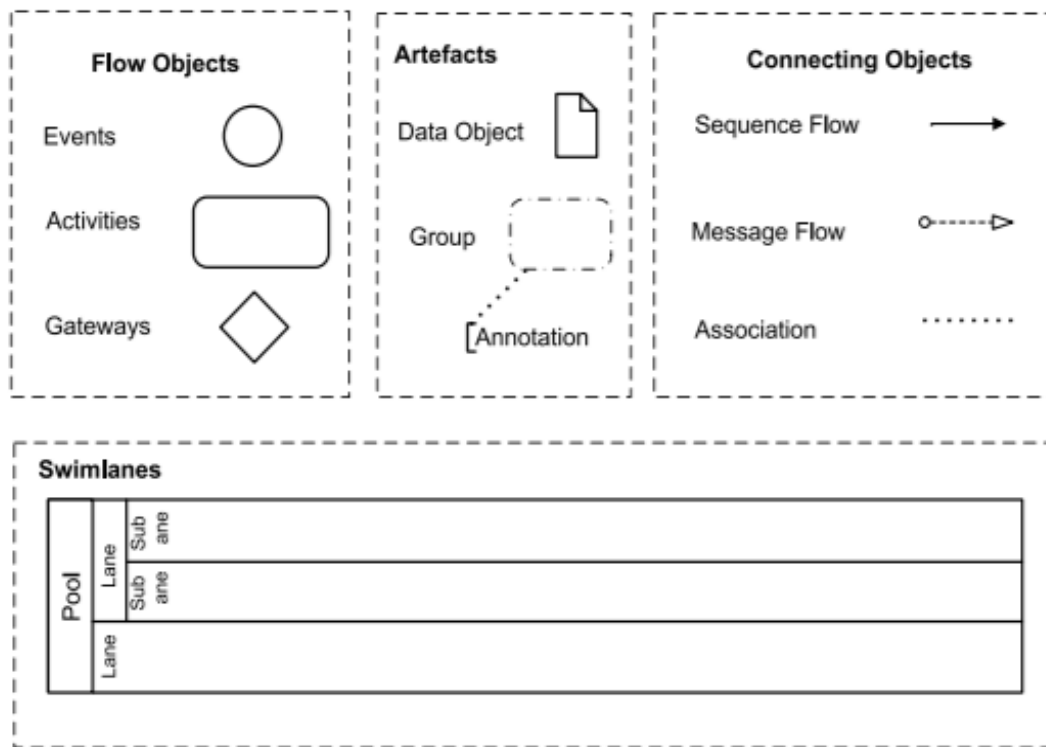
Figure 42: Various categories of BPMN elements [86]

- Artifacts: Artifacts are represented in BPMN as shown in figure 42. Various flow objects are data objects: elements which provide information for the process but does not effect it, group: used to cluster process elements informally, annotation: provide additional information for the process.
- Connecting objects: Various connecting objects are used in modeling process by BPMN, such as:
  1. Sequence flow: used to depict sequence of activities occurred,
  2. Message flow: depicts flow of messages between entities,
  3. Association: used to depict the flow of information elements like data, texts along the process.
- Swimlanes: Participants like business entity or role etc. are represented as pool, which can be further divided into lanes, and sub-lanes.

Business process modeling can be used to communicate various types of organization information for varied variety of audiences. BPMN helps organizations evaluate their process and its elements, which enable the organization to make process segmentation or end to end business processes or business process re-engineering. Using the concepts of business process modeling, and re-engineering, Ford analyzed in 90's their recently acquired organization, Mazda, for how their account staff with no invoice system was

working with just 5 people (75 people relatively taking into account the economy of scale and size of Mazda with ford). Mazda company's receiving department checked the delivery against the original purchase order—if everything was fine, then the payment was processed—i.e., without an invoice. Compared to this, there were 400-500 employees in Ford, to which, they later adopted to re-engineer their process like Mazda with no invoice. They were able to reduce their staff to 125 people with more efficient work process. Thus, BPMN allows easy evaluation/ analysis of a business process in an organization to optimize it [85], [87].

4.2.2 Business process automation

Business process automation (BPA), also called workflow management, enhance the productivity of business activities by analyzing precise description of business processes. It is achieved basically by human components and software applications. BPA can provide process mining by effectively tracking or monitoring process events. BPA enables completion of workflows around multiple applications and processes, by a relation between the design of processes and application integration services. Apart from automation of business activities, BPA enhances and simplify workflow for an organization [88]. BPA are set of strategies which bring an organization optimization with respect to cost, time and effort, providing them competitive advantages and faster services for their customers. The example of applications that can provide BPA can be, SAP R/3, Microsoft office, TALLY etc. BPA can possess certain control systems to provide efficiency and effectiveness of business objectivities. The main goal of these controls can be authorization, validation, segregation of duties, securities and handling errors etc. Apart from these, the main objectives of BPA are to provide confidentiality to data, integrity, availability and timeliness [89]. Automation for a business process can be measured in three steps as 'insulated efficiency gains', 'collaborative efficiency gains', and 'universal efficiency gains'. Thus, BPA provides following benefits to an organization [90]:

- Enhancement in efficiency: Automated procedures results in the removal of unnecessary and redundant activities from the process.
- Improved process control: Audit trials, standardization etc. possible via process automation allows better control over the process.
- Better customer service: Consistency achieved for a project, results in meeting required efficiently by automation, creating better relationships with customers.
- Flexibility: With lesser steps in the workflow and human interactions, the flexibility of business process is enhanced to growing changes in business requirements.
- Improved procedures: Procedures are required to be executed in the same way, again and again, to meet regulatory needs. This is achieved easily by business automation.
- Improved assignments: Assignment of work can be automated along with steps, resulting in more productive work. Moreover, users can dedicate their time on more value adding activities.

4.3 Forecast analysis

Forecast analysis is a plan optimization techniques used in the diverse field of activities. The ability of decision makers or researchers to anticipate future events or outcomes based on past actions in a convoluted environment can prove to provide vast success to implement plans for their projects. An organization is bound to face waste of time and

resources if management carries inadequacy for evaluating randomness, uncertainty, and forecasting future events [91]. Forecasts are often carried out in the diverse field of activities such as finance, governmental, insurances, and consulting firms etc. for effective decision making. This technique is basically used to carry over three elements in business and economics: event outcome or event timing, and time series [92].

It is important to understand the difference between forecast and prediction. Forecasting can be defined as a process of future event estimation through evaluating and casting forward past data. By a predetermined manner, the past data are systematically combined in order to forecast the future event. Prediction can be defined as a process for future event estimation on the grounds of subjective considerations part from past data. By predetermined manner, the subjective considerations are combined and evaluated. It is to be noted that forecasting techniques cannot provide always a certainty in the estimation of values for a future event, and thus, it is distinguished as poor or good forecast [93].
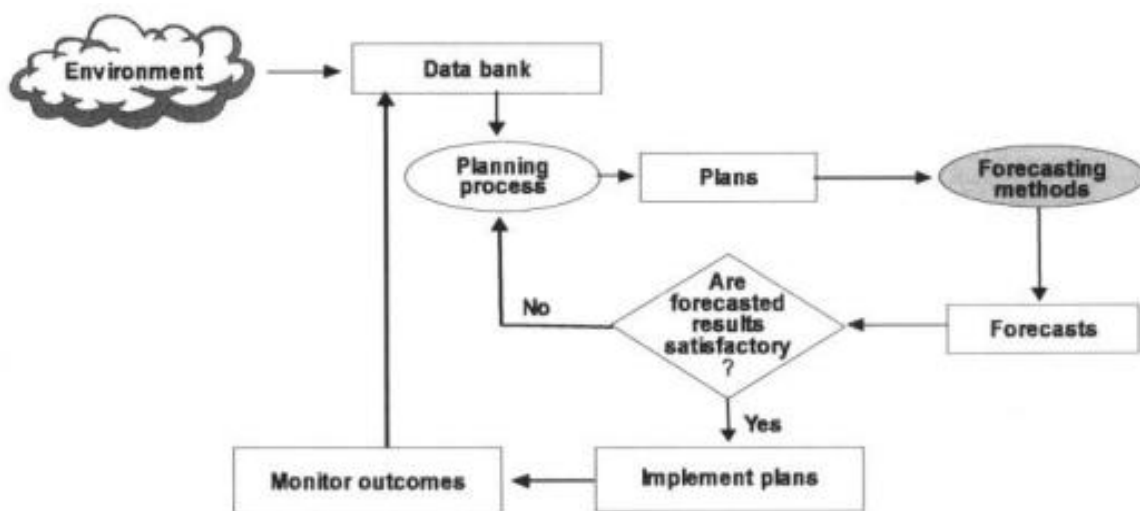
Figure 43: Forecasting and planning framework [94]

Elements that need to be fulfilled, in order to obtain good forecast are discussed below [95]:

- Timeliness: A good forecast should carry good time horizon. For example, situations like demand, or inventory cannot change in instant, but require certain time period in order to be estimated well.

- Accuracy: A good forecast is accurate. When evaluating forecast, accuracy should be mentioned in a quantified manner, so that user can know, the accuracy for future, and look for alternative techniques for forecasting if one is inadequate with accuracy.

- Reliable: A good forecast is reliable. Chosen forecast techniques should provide consistent results, which are not variating too much each time in accuracy and cause a user to feel insecure.

- Reasonable Units: Forecast should estimate future events based on relevant unit appropriate to a user, thus providing meaningful information.

- Writable: Forecasting results should be written and documented for concerned individuals.

- Simple to use: Complex forecasting techniques provide individuals challenge and often under-confidence in using for various activities. Hence, a simple technique is recommended according to the knowledge of a user.

There exist several techniques for forecast analysis, some of which are qualitative and some are quantitative. The classification of various techniques is shown in figure 44.

To make forecast analysis, following steps are undertaken [95]:

- Evaluate the goal of the forecast. Defining required target or time of need, help to determine, the level of detail, resources required, and level of accuracy needed.
- The accuracy in forecast usually decreases with the large time period. Next step, is to determine appropriate time horizon for forecasting.
- Choose the right appropriate forecasting technique. It is chosen on the basis of evaluating data. For example, for time series, the nature of curves is studied to choose right technique.
- Formulate the forecasting.
- Measure the accuracy of forecasting done. If found unsatisfactory, re-evaluate the data or chose a different technique.
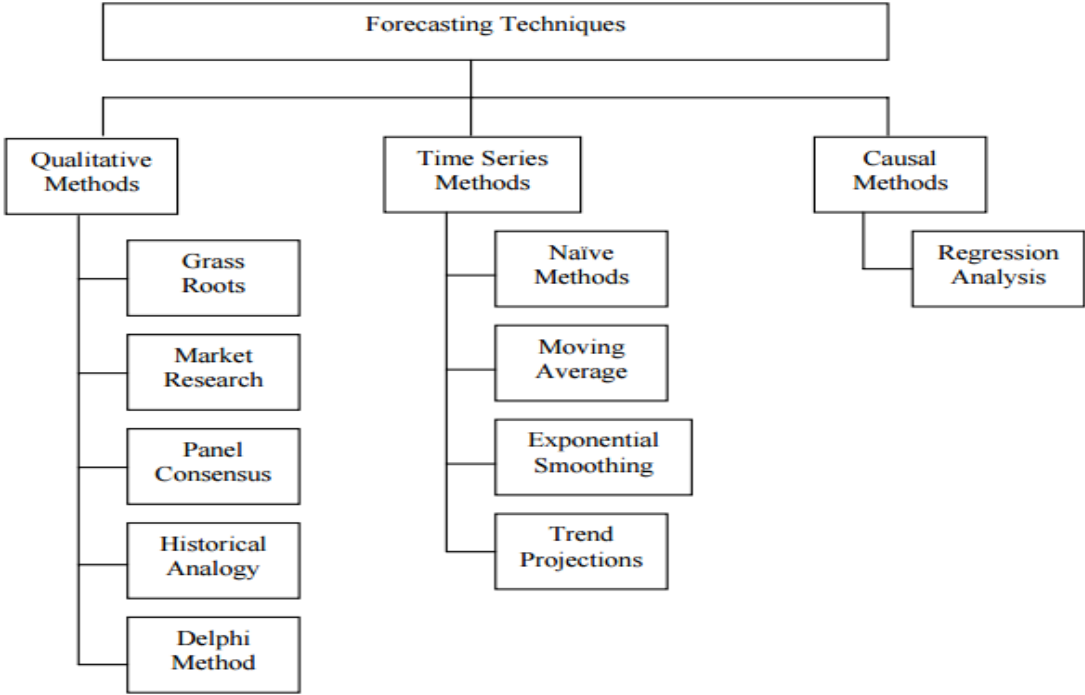
Figure 44: Forecast analysis techniques [93]

4.3.1 Time series patterns and graphs

Over consecutive time periods or points of time, the various sequence of measurements is made for a variable to produce observations called time series. These observations can be on the basis of hourly, weekly, monthly, yearly etc. The measurements may be taken every hour, day, week, month, or year, or at any other regular interval. The pattern of the data is an important factor in understanding how a time series has behaved in the past. If such behavior can be expected to continue in the future, we can use the past pattern to guide us in selecting an appropriate forecasting method. In order to examine the characteristic of recorded data, time-series graph is constructed. The variable values are marked along the vertical line of the graph and the time intervals on horizontal, based on

which data can be evaluated for future [96]. Various important patterns of data observed in these graphs are listed below [96]:

- Horizontal random patterns: When data revolves around some constant average, it follows horizontal patterns.
- Trend patterns: A time series variable value may increase gradually to relatively higher value or lower over long time period.
- Seasonality patterns: Seasonal patterns involves gradual increase or decrease in variable values in short period of time, repeating over every regular interval of time.
- Trend and seasonal patterns: Another form of patterns can be a combination of trend and seasonal patterns, which have increasing or decreasing curve over long time period, with repeating instant changes in regular intervals.
- Cyclic patterns: These occur when the time series plot depicts consecutive points alternating in nature, below and above around the trend line, over the long time period.

4.3.2 Forecasting analysis methods for time series

Various methods for forecast analysis exists based on time series, such as [93], [95]:

- Naïve method: Naïve method is the simplest in the time series forecast analysis. It takes the most recent value of time series and ignores all of the old values to estimate future event value. The problem with this method is that, since only one most recent value is considered, the high random variations in each time instant would produce inappropriate results.
- Moving average method: Moving average overcome the limitation of the Naïve method by considering longer time period for calculation of future event values. It takes a simple average of the most recent set of values. By using a simple average, the future event value is calculated. Greater the amount of observed values of the variant are taken, greater the smoothing of the curve is obtained, which is appropriate to do if curves are less fluctuating and smoother. However, with his method, some values like old ones are ignored, and each recent value considered for calculations, get equal priority, which is not always an appropriate technique.
- Weighted moving: Weighted average overcomes this limitation, by giving proportional values of weightage to each value chosen for average from time series. The weightage is usually given more to the most recent values and decreases with older values.
- Exponential smoothing: This method automates assigning the weightage of values for average over the whole time series curves, and thus, is well structured and simple method to use.

4.3.3 Exponential smoothing

Exponential smoothing is used widely in industry for being more accurate than other forecasting types discussed. There exist three types of exponential smoothing based on the nature of time series graph, given below [97], [98], [99]:

- Single or first exponential smoothing: For forecasting brief ranges such as one month, this type of simple smoothing method is utilized. The variable in time series graph fluctuates about a plausible stable average. It can be applied using following equation:

$$S_t = \propto X_t + (1 - \propto) \, S_{t-1}$$

Where S_t = forecast for upcoming period, $\propto$ = smoothing constant with value between 0 and 1 (lesser the value lesser the damping for smoothing), X_t = latest value known in time series, and S_{t-1} = most recent forecast made.

- Second or double exponential smoothing: It is utilized for forecasting time series with a trend in its curve nature. Thus, this smoothing requires not only updating level in previous period for each time but also trend. The level is a smoothed estimate of the value of data at the end of each period. For the end of each time, smoothing the average growth estimate is called a trend. It can be applied using following equations:

$$S_t = \propto y_t + (1 - \propto) \, (S_{t-1} + b_{t-1})$$
$$\text{Where, } b_t = \gamma(S_t + S_{t-1}) + (1 - \gamma) \, b_{t-1}$$

Where, $L_{t-1} = S_t$, in case of first exponential smoothing, is level for average estimate at t, T_t is trend estimate for t, and β act same like $\propto$ in case of second order smoothing with values between 0 and 1.

- Third or triple exponential smoothing: It is utilized for forecasting time series with trend and seasonality in its curve. It can be further divided to be applied for multiplicative type seasonality or additive type seasonality. Formulae for both are given below:

Multiplicative seasonality:
$$L_t = \propto \frac{y_t}{S_{t-L}} + (1 - \propto) \, (L_{t-1} + T_{t-1})$$
$$T_t = \beta \, (S_t - S_{t-1} + (1 - \beta) \, T_{t-1}$$
$$S_t = \gamma \, (y_t / L_t) + (1 - \gamma) S_{t-L} \, .$$

Forecast: $y_t = (L_{t-1} + T_{t-1} \, \text{k}) S_{t-L}$,

Additive seasonality:
$$L_t = \propto \, (y_t - S_{t-L}) + (1 - \propto) \, (L_{t-1} + T_{t-1})$$
$$T_t = \beta \, (S_t - S_{t-1} + (1 - \beta) \, T_{t-1}$$
$$S_t = \gamma \, (y_t - S_t) + (1 - \gamma) S_{t-L} \, .$$
Forecast: $y_t = L_{t-1} + T_{t-1} + S_{t-L}$,

Where, L_t = Deseasonalized level estimation, T_t = Trend estimation, S_t = Seasonal component estimation, $0 < \propto, \beta, \gamma < 1$ (level, trend, smoothing constants), L = Seasonal factor (time period for repetition of curve), y_t = Most recent value observed, S_{t-L} = Seasonal factor for period T evaluated one season with L periods before.

One of the most plausible ways to measure which exponential smoothing method type suits the time series graph for a particular case can be evaluated by calculating errors in smoothing the curve. This is done by Mean Absolute Deviation (MAD), Mean Squared Error (MSE), Mean Absolute Percent Error (MAPE), or some other similar metrics.

4.4 Quality gates

According to industrial experts, there can be three different types of gates processes for the projects for enhancement of the quality [100]:

- Stage gate: Stage gate are used for the development of product-related processes. In this type of gate system, the project is divided into various stages and gates. The various stages can be: scoping, build a business case, development, testing and validation and launch of the final product. The gates are decision points for each stage.
- Process gate: For the purpose of project management, processes for making a decision, process gate are used. It prevents the passing of defects through the project, by evaluating the various stages of the project.
- Quality gates: Quality gates has been used in various industries to improve their production quality, software quality etc. It was first originated from Europe and was initially implemented in product development processes, eminently in the quality processes of automotive industry, and was further implemented typically for project management and quality processes [101]. Quality gates concept is based on stage gate system. It involves converting a project or a process into various different phases. Checkpoints for quality phases are called the gates, are then placed between phases, for evaluating the status of the project to the plan or the quality of the product in the manufacturing process. After the completion of each gate, project proceeds from one phase to another [102]. Quality gates are part of milestones planning and help to make various decision in the project management. Gates are the decision-making points in the project, at which, gatekeepers (usually the part of quality management for the project), evaluate based on the status of the project to the target, to proceed the project or not, which help management uncover situation of the project to take required actions [103]. Moreover, the quality gates do not have to run one after another, but can also run parallel to each other [102]. According to John Aaron, '' A quality gate is a collection of completion criteria and sufficiency standards representing a satisfactory execution of a phase of a project plan'' [103].

Using the concept of quality gates, a project manager can build their project in such a manner, which could help them to deliver integrated reporting, monitoring the schedule, and evaluate the project progress with respect to not only the quality criteria but also the completion criteria, along with the whole life cycle of the product. Apart from providing a clear warning to project managers, if the desired progress for the project is made or not, as expected by various stakeholder, it acts as a ''forcing function''. This is because it prevents the project from various issues by making sure that project members complete the right things at a right time, protecting the project teams from the risk of performing work in incorrect sequence [103]. To understand the quality gate process, consider the diagram in figure 45.

4.4.1 Types of quality gates

Based on the level of decision in an organization, there are two types of quality gates [100]:

- Quality control gates: These gates are used for the operational level of an organization. The operational process is broken down into several activities, at the end of each activity, the quality check is made.
- Quality management gates: These gates are used for the management level of an organization. These gates help to make decisions which will influence the results of a strategy, project or program.

Various components of quality gates are as following [104]:

- Location: It defines in which part of an organization quality gates are established: operational or management.
- Content: Based on this element of quality gates, at any specific point of the process, quality can be evaluated, based on predefined criteria. It further consists of the following dimensions:
    1. People: Needs with respect to the level of skills, experience, and commitment of individuals involved in the activities of the process. It, thus, defines the resource requirements.
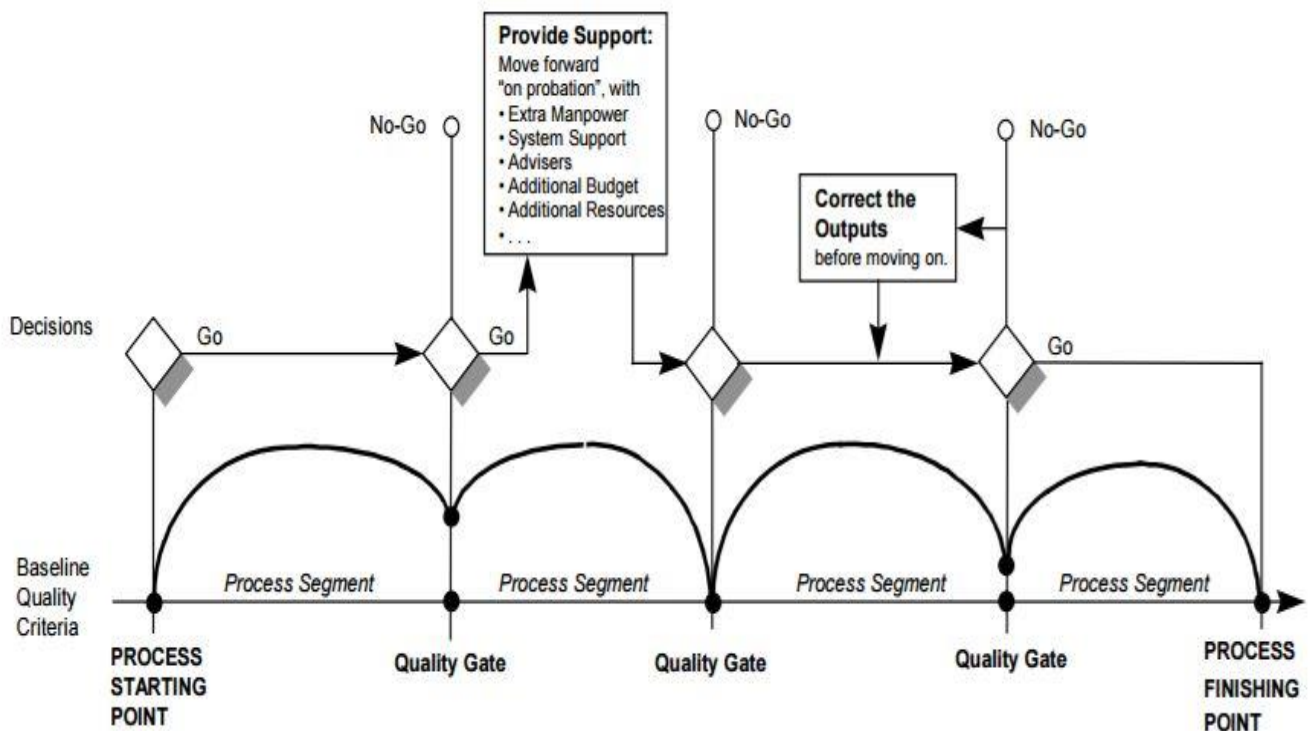
Figure 45: Quality gate process [100]

2. Criteria: Various attributes for product or process or service which are expected at a given point in process or program are defined under this element. It defines various output requirements based on some procedure.
3. Metrics: It is used to measure the performance of project or program, based on performance metrics.
- Process: How to control and make decisions in the process are defined under this component of the quality gates. It defines what happens if a product does not meet predefined criteria, based on various requirements. Rules for making decisions can differ based on the different management level.

Without the application of quality gate in a process, a process would be possible in a number of different ways, causing dramatic variation in time, speed, or performance for a process. However, implementing quality gates in a process would allow fewer variations in the process. It can allow various predictions such as if a process can pass through next quality gate or not.

4.4.2 Method for implementing the quality gates

Following elements can provide a good overview of implementing quality gates method for an organization project [104]:

- Satisfactory criteria of quality for products, output, and activities with various executive, departmental and operations level.
- The decision for policy for management, which ensures the correct application of the method.
- Well worthy metrics, for evaluating the degree of program success and degree of performance for the operation.

Following four phases of implementation for quality gates develop, implement and confirm the above-discussed elements [104]:

- Phase 1 – Set up the gates: For management of program or quality of a project, the criteria of quality and various gates for quality management are defined. Various potential needs for enhancement of quality are determined and related. It involves various task with management team like the current method for quality gates is evaluated and understood, various goals and milestones are determined with complete agreement and acceptance. Various commitments are derived with regard to planning, scheduling, defining roles, responsibilities, and requirements for performance, goal or expectations. A steering committee is assigned along with required resources. Various preparations for training are made for team members.
- Phase 2 – Train the team members: Team members are informed about the expected goals and structure of project and planning, to confirm their expectations. Various tools are provided for the use in the project.
- Phase 3 – Apply the method: Management makes meetings to analyses the performance or results of the project. Management then makes decisions according to predefines criteria. If the requirements are not met, responsibilities and corrective actions are determined and the solutions are documented.
- Phase 4 – Implement the corrective actions: Various measure are applied for the progress and to enhance the quality of the project. Improvements made are then analyzed and documented in a report.

There can also be certain disadvantages of gate processes. Implementing quality gates concept for project can create various challenges, such as, too many paperwork, additional efforts for team members to check the documents before they are passed for next step, extra time for team members to learn new tool and gather new information as a part of quality criteria of quality gate concept etc. [100].

Chapter 5: Applying DMAIC: Optimization of dataset management for calibration process at AVL

5.1 Introduction

The fundamental framework for research for the calibration process at AVL in this thesis encompasses DMAIC methodology from Six Sigma, including its various tools and various technical concepts to overcome current obstacles involved in the process. During the introductory period of research, it was realized that the process required improvement for quality. Considering the Juran trilogy graph, it was realized that the current process was in a situation that there was a rapid spike increase in the cost of poor quality control of dataset quality in the calibration process of AVL. Thus, among three sections of quality in the Juran trilogy, quality improvement was felt to be required for the process. Among, various methodologies of quality improvement, it was realized that Six Sigma DMAIC methodologies would suit the best in above situation of the process considering the reasons for applying DMAIC discussed in literature review. The process quality required improvement under growing complexities in the calibration projects, following the increasing complications in ECU software from suppliers and customer demands. For such complexities, the current solutions for dataset validation or data management were found to be insufficient for appropriate quality control. In order to make improvements in the calibration process, research work was required to:

- Observe for major areas in process for analysis, from wide range of sub-processes or activities involved in the process,
- Measure for ongoing quality in order to assess the advancement or progress in quality,
- Evaluate or study various significant complexities in the process with views of different detailing levels. This include carrying overall view to detailed view on each segment of process,
- Identify appropriate recommendations or feedbacks for improvements by working together with concerned calibration stakeholders who are working actually in the field, and
- Determine plan for further controlling in the process.

Considering above aspects and timeline of the thesis, it was realized that DMAIC methodology would fit perfectly in above scenario. The intentions were not to achieve Six Sigma by utilizing DMAIC, but to identify the solutions for a high degree of quality improvement, which could not be realized without precisely analyzing process from various levels. By levels of analysis, it means that the whole calibration process can be broken down into level 1, as a brief overview, to level 2, as a detailed overview of various sub-processes, and level 3, as a detailed overview of activities and tasks of each sub-processes which were found in level 2 detailing of the process. These levels were shown by BPMN diagrams, as flowcharts were felt to be providing insufficient information with respect to the complexities involved in the calibration process. On the level 3 detailing of calibration process using BPMN, 5 Whys analysis was made on activities and tasks of various selected sub-processes chosen to be improved. The root causes found on level 3

detailed analysis of various activities and tasks of sub-processes, were written down in a single pictorial diagram called Ishikawa root cause diagram. In order to study the calibration process for dataset management at AVL, following activities were carried out:

- Reading documents: Various process flowcharts, official documentation, and standards involved were comprehensively studied,
- Workshop: AVL CRETA tool, which possesses a significant role in dataset management for calibration process, was learned from various calibration pilots. Introductory knowledge of other software was gained from calibration engineers, such as for ETAS INCA and AVL CONCERTO.
- Interviews: A large number of interviews were carried from various calibration process stakeholders, involved directly in the process. This includes extensive interviews for analyzing the process and obtaining feedbacks from numerous calibration engineers, calibration pilots, lead engineers, quality management professionals, and other concerned employees. The measurements involved were defined by several years of experience of employees.
- Literature review: A large amount of literature was studied for the methodology. It includes various tools and various concepts involved for process improvement from various journals, books etc. This helped to know how others implemented the methodology and its tools in their organizations.
- Observations: Based on the literature review, the calibration process at AVL was observed and re-evaluated thoroughly,
- Results: Concepts behind recommendation actions for quality improvement in process, were proven by software prototype built with Microsoft GUI using C# by Visual Studio with CRETA API. Apart from the prototype, other concepts were proved by actual forecast analysis calculations on Microsoft Excel, and case studies for node-link diagram.

DMAIC is carried out with its each phase with the tools which are most suitable to the requirements. The main focus is on customers, for whom, the entire research is carried out for process improvement to provide the datasets with enhanced quality before delivering to the external customers like ECU suppliers or OEMs. The first stage is discussed below:

5.2 Define

The content of define phase utilizes the literature provided in 2.2.1.1.

5.2.1 Brief overview of entire calibration process: SIPOC

The calibration process is a wide process composed of various sub-processes. It is important to define the scope of the project due to time constraints. The entire project revolves around the customers of the calibration process. Firstly, a high-level view (level1) is carried out on entire calibration process using SIPOC to define overall circumstances of the process. After which, the view moves to next detailed level 2 to choose areas to improve, after which the further analysis of level 2 detailing is carried out to next detailed level 3 view in the analysis phase of DMAIC, for each sub-process chosen for improvement in level 2 view. The customers of calibration process are internal employees responsible for calibration as a part of the calibration process, and external clients who expect optimized dataset with no defects from the process. To provide above overall information

SIPOC table for calibration process is provided in figure 46. It is made according to the internal information provided by AVL officials.

5.2.2 Identification of problems from SIPOC brief overview (level 1 overview)

SIPOC provide a brief overview of the entire process. From above diagram, the first block of the process, which deals with suppliers creating ECU software and base dataset with basic optimizations are not covered in this research work. It basically deals with later blocks cases, especially, how calibration engineers manage datasets between various engineers, customers, and lead engineers since calibration process is a collaborative work between companies. How they optimize the datasets, covered in the third block, using test-beds, chassis dynamometer, hardware-in-the-loop or software-in-the-loop etc. with various tests using tools such as ETAS INCA, AVL CONCERTO is not covered in the research work either. But how calibration engineers manage the datasets, between each other, their lead engineers, and customers with the help of calibration pilots is a major concern of the research work, which covers the majority of calibration process work. While managing the datasets, several errors can occur from calibration engineers while producing changes to datasets with ETAS INCA or AVL CONCERTO. Thus, the second and third blocks of calibration process defined in SIPOC are of major concern. This includes iterative tasks with several calibration engineers working together on work divided between them, with lead engineers, customers, and calibration pilots.
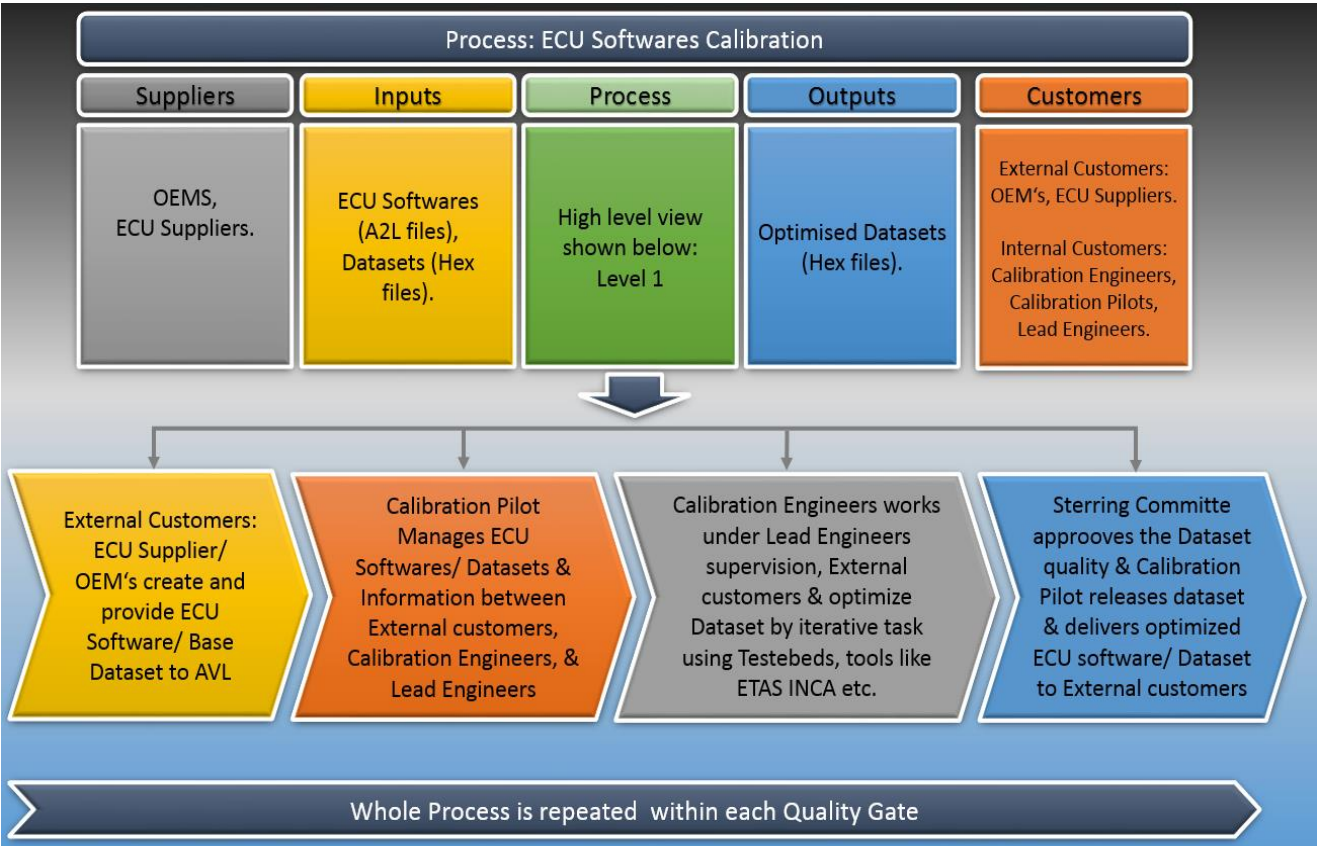

Figure 46: SIPOC diagram for ECU software calibration process at AVL.

The last block has little influence on calibration quality, as it possesses not much manipulation of datasets. The major causes of the decrease in quality of calibration process are:

- Errors in datasets (wrong values in labels of datasets),
- Mismanagement of information between calibration engineers, and
- Limited possibilities to create knowledge from data.

There are many reasons behind the occurrence of above issues in the process. They can be understood by studying process in more detail. Thus, the four blocks of process shown in SIPOC, are transformed to swim lanes in BPMN detailed level 2 view with various sub-processes which together build the blocks. By doing so, the major areas which possess high risks of delivering dataset with defects, in the last three blocks of SIPOC, can be identified. These major areas of dataset validation and management face above causes which can decrease calibration dataset quality delivered by AVL to its customers.

