Planiranje razvoja proizvoda u automobilskoj industriji

Mandić, Marko

Master's thesis / Diplomski rad

2016

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Mechanical Engineering and Naval Architecture / Sveučilište u Zagrebu, Fakultet strojarstva i brodogradnje**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:235:956598

UNIVERSITY OF ZAGREB
FACULTY OF MECHANICAL ENGINEERING AND NAVAL ARCHITECTURE

MASTER'S THESIS

Marko Mandić

Zagreb, 2016

UNIVERSITY OF ZAGREB
FACULTY OF MECHANICAL ENGINEERING AND NAVAL ARCHITECTURE

A Tool for Product Development Planning in the Automotive Industry

Mentor: Student:

Prof. Dorian Marjanović, PhD Marko Mandić

Zagreb, 2016

I declare that this thesis was written by myself using knowledge acquired during my study and from the listed literature.

Marko Mandić

ACKNOWLEDGEMENTS

I would like to express great appreciation to my mentor at the company, Mr. Thorsten Scholtz for providing me with great help for the thesis, but also for the advices on business methodology, product development, management and life in general. I would also like to express great gratitude to my mentor at the faculty, prof. Dorian Marjanović, for being the long distance support and for helping me focus on what is important while doing the research for my thesis.

I also wish to express gratitude to all the people at the company who helped me with my stay at the Germany and with the thesis. Thank you for having so much patience with my everyday questions.

A special thanks goes to my friends that were with me on my path through education, from the first day to the finish line.

I wish to express love and gratitude to my family that had nothing but love and endless support for me through my studies. Finally, I would like to express love to my girlfriend for standing next to me and pushing me forward, A. this is your work also.

SVEUČILIŠTE U ZAGREBU
FAKULTET STROJARSTVA I BRODOGRADNJE
Središnje povjerenstvo za završne i diplomske ispite
Povjerenstvo za diplomske ispite studija strojarstva za smjerove:
procesno-energetski, konstrukcijski, brodostrojarski i inženjersko modeliranje i računalne simulacije

Sveučilište u Zagrebu Fakultet strojarstva i brodogradnje	
Datum	Prilog
Klasa:	
Ur.broj:	

DIPLOMSKI ZADATAK

Student: : **Marko Mandić** Mat. br.: **0035176951**

Naslov rada na hrvatskom jeziku: **Planiranje razvoja proizvoda u automobilskoj industriji**

Naslov rada na engleskom jeziku: **A Tool for Product Development Planning in the Automotive Industry**

Opis zadatka:

Product planning defines the sequence of product development tasks. The automotive industry is recognized for its complexity in product development processes. The task of product developers is to find an optimal solution to customer requirements. Due to frequent changes before design freeze, the optimum is continually evolving. The proper cost estimation of product development is essential for the successful product development process. Based on customer requirements it is difficult to estimate the duration and the cost of the product development phases, especially in the early stages of the project. Therefore data from previous projects must be utilised.

A good estimation of resources engagement distribution is a basis for successful product developmenty, and its accuracy could be improved with the usage of history data from previously completed projects.

For the development of fuel tanks for the automotive industry, the estimation tool for the resource engagement in new product development planning should be developed. The tool should help in preparation of the future, new product development plans.

The candidate should:

- Analyze and evaluate current model for resources estimation in a fuel tank development projects;
- Identify key influence factors, based on the product specifications that affect the resources distribution throughout the project;
- Propose improvements to an existing model, design the tool for parametric resources cost estimation based on available history data and create an implementation plan;
- Propose a model for data gathering and updating considering real-time project data that could be used for tracking the project development;
- Analyze the possibility of introduction of the fuzzy methods and algorithms for utilisation of data from completed projects in resources estimation and compare it with the proposed parametric model.

Zadatak zadan:	Rok predaje rada:	Predviđeni datumi obrane:
10. ožujka 2016.	12. svibnja 2016.	18., 19. i 20. svibnja 2016.

Zadatak zadao

Prof.dr.sc. Dorian Marjanović

Predsjednica Povjerenstva:

Prof. dr. sc. Tanja Jurčević Lulić

“...Adieu, dit le renard. Voici mon secret. Il est très simple : on ne voit bien qu’avec le cœur. L’essentiel est invisible pour les yeux...“

Antoine de Saint-Exupéry, Le Petit Prince

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

OEM	Original equipment manufacturer
PFT	Plastic fuel tank
SCR	Selective catalytic reduction
NGFS	Next generation fuel systems
ECE	Engineering cost estimation
ED&D	Engineering design and development
BOM	Bill of material
DVP	Design verification process
PM	Program manager
PE	Product engineer
PQ	Product quality
TPL	Technical program leader
S	Sales
C	Customer
WBS	Work breakdown structure
SOP	Start of production
ANN	Artificial neural network
CER	Cost estimating relationship
EnC	Engineering change
PLM	Product lifecycle management
IF	Influence factor
BMM	Blow molding machine