5.2.3 Identification of problems by detailed BPMN (level 2 detailed overview)

To identify which section to study in above three blocks of SIPOC for improvement, a thorough process diagram was made using BPMN. Processes do not exist in the physical state, and thus while making discussions, it becomes hard to explain about which parts of these complex processes produce issues. But with BPMN, the image of the process that people have in their mind was drawn into papers, and they could easily point out, which part need to be improved. Just like a ruined machine, where users can easily point out that this part of the machine is ruined. A flowchart was felt to be inappropriate for above situation, because dataset calibration is a complex process, involving various database, data, information exchange, and various complex sub-processes for dataset validations with various activities and tasks etc. The BPMN diagram is shown after some pages.

It is to be noted that above thesis considers calibration process for passenger cars. The calibration process for heavy automobiles like trucks, or heavy equipment machines for constructions etc. is different than these. Calibration project for passenger cars at AVL is carried out with 4-5 quality gates. The processes in each quality gate repeat themselves until required quality of dataset is achieved at the end of a gate. The BPMN shown is a very general representation of complex calibration project carried out in a phase, which possess the highest number of activities to be carried out to validate a dataset. The most important sub-processes and activities or tasks are well covered in it. It is to be noted that many activities and sub-process are not drawn to it, as it would make BPMN extremely compley and hard to read, and hence, irrelevant.

To understand the calibration process for level 2 detail, shown in BPMN in figure 47 (after next two pages), various roles of different groups of employees should be explained. Knowing their roles, various steps involved, and studying the process flow from BPMN can provide a well-structured approach toward understanding the process.

5.2.3.1 Calibration managers

This thesis considers calibration administrator and calibration pilots as calibration managers to simplify the process explanation.

The administrator has responsibility for managing AVL CRETA database. They set project when a customer sends project files such as ECU software and its base datasets. They then assign calibration pilot for the project. It is calibration pilot who has responsibility for effectively managing the datasets for calibration projects. They can also be nominated by

lead engineers together with other process managers. Calibration managers have following responsibilities:

- Manage datasets for calibration process (HEX file),
- Determine dataset merging process for particular project (DCM files with changes),
- Add users to project into CRETA,
- Load variant data structure into CRETA (on left of software with tree structure), software versions, datasets etc.,
- Assign work-packages to labels/parameters of ECU software (A2L file) and label responsibilities, (after discussion with project managers)
- Ensure conflict free merging of changes to datasets from customers and calibration engineers, and
- Organize meetings.

5.2.3.2 Calibration engineers

Calibration engineers download software/dataset (A2L & HEX files) from CRETA and carry out changes to their labels (assigned to them) using tools like ETAS INCA by making tests on chassis dynamometer or test beds etc. Labels can be in the form of curves, maps, or arrays etc. They are in the responsibility of changes to their labels, scoring them (providing labels/calibration maturity) their quality maturity and document their work using CRETA to enable traceability and understanding of changes.

5.2.3.3 Process managers

Process managers in this thesis are considered to include project managers, steering committee which further consists of lead engineers, project leader, skill team leader, and skill area managers.

Project managers bring an agreement with customers over label responsibilities, and later between internal calibration engineers at AVL. They maintain a close relationship with calibration pilots to manage various meetings and prepare reports of work status. Lead engineers work together with project managers for label responsibilities. Apart from responsibilities, they assign a particular work-package to related ECU software label. They have complete responsibilities for approving calibration for their work-packages and approve desired quality for a particular quality gate.

Steering committee gives final approval for work-packages at quality gates milestones in quality gates review meetings. Skill team leaders lead steering committee and its final decisions. Lead engineers approve work related to their work-packages only, project leaders are responsible for calibration team of project and satisfactory project execution, and skill area managers are responsible for every SOP projects.

5.2.3.4 Calibration process workflow: BPMN (level 2)

Final dataset release date for the customer are agreed with customers and internal employees at project beginning. Calibration pilot defines merging process and its timelines to make the merging the changes efficient. A conflict free dataset merging is a central

aspect of efficient calibration dataset validation and management. It means for calibration engineers following aspects:

- Calibration engineer uploads their calibrated labels with DCM files to central repository with required and relevant documentation,
- One calibration engineer communicates with others for their intentions in label changes before releasing the changes. This is because their labels may influence one another,
- They give each other feedbacks, before merging that if changes are fine for their situations. If a conflict occurs, they inform calibration pilot.

A generalized detailed overview of the process can be useful to understand the whole process as shown in BPMN (Level 2) diagram at next page. The customer creates an ECU software with the base dataset and sends it to AVL. As soon as calibration pilot receives the dataset and software, they upload it to the central repository and inform the process managers and calibration engineers. They study the software and dataset by downloading A2L and HEX files from a central repository using CRETA. Process managers then organize meetings, where label splits for responsibilities, between AVL and customer is performed. After this, labels are divided into work-packages. These work-packages belong to specific lead engineers, who assign labels to their team of calibration engineers. The above information of label work-packages and their responsibilities are updated in CRETA by calibration pilots. Calibration engineers with a well-structured project, undergo regular meetings for conflict-free data merging issues, as discussed on the last page with three points. This can result in separating several labels to handle them differently for dataset validation. For example:

- Switch labels: Some labels are switches with binary values. They generally require no test for a correct value, but calibration engineer's experience to define correct value. They are sorted from all labels in the dataset in such meeting, and their values are saved in Excel files, saved in project drive.
- VIP labels: These are labels which are highly sensitive and have a huge influence on other labels. Their value, range, and other data are defined in such meetings and are monitored by calibration engineers and calibration pilots.

Calibration engineers then work on their labels using tools like ETAS INCA upon test beds, chassis dynamometer etc. The changes to their labels with mutual agreements between them result in uploading changes as DCM files to CRETA. Calibration pilots then carry out dataset validation procedures which are vital for conflicts free dataset management, such as:

- Switch labels values change check,
- VIP labels values change check,
- Check for changes in AVL responsible or customer responsible labels which have a high impact on AVL responsible labels, by checking received and delivered datasets etc.,
- Manage any conflicts by informing customers or calibration engineers,
- Create new dataset revision, by merging DCM files produced by changes made by calibration engineers to their responsible labels etc.,
- Deliver new dataset revision to customer,
- Receive new dataset revision from customer and carry out checks in above points again,
- Create new customer based dataset revision on CRETA,

- Inform customer or calibration engineers for faults if found by validation methods.

The process of making changes to dataset continues with several such iterations between customers, calibration pilots, and calibration engineers until an optimal revision is created. This revision undergoes real life tests, one of which is durability run (others not considered in studies).

Durability run is composed of many repeated test runs for checking or proving that calibration of dataset meets customer requirements (discussed in analysis phase). If these tests do not produce satisfactory results, for example, satisfactory emissions of NOx, CO_2 values etc., the dataset is again sent to calibration engineers for further changes. However, if results are found within satisfactory ranges, the dataset is sent to quality gates review meeting to finalize the decisions using quality gates documentation made by using Microsoft Excel macros. Calibration process maturity can be monitored for current status with respect to target maturity for quality gate, using 'CRETA quality gate export' tool

After studying the calibration process using BPMN (to level 2) in figure 47, various sub-processes/ activities were found to cause a reduction in process quality. Sub-processes or activities in various swim lanes and their sub-lanes, which cause major problems are listed below: (Problem definition)

1. Under calibration pilots responsibilities:
   - Switch labels check
   - Check for received or delivered datasets
   - VIP labels check
2. Under calibration engineers responsibilities:
   - Durability run
3. Under process managers responsibilities (lead engineers, steering committee etc.):
   - Monitoring calibration process
   - Making decisions during quality gates review meeting using Microsoft Excel sheets.

It is to be noted that there are numerous quality issues due to a large number of errors found in sub-process under calibration engineers called testing datasets on test-beds. But it was excluded from the thesis, due to limited time issues, and to make DMAIC project not too long. Hence, quality optimization on test bed was not covered in this research work.

5.3 Measure

The content of measure phase utilize the literature provided in 2.2.1.2

The main goal of this section is to determine current calibration process quality for various selected sub-processes found to be highly significant and prone to errors during process execution. Other sub-processes would also be compared as well, which would prove the significance of selected calibration process elements at AVL. AVL did not maintain any history or statistics of a number of errors occurred in the calibration process. Hence, it was decided to use the know-how/ experience of employees to measure the quality of selected sub-processes by using FMEA. However, to utilize FMEA, risk priority number (RPN) measure has to be developed specific to the calibration process, which would provide satisfactory information about the significance of quality of sub-processes. RPN was felt
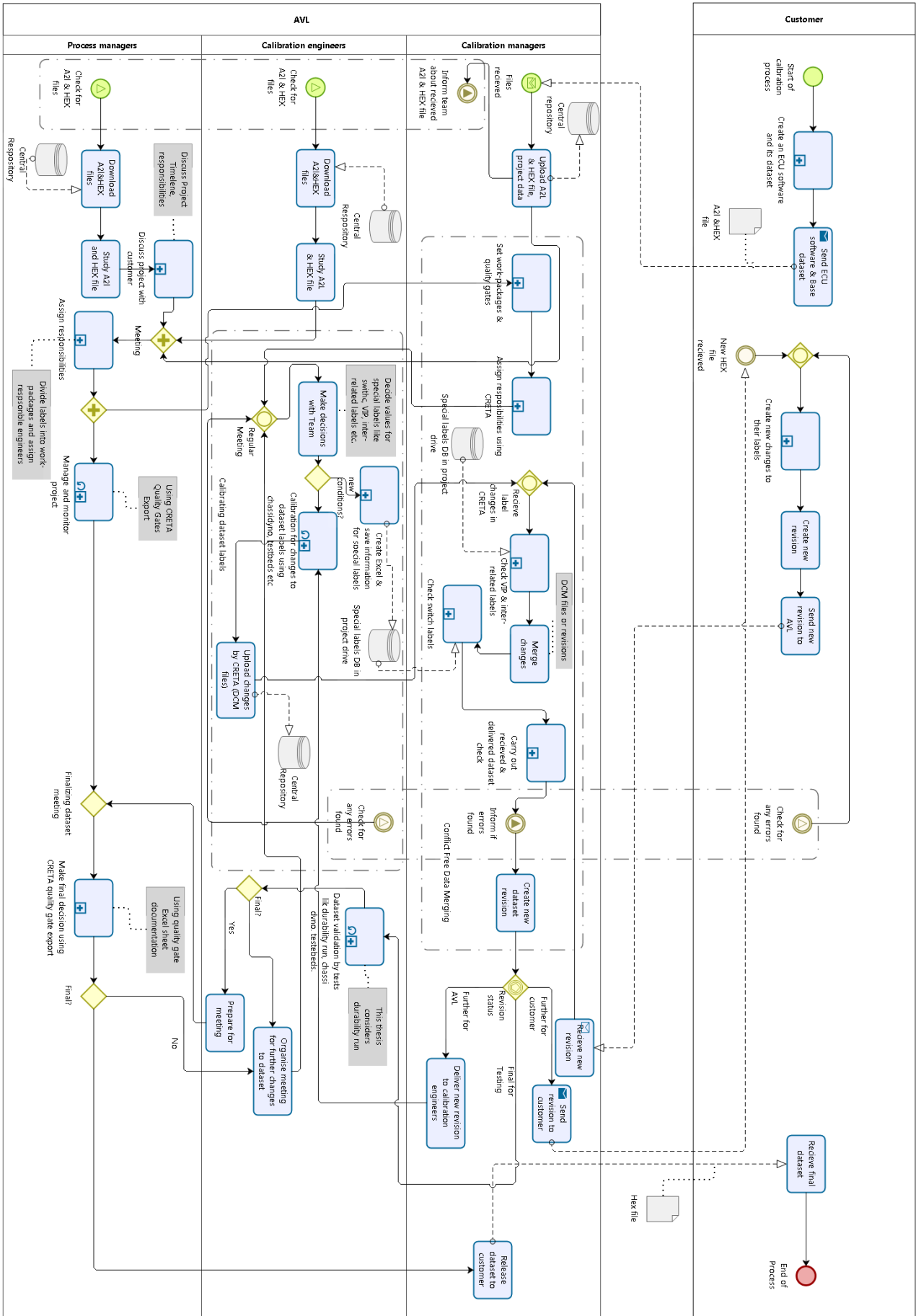
Figure 47 : BPMN of calibration process (level 2 detailing)

Page: 74

to be very effective measure, as it is based on the severity of poor quality or defects, probability/risk of occurrence of quality issues or defects, and detectability of occurrence of these issues through the process. These are suitable points to the quality issue of the calibration process, according to calibration team at AVL. Thus, firstly the components of RPN value are developed, based on which, the current process quality is derived in measurable numbers.

5.3.1 Development of RPN components value table

The suggested components of RPN value defined in SAE standard are not applicable to the calibration process but can be used to get a good idea for how to compose good RPN components for it. These were prepared in accordance with the quality issues in calibration process under the guidance of AVL employees. It is important to check the detectability of failure mode or inaccuracies or their mechanism, as further the error travels in the process, more are the losses to calibration teams. Currently, the quality issues are detected by some process check steps, ad hoc solutions, or have no solid solution at all for validation of datasets. Considering the current situation, table for detection evaluation criteria for FMEA was developed, as shown in figure below:

Calibration process detection evaluation criteria		
Detection	Likelihood of detection by current process design/ tools	Ranking
Absolute uncertainty	Current calibration process will not or posses no tool/ control system that can detect failure mode/ potential cause/ mechanism	10
Very remote	Current calibration process will very remotely detect failure mode/ potential cause/ mechanism	9
Remote	Current calibration process will remotely detect failure mode/ potential cause/ mechanism	8
Very Low	Current calibration process will detect failure mode/ potential cause/ mechanism with very low chances	7
Low	Current calibration process will detect failure mode/ potential cause/ mechanism with low chances	6
Moderate	Current calibration process will detect failure mode/ potential cause/ mechanism with low chances	5
Moderate high	Current calibration process will detect failure mode/ potential cause/ mechanism with moderate high chances	4
High	Current calibration process will detect failure mode/ potential cause/ mechanism with High chances	3
Very high	Current calibration process will detect failure mode/ potential cause/ mechanism with very high chances	2
Almost certain	Current calibration process will detect failure mode/ potential cause/ mechanism with almost certainty	1

Figure 48: Calibration process detection evaluation Criteria.

Higher the probability of occurrence of defects/ inaccuracies in the calibration process, lower the reliability for dataset produced by the calibration process. Hence, the occurrence of elements that cause quality issues possesses great significance for calibration process quality. As discussed, since no statistics or history of defects/ inaccuracies were available, figure 49 shows criteria that were developed under calibration team to effectively study the occurrence of defects/ inaccuracies.

Calibration process defects/ inaccuracies evaluation criteria		
Failure probability	**Likely failure rate for particular calibration sub-process or acitivity**	**Ranking**
Very high: continual failures	More than 100 per thousand events	10
	50 per thousand events	9
High: frequent failures	20 per thousand events	8
	10 per thousand events	7
Moderate: occasional failures	5 per thousand events	6
	2 per thousand events	5
	1 per thousand events	4
Low: relatively few failures	0.5 per thousand events	3
	0.1 per thousand events	2
Remote: failure is unlikely	Less than 0.1 per thousand events	1

Figure 49: Calibration process defects/ inaccuracies occurrence criteria

Calibration process severity evaluation criteria		
Effect	**Severity of effect**	**Ranking**
Hazardous with no warnings	Calibration process results in producing ECU software for vehicle which violates governmental regulations without any warning, effecting customer business by law violations	10
Hazardous with warnings	Calibration process results in producing ECU software for vehicle which violates governmental regulations with some warning /causes long project delay, effecting customer relation seriously	9
Very high	Calibration process results in producing ECU software resulting in inoperable vehicle, like due to loss of main function or causes project delay, effecting customer relation	8
High	Calibration process results in producing ECU software resulting in operable vehicle but with reduced performance level, or short project delays, enough for customer dissatisfaction	7
Moderate	Calibration process results in producing ECU software resulting in operable vehicle, but causes comfort issues, or project timing delivery issues and still makes customer dissatisfied	6
Low	Calibration process results in producing ECU software resulting appropriate results or project timings but customer faced often inconveniences	5
Very low	Calibration process results in producing ECU software resulting appropriate results or project timings but customer faced some inconveniences	4
Minor	Calibration process results in producing ECU software resulting appropriate results or project timings but customer faced minor inconveniences	3
Very minor	Calibration process results in producing ECU software resulting appropriate results or project timings but customer faced very minor inconveniences	2
None	Calibration process results in producing ECU software resulting appropriate results or project timings completely satisfactory to customer	1

Figure 50: Calibration process severity evaluation criteria

Severity would define how significant are these anomalies in the calibration process. It would help distinguish, which elements need more attention for quality optimization. Considering the current situation of the calibration process, the criteria was developed to analyze the severity of various elements as shown in figure 50.

5.3.2 FMEA for various sub-processes (based on level 2 detail of BPMN diagram)

Considering the above RPN components, the FMEA analysis was made on sub-processes in level 2 detailing of the calibration process in BPMN at AVL. The results of analysis are shown in table constructed in figure 51. It was decided that any failure mode throughout the whole DMAIC with RPN values more than 400 would undergo improvement process in above thesis. Thus, in this FMEA various sub-processes were confirmed to be worked on, with more than 400 RPN value. There were other sub-processes which were found to be inappropriate to work on this research work, also in define phase. Although, they possess some threat to process quality for delivered dataset to the customer, but found to have much less significance in quality improvement. This is because they possess low severity and RPN value, for example, sub-processes for assigning responsibilities, as shown in FMEA table. Thus, RPN value of 400 was deciding point for deciding to carry out improvement throughout the DMAIC for selected sub-process or activity of calibration process.

Another major goal for representing calibration process quality work with performance measure like RPN was that it would reflect how much improvement was made after carrying out improvement stage of DMAIC. The FMEA in figure 51, shows various areas to improve (Potential Failure Mode), their effect or result with the customer as focus, main causes of failure behind areas to improve the process, and current solutions available in the calibration process to deal with various areas defined in Potential Failure Modes section.

5.4 Analysis and Improvement

The content of Analysis and Improvement phase utilize the literature provided in 2.2.1.3 & 2.2.1.4

The next phase of DMAIC was to analyze these chosen sub-processes thoroughly. To do so, root cause analysis was made using 5 Whys and brainstorming on activities and tasks composed in selected sub-processes in level 3 detailing of BPMN, in order to gain knowledge about main reasons why these sub-processes of calibration process at AVL are unable to perform efficiently. The root cause analysis was done by interviewing officials involved in above sub-processes, with years of experience. Brainstorming session was made with everyone involved in them, in order to ensure that no aspect related to the process are left out. Furthermore, 5 Whys analysis was made in order to reach through the main reason why various activities or tasks that make up the sub-processes are unable to produce efficient results. These were all sorted out in a well-structured manner in an Ishikawa/Fishbone diagram. This enabled team to view all the elements of all the sub-process to be viewed in a single picture, which further helped to team communicate better in order to move further with the improvement of dataset validation and data management for calibration process quality.

			FMEA: Level 2 detailing								
Item: Calibration process			**Responsibility:** Amritpal Singh							**FMEA number:** 1	
Model: Current			**Prepared by:** Amritpal Singh							**Page:** 1 of 1	
Core Team: Amritpal Singh, Lead engineers, Calibration engineers, Calibration pilots										**FMEA Date (Ori** 10.10.2016	**Rev:** 1

Process Function	Potential Failure Mode	Potential Effect(s) of Failure	Class/Sev	Potential Cause(s)/Mechanism(s) of Failure	Occ	Current Process Controls	Det	RPN	Recommended Action(s)	Responsibility and Target Completion Date	Actions Taken
Calibration process	Unreliable switch label check.	Customer recieve dataset with incorrect switch labels values.	8	Incorrect dataset validation by Excel sheet due to various steps involwed in procedure.	7	Ad hoc Excel validation file.	10	560	To be found out.	Amritpal, before end of thesis.	To be found out.
	Unreliable VIP label handling.	Customer recieve dataset with VIP labels values which are incorrect or may influence other labels in wrong manner.	8	Too many notifications by CRETA tool leading to oral communication for managing VIP labels.	10	Process available but not used by calibration Team (tool usability issue).	10	800	To be found out.	Amritpal, before end of thesis.	To be found out.
	Unreliable recieved & delivered dataset check.	Customer recieve dataset with labels values	8	Current dataset validation procedure only take care of AVL responsible Labels. Customer labels that can influence AVL labels go undetected.	7	Ad hoc Excel validation file.	10	560	To be found out.	Amritpal, before end of thesis.	To be found out.
	Unreliable durability run test data.	Customer dissatisfaction with wrong dataset and test results delivered.	8	Wrong naming assigned to dataset from 'INCA recorder' after test.	10	Ad hoc Excel validation file.	10	800	To be found out.	Amritpal, before end of thesis.	To be found out.
	Incorrect monitoring for project status.	Effecting customer relationship with project delays.	8	Unreliable maturity rating for labels quality with no solution to measure future or past maturity.	10	CRETA quality gate export function.	10	800	To be found out.	Amritpal, before end of thesis.	To be found out.
	Quality gate template.	Customer missinformed about work-package work updates.	5	Incorrect informations in quality gate update Excel sheets.	8	Ad hoc Excel validation file.	10	400	To be found out.	Amritpal, before end of thesis.	To be found out.
	Assign responsibilities.	Wrong labels assigned causing problems in smooth workflow.	5	Wrong interpretation of decided responsibilities.	7	Responsibilities decided in meeting & assigned by CRETA	10	350			

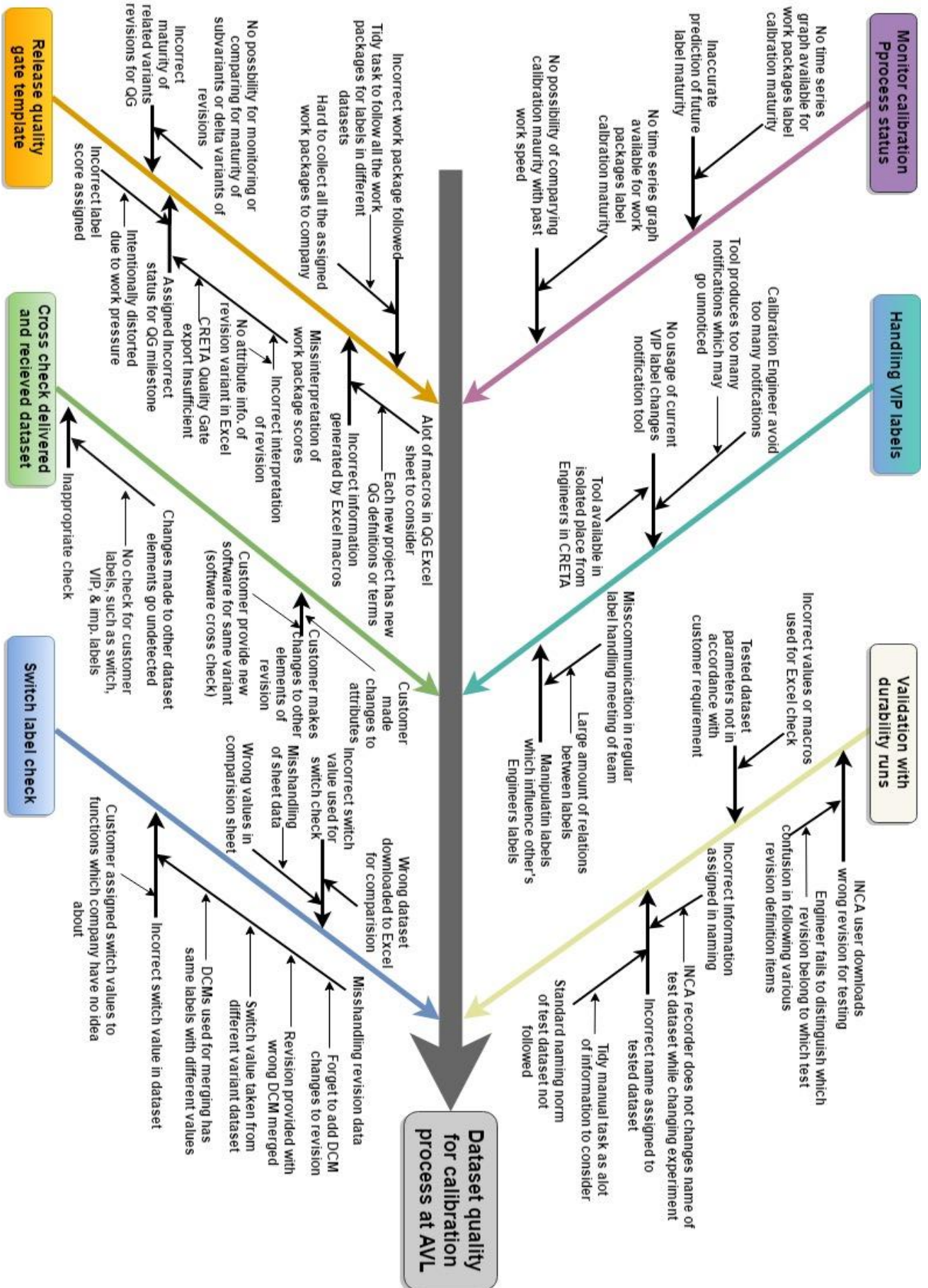Figure 51: FMEA ratings for current sub-processes at AVL

Figure 52: Ishikawa diagram for root cause analysis of activities and tasks for level 3 BPMN modeling detailing of calibration sub-processes at AVL

5.4.1 Sub-process 1: Switch labels check for dataset validation

Switch labels check is one of the important aspects of dataset validation for the calibration process. There are certain labels/parameters inside the ECU software dataset which are called switches, possessing binary values. They are checked by calibration pilot if their value is set to correct value or not, every time a new revision is received from the customer or delivered from AVL to the customer. Following are some important aspects of knowing about switch labels:

- These labels have value as integer type, varying as 0 or 1 only,
- It is not possible to distinguish them from other labels present in the ECU software, as there exist no standard definitions for their properties such as naming, size etc. when ECU suppliers create them. While carrying out switch check, calibration pilot has a list of names of switches in LAB file, based on which they know that among all the labels in software, which are switches. LAB files are the files which contain only the names of dataset labels. This list is prepared by calibration engineers at the beginning of the project when they study their labels at the beginning of the project. But however, some ECU supplier have some switch naming convention, for example, they add '_swt' to their label name to distinguish them from other labels.
- Incorrect value can cause the undesirable behaviors of ECU, or faults while testing the dataset. For example, one switch can alter the functioning of several other labels related to it. This can result in various performance issues of ECU software. A wrong switch which had to be turned on/off sends wrong results of various physical parameters like pressure, fluid inflow etc. while carrying out dataset tests, causing undesirable or unusual test results.
- Their values can be decided without any tests in test beds. This is because it is quite obvious that switch value is decided with 'yes or no'. This is, thus, done at the beginning of the project by calibration engineers. When they study their labels, they prepare a list with names for which labels are switch labels and decide their values based on their know-how. This list from each engineer is merged together to form single Excel sheet with all the switch labels of ECU software dataset with correct value.
- Switch check for dataset validation is carried out, generally, for the last dataset revision present in the variant in CRETA variant tree-view explorer. However, this is not always the case, as according to calibration pilots, it can sometimes be any revision present under the variant tree (a concept discussed later in the report).
- A specific vehicle variant is formed of specific features or elements called its attributes. For example EU6, rear/ front/ all-wheel drive, automatic/manual transmission etc. For those specific attributes forming specific variant, switch label value is always same. Thus, this forms the basis of the new concept. For example, a variant with attributes Euro 6, automatic transmission will have same value everywhere in dataset revisions belonging to a particular ECU software, but a variant with Euro 5, manual transmission and variant with Euro 6, automatic and 4 cylinders (more attributes, hence a different variant) will have different values of switches.

- Generally, when switch check is made, their values are mostly wrong. Hence, it is extremely important to check these values.

Before further analysis for improvement of switch labels check sub-process, it is important to know various aspects related to CRETA. When a project is created inside AVL CRETA, various variants folders for different vehicle variants are created in the tree-view on left for the project, as shown in figure 53 with dummy project data. This is done my loading Excel sheet from the customer, which contains this variant tree structure for the project. Thus, variant tree-view on the left of CRETA interface is structured on the basis of customer needs. This causes major issues for handling project data, since tree structure of variant in CRETA has no relationship with needs or how AVL handle project data. Various software and their dataset revisions are assigned to vehicle variants folder in tree view structure, which are created after making various changes to datasets, time after time. The calibration engineers download a dataset revision from CRETA to work on them. Since the variant tree on the left of CRETA is made according to the needs of the customer, calibration engineers have to manually sort through hundreds of variants folders in tree view to find out required variant for their work. The names of these variants folders cannot be changed since they are used for communicating with the customer for sharing information. Calibration engineers provide the changes to the dataset for optimization of revision and upload their changes to CRETA as DCM file. It contains only the labels with their name and changed values. To create a new revision of dataset, calibration engineers work together to create many DCM files, which all are merged by calibration pilots one after another until a new dataset revision is created. It is to be noted that different DCM files of changes to labels, cannot have same labels with different values. CRETA has functionality already available to check such anomalies. Normally, the DCM files from old revisions are deleted from CRETA as soon as the new revision is created.
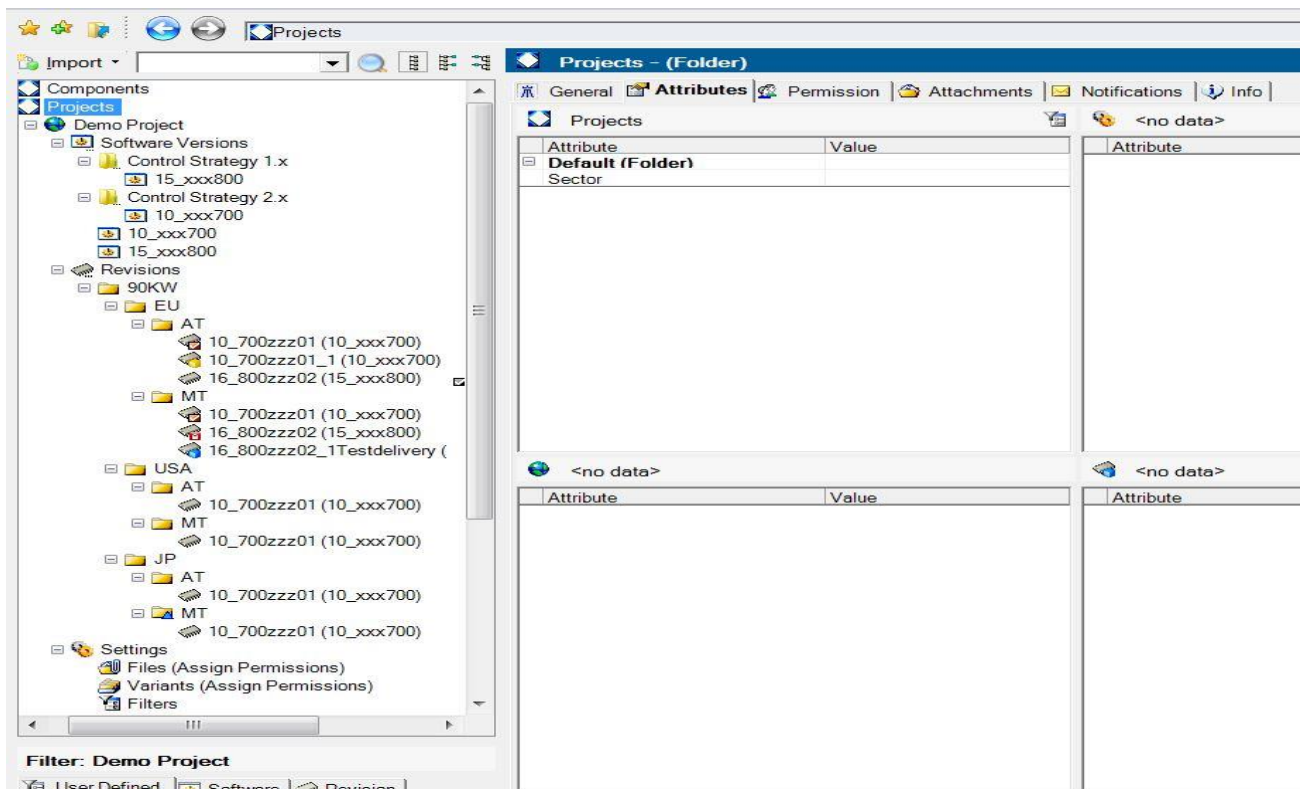

Figure 53: A basic tree structure for a dummy calibration project in CRETA.

The structure of variant tree depends upon a project to project, from company to company. As discussed, this causes data handling issues, as it is not standardized. The folder should be categorized with variant folder properties with attributes like Euro 6, automatic or manual (AT/MT), rear wheel drive (RWD), front wheel drive (FWD) etc. Secondly, the list of labels in datasets are random, i.e. no categories are assigned by data tags if they are switch labels, VIP labels etc. This is the main room of improvement for data management in CRETA.
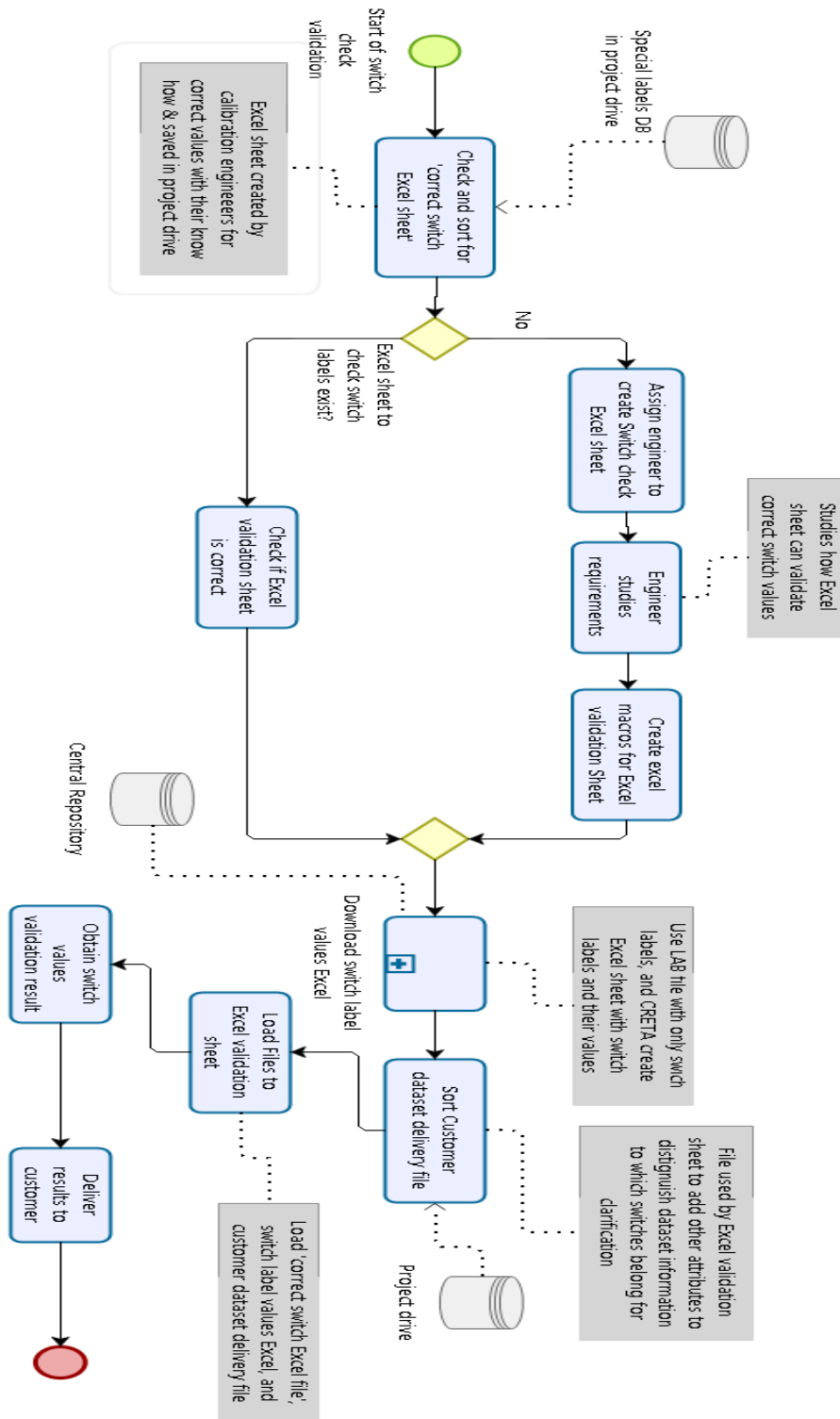

Figure 54: Switch check sub-process expanded from level 2 detailing to level 3 with various activities and tasks

As you can see in BPMN of switch labels check sub-process with various activities and taks in figure 55, current switch labels check validation require a lot of efforts. Above BPMN diagram can be understood as follow:

1. User first go to the calibration project tree-view structure on CRETA,
2. Look through the series of variants folder using correct variant path,
3. Sort through the correct variant folder, the required dataset revision among many dataset revision present for the variant. They choose a dataset revision in which switches values are to be checked,
4. User downloads the dataset revision, which contains switch labels, and save it to drive as an Excel file,
5. Then sort through project drive, 'correct switch Excel sheet' prepared by calibration engineers. This file contains correct switch label values defined by the know-how of calibration engineers.
6. After that, they sort through the project drive for customer datasets delivery report, which contains more attributes defined for variant dataset delivered by customer,
7. Then, calibration pilot loads these three files to Excel validation tool prepared by Excel macros. This Excel validation tool is created with no relationship with CRETA API. CRETA API are programmable functions to access CRETA database. Thus, current Excel validation tool has no connection with CRETA database and is stand-alone project specific ad hoc solution.
8. Excel tool compare/maps the dataset switches labels values to the Excel sheet containing right values, and prepare a template which notifies the customer about the wrong switches values.

Before developing improvement for switch check, it is important to analyze how calibration engineers construct the Excel sheet named 'correct switch Excel sheet' and how calibration pilot prepare other two files required for Excel validation tool file for switch labels check. This is an extension of sub-process in level 2 BPMN named: Create Excel and save information of special labels. Firstly, calibration engineers go through the project. They study its software, dataset, labels, and variants with different attributes etc. Calibration engineers get their labels assigned to them by their lead engineers. Each calibration engineers, decide their switch label value based on various variant type by their know-how. They know that variant with which attribute such as EU6 or EU5 etc., would have which switch value. Then, they bring together switch labels values data of entire team of calibration engineers and prepare such Excel sheet as shown in figure 55. It contains right switch values of the switches suggested by engineers with respect to various vehicle variants attributes. Basically, the main variant attributes which decide the calibration switch label values for a particular variant of the vehicle, are not always fixed, but are mostly as following:

1 Vehicle name,
2 Engine name,
3 Legislature, for example Euro 5, Euro 6 etc.
4 Transmission system, for example: automatic or manual.
5 Train type, for example: Front Wheel Drive, Rear Wheel Drive or All-Wheel Drive.

In the above example of switch label Excel sheet prepared (correct switch Excel sheet), the label column contains all required switches under consideration from each dataset from ECU software. Their list is decided by calibration engineers and contains labels/switches which are not even in the responsibility of AVL. The second column, which defines switch label responsibility and also contains customer responsible labels. Now, the value of each switch label is decided based on the vehicle variant they are being used in. For this reason, their values are categorized in other columns on right, based on their vehicle variant defined by various attributes, for example, vehicle name, engine name, legislation for emission etc. In the table in figure 55, vehicle variant with engine name 'Engine_1' with legislation standard 'EmissionLegi_EU3' has a different value of switches than variant with attributes for engine and emission legislation 'Engine 2' and 'EmissionLegis_EU5'. Please notice that more specific vehicle variant, with same attributes as another variant, but also contains more attributes for its definition, would have different switch value. For example, switch value for the variant with engine name, emission legislation, train type named- 'Engine_1 EmissionLegis_EU3_Traintype_1 (Automatic)' would have a different value than vehicle variant with a variant with same but even more attributes defines for it- 'Engine_1 EmmisionLegis_EU3_Traintype_1_Transmission_1 (transmission is an additional attribute).

	Label	Responsibility	Nr.	Engine_1 EmissionLegis_EU3	Engine_1 EmissionLegis_EU3_Traintype_1	Engine_2 EmissionLegis_EU5	Engine_3 EmissionLegis_EU6	Engine_4 EmissionLegis_EU6_Transmission_1	Engine_5 EmissionLegis_EU3
1	Label	Responsibility	Nr.	Engine_1 EmissionLegis_EU3	Engine_1 EmissionLegis_EU3_Traintype_1	Engine_2 EmissionLegis_EU5	Engine_3 EmissionLegis_EU6	Engine_4 EmissionLegis_EU6_Transmission_1	Engine_5 EmissionLegis_EU3
2	Switch_1	AVL	1.	1	1	1	1	1	1
3	Switch_2	AVL	2.	0	0	0	0	1	0
4	Switch_3	AVL	3.	0	0	0	0	0	0
5	Switch_4	AVL	4.	0	0	0	0	0	0
6	Switch_5	Customer	5.	0	0	0	0	0	0
7	Switch_6	AVL	6.	0	0	0	0	0	0
8	Switch_7	Customer	7.	1	1	1	1	1	1
9	Switch_8	Customer	8.	0	1	0	1	1	1
10	Switch_9	Customer	9.	1	1	1	1	1	1
11	Switch_10	AVL	10.	1	1	1	1	1	1
12	Switch_11	AVL	11.	0	0	0	0	0	0
13	Switch_12	Customer	12.	1	1	1	1	1	1
14	Switch_13	Customer	13.	1	1	1	1	1	1
15	Switch_14	Customer	14.	0	0	0	0	0	0
16	Switch_15	Customer	15.	1	1	1	1	1	1
17	Switch_16	AVL	16.	1	1	1	1	1	1
18	Switch_17	Customer	17.	1	1	1	1	1	1
19	Switch_18	Customer	18.	0	0	0	0	0	0
20	Switch_19	Customer	19.	1	1	1	1	1	1
21	Switch_20	Customer	20.	1	1	1	1	1	1
22	Switch_21	Customer	21.	1	1	1	1	1	1
23	Switch22	Customer	22.	0	0	0	0	0	0
24	Switch_23	Customer	23.	0	0	0	0	0	0
25	Switch_24	Customer	24.	0	0	0	0	0	0
26	Switch_25	Customer	25.	1	1	1	1	1	1
27	Switch_26	AVL	26.	1	1	1	1	1	1
28	Switch_27	AVL	27.	1	1	1	1	1	1

Figure 55: Excel sheet prepared by calibration engineers for the switch values from their calibration know-how.

Whenever customer deliver their datasets to AVL, they provide the datasets information in an Excel sheet which contains information about which variant folder (in AVL CRETA variant tree-view) it is contained. For the last and the third file, the calibration pilot makes a LAB file, which is a file containing only the names of labels of datasets. The required LAB file contains only the switch labels names, which are required to be checked. The calibration pilot then prepares a report called revision difference as follow:

1. Click on 'report' on CRETA interface banner, and choose dataset revision difference (enthropie) report,
2. Select the dataset revisions by going through variant tree-view structure manually to each dataset/revision,
3. Then select the LAB file (containing switch labels name) to make a report only for switch labels value and press start to get the report by CRETA.

CRETA then creates a file containing the switch label names and their values in selected dataset revisions in different columns, in one Excel file. When calibration engineer saves it, it looks like an Excel sheet as in figure 56. This Excel sheet structure is standardized by CRETA. But the Excel sheet which customer delivered with their dataset information is not standardized, as every customer have their own way of preparing such files. The current validation process uses an Excel tool made by using various Excel macros. These macros read the Excel file made by CRETA with switch values of all the revisions selected (as in figure 56) and also reads the Excel file provided by the customer for information of delivered dataset revisions.

Figure 56: CRETA Variant Revision Compare Excel sheet, made by listing various concerned revisions and LAB file

It is interesting to know how Excel validation file macros work. The customer Excel sheet is loaded and converted into the form shown in figure 57. In the first column, there are terms, which is unique identity of each dataset revision. So, a dataset revision made by customer or company both have this unique term in its name. It is coded so that it provides

variant folder path information in CRETA variant tree-view, in final delivered sheet for results of switch check for the customer. When the company makes changes to customer delivered dataset revision, they change its name, except this unique term. This unique term is like a SUID in dataset name generated by the customer. The Excel sheet in figure 57 contains 'columns with attributes names' for vehicle variants, related to columns with 'attributes names for vehicle variants' in figure 55, and helps match/map information of customer delivered dataset sheet to 'correct switch Excel sheet' prepared at very beginning by calibration engineers with their know-how. The Excel dataset validation tool then maps the Excel file with dataset revisions delivered by customer and 'correct switch Excel sheet' with correct switch label values. The final result file contains corrected switch label values and additional information about variant folder path in CRETA variant tree-view defined by various attributes, from customer delivered dataset sheet. This final result file made by Excel switch label validation tool can be delivered to the customer as results of switch check, which contains a table of switches with wrong values in the dataset.

There is one important aspect to note about how calibration engineers code the 'correct switch Excel sheet'. They code it such way that its table is made compact/small as possible. This can be explained with an example in figure 58. The diagram shows the set of vehicles variants with same values for switch labels. All the variants which have attributes engine and emission legislations – 'Engine_1 and Emission EU3' will have same switch value, like variants as follow:

- Engine_1 EmissionLegis_EU3 Train_AWD (additional train type attribute),
- Engine_1 EmissionLegis_EU3 Train_FWD (additional train type and transmission type attribute),
- Engine_1 EmissionLegis_EU3 Transmission_AT (additional transmission attributes).

	A	B	C	D	E	F	G	H	I
2	Variant Folder_1	Engine_name_1	Engine_1	EU3	EU3	Vehicle Model_1	AWD	AT	1
3	Variant Folder_2	Engine_name_2	Engine_2	EU5	EU5	Vehicle Model_2	AWD	AT	1
4	Variant Folder_3	Engine_name_3	Engine_3	EU6	EU6	Vehicle Model_4	AWD	AT	1
5	Variant Folder_4	Engine_name_4	Engine_4	EU6	EU6	Vehicle Model_6	AWD	AT	1
6	Variant Folder_5	Engine_name_5	Engine_5	EU3	EU3	Vehicle Model_3	FWD	AT	1
7	Variant Folder_6	Engine_name_6	Engine_1	EU5	EU5	Vehicle Model_5	FWD	AT	1
8	Variant Folder_7	Engine_name_7	Engine_1	EU6	EU6	Vehicle Model_1	FWD	AT	1
9	Variant Folder_8	Engine_name_8	Engine_2	EU6 Japan	EU6	Vehicle Model_3	FWD	AT	1
10	Variant Folder_9	Engine_name_9	Engine_3	EU3	EU3	Vehicle Model_1	AWD	MT	1
11	Variant Folder_10	Engine_name_10	Engine_4	EU5	EU5	Vehicle Model_2	AWD	MT	1
12	Variant Folder_11	Engine_name_11	Engine_5	EU6	EU6	Vehicle Model_4	AWD	MT	1
13	Variant Folder_12	Engine_name_12	Engine_1	EU3	EU3	Vehicle Model_6	FWD	MT	1
14	Variant Folder_13	Engine_name_13	Engine_2	EU5	EU5	Vehicle Model_1	FWD	MT	1
15	Variant Folder_14	Engine_name_14	Engine_1	EU6	EU6	Vehicle Model_5	FWD	MT	1
16	Variant Folder_15	Engine_name_15	Engine_1	EU3	EU3	Vehicle Model_6	AWD	AT	1
17	Variant Folder_16	Engine_name_16	Engine_1	EU3	EU3	Vehicle Model_3	AWD	AT	1
18	Variant Folder_17	Engine_name_17	Engine_1	EU5	EU5	Vehicle Model_5	AWD	AT	1
19	Variant Folder_18	Engine_name_18	Engine_2	EU5 Belgium	EU5	Vehicle Model_1	AWD	AT	1
20	Variant Folder_19	Engine_name_19	Engine_3	EU6	EU6	Vehicle Model_3	AWD	AT	1
21	Variant Folder_20	Engine_name_20	Engine_4	EU6	EU6	Vehicle Model_1	AWD	AT	1
22	Variant Folder_21	Engine_name_21	Engine_5	EU6	EU6	Vehicle Model_6	AWD	AT	1
23	Variant Folder_22	Engine_name_22	Engine_1	EU3	EU3	Vehicle Model_3	RWD	AT	1
24	Variant Folder_23	Engine_name_23	Engine_2	EU3	EU3	Vehicle Model_4	RWD	AT	1
25	Variant Folder_24	Engine_name_24	Engine_5	EU5	EU5	Vehicle Model_4	RWD	AT	1
26	Variant Folder_25	Engine_name_25	Engine_1	EU5 Belgium	EU5	Vehicle Model_4	RWD	AT	1

Figure 57: Sheet used to map customer variant information with Excel sheet on figure 55

However, if in 'correct switch Excel file' (Excel sheet with correct switch value), a separate column is made for switch values for vehicle variant defined with additional attributes like Engine_1 EmissionLegis_EU3 Train_RWD (same attributes as 'Engine_1 and Emission EU3' but with one additional attribute i.e. train type), then the switch values for these variants will be different as shown in the set diagram in figure 58. The sets switch values of these variants are concurrent but not same. The other condition is obvious that, if engine name and legislation are different in Excel sheet, then switch values will be different.
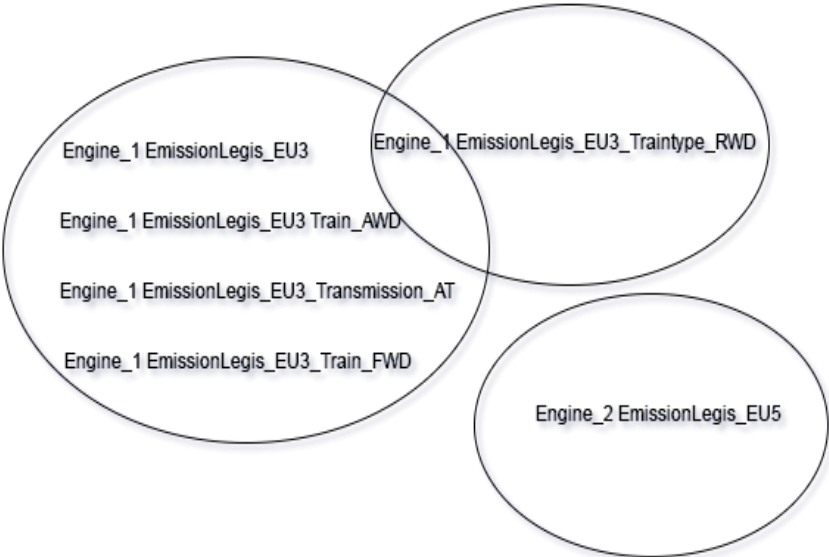

Figure 58: Shows the 'same/not same' set of switch values for different variants.

Calibration engineers code the names of engine attributes for 'correct switch Excel sheet'. For example, 'Engine_1' attribute name could be same for engines with name Engine_name_1, Engine_name_6, and Engine_name_7 etc. They did this to cluster the columns in the table of Excel sheet with correct switch values, to make it as compact as possible. This can also be considered as data classification of variant, as the variants with same switch values are grouped/categorized together. The figure 57 of Excel sheet shows how the coded names of switch variant are de-coded in order to map the information to the table of Excel sheet with correct switch values.

5.4.1.1 Various failure modes in switch labels check validation method

Switch check involves the use 'correct switch Excel sheet' prepared by calibration engineers, which contains the correct switch label values as discussed before. It is possible that this sheet might possess wrong values, due to various miss handling of data inside it. For example, typing mistakes, or confusion in choosing correct cell for a particular value etc. This can cause the sheet to contain some wrong values, as a consequence, when the switch check is made, it is made with wrong values. This would result in delivery of switch check sheet to the customer with wrong values, causing the customer to feel less reliable on the company. There is no step in the current process to detect such issues before the dataset is delivered to the customer. In current processes, when wrong values are delivered to the customer for switch labels, they detect it and give feedback to AVL. Moreover, until the customer corresponds, the wrong values would be used for switch

check in the calibration process. Incorrect switch values used for switch check involves another issue that the calibration pilot downloads wrong dataset revision for switch check. However, this would be less severe as it can be found by calibration pilot itself in the current Excel validation process that the wrong revision dataset was chosen, by Excel switch value validation macros (macros does this by using Excel sheet of delivered dataset provided by customer, by comparing its data with by looking at downloaded datasets variant path in a variant tree-view by unique term assigned to dataset name). However, the current process consists of several manual steps in order to complete the process of switch check. This causes the process to be slower, as there arise requirement of rechecks, again and again, causing a lot of rework. This further results in late delivery of switch check results to the customer.

There are many reasons why there exist incorrect values of switch labels in the datasets or revisions. Some of them are listed below, as further explanation of points in Fishbone diagram:

1   Calibration pilot, either from customer side or company, forgets to add label changes with DCM file to dataset revision. A dataset revision can contain many DCM changes attached to it. In CRETA, they are marked checked when merged.
2   Calibration engineer uploads/provide wrong DCM file (file with label changes) for dataset revision. The calibration pilot unknowingly merges the wrong DCM file to the dataset revision to create new dataset revision, causing new dataset revision to contain wrong values.
3   DCM file can contain same labels with different values. This can cause wrong values of switch labels merged to revision. CRETA can already detect such issue, but, at the customer end, this might not be possible because they might be using calibration dataset management software with no such functionality.
4   It can happen that wrong files are merged from different dataset revision, resulting in wrong values of switch labels in new dataset revision.

All such issues are part of mishandling of dataset revision data as mentioned in Fishbone diagram. It often happens that customer has set switch label values for various functions, to which company's calibration engineer are not concerned about. This is due to lack of communication between the customer and AVL (low process integration). This results in setting the wrong values of switch labels in 'correct switch Excel file' in figure 55. This issue is resolved when AVL delivers their resultant switch check sheet to the customer. Customer finds the wrong values (which calibration engineers of AVL thinks are correct), and resolve the issue by explaining AVL calibration engineers about correct switch values according to them. However, it takes times and such error can surpass in other dataset revision switch check until the feedback of customer is received. Faster the correct switch values are known, better would be the results.

5.4.1.2 FMEA analysis for switch labels check sub-process

Based on the analysis of Fishbone diagram, developed by various brainstorming and 5 Whys methods on level 3 BPMN process modeling detailing, various failure mode and causes of failure were written down to FMEA table. Their effect with respect to calibration process customers is written down related to a failure mode in switch label check. Later,

the current process controls or validation solutions are explained in the table, related to the failure mode. The table is chosen from SAE standard J1739 2002. The various effects of failures and RPN values were decided by a team of employees involved in switch labels checks. RPN value help to distinguish significance of each failure mode and helps decide team to work for its improvement or not. The table is shown in next page in figure 59 and after which, explanation of recommendation action is provided.

5.4.1.3 Improvement/ recommendation actions for switch labels check sub-process

After studying the current process of switch check carefully with respect to various defects that can occur, along with their failure mechanism, many important factors can be determined. These factors would help to further improve the process to make it more reliable by applying process automation to the process.

A good solution to solve the current process problems is to improve the data quality. Data quality can be improved by improving the data management. This can be achieved by classifying data and creating a program that can handle the data based on the classifications and relationships between each classified data. By doing so process automation would be achieved. Thus, for switch check, classifying vehicle variants folders in CRETA tree-view by assigning them attributes like engine name, emission legislation, train type etc. (data tagging), data classification can be achieved. This can be done by just selecting a variant folder in CRETA variant tree-view, and assign them the required attribute. Such data classification would enable grouping of similar variant folders and help sort/ search variants by their attributes (data tag) by an automated program.

Apart from variants folder, data classification shall be applied to labels list in the dataset as well. There can be up to 100,000 labels for a particular variant in a dataset. Switch labels can be distinguished by selecting them from switch label list of the dataset, and saved in CRETA database. This would result in data classification of dataset labels as well. Such classification can enable distinguishing dataset labels in groups as switch labels, VIP labels etc. Automation program will then enable searching for variant by their attributes (data tag), and decide switch value for a particular variant dataset with respect to their attributes, as discussed before, for all the switch labels defined among thousands of labels. The detailed description of various recommended action to overcome various factors causing various defects in the calibration process for switch label check for current validation process are discussed below:

- The current validation process encompasses series of manually done steps to cross check the switch values in the dataset revisions with 'correct switches Excel sheet'. However, if 'correct switches Excel sheet' used for cross checking the switch values itself contains wrong values due to certain reasons, the cross check would be invalid until customer catches the incorrectly set switch values. To overcome it, there shall be a possibility that calibration engineers can double check the 'correct switches Excel sheet' with a certain amount of datasets, before officially releasing it. It would be possible for calibration to confirm the correct values themselves if they can instantly cross-check the values for dataset revision with help of automated solution (wrong switch values can be highlighted as detection type Poka-Yoke). With current validation process which requires many steps being

manual in nature, such checks are not possible instantly. Moreover, using same Excel validation file is not possible if the project is new, as the Excel macros are project specific, and required to be changed as soon as the project changes. Thus, currently there exist no such possibility to confirm the Excel sheet with correct switch values before it is released.

FAILURE MODE AND EFFECTS ANALYSIS

Field	Value	Field	Value
Item:	Switch label check sub-process	Responsibility:	Amritpal Singh
Model:	Current	Prepared by:	Amritpal Singh
Core Team:	Calibration pilots, Calibration engineers, Lead engineers, Amritpal Singh	FMEA number:	2
		Page:	1 of 1
		FMEA Date (Orig.) 17.10.2016	Rev: 1

Process Function	Potential Failure Mode	Potential Effect(s) of Failure	Sev	Class	Potential Cause(s)/ Mechanism(s) of Failure	Occur	Current Process Controls	Detec	RPN	Recommended Action(s)	Responsibility and Target Completion Date	Actions Taken	Sev	Occ	Det	RPN
Switch label check sub-process	Incorrect switch values in 'correct switches Excel sheet' used for switch check.	Late detection of wrong values resulting in customer to feel less reliable on company.	8		Wrong values decided in 'correct switches Excel sheet' due to data mishandling.	6	No process available to check 'correct switches Excel sheet' causing to customer to detect and fix wrong values.	10	480	Possibility of double checking results by company engineers due to faster and automated software tool (Concept: Process automation achieved by data classification through data tagging)	Amritpal Singh, to be completed before end of thesis.	Creation of prototype using C# with CRETA API, to proove concept behind recommendations.	4	7	3	84
	Incorrect dataset revision used for switch check.	Requirement of recheck causing rework and late delivery of switch check test results.	5		Wrong dataset chosen to Excel switch check validation sheet to compare.	8	Revisions manually chosen from variant folders in CRETA variants tree-view.	10	400	software tool automatically search required revisions for switch check (Concept: Process automation achieved by data classification and data tagging)	Amritpal Singh, to be completed before end of thesis.	Creation of prototype using C# with CRETA API, to proove concept behind recommendations.	4	7	3	84
	Incorrect values of switches in responsible dataset delivered by customer.	Many company switches delivered by customer incorrect, requiring rework	4		Mishandling of data in revision dataset by customer.	10	Checks using various manual process steps on Excel, resulting in switch checks being made not very often, causing dataset having wrong switch values in meantime.	10	400	Possibility of instantaneuous check as soon as changes are made to revision using software tool (Concept: Process automation achieved by data classification and data tagging)	Amritpal Singh, to be completed before end of thesis.	Creation of prototype using C# with CRETA API, to proove concept behind recommendations.	4	7	3	84
	Incorrect switch values in 'correct switches Excel sheet' used for switch check.	Customer dissatisfaction	8		Customer Assign switch values to function, for which company have no idea about causing wrong values in 'correct switches Excel sheet'.	6	Correct switch values decided by engineers with limited involvement of customer.	10	480	Customer frequently involved by possibility of faster switch check (Concept: Process automation achieved by data classification and data tagging)	Amritpal Singh, to be completed before end of thesis.	Creation of prototype using C# with CRETA API, to proove concept behind recommendations.	4	7	3	84

Figure 59: FMEA analysis of switch label check sub-process for defects found by 5 Whys analysis of activities and tasks in it (level 3 BPMN detailing)

- To check the dataset revision's switch label values, the dataset revision's variant folder has to be manually searched through CRETA tree-view structure. As discussed, using automated solution, which classifies variants in tree-view using data tags, variants can be easily be searched using data tags/attributes. This would reduce the need of manual search, and also the time needed to search the variant with required revision. The instantaneous results produced would improve customer trust on the company. This solution is prevention type of Poka-Yoke, as users are prevented from searching the wrong dataset manually.
- The main function of switch check is to find out incorrect values present in the switch labels of dataset revisions. There shall be a possibility for instantaneously checking of switch values as soon as label changes are made/merged to dataset revision using DCM file. Wrong switch label values can be highlighted to user. This can be a detection type Poka-Yoke. Currently, due to various steps involved in the switch check process, it is not done. New solution should enable wrong values notification in an automated manner. If customer changes switch values in dataset revision to wrong, with new DCM file merging, the changes could be instantly notified and corrected (data cleaning). Moreover, correction can be made to 'correct switches Excel sheet' if the customer says the label changes and the switch values decided by them are correct or customer can be notified if they are wrong in setting the switch values in new changes to dataset revision. Hence, providing good communication between teams.
- With an automated solution, the customer can be instantly be involved with switch checking process, the problems like switch labels value setting for functions that company doesn't know about, can be overcome faster in the process, causing the process to be more reliable.

The solution of data tagging for data classification the variant folders in the variant tree view and labels lists should follow the procedure of PDCA for data classification as discussed before, for efficient classification in the process. Data cleaning for datasets labels value in CRETA database and wrong label values detection should go hand in hand. The detection of wrong switch value based on data classification with attributes on variants folders and labels lists, notifying wrong switch label values, should instantly undergo data cleaning. The whole solution should ensure various requirements, principles, and dimensions of good data classification, data cleaning, and data quality as discussed before in section 3.1 in the 3rd chapter. These concepts should be used in other sub-processes improvement also, as discussed next.

5.4.2 Sub-process 2: Dataset validation with durability run

Durability runs are made for vehicles for validation of datasets produced by the calibration process. These are performed generally in last parts of calibration project in late quality gates when calibration engineers finalize their datasets for real life tests. Thus, any defects found in this testing would cause greater problems with the customer. They are generally performed with the set of 6-7 times repetition for running vehicle in various long tracks under real life conditions. The main goal is to prove that vehicle performance parameter like NOx emission or CO2 emission etc. using calibration datasets, are in required limits as per requirements of the customer. The optimal parameters test results are based on optimal calibration done for related labels in the calibration process. It is performed using ETAS INCA calibration software on the required dataset revision. Customer sends a request for performing a test on particular dataset revision with an Excel sheet with dataset

revision defined by attributes like emission type, train type or transmission type etc. to which the variant of dataset belongs.

Calibration engineers at AVL download the dataset revision based on this Excel sheet request provided by the customer with attributes defining the path of dataset variant folder in variant tree-view in CRETA. For a particular dataset revision, within ETAS INCA, settings of the experiment are created. INCA then measures all the values of parameters, required to check for customer requirement. Then, calibration engineers using INCA, responsible for dataset validation by durability run, prepare a certain set of values recorded for each tested parameter based on various test runs. Calibration engineers have direct access to CRETA and can directly download the dataset revision for testing themselves. Before testing is started, a naming convention is decided by calibration engineers for a dataset revisions for validation test. They set the naming standard from project to project, based on various information required for identification of revision dataset based on attributes. This helps to know to which vehicle variant the tested dataset revision belong, and to which dataset revision the test parameters belong. ETAS INCA has a tool called Recorder, using which users can set name for the file to save after test.

The dataset name in Recorder stays the same, even if a user changes the experiment with different dataset. So, if the name in INCA's Recorder tool is not changed, when the experiment is changed, the default name or previous dataset name for dataset under test would be assigned while saving the file for test results. The dataset revision under test would have a different name than the right name for it. An example of standard naming given below:

''2016_Enginename_derivat_AT_RWD_vehiclename_emissiontype_purpose_text_date_ name_04''.

Above example dataset revision name for durability run test is composed of the year, engine name, derivative, train type, transmission type, vehicle name, emission legislation type, the name of the purpose of the test, date of test, the name of tester, and some random digit, all separated by an underscore. The concept is similar to switch check, as the dataset revision in both cases is distinguished by various variants attributes such as emission, transmission, train type attributes. Sub-process of durability run with concerned elements for the thesis is shown in BPMN diagram in figure 60. It shows two versions of durability run. One is official, which is done for official testing for dataset revision release for the customer. Other is called customer-oriented testing, which deals with experimenting and optimizing dataset revision without a request of the external customer, but as efforts from AVL calibration engineers to provide further possibilities for better calibration results than expected by the customer. It is done more often than official durability run, and thus, face more errors. The BPMN diagram is quite self-explanatory, explaining general durability run test for dataset revision.

5.4.2.1 Various failure modes in durability run test

The above naming convention used for durability run tests for a particular vehicle is a source of major errors in the process. Especially, in the case of durability run test of type ''customer-oriented testing''. This testing is performed within the team of calibration engineers who perform durability run test, without the involvement of customer or calibration pilot at AVL. Their goal is to provide better results to the customer than they

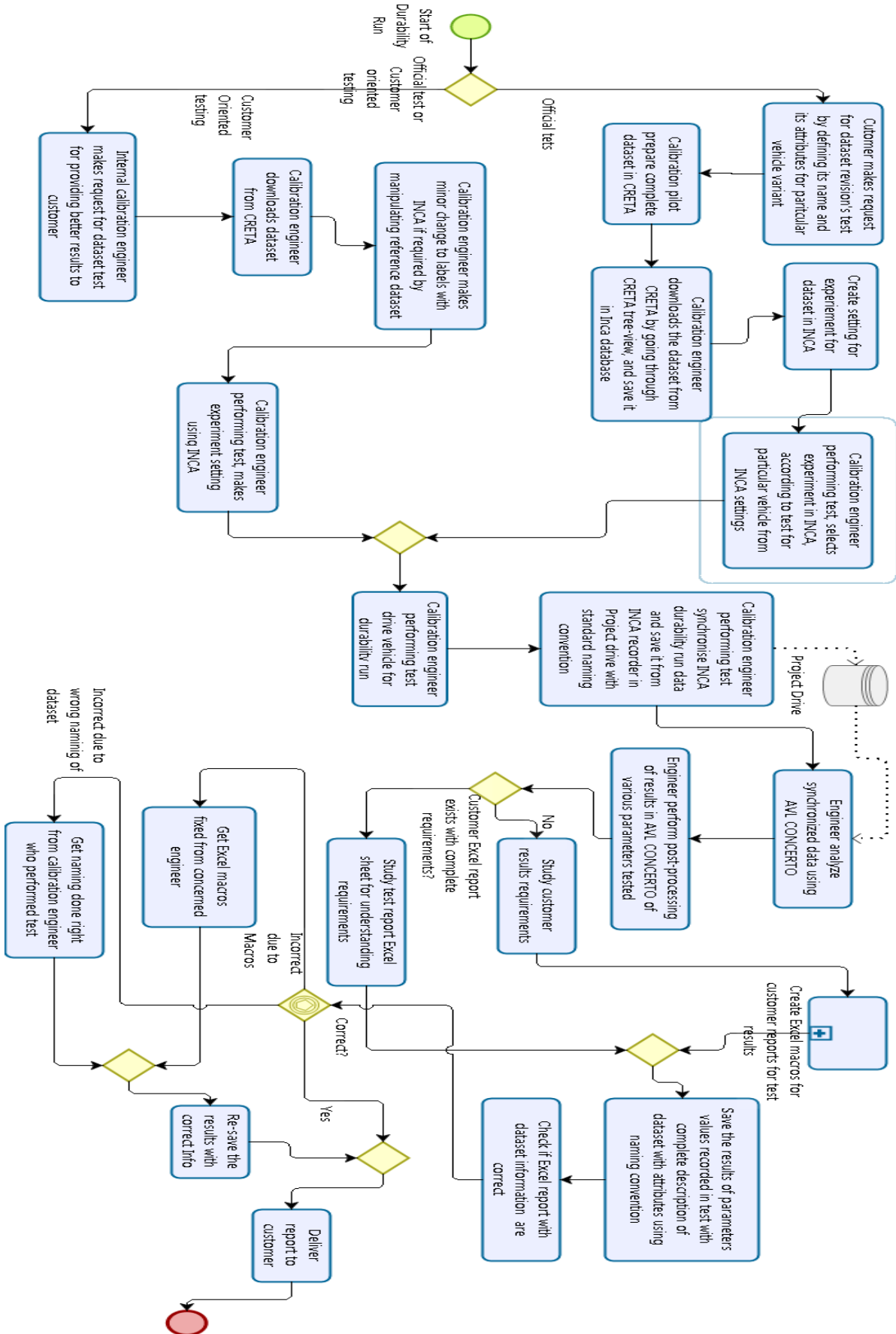expect. This type of testing is performed independently by engineers with less contact with

Figure 60: Durability run dataset validation sub-process BPMN to level 3 detailing with various activities and tasks that together compose it

calibration pilot, who ensure by making several checks that the right dataset revision is delivered for testing. It is performed often, i.e. thrice time more than official durability runs testing. Due to this reason, it is more usual for this type of durability run test that after the test is made, the dataset revision tested is unintentionally named wrong. There are more than 50% chances for the wrong name is written for dataset revision after this type of test. The two main reason behind this are:

- INCA user (calibration engineer who performs the test) did not follow the naming convention. For example, they forget to put the name of one of an attribute such as emission type for revision, or vehicle name etc. This often happens, as usually naming convention requires several items to be mentioned, in a certain order, which becomes a tidy task.

- INCA user might follow right naming convention but may write a wrong item for the attribute name. For example, they may write wrong engine name or vehicle name, the revision may belong to. This is because INCA user makes several tests and require several items to consider to write the name of revision, making naming revision a tidy task.

Figure 61 shows INCA experiment information interface. On the left of the tree is experiment being done using INCA, below them with green symbols are ECU software names for experiments containing datasets, and below them, with red and blue symbols are test vehicles setting called workspace. When user selects different vehicle, on the right of interface, dataset (downloaded from CRETA), shown as AS (dataset which can be manipulated during test in INCA) & RS (reference dataset which cannot be manipulated during test in INCA) are shown with same name as downloaded from CRETA (the default name for downloaded dataset from CRETA is same as dataset revision name in the variant in CRETA tree-view). Below it with CAN buses tree, are hardware for the test vehicle. When engineer saves the file after making a test on INCA, using INCA Recorded window, it is set to be saved as DAT file with default name same name as the experiment, unless user changes it (change it to the standard naming convention for durability run).

Other information, such as dataset revision name (same as downloaded from CRETA) for the experiment, and ECU software names etc. are saved as metadata, which can be shown when user takes cursor over saved file DAT file from INCA or when user checks its properties by selecting it and press right click on mouse. There exist no information about attributes of variant of dataset revision downloaded from CRETA using which test is being made on INCA, such as its emission type (EU5, EU6 etc.), train type (rear/front/ all-wheel drive), transmission type (automatic, or manual) etc. This causes problems for efficient data management in the process. If the dataset downloaded from CRETA is saved with default name with all the data tags/attributes used to define its variant folder, like in switch label check, the dataset revision information would be easily provided to CRETA users. Thus, the need of providing standard naming to revision dataset manually after test would be eliminated (prevention type of Poka-Yoke). This is because when calibration engineer saves test file, the name of dataset revision with attributes would be saved as metadata to the DAT file saved by INCA. So the information of dataset revision about its attributes, flow from CRETA throughout the use during INCA, or other software that works with it in the calibration process.

Currently, when calibration engineer (working with the customer) receives the dataset revision test file produced from durability run test engineer from INCA, they check if the

name is written right by taking a short glance. If they find something missing in the naming convention, or a misspelling in the names, for example, misspelling in engine name, vehicle name etc., they call INCA user (calibration engineer who performed the test and saved test file) and check for the required correction. This process is not reliable as, for example, if engine name is not misspelled but a wrong engine name is written, the manual task of checking naming by taking a glance, would not be reliable. If wrong information about variant folder attributes, to which dataset revision belong is provided, the entire testing can become useless.
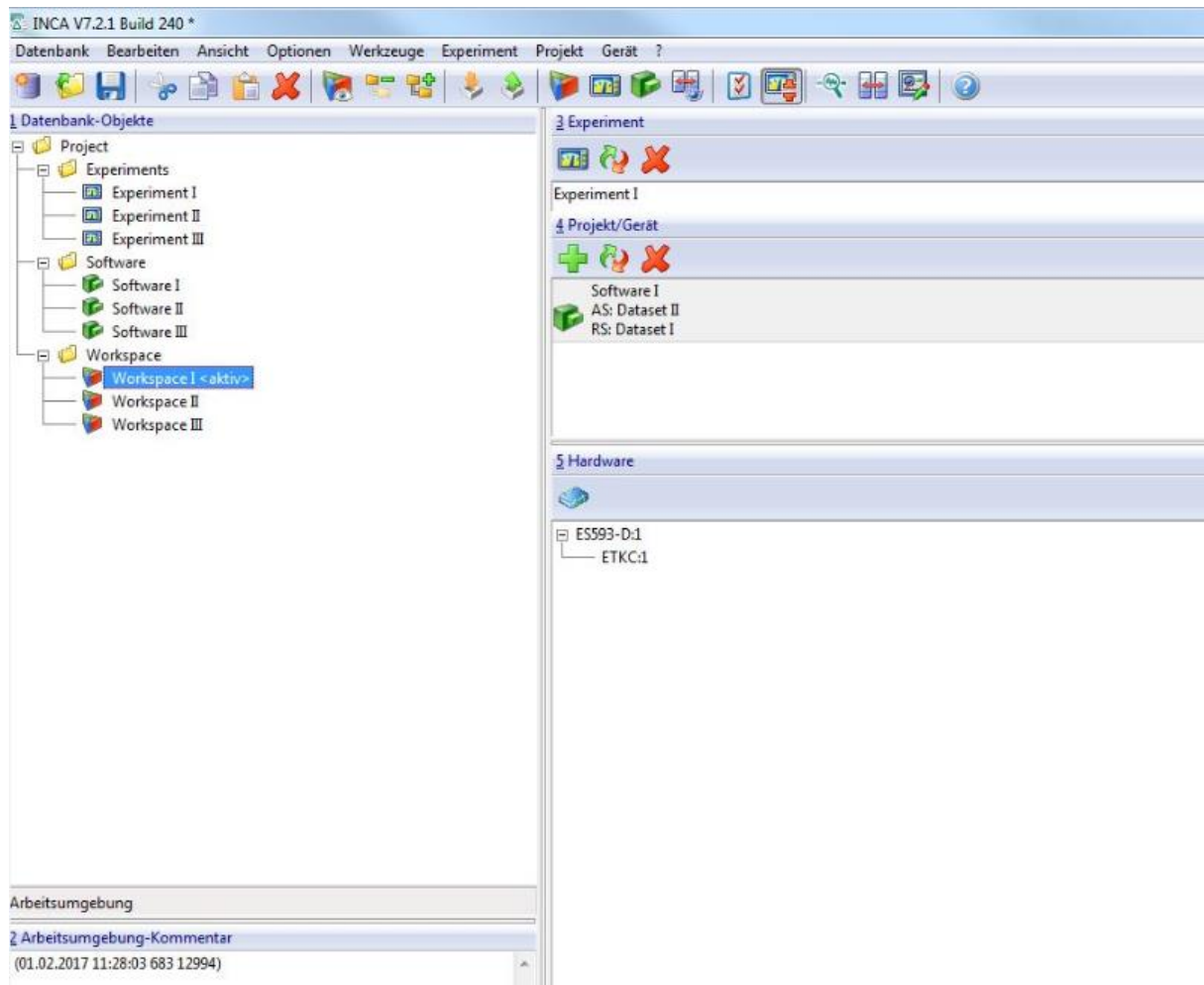

Figure 61: INCA user interface for experiment information

There are several points to consider for developing a standard naming convention for dataset revision for a particular project for its durability run tests. There is an Excel file prepared by project management team, which contains the required dataset revision variants with various attributes names required for information to perform durability run test on customer's vehicles. They read CRETA variant tree-view and decided which vehicle variant attributes are relevant to consider for information for AVL and customer to know where durability run tested dataset revision resides in CRETA variant tree-view. So, out of the CRETA variant tree-view, certain variant attributes names are chosen such as engine name, train type, transmission system, emission type etc., which together makes the name of durability run dataset revision test file. The Excel sheet file for project management is not provided here.

In the end, calibration engineer prepares an Excel sheet file, which contains a description of variant dataset revision using the naming convention of the test. Thus, description of dataset revision contains elements defined in naming convention. Excel files also contain values for each parameter to be tested with durability runs, along with each variant dataset revision. Just for a clear understanding, take an example in figure 62. It shows Excel file that calibration engineer, prepared for the customer for durability run test results. Such sheet is prepared for the customer by calibration engineer to show results of each parameter on right with respect to their concerned vehicle variant defined with various attributes on left: a variant with its engine name (motor), train type, transmission system, and vehicle name etc. These attributes names are present in CRETA tree-view as well. The parameters and their recorded values are on right. The INCA test result file is loaded in a software called AVL CONCERTO. For each dataset revision, there are numerous values for each parameter produced by series of durability run tests. Calibration engineer uses CONCERTO software, and carry out procedures to obtain the unique value of parameter out of numerous values from ETAS INCA. In simple terms, it's like an average. The unique value of parameters from the test is obtained by some statistical calculations of the numerous values of parameters obtained from ETAS INCA. This provides a unique parameter value for each dataset revision for durability run test. The engineer then saves the result for each variant dataset revision in CTF file of CONCERTO. This file contains everything written for Excel sheet in figure 62, for a particular variant dataset revision. Engineer manually copies and paste these values to the Excel sheet shown in figure 62 from CTF file. Some items in Excel like, Comment, V-Number are not present in CRETA tree-view, but they are there for some specific extra information for customer clarifications.

Motor	Getriebe	Antrieb	Fahrzeug	V-Nummer	Abgasnachbehandlung	Datenstand_Referenzseite	Datenstand_Arbeitsseite	Emissionsklasse	Fahrprofil	Strecke / Kommentar	Mileage @ Fahrtbeginn	Mileage aus Com_ISum	Nox Emmision	mittlere Umgebungstempe	CO2 Emmsion
-	-	-	-	-	-	-	-				km	km		°C	
											0	0	0	0	0
											#####	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
											0	0	0	0	0
Engine_1	AT	AWD	Vehicle_1	1234	NSC	Dataset_1	Dataset_1	mc60	FDL	MC60					
Engine_2	AT	RWD	Vehicle_2	1235	SCR	Dataset_2	Dataset_2	EU6-WLTP-TP1	Ueberfahrt	GRZ-OT					
Engine_3	AT	RWD	Vehicle_3	1236	SCR	Dataset_3	Dataset_3	EU6-WLTP-TP1	Überfahrt	GRZ-OT					
Engine_4	AT	AWD	Vehicle_4	1237	SCR	Dataset_4	Dataset_4	EU6-WLTP-TP1	Ueberfahrt	GRZ-OT					
Engine_5	MT	AWD	Vehicle_1	1234	SCR	Dataset_1	Dataset_1	EU6-WLTP-TP1	Ueberfahrt	GRZ-OT					
Engine_1	AT	RWD	Vehicle_2	1235	SCR	Dataset_2	Dataset_2	EU6-WLTP-TP1	Ueberfahrt	GRZ-OT					
Engine_2	AT	AWD	Vehicle_3	1236	SCR	Dataset_3	Dataset_3	EU6-WLTP-TP1	Ueberfahrt	GRZ-OT					
Engine_3	AT	AWD	Vehicle_4	1237	SCR	Dataset_4	Dataset_4	EU6-WLTP-TP1	Ueberfahrt	GRZ-OT					
Engine_1	MT	RWD	Vehicle_1	1234	SCR	Dataset_1	Dataset_1	EU6-WLTP-TP1	Ueberfahrt	GRZ-OT					
Engine_2	AT	RWD	Vehicle_2	1235	SCR	Dataset_2	Dataset_2	EU6-WLTP-TP1	Ueberfahrt	GRZ-OT					
Engine_3	MT	RWD	Vehicle_3	1236	SCR	Dataset_3	Dataset_3	EU6-WLTP-TP1	Ueberfahrt	GRZ-OT					
Engine_4	AT	AWD	Vehicle_4	1237	SCR	Dataset_4	Dataset_4	EU6-WLTP-TP1	Ueberfahrt	GRZ-OT					
Engine_5	AT	AWD	Vehicle_1	1234	SCR	Dataset_1	Dataset_1	EU6-WLTP-TP1	Ueberfahrt	GRZ-OT					
Engine_6	AT	AWD	Vehicle_2	1235	SCR	Dataset_2	Dataset_2	EU6	Ueberfahrt	GRZ-OT					

Figure 62: Final Excel sheet report for various durability run tests for customer

Excel validation sheet with macros, like the one in figure 62, checks if the parameters values are in accordance with the requirement of the customer. There is a list of parameters and their required values provided by the customer such as limits range for NOx emission, mileage, CO2 emission etc. The durability run tests parameter values have

to be within the expected values. Excel validation sheet, however, contains many other parameters not requested for tests by the customers. AVL engineer provides many such parameters results values from durability run test to provide more reliable information on testing. Excel macros are created in order to check if test parameters are within the customer required values, including the one not mentioned by the customer. The unique parameters values produced by using AVL CONCERTO are copied from CONCERTO CTF file to this Excel validation file to check for the parameter values. A possible error can arise if wrong parameter values from CTF file are copied, for a particular dataset revision name, in an Excel table row or due to the wrong macros developed.

Another problem in durability run sub-process can occur, if the INCA user from durability run test, download unintentionally a wrong dataset revision from variant tree-view from CRETA for making durability run test. This is because, CRETA tree-view can have thousands of variants folders, which make it quite tedious to find a required variant with dataset revision for the test. When wrong dataset revision is used for durability run test on ECU, it might cause ECU to not accept the dataset revision from different vehicle variant or ECU may ask for flashing the data inside it. Flashing removes all the data of previous software in it, which may baffle calibration engineer, causing various delays in above sub-process. In some cases, ECU may accept a wrong dataset revision and whole test results would become useless. Thus, this could cause various rework in the process. The main cause for this failures is because of confusion in following various items in dataset revision name, which defines in which variant in CRETA variant tree-view it belongs. INCA user manually selects the dataset revision by going through several variants in CRETA tree-view structure. The defects in such cases are certain.

5.4.2.2 FMEA analysis: Durability run results validation

Based on the analysis of Fishbone diagram, developed by various brainstorming and 5 Whys methods on level 3 BPMN process modeling detailing, various failure mode and causes of failure were written down to FMEA table. The table is chosen from SAE standard J1739 2002. The various effects of failures and RPN values were decided by a team of employees involved in durability run test. The table is shown in figure 63 and is later further explained for a recommendation for improvements.

5.4.2.3 Improvement/recommendation actions for with durability run test sub-process

After studying the current process of durability run carefully with respect to various defects that can occur, along with their failure mechanism, many factors can be determined in order to further improve the process and make it more reliable as a part of a software solution tool to be developed. The detailed description of various recommended actions to overcome various factors causing various defects in the calibration process for durability run test results for current validation process are discussed below:

- The name for tested dataset revisions saved after making durability run test for various parameters on ETAS INCA has to be in a correct standard form. This is can be achieved if information of various attributes/data tags for variants folders of the

variant dataset could be available to INCA USER. Software tool should, therefore, make it possible to save dataset revision from CRETA with default name with its data tags/ attributes assigned to its variant folder in CRETA variant tree-view along with name from customer. Since, dataset revision downloaded would be saved with attributes names such emission type, train type, transmission type etc., the manual requirement of writing name would be eliminated providing error prevention (prevention type Poka-Yoke) than the manual task of detection by taking glance by eyes for correct naming.

FAILURE MODE AND EFFECTS ANALYSIS

Item: Durability run test sub-process	Responsibility: Amritpal Singh	FMEA number: 3
Model: Current	Prepared by: Amritpal Singh	Page: 1 of 1
Core Team: Calibration pilots, Calibration engineers, Lead engineer, Amritpal Singh		FMEA Date (Orig) 25.10.2016 Rev: 1

Process Function	Potential Failure Mode	Potential Effect(s) of Failure	Sev	Class	Potential Cause(s)/ Mechanism(s) of Failure	Occur	Current Process Controls	Detec	RPN	Recommended Action(s)	Responsibility and Target Completion Date	Action Results: Actions Taken	Sev	Occ	Det	RPN
Durability run test sub-process.	Incorrect name assigned to tested dataset revision with INCA Recorder.	Tested parameters show incorrect values for actual tested dataset in Excel validation sheet.	8		Standard naming norm for tested dataset revision not followed.	9	Manually checking the naming by looking through the tested dataset revision name.	10	720	Software tool to save dataset with name with its attributes, so that there is complete information flow through ETAS INCA or other softwares (Concept: data classification/ data tagging)	Amritpal Singh, to be completed before end of thesis	Creation of prototype using C# with CRETA API, to proove concept behind recommendations	4	7	3	84
		Requirement of recheck causing rework and late delivery of test results.	8		Incorrect attribute name assigned in naming dataset revision.	8	Manually checking the naming by looking through the tested dataset revision name.	10	640	Software tool to save dataset with name with its attributes, so that there is complete information flow through ETAS INCA or other softwares (Concept: data classification/ data tagging)	Amritpal Singh, to be completed before end of thesis	Creation of prototype using C# with CRETA API, to proove concept behind recommendations	4	7	3	84
	INCA user downloads wrong dataset revision for testing from CRETA.	Lose of time and rework to make durability test by finding right revision again.	8		Confusion in finding the right data revision variant folder in CRETA variant tree-view due to many variants folder in CRETA variant tree-	8	No process available to user.	10	640	Software tool enables engineers to search for required dataset revision variant folder in CRETA variant tree-view by their attributes/data tags (Concept: data classification/ data Tagging)	Amritpal Singh, to be completed before end of thesis	Creation of prototype using C# with CRETA API, to proove concept behind recommendations	4	7	3	84
	Tested dataset parameters not in accordance with customer requirement.	Customer dissatisfaction with wrong parameters values delivered	10		Incorrect value compare, for example, due to incorrect macros or wrong parameter requirements entered in Excel.	4	Excel macros are used to make comparision. The parameter requirements are manually entered.	10	400	Software tool to assign parameters and their value ranges to dataset, which can notified if they does not match with durability run test results from CTF file of CONCERTO (Concept: process automation for target limits cross checks)	Amritpal Singh, to be completed before end of thesis	Creation of prototype using C# with CRETA API, to proove concept behind recommendations	4	7	3	84

Figure 63: FMEA diagram for validation with durability runs tests based on 5 Whys analysis of acitivities and tasks in its 3rd level detailing with the BPMN model

- It is possible that INCA test user downloads wrong dataset revision by going through wrong variants in populous CRETA variant tree-view structure. The software tool should provide the possibility to search the required dataset revision using attributes/data tags assigned to its variant folder, as they distinguish each variant revision from others. This would prevent downloading the wrong dataset for durability run testing, among thousands of variants folder in CRETA variant tree-view.

- There is certain extra information that has to be provided, which sometimes does not exists in CRETA. For example, let say, one of the variant folder (also called node) in CRETA variant tree-view does not contain SOP date for dataset revision or its vehicle number etc. and calibration engineer wants to add such information into the Excel sheet for the customer. It would be easy if such extra information can be data tagged to a variant folder, and saved along with the name of dataset revision if downloaded from CRETA. The extra information about variant dataset revision like SOP date, vehicle number etc., can thus, be provided to both INCA and CONCERTO users. This would provide error prevention as well (prevention type Poka-Yoke).

- The parameters values measured from INCA test results such as mileage, NOx or CO_2 emissions, are studied and final values are derived using AVL CONCERTO. These values are written in Excel sheet for the customer as shown in figure 62. Calibration engineer needs to check if these values are correct or are in certain ranges as required by customers. Thus, software tool should provide possibilities to assign parameters to datasets and ensure that the value of these parameters, is inside the ranges or meets the values as defined by the customer, else be notified for validation failures. This would enhance data quality in durability run test sub-process, considering the dimensions of data quality discussed. The parameters name and their target values can be easily set by software, replacing the need of creating new Excel macros for every new project. It is to be noted that not all parameters (checked by durability run test) are same as parameters customer asked for testing. As discussed earlier, calibration engineers add more parameters names in Excel list, for providing better results to customers. This has to be taken care by the software tool. Also, the name of parameters requested by the customer can be different than written by calibration engineers, although the meaning is same.

5.4.3 Sub-process 3 and 4: Release quality gates template & monitoring calibration project status

For the purpose of efficient project management, the company uses quality gates for handling calibration datasets quality. Thus, the calibration process for ECU software dataset optimization is divided into various phases. The main purpose for the use of quality gates in calibration process are:

1 To analyze or trace the calibration quality at any time of the project (current project status),
2 To display the advances in the calibration process,
3 To perform risk analysis for quality management system,
4 To enable each group, responsible for their group of work, for review meetings to exchange calibration knowledge, and

5 To provide information for assessment meetings at various milestones in the process.

Each project for calibration within the company has to go through quality gates process. There are many conditions set by the company in association to quality gates. The milestones planning for projects for quality gates are decided by the project leaders. Sometimes, the customer may also be involved in the planning, together with project leader if required. The quality gates targets are defined for various sub-work-packages. The Calibration process at the company is divided into various parts called lead groups. The lead groups are further divided into various sub-groups called work-packages, which are further divided into sub-work-packages, which are further divided into various sub-categories as shown in figure 64 (figure from AVL internal documentation developed AVL internal knowledge)

Figure 64: Various groups in company composing various work-packages and their sub-categories.

These sub-categories are further divided into functions, which are composed of groups of dataset labels. A general process at AVL for the management of various data in calibration process with respect to quality gates can be explained as follow:

1 Calibration pilot sets up a project in CRETA and creates a proposal for ECU software labels responsibility to various employees. They assign various labels to respective work-packages. Before this, it is decided by a customer that which labels and work-packages will be responsible to AVL.

2 Project leader then goes through the proposal and defines various milestones for quality gates together with lead engineers, as lead engineers have responsibility

for work-packages assigned to them. Calibration pilot then updates the dates decided for milestones to CRETA.

3 Finally, customer agreement on labels responsibilities finalizes the plan for the project. The agreement is made for work-package responsibilities and various other project factors like dates for milestones. Lead engineer or calibration pilot then uploads finalized project data to CRETA database and project drives and defines various important labels (VIP labels), and differences between various variants.

4 Project leader then shares various information of project with the employee involved in it, such as label and work-package responsibilities, times to complete various tasks with respect to milestones of quality gates, various important labels, and defining maturity (label scores) of various labels etc.

5 Labels and work-packages responsibilities, and dataset management process are discussed time to time in project meetings about after 2 months, and then after months from project start for smooth workflow in the project.
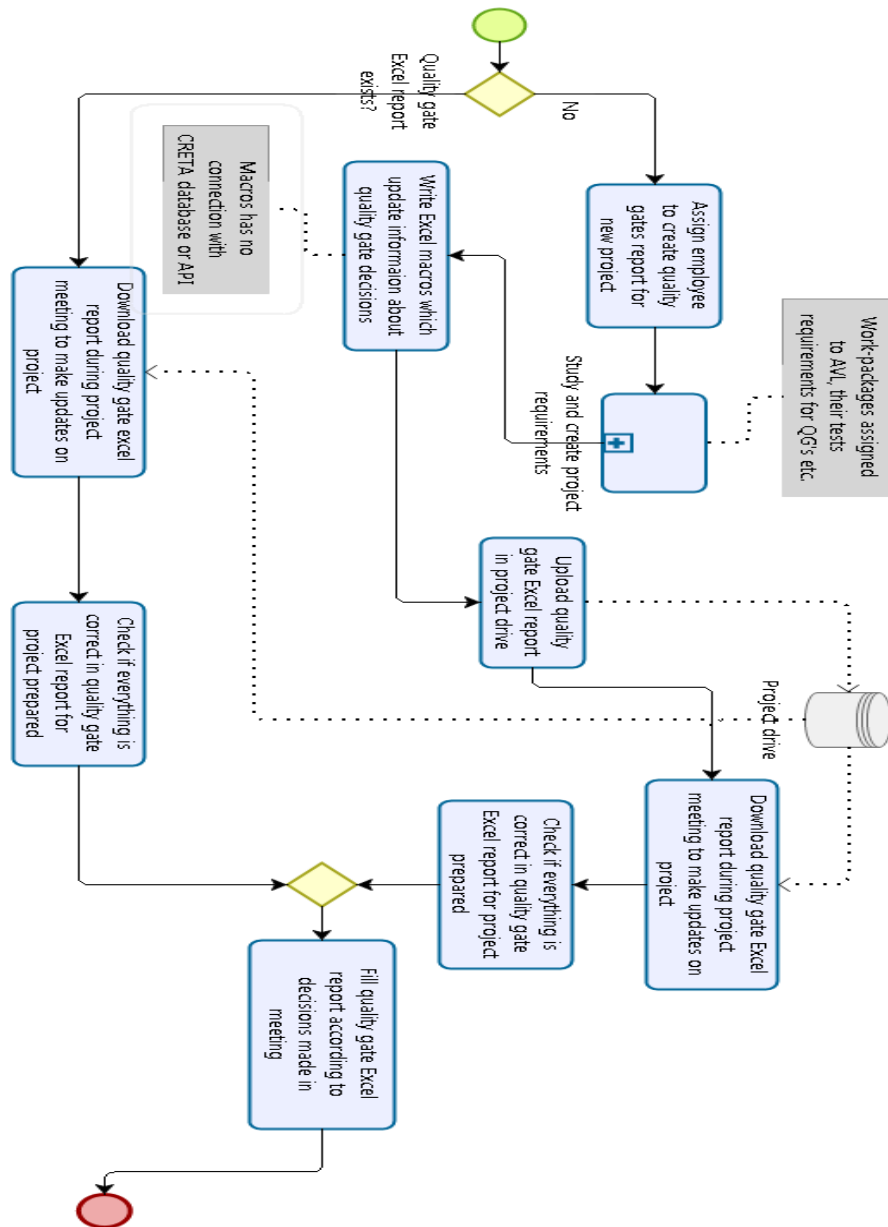

Figure 65: BPMN modelling (level 3 detail) for general quality gate template update sub-process

The quality gate process is composed of various members of the team, associated with certain roles for the quality gates process. Calibration engineers do the calibration, create documentation and assign a rating of labels. Rating helps to measure the maturity of labels calibrated. As labels are assigned to work-packages, the label scores belonging to particular work-package further determine the scores or maturity of the work-package as a whole. Calibration pilot takes care for any label responsibility violation or missing scores after the calibrated labels are delivered. Department steering committee approves targets for every quality gate during the meeting for project review. The steering committee is composed of following members:

- Skill team leader: Targets for a particular gate undergo final approval from them,
- Lead engineers: Each work-package belongs to a particular lead engineer, who approves if their work-package has met targets, based on the score of various labels inside it and various required tests, and
- Project leader specialized: Responsible for the calibration team for the project.

The quality gate process for calibration consists of generally five phases. Each phase consists of various tasks to be completed, and a certain level of labels maturity to be achieved, in order to get approval for completion to move on to next phase of the process. The tasks involved are defined with certain goal description and criteria of measurements. The goal descriptions and measurements for various tasks, all depends on project to project. However, for most of the projects the average label scores from a particular work-package (work-package maturity) are same as follow (from AVL internal documentation developed from AVL internal knowledge):

- Gate 0: Start- A team is defined, tasks are defined, label responsibilities are assigned, customer-specific requirements are fixed, CRETA database is set-up, and reference dataset is defined.
- Gate 1: Runnable – Average target maturity for a particular work-package to pass this gate is generally 25%, apart from various tests requirements. This gate involved various activities such as under standard environment conditions, safe engine and vehicle operations are achieved. Hardware is fixed, and cold start is achieved in 2nd attempt.
- Gate 2: Redefined – Average target maturity for a particular work-package to pass this gate is generally 50%, apart from various tests requirements. In the case of OBD, this gate involves achieving legal limits of emission with nominal hardware. Calibration reaches to a level suitable for the start of durability testing. Smoke limitations and cold starts are attained on target.
- Gate 3: Targets Achievement – Average target maturity for a particular work-package to pass this gate is generally between 50% and 75%, apart from various tests requirements. Calibration is completed for emissions, along with performance and drivability at all specified environmental conditions with nominal hardware. Cold start engineering targets are met and calibration suitable for emissions durability testing is attained.
- Gate 4: Robust – Average target maturity for a particular work-package to pass this gate is generally 75%, apart from various tests requirements. Robustness

verification and calibration for PPV / c-sample vehicles are made. Emissions diagnostics is complete, but not verified.

- Gate 5: Validated – Average target maturity for a particular work-package to pass this gate is generally 100%, apart from various tests requirements. This stage includes various activities such as calibration for MVB / pre-series vehicles are made, calibration is validated with PPVs / c-sample Hardware, diagnostic operation is evaluation completed, certification is attained, and in the end, label review meeting is completed.

As discussed, generally the average of total label scores belonging to particular work-package defines the score or maturity of a work-package in a project. This average label scores/maturity is a simple weighted average of scores of all the labels in work-package. It is to be noted that if the average of label scores is less than the quality gate target for a work-package, it can still be approved. This is because if certain labels are untouched and have a null score, the average would cause to judge the wrong maturity of the entire work-package in a process, as weighted average possess certain limitations. Or another case could be that the average of total labels score is more than the target score for the quality gate. This can happen if many labels in the beginning in work-package gets a full score, causing overestimation for an average score of labels. Thus, the quality gate would still be marked as not completed even if the average score is more than quality gate target. Sometimes, scores of labels are considered to reach the quality gate target, but the various tasks involved in the processing of quality gate are not completed. Then, the quality gate is said to be achieved with conditions. When the conditions are successfully accomplished, the lead engineer passes the gate later. Marking labels scores is compulsory and the scoring convention is standardized internationally for the calibration process. Various label scores involved in measuring label maturity for measuring calibration quality in a process are as follow (from AVL internal documentation developed from AVL internal knowledge):

- Untouched: The labels are marked with 0% score. A first software release consists of 0% maturity rating, as the labels are untouched.
- Prelim Calibrated: The labels are marked with 25% score. It consists of desktop calibration from engineer's know-how, and thus, no verification by various simulation or measurements.
- Calibrated: The labels are marked with 50% score. Calibration for labels is made in standard environmental conditions.
- Checked: The labels are marked with 75% score. Calibration for labels is made under all environmental conditions.
- Completed: The labels are marked 100% after final validation on pre-series of engine or vehicle. Thus, completing the final calibration for the label.

These can be compared to quality gates scores discussed before. There are several reports sheets prepared related to quality gates evaluation in review meetings such as in figure 66 and 67. One of them is called 'Minutes of quality gate review'. The report is prepared in Excel sheet using various macros. It consists of a number of cells containing various information about particular work-package like goal description, goal measurements etc. There are options for filling remarks in the box if in case, a quality gate is achieved with conditions. It consists of approval checkboxes for a particular work-

package with yes, no or N/A. There exists a final sheet in the Excel report on the first page which consists of remarks and status in brief for all the information filled by a lead engineer in meeting for work-packages filled in different pages of Excel report as shown in figure 66. Based on the remarks and status in this sheet, and discussing various situations, the result is decided at the end of the meeting if the quality gate for whole work-packages is achieved or not. An example can be considered in order to understand it. Excel report for work-packages named sensor calibration and air and path control are shown in figure 66.

Figure 66: Example of information to be filled by lead engineer for work-package

The macros work in such a manner that when remarks or checkbox are filled in an each cell belonging to particular work-package by different lead engineers in a sheet like in figure 66, the final sheet at first page is automatically filled for brief overview of report for each work-package marked in different sheets, which looks like in figure 67. The maturity of work-packages can be judged by a report produced by CRETA called quality gates export as shown above in figure 68. It summarizes maturity of each work-package in a single sheet, by calculating average maturity of each lead group from weighted average scores of each work-package from its labels scores. The weighted average score of each work-package is equal to the sum of each label scores, divided by a number of labels present in work-package. As discussed before, it is not an appropriate way of judging the overall maturity for the work-package. The red cells show the unreached targets of average maturity based on current status of the project, while the green one shows the ones which have reached target successfully. The comparisons are made with blue cells which show target scores/maturity to reach.

Minutes of Quality Gate Review
Diesel Calibration Gate 1 – Runnable

Participants

Role	Name	Role	Name
Skill Team Leader DES		Lead Engineer – P&E:	
Skill Team Leader DEC		Lead Engineer – EAS:	
Lead Engineer – Air Pa		Lead Engineer – OBD:	
Lead Engineer – TnF		Project Leader Specializ	

Decision

☐ Quality Gate 1 achieved **without** conditions.
☐ Quality Gate 1 achieved **with** conditions.
☐ Quality Gate 1 **not** achieved.

GREEN

Listing of Conditions (incomplete preconditions and work packages)

	Work Package	Remarks / Next Steps	Due Date
Air Path Mod. & Contr.	Air system modellin	Not reviewed.	
	Sensor calibration	Not reviewed.	
	Air path controllers	Not reviewed.	
	EAS Controllers	Not reviewed.	
	Thermal energy Ma	Not reviewed.	
P&E	Full load & comp.	Not reviewed.	
	Cycle emis. & FC	Not reviewed.	
	Engine map cal.	Not reviewed.	
	RDE	Not reviewed.	
	Torque structure	Not reviewed.	
	Injection System	Not reviewed.	

Figure 67: Final sheet to be filled by macros automatically on the basis of data in sheets in figure 66.

5.4.3.1 Various failure modes in quality gate process and for monitoring calibration project status

Quality gate process involves creating maturity report of the entire project of the calibration process for a customer. Thus, it involves handling a large amount of data related to it. The process currently involves various Excel sheets to report the data prepared by Excel macros. It is not only time consuming, as one minute of quality gate review file takes three working days to be prepared, but also causes various sort of errors. Various failure modes as a part of FMEA analysis are discussed in detail below:

- Incorrect information generated by Excel macros: The minutes to quality gate review file contains a lot of macros. It contains various information as shown in figure 66 and 67, for each work-package assigned to AVL. The macros are prepared for each new project, as each new project can have a different amount of work-packages and their description. This can cause various effects on the project management, as wrong macros can generate wrong information for final project update report prepared for the customer, which further could cause customers to lose their reliability on the company.
- CRETA quality gates only provide current project status with respect to the target. No information about the past or future labels/work-package's scores/maturity is

provided. This lead to lead engineers to make a wrong prediction about the future status of the project. Moreover, they lack knowledge about patterns of past status of the project.

Work-packages	Creator	Sub-Workpackages	G-CG1 [T-23months] Q-target	Q-stage	P-stage	G-CG2 [T-16months] Q-target	Q-stage	P-stage	G-CG3 [T-9months] Q-target	Q-stage	P-stage	G-CG4 [T-7months] Q-target	Q-stage	P-stage	G-CG5 [T-2months] Q-target	Q-stage	P-stage
Base Calibration	122	Air Path	50,00	0,00	0,00	60,00	25,00	28,65	75,00	75,00	75,00	100,00	96,00	99,34	100,00	90,00	99,34
	503	Fuel Path	50,00	0,00	0,00	50,00	0,00	43,15	75,00	75,00	76,19	100,00	100,00	100,00	100,00	100,00	100,00
	447	Torque Path	50,00	0,00	0,00	50,00	0,00	52,18	75,00	75,00	75,1	100,00	75,00	99,50	100,00	75,00	99,50
	1781	Sensor Calibration	50,00	0,00	0,00	50,00	0,00	55,60	75,00	75,00	75,42	100,00	90,00	59,01	100,00	90,00	99,0
	197	Start & Cold Start	50,00	0,00	0,00	50,00	25,00	50,63	75,00	75,00	75,00	100,00	100,00	100,00	100,00	100,00	100,00
	198	Hybrid functionality	25,00	0,00	0,00	50,00	25,00	73,36	75,00	75,00	75,00	100,00	100,00	100,00	100,00	100,00	100,00
	702	Energy management	25,00	0,00	0,00	50,00	0,00	71,72	75,00	80,00	74,54	100,00	90,00	98,22	100,00	50,00	98,22
total	3950	Average BaseCalibration	44,30	0,00	#DIV/0!	50,00	3,27	53,78	75,00	70,56	75,19	100,00	90,58	99,44	100,00	90,58	99,44
P&E	184	Base Combustion	50,00	0,00	0,00	50,00	0,00	48,30	75,00	75,00	75,00	100,00	100,00	100,00	100,00	100,00	100,00
	503	Combustion Corrections	25,00	0,00	0,00	50,00	0,00	47,30	50,00	75,00	75,00	75,00	75,00	99,50	100,00	75,00	99,50
	166	State Machines & Transitions	25,00	0,00	0,00	50,00	25,00	52,44	50,00	75,00	75,00	75,00	90,00	99,58	100,00	90,00	99,58
	262	Limits & Thresholds	90,00	0,00	0,00	50,00	25,00	58,37	75,00	75,00	75,00	75,00	90,00	99,69	100,00	90,00	99,69
total	1115	Average P&E	35,00	0,00	#DIV/0!	50,00	9,80	51,88	50,00	75,00	75,00	73,17	84,66	99,69	100,00	84,88	99,69
Driveability	407	Vehicle Functions	25,00	0,00	0,00	50,00	25,00	74,53	50,00	75,00	76,81	75,00	90,00	95,14	100,00	90,00	95,14
	185	Vehicle Driveability	25,00	0,00	0,00	50,00	25,00	67,08	50,00	75,00	75,00	75,00	90,00	93,35	100,00	90,00	93,35
	1470	Engine Controllers	50,00	0,00	0,00	50,00	0,00	47,2	50,00	75,00	75,03	75,00	75,00	96,31	100,00	75,00	96,3
total	2062	Average Driveability	42,84	0,00	#DIV/0!	50,00	7,16	62,97	50,00	75,00	75,55	76,00	79,31	94,93	100,00	79,31	94,93
EAS	680	DPF Temp. Management	50,00	0,00	0,00	50,00	0,00	51,40	75,00	75,00	76,79	100,00	90,00	99,95	100,00	90,00	99,95
	0	LNT Rich Combustion	50,00	0,00	0,00	50,00	0,00	0,20	75,00	0,00	0,00	100,00	100,00	100,00	100,00	100,00	100,00
	0	LNT State Machine	50,00	0,00	0,00	50,00	0,00	0,00	75,00	0,00	0,00	100,00	100,00	100,00	100,00	100,00	100,00
	409	Exhaust Model	50,00	0,00	0,00	50,00	0,00	70,8	75,00	75,00	75,24	100,00	90,00	97,03	100,00	90,00	97,03
	256	DPF Soot Loading Model	25,00	0,00	0,00	50,00	0,00	25,09	75,00	75,00	91,43	100,00	90,00	99,72	100,00	90,00	99,72
	2	SCR Calibration	25,00	0,00	0,00	60,00	50,00	50,00	75,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
total	1357	Average EAS	44,80	0,00	#DIV/0!	50,00	0,07	39,35	75,00	75,04	85,87	100,00	90,01	99,46	100,00	90,01	99,46
OBD	975	Engine-EAS Base & Electrical Diagnostics	50,00	0,00	0,00	75,00	25,00	44,3	75,00	50,00	71,08	75,00	90,00	94,79	100,00	90,00	94,79
	477	Engine OBD	25,00	0,00	0,00	50,00	0,00	54,47	50,00	50,00	70,42	75,00	50,00	93,54	100,00	50,00	93,54
	451	Exhaust After-Treatment System OBD	25,00	0,00	0,00	50,00	0,00	27,90	50,00	50,00	69,75	75,00	75,00	94,93	100,00	75,00	94,93
	5813	Fault Manager & Scan Tool	50,00	0,00	0,00	50,00	25,00	50,22	50,00	50,00	74,06	75,00	80,00	90,03	100,00	90,00	98,03
	347	Engine-EAS End of Line & Workshop Testing	0,00	0,00	0,00	25,00	25,00	54,39	50,00	50,00	64,91	75,00	90,00	96,40	100,00	90,00	96,40
	1016	Engine Auxiliaries & Vehicle Diagnostics	50,00	0,00	0,00	75,00	25,00	52,57	75,00	50,00	74,30	75,00	90,00	95,38	100,00	90,00	95,38
total	9079	Average OBD	45,50	0,00	#DIV/0!	59,53	22,44	47,29	59,48	50,00	70,38	75,00	87,15	94,18	100,00	87,15	94,18
AVL total	17563	Average AVL	44,205147	0,00	#DIV/0!	52,34015	13,80	48,07	61,02	61,08	76,61	82,81615	87,08	98,19	100	87,08	98,19

Figure 68: CRETA quality gate export sheet showing current label maturity compared to target maturity

- Incorrect work-packages added to Excel sheet: During the meeting for approving quality gates, it can happen that there are wrong work-packages added to the Excel sheet, which are in the responsibility of AVL. This is, however, could be detected easily during the meeting but causes certain inconveniences to employees involved in process management. However, it can also happen that a work-package

assigned to AVL is missing from the Excel sheet for the evaluation. If a lead engineer is responsible for few work-packages, they can detect that in the meeting. However, if the lead engineer has the responsibility of a lot of work-packages, the missing package would go undetected in the meeting and hence unevaluated. Moreover, it is a tidy task for employee preparing Excel macros to take care that each work-package assigned to AVL is present in the Excel sheet report, as work-packages are assigned to thousands of labels present inside the dataset revision. An unreported work-package in the final report sheet prepared for the customer could create certain inconveniences with the customer and unreliable updates on a project for the process management.

- Assigned incorrect status for QG milestone: When the calibration engineers work on their label responsible to them, they assign the label score for defining the maturity of a project as a mandatory task. It can happen that score assigned to the labels are incorrect, as due to the work pressure the calibration engineer can assign more score than the actual status. The score is defined by the know-how of the calibration engineer responsible and thus goes undetected. Moreover, the customer can be easily convinced by the engineer for an explanation for the score. Thus, it is not a significant issue in the process. But, if it is done in large scale, which has never happened before at AVL, it can cause certain problems for lead engineer responsible for work-package. Thus, this point cannot be ignored for future cases, but improvement action for this point was decided to be untreated for above thesis due to low RPN value.

- The CRETA quality gate export is inappropriate for judging the maturity of the entire project. The file defines the maturity of each work-package as a weighted average score of all the labels present in them, which is not an appropriate solution. There can be a lot of outliers in calculating the average. For example, some of the labels which can get high score up to 90%, just in the first quality gate, can create overestimation of maturity for a work-package, or large amount of labels can have low score for example 25% up to quality gate 4, which can cause underestimation of maturity of work-package.

- There is another little issue associated with current minutes to quality gate review file. It provides no information about the variant of revision. In the Excel sheet, it only contains the name of dataset revision to be evaluated for the quality gate, which contains no information such as various attributes associated with the variant. This may cause confusion to lead engineers, as they might require discussing on the various attributes like engine type, emission legislation, train or transmission type etc. during a meeting. Such data is required to be made available automatically for a particular dataset revision being evaluated for the work-package for the quality gate (process automation).

- An incorrect maturity of related variants revisions in QG: Current minute to quality gate review file contains no information on related sub-variants or delta-variants of a lead variant. A lead variant is a variant, whose revision is evaluated for the quality gate. A sub-variants or delta-variants has very fewer differences than lead-variants,

and are required to possess maturity nearly same to a lead variant, as they are worked along with them. Thus, current Excel file provide no space for evaluating and comparing sub-variant or delta variant maturity, which could further cause the an insufficient growth of project than expected and planned by the company and customer. This is required to be provided automatically to lead engineers for their work-packages (process automation).

5.4.3.2 FMEA analysis for quality gates and monitoring calibration project status

Based on the analysis of Fishbone diagram, developed by various brainstorming and 5 Whys methods on level 3 BPMN process modeling detailing, various failure mode and causes of failure were written down to FMEA table. The table is chosen from SAE standard J1739 2002. Various effects of failures and RPN value were decided by a team of employees involved in above sub-process. The table is shown in next page in figure 69. Next paragraph contains further explanations for recommendation actions in detail.

5.4.3.3 Improvement/ recommendation actions for release quality gates report template and monitoring calibration project status

1 Currently, CRETA quality gate export tool provides only present status of the project. Calibration labels score/maturity records should be saved in the database along project timeline, to enable CRETA to build time-series graph of calibration maturity of work-packages for projects. It can be weekly or monthly basis based on project requirement. Forecast analyses should then be applied using time-series graph, to estimate future calibration maturity of the project in a quantitative form based on past data of label scores of various work-packages. Moreover, past data records from time-series would help lead engineers to compare the performance of engineers under them.

2 The ECU software on which whole calibration team works, are assigned responsibilities at the very beginning of a project such as for labels, and their work-packages etc. If an owner of the label assigned belongs to AVL, it means the 'work-package' assigned to the label inside a dataset revision file, belong to AVL. Since, in CRETA, calibration engineer responsibility to a label is assigned only if it is belonging to work-package of AVL. Hence, by reading if an owner for a label belongs to AVL, the 'work-packages' assigned to the labels inside the dataset revision can be easily distinguished if they belong to AVL or to the customer (such as by process automation solutions). This would replace the manual task of distinguishing work-packages assigned to thousands of labels if they belong to AVL. It would create a more reliable solution for the quality gate evaluation, as all the work-packages assigned to AVL will be automatically known.

3 It often occurs in calibration process that calibration engineers assign wrong maturity score to their responsible label. Generally, they provide no scores at the beginning of quality gate, taking their work politely. When the end of quality gates arises, work pressure is produced for calibration engineers to reach require maturity targets. Nearly all calibration engineers start assigning the scores in last two weeks of quality gate review meeting. This leads to an instant rise in work-package

Item:	Quality Gates Template & Monitoring Project status		Responsibility:	Amritpal Singh	FMEA number:	4
Model:	Current		Prepared by:	Amritpal Singh	Page:	1 of 1
Core Team:	Lead engineer, calibration pilots, calibration engineers and Amritpal Singh.				FMEA Date (Orig) 01.11.2016	Rev: 1

FAILURE MODE AND EFFECTS ANALYSIS

Process Function	Potential Failure Mode	Potential Effect(s) of Failure	Sev / Class	Potential Cause(s)/ Mechanism(s) of Failure	Occur	Current Process Controls	Detec	RPN	Recommended Action(s)	Responsibility and Target Completion Date	Actions Taken	Sev	Occ	Det	RPN
Monitoring calibration process status	No possibilities to know past or future calibration maturity or for a project.	Delays in calibration projects for external customers, providing unsatisfactory results.	8	No tools available to know future or past maturity for calibration project.	9	No process available.	10	720	Create time series graph for calibration maturity record and perform forecast analysis based on 2nd order exponential smoothing).	Amritpal Singh, to be completed before end of calibration.	Exponential smoothing performed on Excel with dummies data of calibration.	3	8	5	120
Quality gates Excel sheet template	Incorrect work-packages of project added to Excel sheet.	Missing data for quality gates causing inconvenience to lead engineers concerned.	5	A lot of work-packages and there related data to be added to Excel sheet.	8	No process available to check missing work-package or data in Excel sheet except manual CRETA data check.	10	400	Replace Excel sheet with software tool that can detect and assign work-packages to AVL and customer before end of (Concept: process automation).	Amritpal Singh, to be completed using C# with CRETA API, thesis.	Creation of prototype using C# with CRETA API, to proove concept behind recommendations.	3	1	1	3
	Assigned incorrect results for lead engineers and Customer.	Incorrect label maturity results for lead engineers and customers.	6	Intentionally wrong label scores assigned by responsible calibration engineers due to work pressure.	5	No process available as scoring is decided by calibration engineers with there know-how.	8	240	Low RPN	Amritpal Singh, to be completed before end of thesis.					0
	Assigned incorrect status for QG milestone.	Incorrect label maturity results for lead engineers and customers.	8	Missinterpretation of work-package scores due to outliers in data, while making weighted average of all label scores.	6	weighted average of all labels, Outliers information is written manually on Excel file for lead engineers.	10	480	Software tool provide automatically an improved version of CRETA quality gate export with standard deviation for label scores (Concept: process automation).	Amritpal Singh, to be completed before end of thesis.	Exponential smoothing performed on Excel with dummies data of calibration maturity.	3	8	3	72
	Poor status for maturity of related variants revisions in QG.	Maturity of sub-variants or delta-variants not in chorence with Lead-variants maturity.	8	No possibility for monitoring or comparing for maturity of sub-variants or delta-variants of dataset revisions for QG.	8	No Excel macros added in current process file to interpret it. Manually see maturity of delta or sub-variants (Concept: process automation).	6	384	Improved version of the quality gate export from CRETA includes option to add column to show maturity of delta or sub-variants (Concept: process automation).	Amritpal Singh, to be completed using C# with CRETA API, thesis.	Creation of prototype using C# with CRETA API, to proove concept behind recommendations.	3	5	3	45
	Incorrect information generated by Excel generated macros.	Incorrect QG data generated for lead-engineers, customer, and skill team leaders.	6	Large amount of macros to be generated for various sheets from each work-package for each new project.	9	No process available to check correctness of Excel macros.	10	540	Replace Excel sheets for meeting with software tool (Concept: process automation).	Amritpal Singh, to be completed before end of thesis.	Creation of prototype using C# with CRETA API, to proove concept behind recommendations.	3	1	1	3

Figure 69: FMEA for quality gate template and monitoring project status based on 5 Whys analysis of acitivities and tasks in its 3rd level detailing with the BPMN modelling

maturity in last days of quality gate stage. For example, at first 3 and a half months of quality gate 1, maturity would rise slowly from 0% to 12 %. It rises near to 25% (target) instantly in last two weeks. This thesis does not consider controlling such issues in the process. This is because, the main goal of this studies is to improve the process with respect to the customer, and the customer is easily convinced by calibration engineers for scoring provided by them to their label, which causes no problems with customers.

4 The current process involves a high probability of assigning an incorrect status for the quality gate milestone due to incomplete information provided by a weighted average of label score for work-package maturity. There should be data of outlier scores provided in CRETA quality gate export report such as in the form of standard deviation. This would overcome the problem of the incorrect decision made by a lead engineer because of relying just on weighted average without knowledge of how many scores have too less or too much score as outliers. This can be provided automatically by providing deviations from the mean average point.

5 The calibration maturity of delta-variant or sub-variant datasets shall be made available for comparing their maturity with their lead variants, automatically. Apart from this, attributes/data tags of dataset variant information should be provided, so that, the reader has a clear understanding about which dataset revision variant maturity they are studying. Currently, such issues are handled manually by Excel sheet, which needs to be replaced by process automation.

6 The current process involves Excel file for preparing a quality gate review meeting report for lead engineers and the final report for project status for the customers. It shall be replaced by software, which could generate a report for each work-package for each new project automatically, and save the report in the CRETA database, to which all the concerned employee to the quality gates can have access anytime unlike Excel file saved isolated in project drive. This would replace various macros to be prepared for report file, providing process automation. Moreover, it would reduce the working efforts to prepare the minutes to quality gate review file each time for a new project (replacing work of up to 3 days by process automation proved by thesis software prototype). The solution would be fast and more reliable than the conventional Excel file using macros.

All of above points would provide better data quality for process management considering various dimensions of data quality discussed. Proves for the effectiveness of various concepts like business process automation, forecast analysis follows in next chapter.

5.4.4 Sub-process 5: Validation for cross-checking delivered and received dataset

Whenever a dataset is received from a customer to AVL, it is cross-checked using revisions comparison tool in CRETA. This tool helps to check if a customer has made any changes to the following elements of labels which are in the responsibility of AVL:

- Label-value: Customer may change label values responsible to AVL, intentionally or unintentionally. Labels values can be changed unintentionally by mistakes and can go undetected by AVL for such changes. This is possible due to poor process integration between customers and AVL. If the values are not set back to the original values, it may cause various trouble to AVL engineers. Sometimes, the

customer may change AVL responsible label values intentionally to some specific, and may not share the information with AVL, which may cause certain issues for AVL employees. Thus, such changes are checked by current validation system by revision comparison tool of CRETA.

- Label-scores: Customer may unintentionally change the score of labels which are not in their responsibility. This would lead to misinterpretation of maturity for the label. Thus, such issues are well-taken care by current validation system.

- Label-comments: Comments are added to dataset labels to deal with various issues associated with the labels, such as, to provide information about reasons behind changes made to labels between teams and their members. It is a mandatory task for calibration process by standard. Thus, they are important element related to labels. Customer may unintentionally manipulate the comments. Sometimes, they may also intentionally change it while working on their tests and making certain changes to label values or scores etc. If AVL engineers are not informed about these changes to comment of their labels, there can be various confusions in the process.

- Label-errors: There can be certain errors added to labels data from customer system, for example, corruption of dataset file. They might cause various problems during working with labels. CRETA can check such errors added to dataset file by the customer, and it is checked for every time the dataset is delivered from a customer to AVL.

- Label-dimension: Labels can have certain dimensions. For example, a label may contain a set of values in the form of a table instead of single value. The table can have a certain set of rows and columns. Each cell in the table would have certain numeric values in them. Let say, if the table has 5 columns and 7 rows, it would be said to have 5 x 7 dimensional label. It may happen that customer may unintentionally or intentionally, in the same way, such as label value change, make changes to the number of rows or columns by adding or deleting them. This would lead to dimensional changes of labels and various issues in the same way as changes in label values.

- Label-type: Calibration dataset is composed of various types of labels, such as value type label, standard curve, standard map, group map, value block, value matrix etc. Due to same reasons, as discussed in the label value changes, the label type data may be changed intentionally or unintentionally. AVL calibration engineers have to be informed of such changes in the calibration process for their responsible label.

These are various elements of labels for the dataset, which are regularly checked for a dataset that is received or delivered from/to a customer. The basic process of cross checking the dataset, which is received or delivered to or from a customer is shown in figure 70.
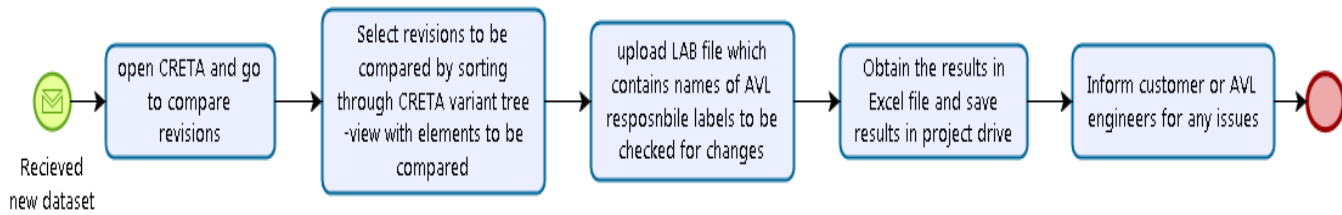
Figure 70: Steps in cross-checking received and delivered datasets

5.4.4.1 FMEA for cross-checking received and delivered dataset from customer

Based on the analysis of Fishbone diagram, developed by various brainstorming and 5 Whys methods on level 3 BPMN process modeling detailing, various failure mode and causes of failure were written down to FMEA table. The table is chosen from SAE standard J1739 2002. Various effects of failures and the RPN values were decided by a team of employees involved in above sub-process. FMEA table is shown in figure 71 and is later further explained along with points in Fishbone diagram in related to cross-checking delivered and received datasets. The main result of carrying out FMEA for this sub-process was to realize the need of adding other activities to it like switch labels check, VIP label check etc. (these labels can include labels which are not in the responsibility of AVL). Doing so would make the process more reliable and would provide betterment in the integration of calibration process between customers and AVL.

				FAILURE MODE AND EFFECTS ANALYSIS								
Item: Cross-checking of dataset				**Responsibility:** Amritpal Singh				**FMEA num** 5			
Model: Current				**Prepared by:** Amritpal Singh				**Page :** 1 of 1			
Core Team: Lead Engineer, calibration pilots, and Amritpal Singh.								**FMEA Date** 15.11.2016	**Rev:** 1		

Process Function	Potential Failure Mode	Potential Effect(s) of Failure	Sev	Class	Potential Cause(s)/ Mechanism(s) of Failure	Occur	Current Process Controls	Detec	RPN	Recommended Action(s)	Responsibility and Target Completion Date	Actions Taken	Sev	Occ	Det	RPN
Cross-checking delivered and recieved dataset	Customer made changes to labels in their responsibilities but influenced AVL assigned labels with such changes.	Various issues in calibration process or various tests for AVL Calibration engineers such as rework like switch check/VIP labels check, rechange labels values to new settings.	8		Changes made to VIP labels, switch labels, important labels etc. which are not checked by current process procedure, as it only check for AVL assigned labels only.	5	Cross check made by manually selecting various elements using compare revisions command in CRETA.	10	400	Cross-checking procedure of datasets in software tool to include other labels which are assigned to customer, but may influence AVL assigned labels (Concept: process automation).	Amritpal Singh, to be completed before end of thesis	Creation of Prototype using C# with CRETA API, to proove concept behing recommendations	4	3	7	84
	Customer made change to other elements of revision such as attributes set by AVL to handle them in new purposed solutions.	Errors in handling files for process at AVL	8		Customer unintentionally made changes	8	No process available to check such changes	10	640	Cross-checking procedure of datasets in software tool should be able to check changes in attributes/data tags assigned to them (Concept: process automation)	Amritpal Singh, to be completed before end of thesis	Creation of Prototype using C# with CRETA API, to proove concept behing recommendations	4	3	7	84

Figure 71: FMEA for cross checking delivered and received datasets based on 5 Whys analysis of acitivities and tasks in its 3rd level detailing with the BPMN modelling

5.4.4.2 Improvement/ recommendations actions for cross-checking received and delivered dataset

Various labels like labels among switch labels, certain important labels (which are related to many other labels), VIP labels, can be in responsibility of not only AVL but also in the

responsibility of the customer. Since, current procedure of cross-checking only checks for AVL responsible labels, it is recommended to make cross-checking for switch labels, important labels, and VIP labels as well. Their names should, thus, be added to a LAB file along with just AVL responsible labels. Since, dataset revisions attributes have become quite an important aspect for handling calibration data, it is recommended to check for changes made to attributes/ data tags assigned by AVL to a dataset revision, to ensure everything is completely same when receiving dataset from the customer. This process can undergo process automation, such that, when new datasets are received or delivered, these elements are automatically cross-checked, without carrying out cross-checks manually with tools like compare revisions.

5.4.5 Sub-process 6: Handling VIP labels

VIP label is AVL's conventional naming to labels which are considered to possess significant importance in the calibration process. Any change made to them would produce significant changes in the functions of other labels. VIP labels changes are important to be notified, since early quality gates of calibration project. However, after quality gate 3, almost every labels become significant to be notified of changes. Various labels can be considered as VIP, for example, hardware related labels, sensor calibration related labels, injection map, energizing time map, scientific or natural constants, basic setup models- field assignment, basic air mass set point map, DINH- Diagnostic Inhibit Handler Label etc. VIP labels usually are assigned with relevant values at the beginning of the project, and thus, possess high score/maturity at the very beginning of calibration project in early quality gates.

CRETA help handle these labels by notification tool available in it, which notify calibration pilots if any changes are made to VIP labels with respect to changes in various elements of labels such as common axis of label, values, soft and hard limits, if label map is strictly monotonic or not, resolution of label, scores, responsibility, or if new software or dataset revision is added for label to variant. It does so when DCM files with labels changes is uploaded to CRETA.

Besides these, there exists advance range check, which is similar to limit check for changes in values of labels, but with advance options. Normally, minimum and maximum values of a limit check for labels are read from an A2L file (ECU software). These values are equal for each variant. With advance range check notification, user can customize limits, but also the bit masks and check constraints to predefined DCM file (file with label value changes) values.

5.4.5.1 Various failure modes in handling VIP labels

There are certain issues associated with CRETA notification tool discussed above. Firstly, it is isolated from the CRETA main user interface and it becomes quite tedious to use. Secondly, it can produce, a lot of written notifications to read for engineers to use in some situations. For example, for calibration project, more and more labels can be tagged as VIP label as project time move ahead. After the third quality gate, changes in every label is significant. Thus, in those situations, calibration engineers would face too many notifications. Due to these reasons, calibration engineers are not using above tool to handle their important label. Thirdly, the notification is produced when calibration engineers upload changes to their labels by DCM file into CRETA. In reality, calibration

engineer makes oral communicate to work together in related labels. For example, in the regular morning meeting, they talk and share ideas with each other that they would make changes to their labels and how it will affect others labels or how others engineer's labels would affect their labels. Thus, they inform each other for the relationship between their labels. So, the current tool provides notification for changes in VIP labels later than required and is facing usability issue.
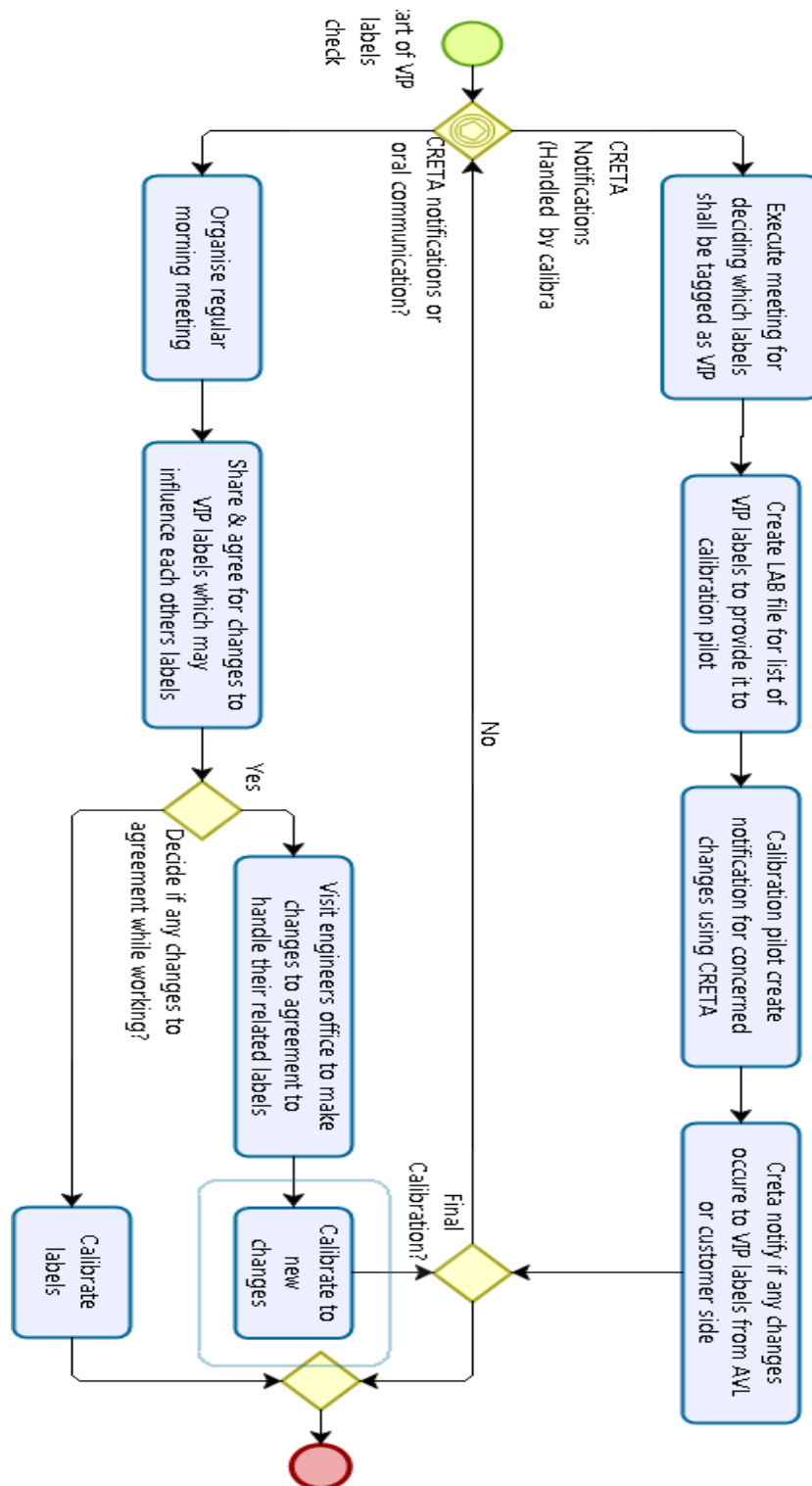
Figure 72: BPMN modeling with activities and tasks in handling VIP labels changes

5.4.5.2 FMEA for handling VIP labels

Based on the analysis of Fishbone diagram, developed by various brainstorming and 5 Whys methods on level 3 BPMN process modeling detailing, various failure mode and causes of failure were written down to FMEA table. The table is chosen from SAE standard J1739 2002. Various effects of failures and RPN values were decided by a team of employees involved in handling VIP labels. FMEA table for VIP labels handling procedure is shown in figure 73. It is later further explained along with points in Fishbone diagram in related to handling VIP labels.

	FAILURE MODE AND EFFECTS ANALYSIS										

Item:	Handling VIP labels changes	Responsibility:	Amritpal Singh		FMEA numb	6
Model:	Current	Prepared by:	Amritpal Singh		Page :	1 of 1
Core Team	Calibration pilots, calibration engineers, lead engineers, and Amritpal Singh.				FMEA Date:	01.12.2016	Rev:	1

Process Function	Potential Failure Mode	Potential Effect(s) of Failure	Sev	Class	Potential Cause(s)/ Mechanism(s) of Failure	Occur	Current Process Controls	Detec	RPN	Recommended Action(s)	Responsibility and Target Completion Date	Actions Taken	Sev	Occ	Det	RPN
Handling VIP labels changes	No usage of current VIP labels changes notification tool from CRETA leading to misscommunication issues between engineers working on related labels.	Poor handling of important labels throughout the calibration process requiring more efforts to be made.For example, engineers unintentionally often makes changes to label which influence many other important labels.	8		Current label changes notification tool posses usability issues because it create too many notification notes, it is isolated from users, and it produces changes notification later than required. This lead to oral coomunications to handle these important labels.	8	Oral communications in meeting or in rear occasions use of CRETA notification tool.	10	640	Data tagging labels VIP in Datasets. Then, using data visualisation with node-link diagrams for monitoring relation of VIP labels with each other and highlighting changes made. Written notifications only for 'number of changes' violated (Concept: Data tagging and process automation).	Amritpal Singh, to be completed before end of thesis.	Creation of Prototype using C# with CRETA API, to proove concept behind recommendations & case studies to prove that node-link diagram are more efficient than written notifications.	4	3	7	84

Figure 73: FMEA for handling VIP labels based on 5 Whys analysis of acitivities and tasks in its 3rd level detailing with the BPMN modelling

5.4.5.3 Improvements/ recommendations actions for handling VIP labels changes

The notification tool provides notification in windows form with notification list for each label changed. Let say, a VIP label has a relationship with more than 30 labels and its values are changed which violates 30 labels. Receiving 30 written notifications would be tedious for engineers. Beside notification list of changes, it provides e-mail for changes to engineers, which is furthermore tedious. Thus, it was decided to develop a software tool that can provide notification in the form of visual relationship diagram. This would enable engineers to see the relationship of VIP label with other labels in a single picture. In new software tool, each calibration engineer should be able to mark a label VIP from labels list in dataset revision. They should be able to mark it by just right click and should also be able to define relationship for which change of VIP label would influence related labels in

which manners, such as changes in common axis of label, values, soft and hard limits, if label map is strictly monotonic or not, resolution of label, scores, responsibility, or if new software or dataset revision is added for label to variant. Software tool should save this data and should automatically create a node-link diagram, as a data visualization tool, showing relationships of all the labels marked by each calibration engineer team. Labels can be nodes and relationships between them can be arrows. The label size in the diagram can be proportional to the number of relationships related to the label. Label with more relationships should be shown bigger in the node-link diagram, as it would be VIP label. Since VIP labels are related to many other labels, their relationship to various other labels can be visually evaluated. Any violating change made to VIP label, affecting any label among related labels, can be visually highlighted with arrows of the node-link diagram, instead of providing written notification. Such VIP labels relationships can be saved in database as data table and a network diagram with other labels can look as follow:
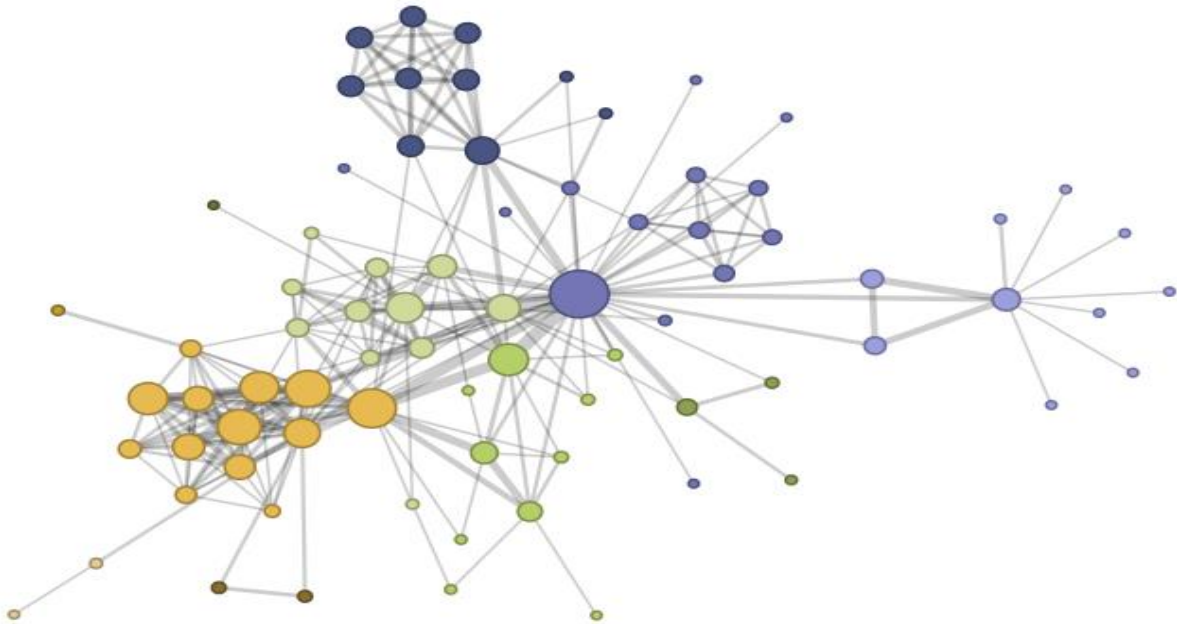
Figure 74: Node-link diagram for calibration labels relationships to each other [105]

The control phase deals with results. Thus, the control phase of DMAIC is discussed in next chapter.

Chapter 6: Prototype, results, and potential for recommended concepts to support research work

In order to evaluate the recommendation actions defined in various FMEAs in analysis phase precisely and to provide them newer RPN values for each, it was decided to build a prototype with Microsoft Windows form application by C# programming language using AVL CRETA API. Moreover, prototypes are very important to convey the idea behind the solutions in the development process. Ideo, one of the most innovative organization in the world from Silicon Valley, recommends to always take a prototype of a product for meeting with clients to enable them to understand ideas in an efficient manner [106]. Every row in the CRETA database is identified by a unique identifier called SUID (String Unique ID). The data type of this column is VarChar2 (250) in ORACLE and Text(250) in MS – Access. This field is never empty!

Example: „65ED59M-7F9D4BD1A55B425DAD162646D7F4F050"

The SUID is basically an UUID / GUID. The API uses CRETA – Tasks and therefore CRETA must be installed on the same machine including all required support files for CRETA. All items in the variant tree-view (projects, revisions, software version, change packages etc.) are called nodes and are stored in APP_NODE. The API handles read/write to the table. Multiple return values are separated with tabulator (CHR(9)).

It was decided to build a prototype for proving not all recommendation actions, because of limited time, but the concepts behind them, such as:

- Data classification: Data tagging (proved by Microsoft GUI prototype),
- Forecast analysis (proved by exponential smoothing forecasting in Excel),
- Data visualization (proved by visual network diagram and case studies),
- Business process automation with software tool such as automation of data cross-checking for data cleaning (proved by Microsoft GUI prototype)

The prototype is made for switch labels check and Quality Gate template, proves all above concept, except forecast analysis, which was proven separately using dummy data of calibration maturity from various projects by applying exponential smoothing using Microsoft Excel. Besides this, data visualization for proving ease of interpretation of labels relationships instead of written notifications is provided later.

6.1 Microsoft Windows GUI prototype

Firstly, AVL CRETA is opened. A dummy project was created using certain ECU software and datasets to create a variant tree-view in CRETA with various variants and their revisions on left side of CRETA main interface. To make it easier for sorting the data from the variant tree-view, to easily search for the required variant, data tagging is done using attributes tab on the selected variant data form in CRETA interface. These attributes are considered as data tags, as prototype searches variants in tree-view using them. The prototype is set to search automatically, for last dataset revision of the variant among various dataset revisions assigned to it, as it is a most used revision for switch checks or checking delivered or received dataset from customer etc. Figure 75, shows the AVL

CRETA interface, with dummy project variant tree-view. The selected dataset revision is tagged with attributes shown on the right, on attributes tab.

Figure 76, shows the windows form view of the prototype. It is divided into two tabs as 'Switch Labels' and 'Quality Gates form'. The working of the prototype is written below:

1   User clicks on the search button and search window pops up called 'Set attribute values to find the Variant '. Its interface was designed to look just like CRETA attribute tab interface for user ease (shown in figure 77).

2   There exists two types of searches by data tags in the search interface of the prototype. First one, called 'Exact Attribute Value Match', which search variant for required dataset. It requires to find dataset revision by providing exactly the same attributes tagged to a variant in CRETA, i.e. same attribute types, and by the same number of attributes assigned. The second search, called 'Approximate Attribute Value Match', requires to input any data tag as the attribute, which was assigned to a variant, and it provides all the variants possessing that data tag. It is just like a filter data tool. For example, all the variants with Euro 6 attribute would appear in prototype after search, but the first one type of search only provides the variant only when each and every attribute assigned to the variant is provided for search. User selects the same data tags used as attributes for the variant in CRETA, in this search form of prototype. Let say, user chooses the attribute 'rear wheel drive', with search type 'Approximate Variant Value Match' to find all dataset with this data tag.
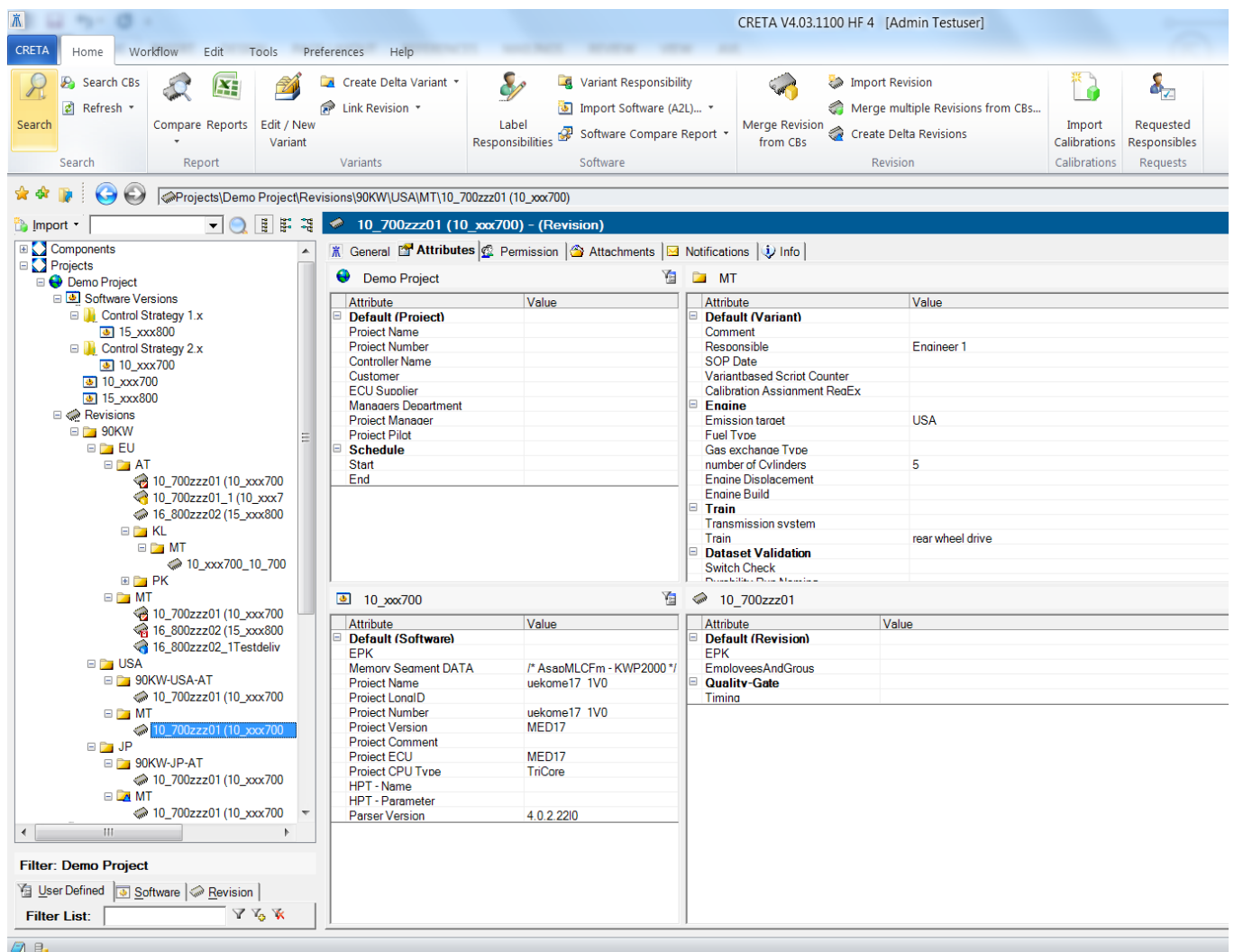
Figure 75: AVL CRETA Interface to show tagging attributes to variant for searching them

3   User then selects project name, and then selects its software's assigned, on the right side of windows. After this, user clicks confirm button and presses ok. The search windows in figure 77 closes.
4   'Fetched Variant Name' panel shows all the datasets with 'rear wheel drive' tag, on the main interface of the prototype (shown in figure 78).
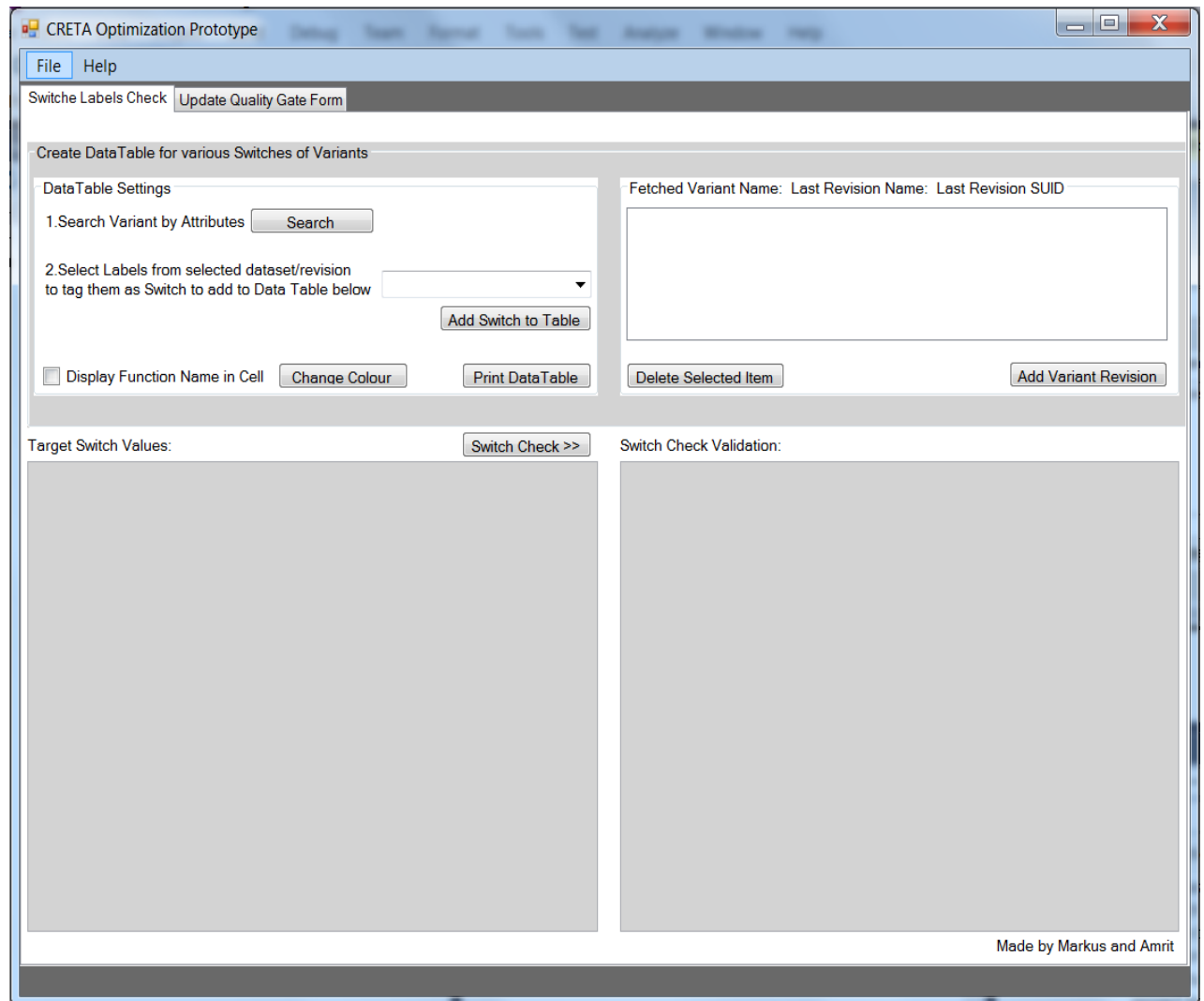
Figure 76: Main Interface of Prototype

5   User then clicks 'Add Variant Revision' button to add fetched variants from data tags to Datagridviews below for making switch labels check (shown in figure 78).
6   User then moves to panel named 'DataTable Settings', and selects switches from switches labels generated automatically from selected ECU Software, on dropdown menu named 'Select Labels from selected dataset/revision' (shown in figure 78).
7   Selected switch labels are tagged as a switch and added to Datagridview for cross checking by clicking button 'Add Switch to Table' (shown in figure 78).
8   User then inputs targets/ correct values of switches on left Datagridview table.
9   User then clicks 'Switch check' button, after which Datagridview on right shows switch labels values for Dataset in CRETA Database (shown in figure 78).

10 The wrong values are highlighted by dark orange color, to visually notify user for wrong switch values, than sending written notifications (detection type Poke-Yoke) (shown in figure 78).
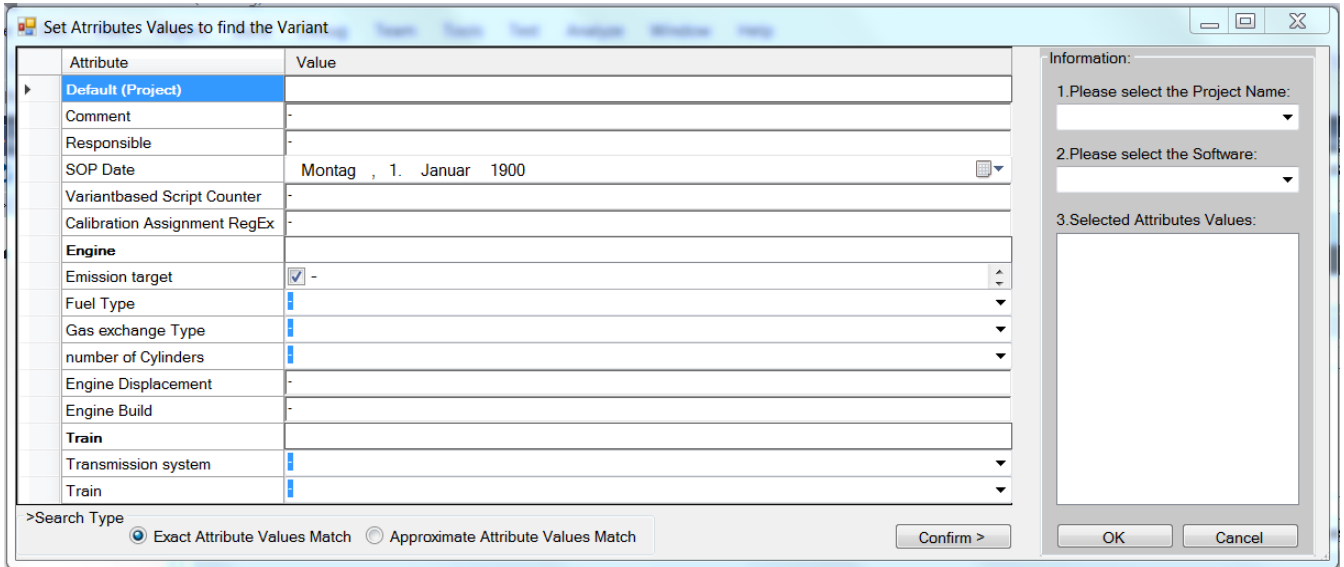
Figure 77: Search Variant Interface of Prototype, opened by clicking search button
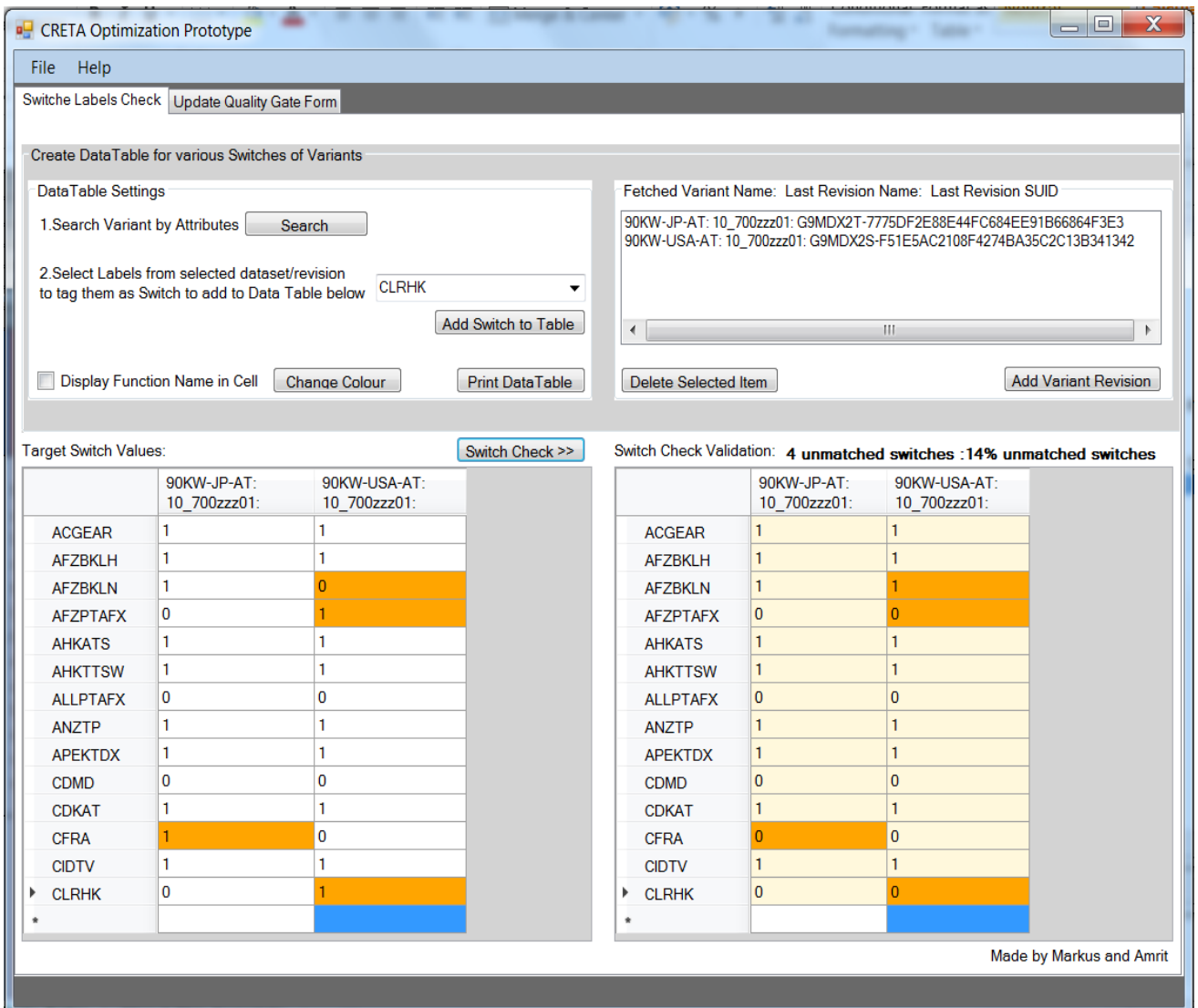
Figure 78: Shows final result of switch check on main interface

The switch labels value check in such a manner, by using automation software tool prototype, proves effectiveness of various aspects such as:

1 Data cleansing for calibration dataset in CRETA database by error detection (Poka-Yoke) for wrong switch values in ECU software dataset,

2 Process automation, as user does not have to go through a numerous number of variants in CRETA variant tree-view. They can sort required dataset revision variant by data tags/attributes with once applied attributed to variants on CRETA interface. Moreover, cross-checking of label values is done in an automated manner,

3 The result of switch labels values checks is displayed in visualized manner, instead of written notifications, with colored blocks for wrong values to user, which shows the efficiency of data visualization in calibration management.

The next tab called 'Update Quality Gate Form' was made for quality gate Excel sheets produced earlier by calibration pilot using Excel macros, to replace them by a Windows application with SQL database, to save all projects status in SQL database and manipulate by Microsoft Windows GUI. Such prototype solution removes 3 days of individual work required to produce Excel report files by macros.
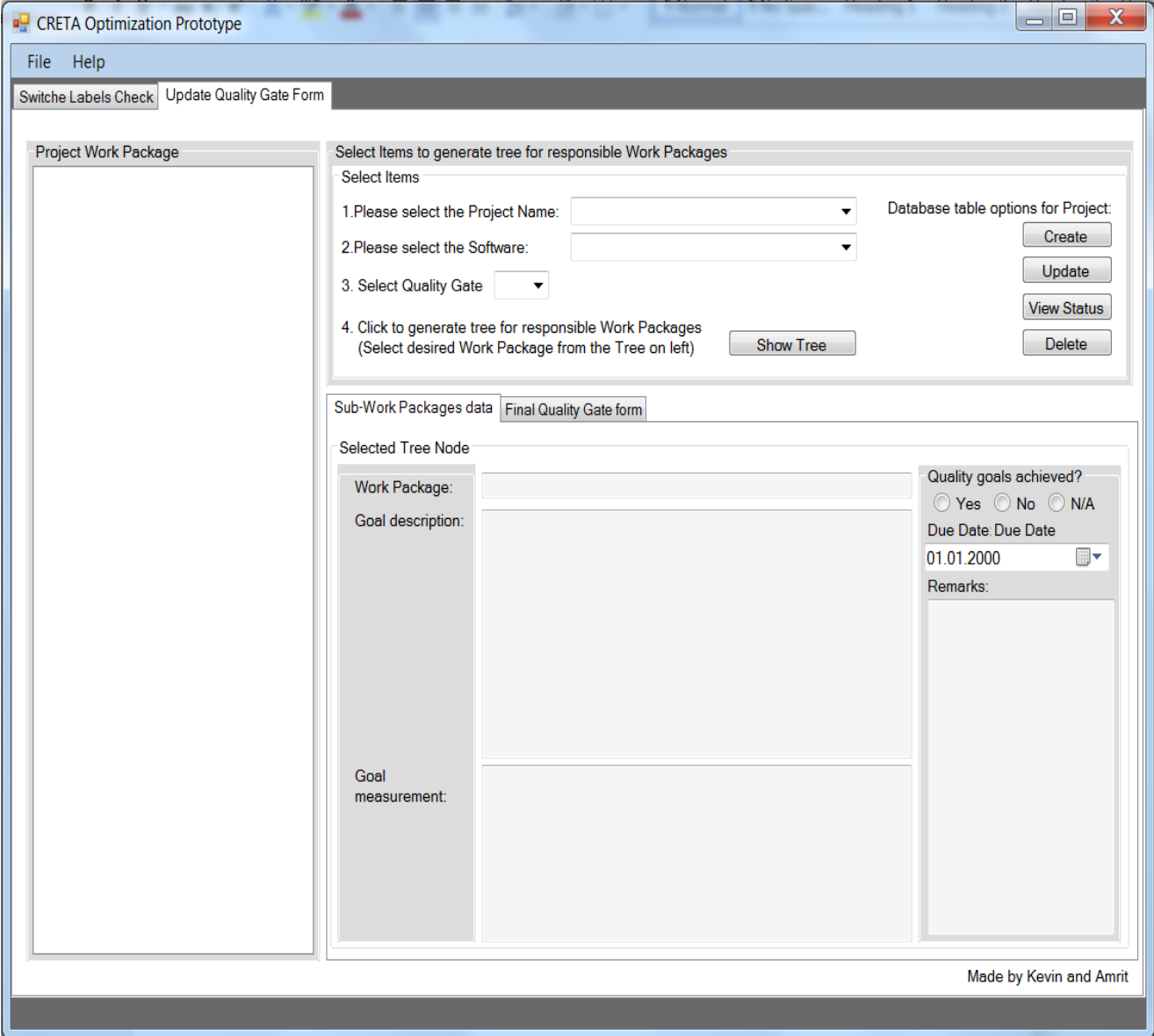
Figure 79: Interface for Quality Gate Status tab Interface

The above section of prototype work as below:

1. User selects project name from dropdown menu from 'Select Items' panel's number 1 option,
2. User then selects software inside project on step 2 of the same panel, and then quality gate for evaluating or manipulating quality gate data,
3. User then clicks the button named 'Show Tree', which generates tree-view for AVL assigned work-packages and sub-work-packages on the left panel of form.
4. User can click on (right side of Select Items panel) 'Create' button to create a database table for particular work-package selected from the tree-view. They can update it if it already exists in the back end by clicking 'Update' button. User can click 'View Status' button to see data saved before for selected work-package or sub-work-package from tree-view on left.
5. After seeing Goal description and Goal measurement for work-package (in 'Sub-Work Packages data' tab), user can click on 'Update' or 'View' buttons, to update status or check status for quality gate achievement with Yes/No/N/A radio buttons. User can set or view 'Due Date' and add 'Remarks', for example, if in case, quality gate is not achieved. This can be seen as below:
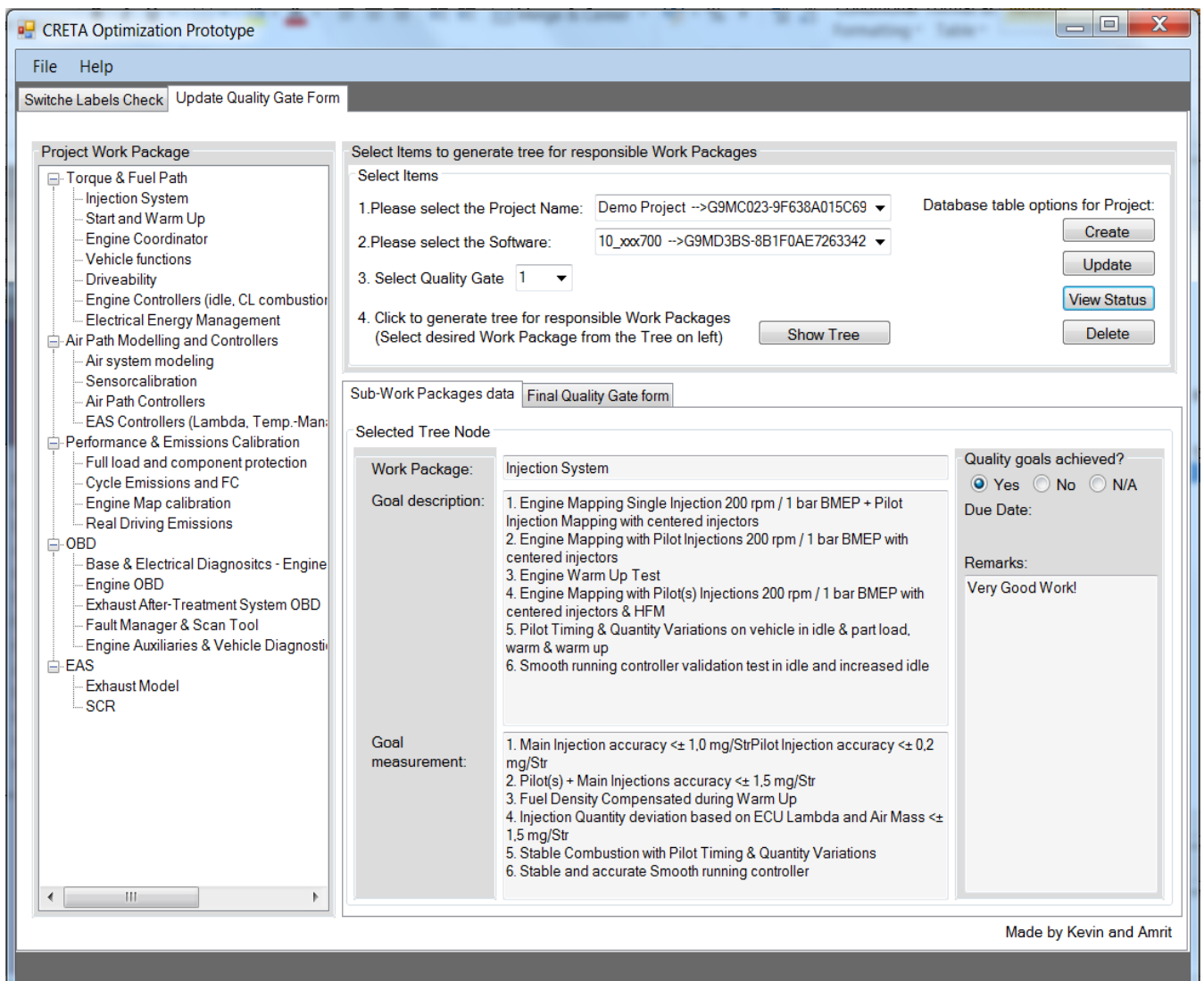
Figure 80: Quality Gate interface showing data saved or updated in prototype

'Due date' appears if user selects 'No' for 'Quality Gate Achieved' question in figure 80 on the lower right panel. 'Due Date' is provided for deciding about what date the sub-work-package targets have to be fulfilled. This status and remarks are then saved in SQL database at the backend of the prototype by clicking 'Update' button. When user clicks 'Final Quality Gate from' tab in the middle of the interface, and clicks generate report button (shown in figure 81), the target achievement status, remarks for every sub-work package assigned in 'Sub-Work Package data', appears. The one which is not achieved for the target are appeared with a due date and can be visualized inside red cells of the report. One which is partially achieved for targets is shown inside gray cell color with a due date for completion. A green cell means that the sub-work-package target is fulfilled successfully. This can be seen in Figure 81.

If user clicks 'Left data button', it shows sub-work packages for which data for status was not filled and updated by user in 'Sub-Work Package data' tab, and thus, is empty in SQL database.

Figure 81: Final report for status sub-work packages data updated

Above use of prototype help proves following concepts:

1. The significance of error prevention type Poke-Yoke, as it prevents selection of wrong sub-work-packages which are not in the responsibility of AVL. Prototype automatically only shows in the interface the sub-work-packages which are assigned to AVL for calibration, by providing process automation.
2. Secondly, it also proves the significance of data visualization as the status of various work-packages can be easily viewed for different projects, instead of written notifications.

The next concept to be proven is forecast analysis application for calibration maturity for the projects. For this, firstly, time-series graphs were generated by studying patterns of maturity for calibration work-packages of various projects of various customers with the help of calibration pilots.

6.1.1 Architecture of the Microsoft Windows GUI prototype
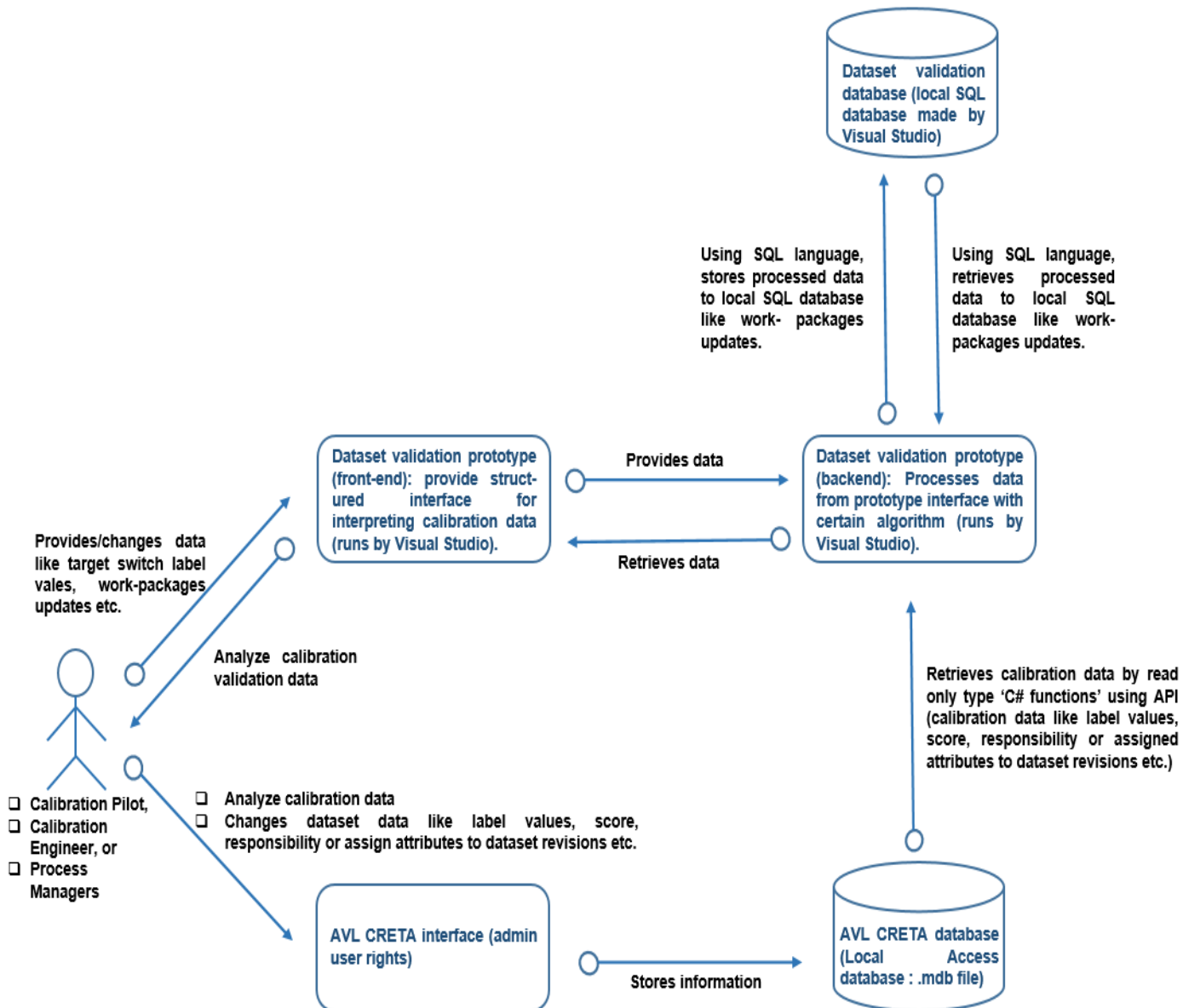
Figure 82: Dataset validation prototype architecture

In order to build the prototype, Visual Studio was used with C# programming language. CRETA API provides various C# functions in order to access database. Thus, the program built using Visual Studio used various API functions which helped to retrieve data related to switch label values, work-packages assigned to the labels from CRETA database. C# program used various C# methods, statements, LINQ queries, and various algorithms etc. to build the lists of required data at the backend. Using Microsoft GUI programming on Visual Studio for front-end, various data is shown to a user in a well-structured manner. The data of work-packages is saved in separate SQL local database made by Visual Studio, which could be changed or retrieved using prototype interface. For the thesis, a CRETA Access database was created locally on PC, from which data was retrieved using read type C# functions only, so that prototype dataset validation tool does not change the CRETA database, but can only analyze and generate information from CRETA database, which could be stored in its personal database. In order to use the prototype, user requires access to CRETA, and Visual Studio installed in the PC.

6.2 Forecasting calibration maturity: forecast analysis

The next step of research work was to prove the concept of forecast analysis for forecasting calibration maturity. For this purpose, a Microsoft Excel sheet was produced. It contained calibration maturity data for typical calibration processes at AVL. The calibration maturity (weighted average of label score of all the labels in a work-package) of various projects was studied and listed down on weekly basis on Microsoft Excel sheet. After plotting a time series graph of two years of calibration maturity for various work-packages considered for the test, it was realized that they follow 'trend' in their graph line. Hence, exponential smoothing of second order was performed using Excel shown in figure 83 as below using following formulae (the content of this method and formulae utilizes the literature provided in 4.3.3):

$$F_{t+1} = \ L_t + T_t \text{ and } F_{t+n} = \ L_t + \text{n } T_t$$
Revised estimates for Period t,
$$L_{t+1} = \propto D_{t+1} + (1 - \propto)\ (L_t + T_t)$$
$$T_{t+1} = \text{ß}(L_{t+1} - \ L_t) + (1 - \text{ß})\ T_t$$
$$D_t = \text{at} + \text{b}$$
$$T_0 = \text{a}, L_0 = \text{b}$$

1. Firstly, regression analysis is performed on data, by going in Microsoft Excel to analysis toolbar in Excel ribbon for data analysis and selecting regression. In the toolbar, Y range is entered as calibration maturity over time, shown in Excel figure 83 in the y-axis and X range is entered as weeks of time in the x-axis of the graph. The output range was selected as random cells on Excel sheet to obtain results of regression analysis.

2. This results in obtaining the 'curve intercept coefficient', and 'X variable coefficient', as a result of regression analysis on Excel sheet. The curve intercept coefficient is taken as the first Level L_t . Thus, this curve intercept coefficient value is entered into cell D2 of Excel sheet in figure 83 as first value of column for Level L_t. X variable coefficient is used for first Trend T_t. Thus, the X variable coefficient value is entered into cell E2 for making exponential smoothing as beginning values.

3. Alpha and beta constant values are first chosen randomly (better are the chosen values for Alpha and beta constants better is the forecast, and these values are optimized later) within their standard range. These alpha and beta values are used for L_{t+1} and T_{t+1} formula shown above. These formulae gives values for column D3

and E3 in Excel. In formula, D_{t+1} is as first Calibration maturity cell values in Excel table of cell C3 i.e. 1. Similarly trend T_{t+1} was calculated using L_t and T_t alpha values in cell D2 and E2. The calculated D3 and E3 cell values are dragged down for further cells up to D48 and E48 for automatically computing all the values (continued to other points at page 128).

Figure 83: Showing Excel sheet data for forecast analysis of typical calibration maturity for work-package at AVL

Summary statistics (right-hand block):

Metric	Value
Alpha	1
Beta	0.7251244
MSE	8.6618514
MAD	2.00214529
MAPE	0.1644956

Annotation (column I): "Calculated from these errors data >>>>>>"

Annotation (green cell near rows for week 45–49): "Forecasted from here"

Data table (column headers): Week | Yearly Time | (Weighted Average of Labels) | Level, L_t | Trend, T_t | Calibration Maturity % For Testing Method | Forecast, F_t | Forecast Error, E_r | Forecast Error, E_{r2} | Error Value | $|F_t - MilIM_t|$

Week	Yearly Time	Weighted Avg (C)	Cal. Maturity (F)
0	1 January 2018	1	1
1	8 January 2018	1	1
2	15 January 2018	2	2
3	22 January 2018	2	2
4	29 January 2018	2	2
5	5 February 2018	3	3
6	12 February 2018	4	4
7	19 February 2018	4	4
8	26 February 2018	4	4
9	5 March 2018	4	4
10	12 March 2018	4	4
11	19 March 2018	6	6
12	26 March 2018	7	7
13	02 Apr 18	8	8
14	09 Apr 18	8	8
15	16 Apr 18	8	8
16	23 Apr 18	8	8
17	30 Apr 18	8	8
18	7 May 2018	9	9
19	14 May 2018	10	10
20	21 May 2018	11	11
21	28 May 2018	12	12
22	4 June 2018	13	13
23	11 June 2018	17	17
24	18 June 2018	20	20
25	25 June 2018	25	25
26	2 July 2018	25	25
27	9 July 2018	25	25
28	16 July 2018	26	25
29	23 July 2018	27	26
30	30 July 2018	28	27
31	06 Aug 18	29	28
32	13 Aug 18	30	29
33	20 Aug 18	30	30
34	27 Aug 18	32	30
35	03 Sep 18	33	32
36	10 Sep 18	34	33
37	17 Sep 18	34	34
38	24 Sep 18	35	34
39	1 October 2018	36	35
40	8 October 2018	40	36
41	15 October 2018	40	40
42	22 October 2018	45	45
43	29 October 2018	50	50
44	05 Nov 18	51	51
45	12 Nov 18	51	51
46	19 Nov 18	52	52
47	26 Nov 18	52	51
48	3 December 2018	53	3
49	10 December 2018	53	4

Page: 126

In figures 84, 85, and 86, blue line represent actual calibration maturity of a work-package which actually occurred in timeline of project. The red line represent the calibration maturity data company had at a particular time. The yellowish-orange line represent forecast, and it can be compared with blue line for forecast accuracy. The green line shows smoothened trend in the curve. The y-axis represent timeline of project, and x-axis represent calibration maturity of a work-package calculated by weighted average of its label scores.
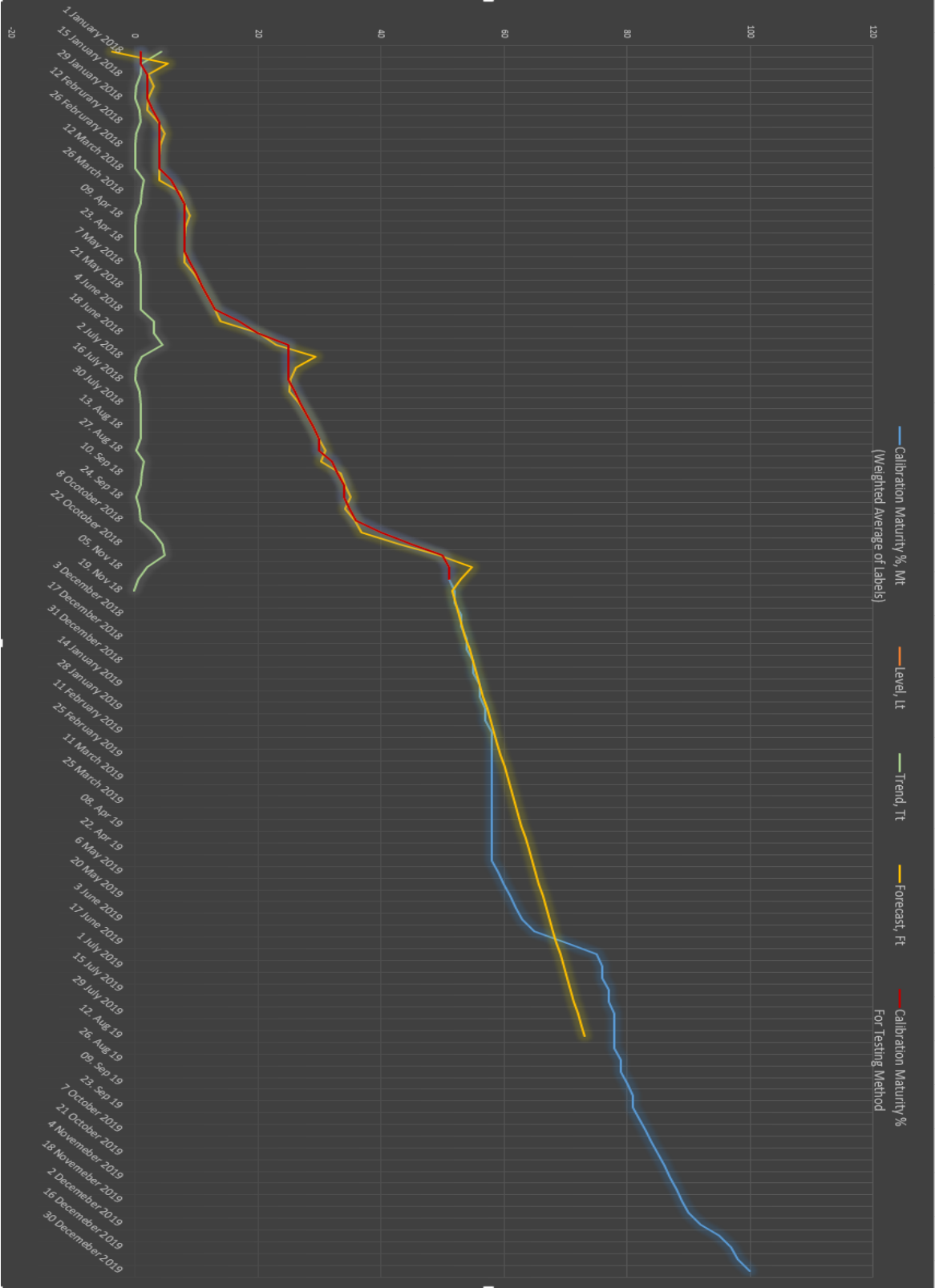

Figure 84: Forecast analysis graph

4   Forecast value is then calculated for cell G3, by adding the previous row L_t and T_t values. It is to be noted that for making the forecast for column G, calibration maturity for cells F3 to F48 were taken, as an assumption that project maturity is here at current time. 'Forecast' for calibration maturity was made, afterwards, from cell G49 to G87. Cell C3 to C107 contains calibration maturity which actually happened, to compare precision of forecast in column G. It was used to compare forecast in graph. Thus, based on data of calibration maturity shown in blue line and the forecast was made for next times shown by yellow line.

5   Forecast is made from G50 cell to later cells in same columns using F_{t+n} formula. The last cells values i.e. D49 and E 49 for values of L_t and T_t are taken to make forecast for rest of forecasted values.

6   Later, the errors are calculated in the forecast. Error values from cell J3 in column J are calculated by difference of cell values of the same row of forecasted F_t in column G, and actual calibration maturity provided in column C. Squared error values in column K, are square of column values at J3 with cells in same row. Absolute error values for column L are absolute values (non-negative) of column J with same row. Values are dragged using cursor for all below cells. Similarly, values for column M are calculated by formula in shown in header cell of column M, where M_t is calibration maturity in cells of column C.

7   Then, MSE is calculated which is an average of squared error values at column K, MAD is an average of absolute error values in column L, and MAPE is an average of column M values.

8   In the end, the calculation for best values for alpha and beta constants were made, which would give best forecasted values possible in the column. This is because, the better values of alpha and beta lowers the values of MSE, MAD, and MAPE. It was done using Excel solver tool. In the tool, set objectives are chosen as the minimum value of MSE, by changing variables: alpha and beta, subjected to constraints as alpha and beta values as between 0 and 1. This gave alpha value to be 1, and beta value as 0,725124 for the best forecast.

Similarly, two more projects were studied and their forecast analysis was performed on their calibration maturity. Both were one year projects with trend type based curve of calibration maturity increase with time. For such projects, forecasting proved to be a very effective method as well.

However, it is to be noted that there exist certain work-packages, for which calibration maturity increases with steps in the end of each quality gate. This is because all of their labels are tested in the end days of the quality gate, for example on test-beds, and all labels get calibration maturity score increase at the same time, resulting in stepwise increase in calibration maturity. For such work-packages, this technique does not make sense to be utilized. These are basically combustion related work-packages. For example exhaust after-treatment work-packages, as all their work-packages get label scores increase in the same day because they can only be tested all-together, not individually on test-bed. However, for many work-packages such as on-board diagnostics, the calibration maturity grows with stable trend based curve in general. The forecasts for more of such work-packages are shown in next pages. Sometimes, calibration maturity for certain project begins with higher maturity than 0 like 25% or 50%. This happens in cases when company take over project which customer was already working on. This case is shown in figure 85. Next to it is figure 86, which shows short 1 year project with calibration maturity beginning from 0% as in normal cases.
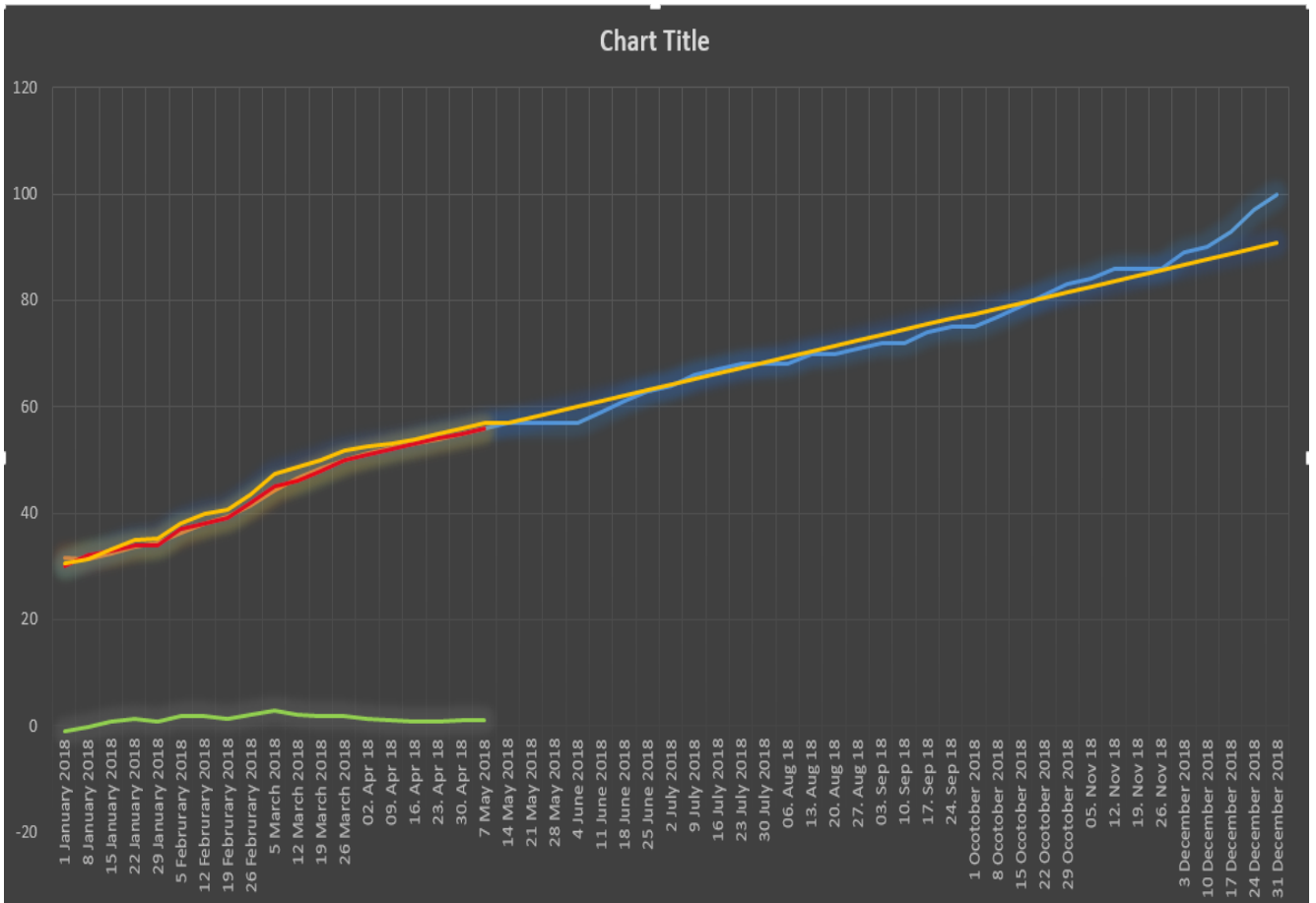
Figure 85: Graph of a short term take-over project beginning with 30% maturity
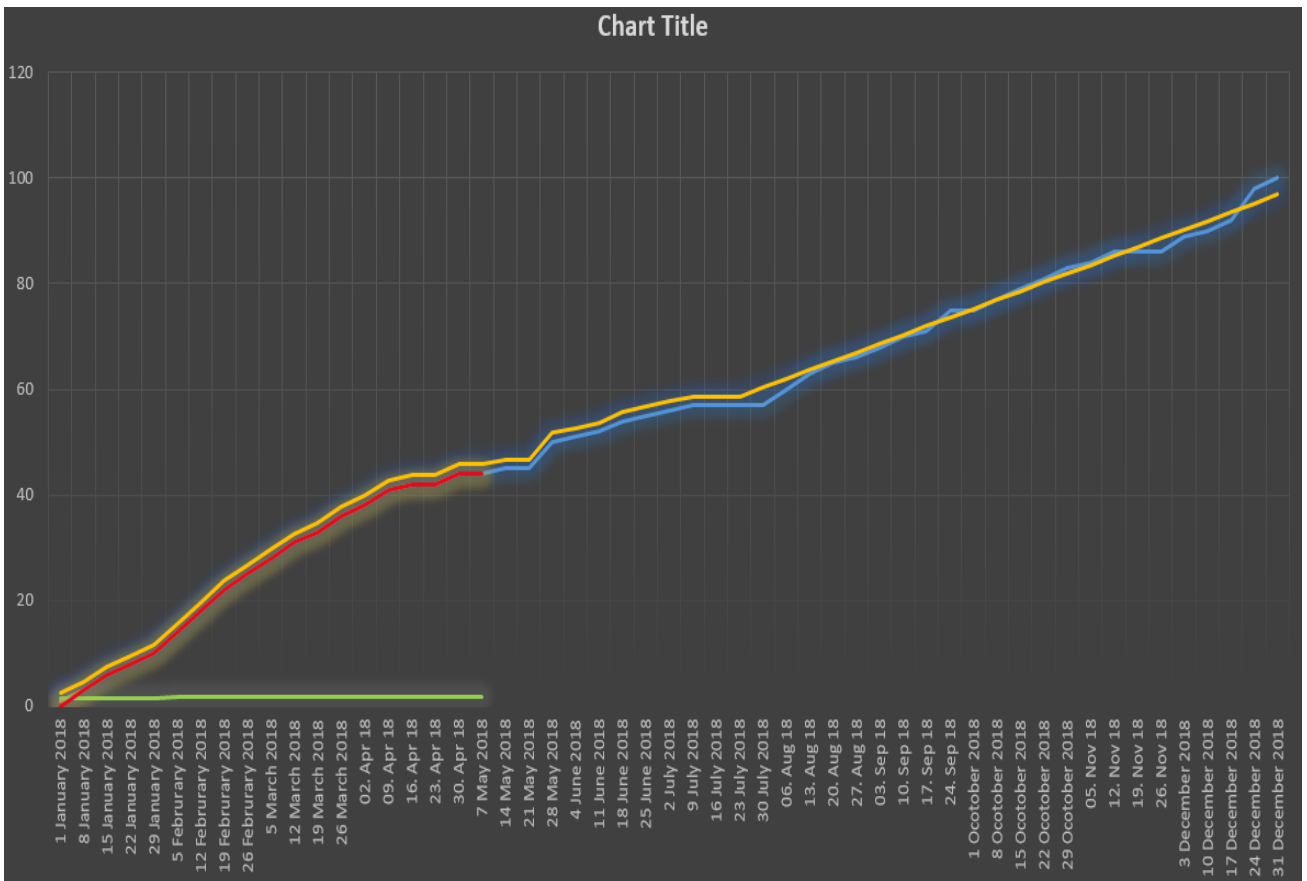

Figure 86: Graph of a short term project beginning with 0% maturity

6.3 Data visualization for changes to VIP labels: visual proof and case studies
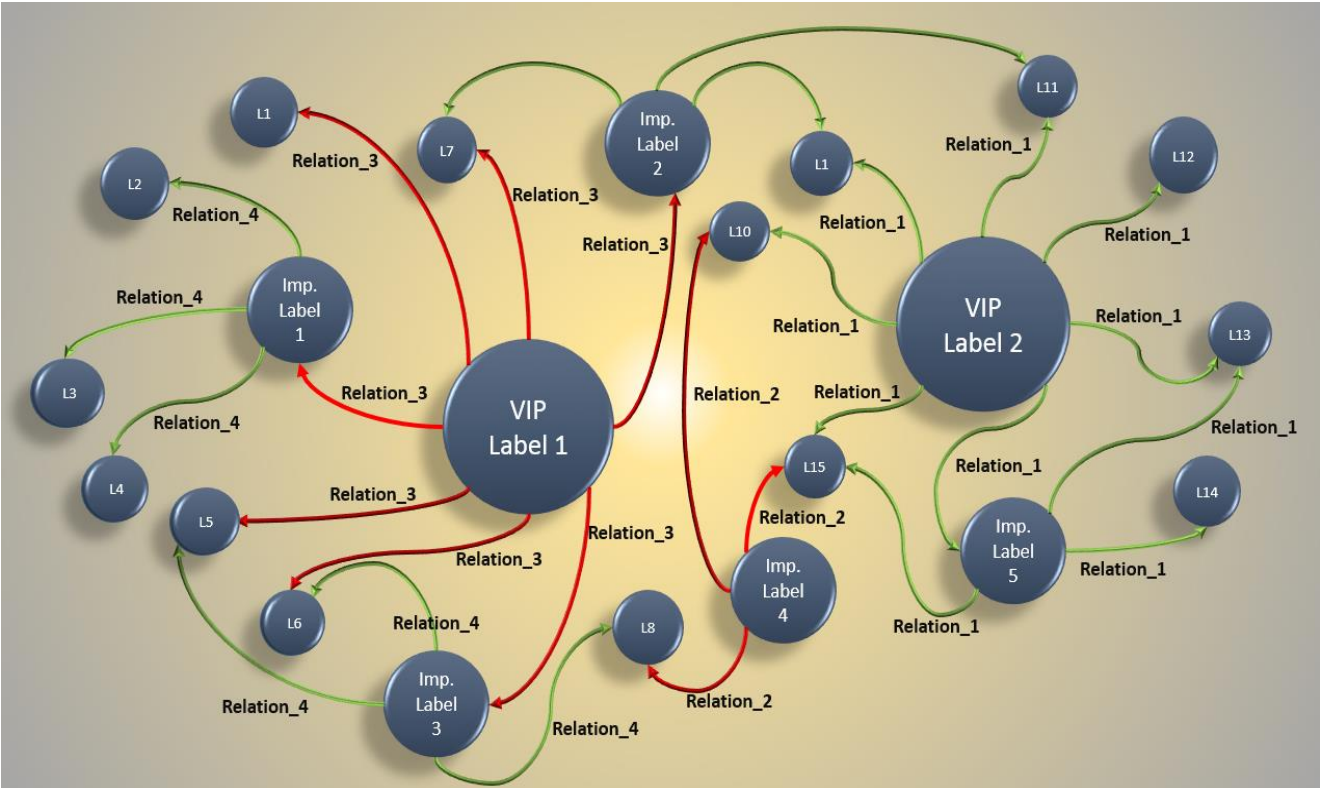

Figure 87: Node-link/ network diagram for calibration label relations

As discussed earlier, receiving a lot of written notifications would be inefficient to update calibration engineers by notifying that labels influencing their labels are changed in the certain manner of concern. To prove this, a node-link diagram was made with dummy data. Node-link diagram can replace conventional way of notification (written notes) by highlighting the nodes which are changed.

Let's consider a case, in which there exist two VIP labels in a small work-package. One called VIP label 1, which influences 7 other labels by relation named Relation_3, which can be dimension change, value change, change of soft/hard limits, and change more than range limit etc. Similarly, VIP label 2 influences 7 other labels with relation Relation_1. There exist several random labels named L1, L2, and L3 and so on. Some labels influence a number of labels with some relation, but not as much as VIP labels can be called 'important labels'. Let's name them to label 1,2,3,4.

A node-link/network diagram can then be formed showing the relationship between various labels. The size of labels can depend on a number of labels they influence, resulting in VIP labels with bigger size than other labels and thus, can be easily recognized. For example, each engineer can make data tags in CRETA for their labels, such that they tag the labels which influence their labels, or how their labels influence other labels with some relationship. Then, this data can be saved in CRETA database. Using this data, the software tool can automatically generate a node-link relational diagram for all the labels marked with relationships by each calibration engineer. This would allow all engineers to see relationships for all the labels in the same single visual diagram. VIP labels with bigger size can be easily and automatically spotted. If some relation is violated, for example, user

changes VIP Labels, which influenced many labels because of change in a relation can be highlighted. In figure 87, arrows of relation_3 of VIP label 1 is violated due to changes in it and thus is highlighted by red color (i.e. 10 written notifications replaced by 10 colored arrows). These arrows would not be highlighted with red color if a violation of labels changes did not occur. The user is easily notified visually. Notifying a user with such 10 written notifications could be tedious and the user might ignore them. It can be easily be understood by looking at figure 87, that visual notification will work much better.

6.3.1 Implementation of node-link diagram in AVL dataset management system

To implement the node-link diagram, the data table can be used named association matrix. Association matrix represents column and rows header as nodes. If there exists a pair of nodes that has a relationship with each other, their cell is filled with a property of relation that exists between them. In a case of an undirected relation between nodes, the values will be repeated. This is for the reason that each node pairing appears twice: one time for the first node as a row and the second time for a column, and then vice versa [80]. Thus, a software tool can be derived based on this. A user can mark a label VIP from label list on the interface. They can fill in the information that which label effects which label. The effecting label can be a row header, and effected label can be a column header. The relation that how they influence each other can be filled in the cell the table. The data for a relationship for various labels in a work-package can then be saved in the database as shown below:
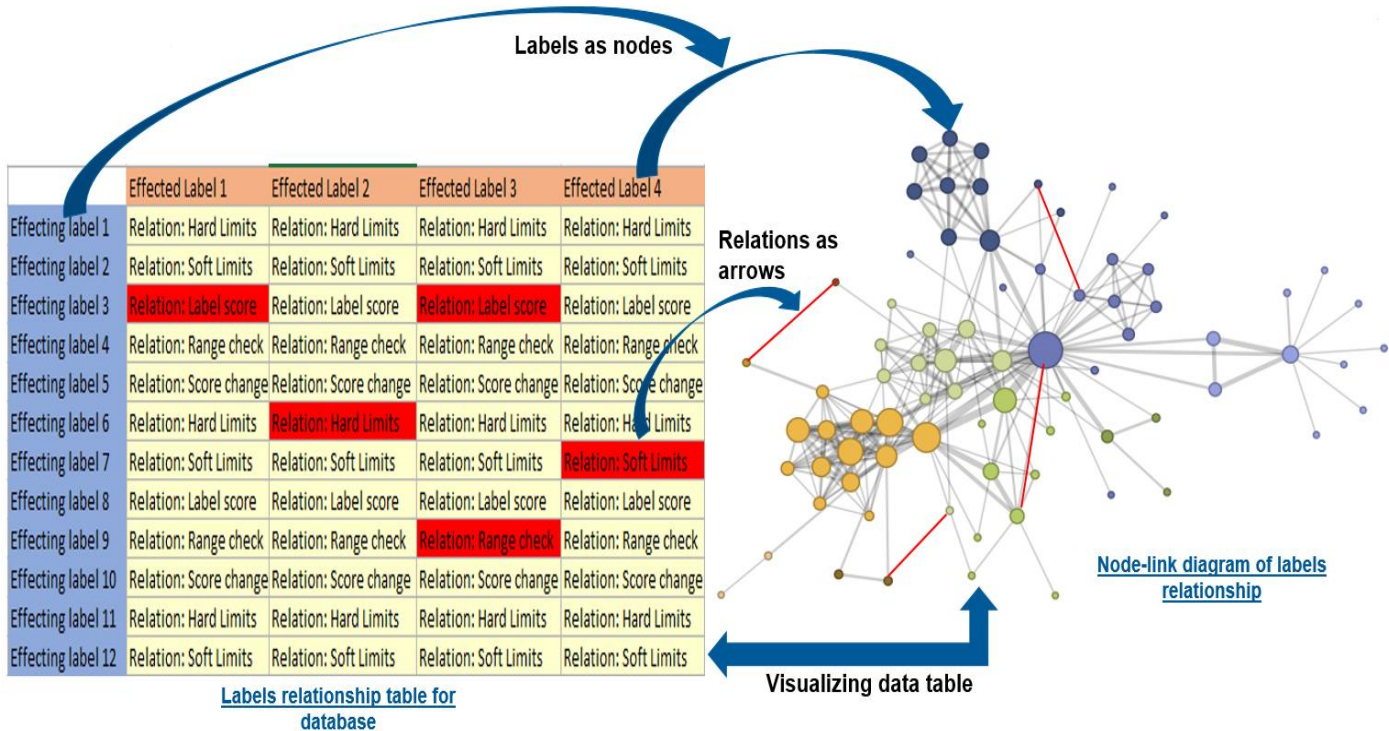
Figure 88: Implementation of node-link diagram for showing relationship between various labels of a work-package

More relationship a label would have with other labels, more would the radius of the circle of its node in the network diagram.This would enable a user to see easily which labels are most important ones. Relationship written in data table cells can be shown as arrows in the node-link diagram. Colour codes can be used in order to show effecting or effected nodes of labels. Any violation of relationship can then be marked as red arrow in order to show the influence of changes of effecting_label. In the diagram, the relationship written in

the cell of data table shall be shown when a user hovers over the mouse on the nodes arrow. Engineers shall be able to make changes to this data table for label values before making actual changes to a label on a HEX file and uploading DCM file to CRETA. In this way, they would be able to communicate for their intended changes to their label before making changes to the HEX file and producing a DCM file for changes, unlike now, where engineer get a notification about label changes when the upload their DCM files. This is the reason they communicate orally because they need to share the ideas for changes to the labels before making changes to their labels on HEX file and producing a DCM file to be uploaded on CRETA. Limitations of this method would be that the node-link diagram would provide a brief overview of relation, and would make the software tool heavier to be used for the machine. One example of a software which converts such data table/ datasets into a node-link network diagram for data visualization is 'Gelphi' [107]. Such a solution would enable simulation type of analysis of the relationship between labels that how they are influenced when making changes to label values. This would also help to generate knowledge out of the network diagram.

6.3.2 Case studies for node-link diagram for VIP labels relations data visualization

Among various data visualization techniques, node-link diagrams has the widest variety of appeal with respect to any kind of data and demand the lowest amount of technical knowledge in order to be understood. Consider the following short case studies [80]. The current CRETA tool for handling VIP labels is going through 'usability' issues in the field. This is because the current solution of written notifications is not well organized, as it produces too many notifications which makes it hard to interpret the results of label changes, and hence is not much appreciated and accepted by the calibration teams. The following case studies can answer why the node-link network relation diagram can overcome these issues. The first case study shows how a structured approach toward data with relationships between its entities can help generate a knowledge out of it, the second case studies shows how well-structured network diagrams help to better interpret the relationship between entities, and third one shows why visual tool are more accepted by humans than solutions in the form of models or written form.

1. In late 2001, Enron Corporation underwent one of the biggest bankruptcy. Its stocks plunged from 90$ per share to 0.60$ only, resulting in, ruining of employee pension and stakeholder investments. FBI came forward for investigation acquiring 300 boxes of documents, 4 Terabytes of data, and about 60,000 emails among important executives. FBI realized reading every single email wouldn't lead them anywhere since in those time phones were used a lot than just emails. Moreover, frauds are often hidden by verbal communications than written like e-mails.

FBI realized apart from reading e-mails, they need to understand the structure of communication in the organization. This would allow them to look for important employees receiving many emails and making more decisions. They could overcome this challenge by creating node-link/ network diagram of emails communication between various employees in the company. This allowed them to see visually, what e-mail were not replied, and which employee received the highest number of e-mails etc., signifying them to be important decision-makers employee for the organization.

2. In the eighteenth century, there existed one popular riddle of Königberg seven bridge problem. In those times in Königberg in Prussia, people used to try a riddle which challenged individuals to cross island having bridges between them, to find a route which

would allow them to pass through all the bridges over Pregel river only once without passing twice. It is shown in figure 89.
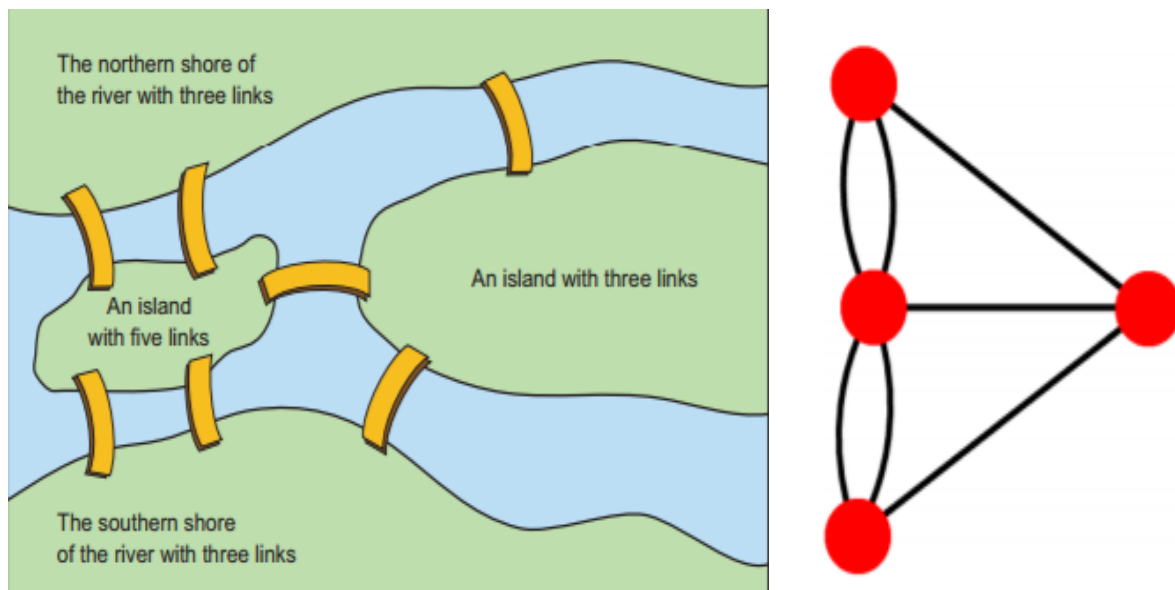

Figure 89: Map interpreted with ease by drawing node diagrams [80]

Leonhard Euler, a Swiss mathematician, solved above challenge by drawing a node-link diagram. It was obvious just by the look that nodes with an even number of bridges can be crossed easily (go and come back). However, for nodes with an uneven number of bridges, it is impossible to pass them once. Since on of node has an uneven number of bridges, and all other even, it proves that it was impossible to cross all bridges at once.

3. In 1909, the model of the atom was developed by Sir Ernest Rutherford which is similar to the node-link diagram, where electrons revolve around proton and neutron nucleus. Later it was replaced by the more accurate model developed by Erwin Schrodinger based on quantum mechanics. However, it received much less acceptability by masses, as even after 90 years, Rutherford's simple visual model is much easier to interpret the model of the atom by individuals.

Above cases answer 'Yes' to following questions:

1. Does visual network diagrams provide ease to interpret data than written forms?

2. Are visual network diagrams accepted more by users than written information?

3. Does visual network diagrams help to solve complex problems in an organized way?

6.4   Control phase (DMAIC)

The content of control phase utilize the literature provided in 2.2.1.5

It was realized after above proofs that certain changes can be executed to the calibration sub-processes drawn in BPMN models. These sub-processes can then be simplified for betterment due to recommended actions, removing inefficient steps, integrating many activities, and automating various activities. This can be explained by considering major issues resolved for each sub-process as below:

1. Switch labels check for dataset validation: Calibration engineers can save correct values for switch labels, they decide in project beginning, in a database table in matrix form with each switch label having its particular row. The values in cells of each row with respect to a different variant with certain attributes/data tags can be assigned once by engineers. New software plugin for CRETA, can then map these switch label values saved in database (by data tagging labels as switches among thousands of labels once, and data tagging variants with vehicle attributed, and then relate them together on the basis of these tags) to any dataset and DCM file uploaded or any dataset, which calibration pilot want to cross check. This would enable easy data cleansing for calibration data, with business process automation achieved. As a result, various steps in current BPMN process diagram of switch labels check can be removed. This will enhance the process quality with minimal errors and fast results.

2. Dataset Validation with durability run: Calibration pilots can make data tags/ assign attributes to variants on variant tree-view of CRETA. Whenever any user would download the dataset for making durability run test, the downloaded dataset name can contain the attributes of the variant with customer defined name. This will result in ease of recognizing dataset in various development stages of the dataset. This removes various inefficient steps involved in current BPMN diagram of current validation with durability runs the process. By doing so, Poka-Yoke technique of error prevention is utilized, where user is prevented from writing dataset name manually with errors. The dataset management process quality would enhance, as there will be an efficient flow of information between various systems.

3. Release quality gates template & monitoring calibration project status: New CRETA software tool would remove need of preparing separate Excel files for quality gates. All the data for quality gates update can be saved in CRETA database, easily accessible to various employees, and final report for the customer can be created in few seconds (instead of 3 days of manual task of creating Excel macros). This is achieved by using Poka-Yoke technique of error prevention, by applying business process automation principle. This would result in removing various inefficient steps involved in current BPMN diagram of these process. Forecasting calibration maturity would enable lead engineers to make better future plans, such as for working hours scheduling for a team, managing human labor, certain agreements with customers, resource allocation etc. Moreover, sharing forecasted data with the customer would enhance customer relationship.

4. Handling VIP labels: The conventional way of notifying user about changes to related labels using written notes or meetings with oral communications can be replaced by CRETA software plugin with single visual diagrams. This can connect various calibration engineers with their labels, and would enable them to efficiently communicate using single diagram. This would replace usability issue of current tools for handling VIP labels in CRETA with written notifications. Moreover, this diagram can be made automatically by programming. For example, user selects labels and mark related labels and relation with mouse clicks, and data can be saved in a matrix table. VIP labels names can be saved in first column cells, and related labels in first cell of other columns. The relationship data can be saved in each cell of matrix table. A software can easily make such node-link diagram using such matrix table with nodes as labels and relationship data in each cell as arrows of nodes.

5. Validation for cross-checking delivered and received dataset: Current process checks only AVL responsible labels for any changes made by the customer. Newly

purposed sub-process would enable cross check including VIP labels, switch labels which are often not in the responsibility of AVL engineers, and enable better data management. Any issue found can be resolved by instant data cleaning, enabled by the concept of business process automation.

6.4.1 Integration of the solutions into the sub-processes

All of the recommendation actions can be applied with the help of a software plugin to CRETA that utilizes all of the recommendation actions to solve various issues using basically the concepts proved in the thesis. Such a solution, developed by studying the calibration process minutely by the methodologies mentioned in the thesis, provides a very practical solution to be used in the field. It would replace or simplify various activities or tasks which together build up the sub-processes optimized. This can be seen if the reader goes back to various sub-processes modeled for each sub-process and whole calibration process. The basic architecture of final solution would be similar to the architecture of prototype developed as below:
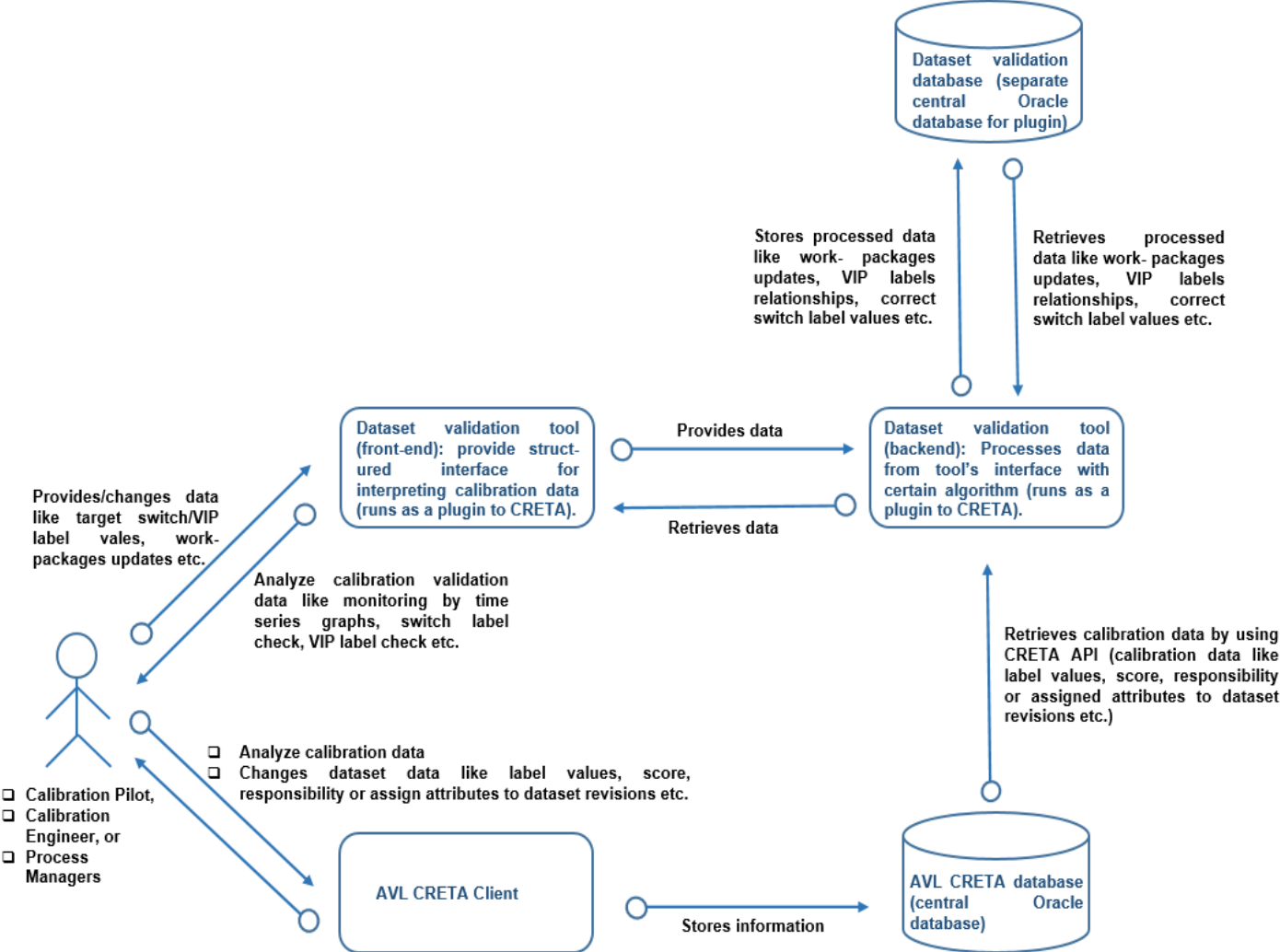

Figure 90: Architecture for actual solution for the calibration process dataset management

In actual situation, a separate oracle database would be provided for data for dataset validation solution.

6.4.2 FMEA of newly purposed sub-processes (results)

Considering the newly purposed sub-process plan, with prototypes and other proofs made, the RPN values were easily and efficiently calculated, reflecting the potential of recommendation actions made using FMEA in 'analysis and improvement' phases of DMAIC. The FMEA table is shown in figure below:

FMEA: Level 2 detailing

Item: Calibration process	Current	Responsibility: Amritpal Singh	FMEA number: 1
Model:	Current		Page: 1 of 1
Core Team: Amritpal Singh, Lead engineers, Calibration engineers, Calibration pilots		Prepared by: Amritpal Singh	FMEA Date (Ori) 12.12.2016 — Rev: 2

Process Function	Potential Failure Mode	Potential Effect(s) of Failure	Sev	Class	Potential Cause(s)/ Mechanism(s) of Failure	Occ	Current Process Controls	Det	RPN	Recommended Action(s)	Responsibility and Target Completion Date	Actions Taken	Sev	Occ	Det	RPN
Calibration process	Unreliable switch label check.	Customer recieve dataset with incorrect switch labels values.	8		Incorrect dataset validation by Excel sheet due to various steps involved in procedure.	7	Ad hoc Excel validation file.	10	560	Data cleaning by business process automation using data classification.	Amritpal, before end of thesis.	Microsoft GUI prototype using CRETA API with C#.	4	4	2	32
	Unreliable VIP label handling.	Customer recieve dataset with VIP labels values which are incorrect or may influence other labels in wrong manner.	8		Too many notifications by CRETA tool leading to oral communication for managing VIP labels.	10	Process available but not used by calibration Team (tool usability issue).	10	800	Data visualisation notifications by node-link/ network diagrams.	Amritpal, before end of thesis.	Data visualization diagram, and case studies.	6	6	4	144
	Unreliable recieved & delivered dataset check.	Customer recieve dataset with labels values	8		Current dataset validation procedure only take care of AVL responsible Labels. Customer labels that can influence AVL labels go undetected.	7	Ad hoc Excel validation file.	10	560	Data cleaning by business process automation.	Amritpal, before end of thesis.	Microsoft GUI prototype using CRETA API with C#.	5	6	4	120
	Unreliable durability run test data.	Customer dissatisfaction with wrong dataset and test results delivered.	8		Wrong naming assigned to dataset from 'INCA recorder' after test.	10	Ad hoc Excel validation file.	10	800	Forecast analysis using exponential smoothing technique.	Amritpal, before end of thesis.	Forecast analysis results in Excel.	5	6	4	120
	Incorrect monitoring for project status.	Effecting customer relationship with project delays.	8		Unreliable maturity rating for labels quality with no solution to measure future or past maturity.	10	CRETA quality gate export function.	10	800	Error prevention by Data tags saved to downloaded dataset names by business process automation.	Amritpal, before end of thesis.	Microsoft GUI prototype using CRETA API with C#.	4	4	2	32
	Quality gate template.	Customer missinformed about work-package work updates.	5		Incorrect informations in quality gate update Excel sheets.	8	Ad hoc Excel validation file.	10	400	Error prevention by business process automation by SQL based database system.	Amritpal, before end of thesis.	Microsoft GUI prototype using CRETA API with C#.	4	4	2	32
	Assign responsibilities.	Wrong labels assigned causing problems in smooth workflow.	5		Wrong interpretation of decided responsibilities.	7	Responsibilities decided in meeting & assigned by CRETA	10	350							0

Figure 91: FMEA analysis for newly purposed sub-process plan with various solutions

6.4.3 Results for company case study (chapter-1)

At the beginning of thesis in introduction chapter, various issues were discussed, with which current calibration process is struggling due to various growing projects complexities. These can now be resolved, as explained below:

- Process workflow: Due to business process automation and integration, achieved by recommendation actions and proved by software prototype, newly purposed dataset validations and management process will have a better workflow. As various inappropriate activities or tasks in each sub-process would be removed. Business process automation in sub-processes like switch/VIP label checks, work-packages status updates, cross checking delivered or received dataset etc. would replace various ad hoc solution with a well-optimized solution developed by carrying a holistic view on the whole process. The increase in efficiency of process workflow can be measured by analyzing the improvement in RPN values in FMEA of various sub-process shown before.
- Collaborations and traceability: Purposed plan for VIP labels handling with data visualization in the single diagram for labels in the responsibility of different engineers, and creating a database for various data for quality gates, switch labels correct values, monitoring project status by time-series graphs using forecast analysis etc. would enable better collaborations and traceability in the calibration process. The increase in efficiency of collaborations and traceability in the process can be measured by analyzing the improvement in RPN values in FMEA of various sub-process shown before.
- Responsibility management: With a better plan for cross-checking delivered and received dataset including validation for switch labels, VIP labels, automated quality gate data management, monitoring project status by forecast analysis, better responsibility management can be achieved in the calibration process. The increase in efficiency of responsibility management in the process can be measured by analyzing the improvement in RPN values in FMEA shown before of various sub-process discussed to this point.
- Conflict free data merging: With better monitoring of VIP labels changes, switch label values, more reliable cross-checking of delivered and received datasets, better data merging can be achieved with newly purposed validation methods and data management. The increase in efficiency of conflict free data merging in the process can be measured by analyzing the improvement in RPN values in FMEA shown before of various sub-process discussed to this point.
- Reporting and documentation: New purposed quality gate data management system can enable teams to instantly produce status/report of the project for a customer, including past, present, and future data of project using time series diagrams. This will result in producing efficient report and documentation for lead engineer and customers. The increase in efficiency can be measured in same manner as with FMEA as discussed in other points.
- Project status control: Better quality gate data management system, monitoring project status with time series graphs with forecasting project status, along with data for past and present status, would enable to lead engineers to perform better control over the project. The increase in efficiency can be measured in same manner as with FMEA as discussed in other points.

- Using know how: Saving switch label values (correct ones), VIP label network diagrams etc. in CRETA database instead of ad hoc Excel file in project drives, for different software and projects, would enable better management of calibration engineers to use their know-how for future. The increase in efficiency can be measured in same manner as with FMEA as discussed in other points.

These results were written after the agreement of calibration process stakeholders working in the process. A disadvantage with changes in the current process would be that a special training would have to be provided to calibration team for new management system. There would be requirement for more efforts for system maintainance in some areas, but making more efforts is often a pre-requisite for better quality management.

6.4.3 Purposed plans for continuous improvement for future

Apart from recommendation actions, there are various other elements that shall be performed for continuous improvement of calibration process at AVL. These are written below:

- While applying DMAIC methodology, statistical evaluation was not possible due to non-availability of error records of data management. Thus, AVL employees such as calibration pilots should prepare a plan to record errors occurred in performing the calibration process. If statistical evaluation can be made for calibration process errors, the quality level can be derived, such as 3 sigma, 4 sigma, 5 sigma or 6 sigma. Providing process quality level information (for example that our calibration process has the quality of Six Sigma) to the customer would make the customer feel more reliable on AVL.
- Various factors should be included in quality gate approval reports. For example: to share a report of quality control, results of sub- processes such as VIP labels relations, to make compulsory cross checking of datasets and data cleaning etc.
- Detailed BPMN diagrams of calibration process with various sub-processes should be constructed. This would enable easy process planning for unproductive areas in the calibration process and would enable faster business process automation at AVL. Thus, fast growing complexities in ECU software and calibration process can be handled efficiently.
- AVL employees involved in calibration process shall be provided with training with quality management tools in order to achieve high-level quality in the calibration process, as quality management is a collaborative process.
- Other important sub-processes like testbed data management, assigning responsibilities etc. which were left untreated in the thesis due to limited time or incapability to evaluate process statistically due to no record of past defects, can be worked on later, as a part of continuous improvement of DMAIC for later cases.

Chapter 7: Conclusion and final statement

The growing customer requirements, strengthened legal laws, intensified competition etc. has led to an increase in the requirements of product differentiation in automotive industry. With limited scope of growth in mechanical systems development, this is being accomplished by electronics and software technology such as ECUs and their embedded software. However, developing the ECU software possess a huge challenge for the industry, especially the last phase of the calibration process. This is because of constantly growing ECU software parameters (also called labels), with limited standardizations in calibration data management systems. ECU software calibration process, being a complex task, often requires OEMs, ECU software suppliers, and research companies to work together. However, due to insufficient developments in calibration data management, this often possess huge challenges in order to achieve end results with efficiency. This leads to process quality management issues with various errors including various project management issues. There exist certain data management solutions like AVL CRETA to handle dataset management between companies and collaboration of work for their engineers for optimization of the process. But due to growing complexities with various issues discussed, such solution need continuous improvement in order to provide better results with ease and optimal quality.

The aim of this thesis is to work out and recommend actions based on complete analysis of critical sub-processes involved in dataset management for ECU software, to overcome challenges being faced by companies for collaborative work to enhance the quality of dataset validation methods and data management process, considering an overall process with a holistic view. The research work began with getting an overall view of the calibration process, which was then studied in detail using BPMN diagrams. The flowchart was felt insufficient to evaluate the complexities of the calibration process. Thus, BPMN diagrams with their well-researched process modeling techniques, providing possibilities to model complex elements involved in the process, enabled well-structured analysis of calibration process at AVL. Just like it would be inefficient to discuss designs of mechanical components without CAD models, it would be inefficient to work together with teams to optimize calibration process without process modeling techniques like BPMN. The individuals involved in the process were able to point out visually the required areas of improvement, which were later studied in more detail using BPMN. BPMN helped to break down the overall process into sub-processes, then into activities that composed the sub.-processes, and then into tasks which together create an activity. Thus, the measurement of current performance in a quantitative manner, and efficient analysis of calibration process as a part of the renowned methodology, DMAIC, in the automotive world with its various tools, were made. This enabled developing effective recommendation actions to enhance the quality of the calibration process using its major tool FMEA. The recommendation action considers ad hoc solution being used at AVL for dataset validation for particular sub-processes and developed a better solution which can fit all of the calibration projects. The recommendation actions include data cleaning of ECU datasets with the help of process automation, visualizing complex elements of development data, and providing possibilities of efficient project management for collaborative work between companies. The calibration data provided by customers can be managed efficiently without making changes to data structures created by the customer (variant tree-view structure of customer) providing better possibilities of communication and managing data between companies. The concepts behind recommendation actions are well proven by Microsoft

GUI prototype, forecasting analysis on Excel, and brief case studies on data visualization. DMAIC methodology with its various tools like SIPOC, Ishikawa diagram, FMEA etc. provides well-structured approach and ease of communicating ideas between various individuals involved in the improvement process, for increasing quality of calibration process. Moreover, the solution developed proves the significance of analyzing processes for developing IT solutions for them. With the quality improvement initiative, the AVL employees are provided with a sense of responsibility with rewarding results by increasing efficiency in their work, which had led to an increase in motivation at work. The business process automation can bring better customer relationship and competitive advantage for AVL. Moreover, the thesis emphasizes the need of making records of defects in the quality process, which could enable the company to measure the quality performance delivered to the customer. If these quality levels are achieved to considerable levels, such as Six Sigma quality level, in future, the customer would feel more reliable on AVL, providing them with more of project work, resulting in the creation of more valuable jobs.

The master thesis carries a holistic approach to the calibration process to increase the quality in collaborative project work, with continuous improvement process for dataset validation and management. Due to this, better calibration process data management between customer and AVL is possible, overcoming the current challenges.

References

[1]   S. Arthur, H. N. Breed, and C. Schmitt-Luehmann, "Report: Shifting car makeup shakes up OEM status quo : Software strength is critical," New York, 2003.

[2]   R. Hegde, G. Mishra, and K. S. Gurumurthy, "Software and Hardware Design Challenges in Automotive Embedded System, BMS College of Engineering, UVCE, Bangalore, India," *Int. J. VLSI Des. Commun. Syst.*, vol. 2, no. 3, p. 165, 2011.

[3]   M. Conrad, "Automotive Software Engineering Teil I : Software-basierte eingebettete Systeme im Automobil, [online] available: https://www2.informatik.hu-berlin.de/~hs/Lehre/2005-WS_SWE2/20051214_AutomotiveSoftwareEngineering_I (M. Conrad).pdf, accessed on: 19 January," no. C, p. 3, 2006.

[4]   "automotive-ECU-market-report, [online] available: https://www.gminsights.com/industry-analysis/automotive-ECU-market-report, accessed on: 20 February 2017." Global Market Insights Inc., Delaware.

[5]   K. G. Sanjay Rishi, Benjamin Stanley, "Automotive 2020: Clarity beyond the chaos, IBM Institute for Business Value, 2008," 2008.

[6]   "automotive-software-engineering-defects-on-the-rise @ blog.parasoft.com [online] available: https://blog.parasoft.com/automotive-software-engineering-defects-on-the-rise, accessed on: 17 February 2017." Parasoft Corporation Headquarters, Monrovia, CA, USA.

[7]   R. Rana, "Software Defect Prediction Techniques in Automotive Domain : Evaluation , Selection and Adoption, ISBN 9789198223712, Doctoral Thesis, IT Faculty, University of Gothenburg, Sweden, 2015," 2015.

[8]   Klocwork, "Three ways to deliver safer and more secure automotive software, Rogue Wave Software, Inc.," p. 9, 2014.

[9]   S. Engineering, "Report:Software Design, Faculteit Wiskunde en Informatica, Technische Universiteit Eindhoven, 2012," pp. 1–127.

[10]   A. I. Mil and H. Approach, "Model Based ECU Development Agenda • Model-Based ECU Development – Today • Challenges of the Traditional Process • Requirements of an Integrated MiL / SiL / HiL Aproach • An ETAS Solutions • [online] available: http://www.testing-expo.com/usa/08conf/pdfs/," pp. 1–26, 2008.

[11]   AVL List GmbH, "Calibration Lifecycle Management: AVL Creta, [online] available: https://www.avl.com/documents/10138/885959/07613_Creta_PA3054E_web.pdf/027b7350-7ce8-4d22-aa44-7bd4ef1ad56a, accessed on: 15 January 2017."

[12]   T. Chandrupatla, "Quality Concepts, Cambridge University Press, ISBN 978-0-521-51522-1, 2009," *Qual. Reliab. Eng.*, p. 5, 2009.

[13]   R. Marzewski, "Chain reaction model," *Phys. Teach.*, vol. 26, no. 8, p. 514, 1988.

[14]   U.-P. D.-I. D. techn. C. Ramsauer(Curator), *Quality management, Institute of production science and management, Graz university of technology,* 7th ed. Graz, 2012.

[15]   J. M. Juran and A. B. Godfrey, *Juran's Quality Control Handbook, ISBN 0-07-034003-X*, 5th ed. McGraw-Hill, 1998.

[16]   Isixisgma, "DMAIC Versus DMADV, [online] available: https://www.isixsigma.com/new-to-six-sigma/design-for-six-sigma-dfss/dmaic-versus-dmadv/, accessed on 2 February 2017," *Isixisgma.* [Online]. Available: https://www.isixsigma.com/new-to-six-sigma/design-for-six-sigma-dfss/dmaic-versus-dmadv/.

[17]   ISO, "International Organization for Standardization, ISO 13053-1:2011 Quantitative methods in process improvement, Six Sigma, Part 1: DMAIC methodology, pg. 32, Geneva, Switzerland,09-2011." ISO, p. 32, 2011.

[18]   H. Urdhwareshe, "The Six Sigma Approach, Published originally in 2000 on web journal of Symphony Technologies, reproduced in Quality World, 2006."

[19]   D. L. Ruffner, "Understanding how Cp & Cpk are used to ensure quality, Isotemp Research Inc.," pp. 6–9, 1996.

[20]   "Six Sigma DMAIC Process - Measure Phase - Measurement System [online] available: http://www.sixsigma-institute.org/Six_Sigma_DMAIC_Process_Measure_Phase_Measurement_System.php, accessed on: 18 February 2017," *sixsigma-institute.* .

[21]   T. N. Desai and R. L. Shrivastava, "Six Sigma - A New Direction to Quality and Productivity Management," *Wcecs 2008 World Congr. Eng. Comput. Sci.*, no. 2004, p. 1047–1052\r1208, 2008.

[22]   E. Commission, L. L. P. Leonardo, and V. Transfer, "Process optimization methods, Cosima, together with European Commission, 2013," 2013.

[23]   P. Donald, "How to scope DMAIC projects, ISSN 0033524X," *Qual. Prog.*, vol. 36, pp. 37–41, 2003.

[24]   P. Jirasukprasert, J. A. Garza-reyes, H. Soriano-meier, and L. Rocha-lona, "A Case Study of Defects Reduction in a Rubber Gloves Manufacturing Process by Applying Six Sigma Principles and DMAIC Problem Solving Methodology," *2012 Int. Conf. Ind. Eng. Oper. Manag.*, pp. 472–481, 2012.

[25]   Go Lean Sixsigma, "DMAIC : The 5 Phases of Lean Six Sigma, [online] available: https://goleansixsigma.com/wp-content/uploads/2012/02/DMAIC-The-5-Phases-of-Lean-Six-Sigma-www.GoLeanSixSigma.com_.pdf, accessed on: 29 January 2017," 2012.

[26]   F. Nygren, "Lean Six Sigma in the service credit business: A case study on how the DMAIC method should be adapted on an already existing process, Teknisk- naturvetenskaplig fakultet UTH-enheten," 2016.

[27]   T. Pyzdek, *The Six Sigma handbook, ISBN 0-07-141596-3*. MacGraw-Hill, 2003.

[28]   D. Montgomery, *Introduction to statistical quality control, ISBN 978-0-470-16992-6*, 6th ed. Danvers, MA: Wiley, 2009.

[29]   Report: Improvement Skills Consulting Ltd, "Defining a process with SIPOC,"
         vol. 44, no. 0. Improvement Skills Consulting Ltd, High Wycombe, 2008.

[30]   Office of Quality Improvement, "Process Flowcharting a Powerful Tool for
         Continuous Improvement." University of Wisconsin System Board of Regents,
         Madison, WI, USA, p. 10, 2007.

[31]   S. Tikkala, "Lean Six Sigma in a Manufacturing Lead Time Improvement
         Project, master's thesis, Department of Innovation Management, School of
         Industrial Engineering and Management, Lappeenranta University of
         Technology School," 2014.

[32]   B. Aronson, "DMAIC for Interprofessional Education: An Industrial Engineering
         Approach to Primary Care Standardized Patient Simulation, [online] available:
         http://ilab.engr.utk.edu/hits/ipe/files/IPE_IE-DMAIC.pdf, accessed on: 15
         January 2017." The University of Tennessee, Knoxville, Knoxville, pp. 1–7.

[33]   A. S. Early and A. S. Possible, "Do the best you can with what you, [online]
         available:
         http://www.okstate.edu/ceat/msetm/courses/etm5291/documents/FMEA_MSET
         M_5291_1b.pdf, accessed on: 19 January 2017," *ASQ Qual. Press*, pp. 1–20,
         1998.

[34]   M. Arivananthan, "Free-flowing creativity for problem-solving, [online] available:
         https://www.unicef.org/knowledge-
         exchange/files/Brainstorming_production.pdf, accessed on: 15 January 2017."
         UNICEF, New York, p. 7, 2015.

[35]   I. Foundation, "Brainstorm - 8 Rules, [online] available: https://public-
         media.interaction-design.org/pdf/Brainstorm.pdf, accessed on: 27 January
         2017."

[36]   "Fundamentals of Quality [online] available:
         http://courses.aiu.edu/Certificate/Strategy and Management
         Quality/Fundamentals of Quality/Leccion 9/Fundamentals of Quality LA 309
         Sesion9.pdf, accessed on: 17 February 2017." Atlantic International University,
         LA.

[37]   M. Sokovi, J. Jovanovi, and A. Vujovi, "Basic Quality Tools in Continuous
         Improvement Process, Strojniški vestnik - Journal of Mechanical Engineering
         55(2009)5," vol. 55, pp. 1–9, 2009.

[38]   Z. C. and interspin CreaLab, "Fishbone Diagram, [online] available:
         http://www.becreate.ch/en/methods/fishbone-
         diagram.aspxhttp://www.becreate.ch/en/methods/fishbone-diagram.aspx,
         accessed on: 19 January 2017." Zukunftslabor CreaLab and interspin CreaLab.

[39]   M. Jan Schriefer, DrPH, Hannah King, "QI Tools: Fishbone, Pareto, Root
         Cause and FMEA, [online] availale: https://public.vtoxford.org/wp-
         content/uploads/2016/04/4_QI_Tools_FMEA_Root_Cause_Analysis_Fishbone
         _Pareto_Schriefer_King.pdf, accessed on: 4 February 2017," pp. 1–11, 2016.

[40]   C. Ltd, "Failure Modes and Effects Analysis (FMEA) and Reliability Centred
         Maintenance (RCM) Stepped Application of FMEA RCM, [online] available:
         http://www.oilandgasfundamentals.com/wp-content/uploads/OGF-AIM-Course-
         4-Module-2.pdf, accessed on 8 February 2017."

[41] P. Parikshit, V. Nair, and P. Nikunj, "A Review on use of Mistake Proofing (Poka Yoke) Tool in Blow Molding Process, ISSN: 2319-7064," *Int. J. Sci. Res.*, vol. 2, no. 2, pp. 93–97, 2013.

[42] H. Robinson, "Using Poka-Yoke Techniques for Early Defect Detection," in *Proceeding of the Sixth Annual Conference on Software Testing, Analysis and Review*, 1997, pp. 119–142.

[43] I. Fantin, *Applied Problem Solving, ISBN: 978-1499122282, Ivan Fantin, 2013.* APS, 2014.

[44] The International Marine Contractors Association, "Report: Guidance on Failure Modes & Effects Analyses," The International Marine Contractors Association, 2002.

[45] SAE, "SAE Standard, SAE J1739-2002 Potential Failure Mode and Effects Analysis." SAE, Warrendale, PA, USA, 2002.

[46] G. M. C. Crysler Corporation, Ford Motor Company, *SAE standard SAE J-1739: Potential Failure Mode and Effects Analysis (FMEA) Reference Manual, Crysler Corporation, Ford Motor Company, General Motors Corporation, 1995*. 1995.

[47] C. S. Carlson, "Understanding and Applying the Fundamentals of FMEAs, 2014 Reliability and Maintainability Symposium, January, 2014," *2014 Annu. Reliab. Maintainab. Symp.*, p. 12, 2014.

[48] N. Ramly, "Six Sigma DMAIC: Process Improvements towards Better IT Customer Support, ISSN 20103654," *Int. J. e-Education, e-Business, e-Management e-Learning*, vol. 2, no. 5, pp. 359–364, 2012.

[49] T. (Operational E. M. Liesener and Z. I. C. Ltd), "PDCA, A3, DMAIC, 8D/PSP – what are the differences? [online] available: https://www.linkedin.com/pulse/20140823033057-34652253-pdca-a3-dmaic-8d-psp-what-are-the-differences, accessed on 6 February 2017." [Online]. Available: https://www.linkedin.com/pulse/20140823033057-34652253-pdca-a3-dmaic-8d-psp-what-are-the-differences.

[50] Vector, "Measurement and Calibration of ECUs [online available]: https://vector.com/vi_embedded_measurement_calibration_en.html, accessed on: 12 February 2017." [Online]. Available: https://vector.com/vi_embedded_measurement_calibration_en.html.

[51] ETAS, "Measurement, ECU Calibration, and Diagnostics – Development Solutions for Automotive Embedded Systems." ETAS.

[52] B. Berger, "Modeling and Optimization for Stationary Base Engine Calibration, doctoral thesis, Faculty of Mechanical Engineering, TU Muenchen, Munich, Germany 2012," Technischen Universitaet Mueunchen, 2012.

[53] S. S. Gokhale and S. S. P. Kulkarni, "Modeling and optimization of Automotive Engine calibration process – A Review, ISSN 2321-5984," vol. 3, no. 1, pp. 74–83, 2015.

[54] M. Kordon, A. V. L. L. Gmbh, N. Keuth, J. C. Wurzenberger, and D. I. G. Vitale, "Model based development : Innovative ways to increase calibration quality within the limits of acceptable development effort, AVL List GmbH, Austria."

[55] H. C. Viswanatha, D. Sundar, H. Padmawar, D. G. Hickman, E. Gmbh, and K. Malpure, "AUTOMATION TECHNIQUE FOR THE VALIDATION AND RE-CALIBRATION OF BASE AIR AND FUEL ON THE VEHICLE."

[56] ETAS, "INCA Tools for Measurement, ECU Calibration, and Diagnostics, [online] available: https://www.etas.com/en/products/inca_software_products.php, accessed on: 25 January 2017." [Online]. Available: https://www.etas.com/en/products/inca_software_products.php.

[57] Vector Informatik GMBH, "Manual: AUTOSAR Calibration- Measuring and calibrating of AUTOSAR Applications with XCP and CANape," Stuttgart, 2016.

[58] AVL List GmbH, "SUCCESS STORY:Reuse of Calibration Components, [online] available: https://www.avl.com/documents/10138/885905/SuccessStory+CRETA+Liebher r.pdf/5111aa01-e5f0-4f1e-87f9-6be49934742a, accessed on: 17 January 2017."

[59] AVL List GmbH, "SUCCESS STORY:Introduce a standardized calibration process, [online] available: https://www.avl.com/documents/10138/885905/SuccessStory+CRETA+PSA.pd f/2188327f-996f-428e-903b-db244ed7cb49, accessed on: 15 January 2017."

[60] H. Pflügl, "Productivity Increase in Powertrain Calibration Development Process : Example calibration at engine testbeds [online] available on: http://www.testing-expo.com/europe/10txeu_conf/pdfs/day_2/1130_wednesday_avl_pfluegl.pdf, accessed on: 9 February 2017," pp. 1–22, 2010.

[61] G. Vitale, "AVL CALIBRATION TECHNOLOGIES [online] available: https://www.avl.com/documents/10138/1986409/01_Overview_Calibration_Virt ual_Solution.pdf, accessed on: 16 February 2017."

[62] AVL List GmbH, "AVL Concerto 4: Product Description, [online] available: https://www.avl.com/documents/10138/885979/AVL+CONCERTO+4™+Produc t+Description.pdf, accessed on: 15 January 2017," *Access*, 2015. [Online]. Available: https://www.avl.com/documents/10138/885979/AVL+CONCERTO+4™+Produc t+Description.pdf.

[63] J.-X. Wang, J. Feng, X.-J. Mao, L. Yang, and B. Zhou, "Development of a new calibration and monitoring system for in-vehicle electronic control units based on controller area network calibration protocol, ISSN 0954-4070," *Proc. Inst. Mech. Eng. Part D J. Automob. Eng.*, vol. 219, no. 12, pp. 1381–1389, 2005.

[64] T. A. and W. Dave Rogers - Automotive Engineer, "Technology focus - ECU Development - Measurement and Calibration, [online] available: http://autoelexblog.blogspot.co.at/2013/11/ecu-development-measurement-and.html, accessed on: 19 January 2017." [Online]. Available: http://autoelexblog.blogspot.co.at/2013/11/ecu-development-measurement-and.html.

[65] G. Seegers, H. Kempff, and C. J. Van Rees, "Data management and data analysis, ISSN 0304422X," *Poetics*, vol. 16, no. 6, pp. 535–552, 1987.

[66] P. Smyth and K. Lee, "An Overview of Data Mining, ISBN 9781509033676," pp. 1–21, 1997.

[67] P. Cupoli, S. Earley, and D. Henderson, "DAMA-DMBOK2 Framework," Chicago and, 2014.

[68] D. Loshin, "Monitoring Data Quality Performance Using Data Quality Metrics (White Paper)," Redwood City, CA, 2006.

[69] F. Sidi, P. H. S. Panah, L. S. Affendey, M. A. Jabar, H. Ibrahim, and A. Mustapha, "Data quality: A survey of data quality dimensions, ISBN 978-1-4673-1090-1," *Proc. 2012 Int. Conf. Inf. Retr. Knowl. Manag.*, no. March 2017, pp. 300–304, 2012.

[70] E. Rahm and H. Do, "Data cleaning: Problems and current approaches, ISBN 9781595938275," *IEEE Data Eng. Bull.*, vol. 23, no. 4, pp. 3–13, 2000.

[71] A. D. Chapman, "Principles and methods of data cleaning: PRIMARY SPECIES AND SPECIES-OCCURRENCE DATA, paper commissioned by GBIF," *Rep. Glob. Biodivers. Inf. Facil.*, no. version 1.0, p. 72, 2005.

[72] H. Müller and J. Freytag, "Problems, Methods, and Challenges in Comprehensive Data Cleansing, Humboldt-Universität zu Berlin zu Berlin, 10099 Berlin, Germany," *Challenges*, no. HUB-IB-164, pp. 1–23, 2003.

[73] Verdasys, "White paper: Data Classification: A scalable approach to managing sensitive data," Verdasys, Inc., Waltham, MA, USA, 2012.

[74] EMC, "Data Classification and Storage Optimization." EMC Corporation, Massachusetts, USA, p. 9, 2003.

[75] D. Kish, "Report: Emerging Technology Analysis : Data Classification Sophistication Will Be a Useful Tool in Data Protection, reference: G00296550," no. April. Gartner, Inc., p. 9, 2016.

[76] Frank Simorjay, "Data classification for cloud readiness," *Microsoft Trust. Comput. Doc.*, pp. 1–19, 2014.

[77] SAS, "Data Visualization Techniques: From Basics to Big Data With SAS® Visual Analytics," *White Pap.*, pp. 2–16, 2013.

[78] D. A. Keim, "Introduction to the Special Issue on Visual Analytics, ACM SIGKDD Explorations Newsletter, ISSN:1931-0145 EISSN:1931-0153," vol. 9, no. 2, pp. 3–4.

[79] T. Taylor, "Principles of Data Visualization - What We See in a Visual [online] available: http://www.fusioncharts.com/whitepapers/downloads/Principles-of-Data-Visualization.pdf, accessed on: 11 February 2017," 2014.

[80] C. L. Lanum, *Visualizing Graph Data, ISBN-13: 978-1-61729-307-8*. 2017.

[81] C. Møller and D. Ph, "Process Management," *Production*, pp. 7–25, 2012.

[82] I. Betreuer, D. Clemens, T. Schaettler, H. Packard, H. B. Mitschang, and X. Li, "Business Process Integration, masters thesis, Diplomarbeit Nr. 2152 Business Process Integration Fakultät Elektrotechnik, Informatik, Informationstechnik, Universität Stuttgart, Stuttgart, Germany 2004," 2003.

[83] G. Decker, R. Dijkman, M. Dumas, and L. Garcia-Banuelos, "The Business Process Modeling Notation, paper by Pennsylvania State University, 2010," *Mod. Bus. Process Autom. YAWL its Support Environ.*, pp. 347–368, 2010.

[84] S. Systems, "The Business Process Model, Sparx Systems [online] available: http://www.sparxsystems.com/downloads/whitepapers/businessProcessModelTutorial.pdf, accessed on: 11 February 2017." pp. 1–10, 2007.

[85] S. White, "Introduction to BPMN, IBM Corporation, ISSN 09636897," *IBM Coop.*, vol. 2, no. 0, pp. 1–11, 2004.

[86] M. Weske, "Business Process Management Concepts, Languages, and Architectures, ISBN 978-3-642-28615-5," pp. 1–10, 2012.

[87] D. Miers, "The Split Personality of BPM, Enix Xonsulting Ltd.," *BPMG Newsl.*, vol. 11, no. 11, pp. 1–22, 2003.

[88] S. Leinonen, "Business Process Improvement in ERP post- implementation context," University of Oulu, 2015.

[89] "The Institute of Chartered Accountants of India, Report: Business Process Automation through Application Software, [online] available: http://www.icaiknowledgegateway.org/littledms/folder1/chapter-5-business-process-automation-through-application-software."

[90] "White paper: Business Process Automation: A revolution in the making, Tiebeam Ventures, 2013 [online] available: http://www.digiprise.com/BusinessProcessAutomation.pdf, accessed on: 14 February 2017," 2013.

[91] U. S. Michael Gilliland, Len Tashman, *Business Forecasting: Practical Problems and Solutions*. Cary, North Carolina, USA.: Wiley, 2015.

[92] F. X. Diebold, *Forecasting in Economics, Business, Finance and Beyond*, 2015th ed. Pennsylvania: University of Pennsylvania, 2015.

[93] R. Mahendra, "Forecasting Techniques, CSIR-NISCAR, New Delgi, India."

[94] J. S. Armstrong, *Principles of Forecasting: A Handbook for Researchers and Practitioners, ISBN 0-7923-7930-6*, 2002nd ed., vol. 18, no. 3. Philadelphia, Pennsylvania USA: Kluwer Academic Publishers, 2002.

[95] M. H. H. Education, "Forecasting, McGraw Hill Higher Education, [online] available: http://highered.mheducation.com/sites/dl/free/0072443901/21708/ste43901_ch03qxd.pdf, accessed on: 23 January 2017," pp. 67–124, 2001.

[96] D. Anderson, D. Sweeney, T. Williams, J. Camm, and R. Martin, "An Introduction to Management Science: Quantitative Approaches to Decision Making, ISBN: 9781111532222," in *An Introduction to Management Science: Quantitative Approaches to Decision Making*, 2012th ed., Cengage Learning, 2011, pp. 1–896.

[97] H. V Ravinder, "Determining The Optimal Values Of Exponential Smoothing Constants – Does Solver Really Work?," *Amercian J. Bus. Educ.*, vol. 6, no. 3, pp. 347–360, 2013.

[98] N. U. S. D. of Commerce., "Engineering Statistics Handbook: Exponential

Smoothing, [online] available:
http://www.itl.nist.gov/div898/handbook/pmc/section4/pmc433.htm, accessed
on: 19 January 2017." [Online]. Available:
http://www.itl.nist.gov/div898/handbook/pmc/section4/pmc433.htm.

[99] P. Kalekar, "Time series forecasting using Holt-Winters exponential smoothing," *Kanwal Rekhi Sch. Inf. Technol.*, no. 4329008, pp. 1–13, 2004.

[100] S. Koul and J. J. Alexander, "Enhancing Project Management – a Quality Gate Usage," 2013, pp. 160–172.

[101] V. Ambartsoumian, E. T. E. D. Lee, T. Meservy, and C. Zhang, "Implementing quality Gates Throughout the Enterprise IT production Process, ISSN 1042-1319," *J. Inf. Technol.*, vol. XXII, no. 1, p. 11, 2011.

[102] J. P. Dourado and R. Silva, "Development of New Products Using Apqp and Quality Gates," *Int. J. Eng. Ind. Manag.*, pp. 79–91, 2015.

[103] T. Flohr, "Assessing the quality of quality gate reference processes," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 4980 LNCS, pp. 207–217, 2011.

[104] PROsys, "INTRODUCING THE PROsys QUALITY GATES METHODOLOGY [online] available: www.prosys-llc.com/Web_Files/qgintro1.pdf, accessed on: 9 February 2017," 1998.

[105] and V. O. Jeffrey Heer, Michael Bostock, "A Tour Through the Visualization Zoo, [online] available: https://homes.cs.washington.edu/~jheer/files/zoo/, accessed on: 4 February 2017," *Stanford University*. [Online]. Available: https://homes.cs.washington.edu/~jheer/files/zoo/.

[106] M. Buchenau and J. F. Suri, "Working Paper: Experience Prototyping, Stanford University 2000," San Francisco, 2000.

[107] Ken Cherven, *Network Graph Analysis and Visualization with Gephi, ISBN 978-1-78328-013-1*. Birmingham, UK.: Packt publishing, 2013.