SAŽETAK

Automobilska industrija jedno je od najzahtjevnijih okruženja u kontekstu razvoja novih proizvoda. Karakterizirana je konstantnim napretkom u proizvodnji i razvojnim procesima, što za dobavljače stvara izuzetno izazovno okruženje te ih prisiljava na kontinuiranu prilagodbu zahtjevima tržišta. Razvoj i proizvodnja spremnika za gorivo je dio industrije koja je pod utjecajem strogih emisijskih normi i uvjetovana je visokim zahtjevima proizvođača automobila. U tako dinamičnom okruženju koje ovisi o globalnim ekonomskim prilikama, od izrazite je važnosti da se projekti provedu efikasno i unutar određenog vremenskog perioda. Projekt se može nazvati efikasnim ako su ciljevi ostvareni na vrijeme i unutar određenog financijskog okvira pri tome da kvaliteta proizvoda ostane na traženom nivou. Tvrtke koje imaju dobro isplanirane i osmišljene projekte imaju prednost nad konkurencijom te imaju su veće šanse da provedu uspješnije projekte. Jedan od najvažnijih aspekata planiranja projekta je izrada financijskog plana te kreiranje kompetitivne ponude prema krajnjem kupcu. Da bi se uspješno izradio financijski plan projekta potrebno je napraviti preciznu procjenu troškova projekta. Postoji nekoliko metoda koje se mogu koristiti za predviđanje troškova, ovisno o kakvom tipu projekta se radi i u kojoj fazi se očekuje izrada procjene. Metode koje se najčešće koriste su intuitivne i spadaju u grupu kvalitativnih metoda. Međutim, za tvrtke koje imaju dugogodišnje iskustvo u proizvodnji specifičnih proizvoda, moguće je primijeniti kvantitativne parametarske metode predviđanja koje se oslanjaju na podatke o prethodnim projektima. Za promatrani slučaj, proizvodnju spremnika za gorivo, parametarsko predviđanje troškova se pokazalo metodom izbora. Parametarske metode obično koriste regresijsku analizu, iako je moguće koristiti i napredne algoritme poput neizrazitih metoda. Izrada alata za predviđanje troškova motivirana je inicijalnom analizom postojećih metoda i uočavanjem nepravilnosti u procjenama. Nadalje, činjenica da odjel program managementa nema standardne procedure u predviđanju troškova, samo doprinosi potrebi za izradu takvog alata. Fokus rada je na analizi metoda za predviđanje i povijesnih podataka te ispitivanje mogućnosti za unaprijeđenje procesa razvoja. Također, cilj je bio napraviti alat za predviđanje troškova primarno za program management odjel te ispitati mogućnosti korištenja regresijskih tehnika u usporedbi s neizrazitim metodama. Predviđanje pomoću alata omogućeno je kombinacijom parametarske metode i utjecaja stručnjaka, iz razloga što pozitivan utjecaj stručnjaka s iskustvom u planiranju projekta može doprinijeti točnosti procjene s obzirom na ograničen uzorak podataka. Alat je

napravljen na način da, ukoliko bude mogućnosti za daljnju analizu podataka i povećanje uzorka, omogući poboljšanje procjene jednostavnim editiranjem programskog koda.

Ključne riječi: predviđanje troškova, parametarska metoda, neuronske mreže, spremnik za gorivo, automobilska industrija, povijesni podaci, višestruka regresija

SUMMARY

Automotive industry has proven to be one of the most challenging environments for new product development. Constant advancements in manufacturing technologies and development processes put out by the OEMs, makes it very demanding for the suppliers to keep up the pace with the ever changing market needs. Fuel tank development and manufacturing is considered to be part of the industry impacted by the strict emission rules for automotive industry thus being heavily conditioned by the OEMs emission requirements. In industry that dynamic and with market that depends on global economy, it is of crucial importance to execute projects with great efficiency in a reasonable period of time. Project could be called efficient if the goals are met in time and within planned financial framework while quality of the product is not compromised. Companies that have well planned and thought out projects have a competitive advantage over others, and they are consequently more likely to have more positive project outcomes. One of the most important aspects of project planning is establishing the financial framework and offering a competitive quote to the customer. Way to achieve that is to estimate total project costs as accurately as possible. There are several methods for cost estimation techniques that could be used, depending on the type of the project and the phase in which the estimation is being done. Most common methods are qualitative intuitive ones. However, for the companies with proven track record and extensive experience with the particular product over the years, it is possible to use quantitative methods in the form of parametric or analytical cost estimations relying on history data of previous projects. For observed case, the fuel tank development, parametric cost estimation method has proven to be the method of choice. Parametric methods usually use regression techniques, although there are also possibilities to use advanced algorithms such as fuzzy methods. Tool creation was motivated by initial observation of the existing estimation methods and the fact that the program management department didn't have any standardized procedures in the cost estimation process. The focus of this thesis is on analysis of the estimation procedures, history data and the possibilities for improvements. Furthermore, the goal was to create the usable tool for cost estimation, primarily for program management department, and to give the overview of using the well known regression techniques as opposed to the advanced fuzzy algorithms. The tool was created in a way that allows influence of an expert on calculated estimation because positive influence can compensate possibly small sample of data. Moreover, the tool was

created in a way that if more data is to be analyzed, it would be easy to improve the estimation accuracy by simple intervention to the code.

Key words: cost estimation, parametric, neural networks, fuel tank, automotive industry, history data, multiple regression

PROŠIRENI SAŽETAK

Konstantan napredak tehnologije karakterističan za automobilsku industriju, stvara izrazito kompetitivno okruženje u pogledu opstanka i napredovanja na tržištu. To se posebice odnosi na dobavljače koji moraju držati korak s najnovijim trendovima koje nameću kako proizvođači vozila tako i sve stroži propisi regulatornih tijela. Da bi tvrtka ostala kompetitivna na tržištu mora zadovoljiti uvjete kvalitete u proizvodnji spremnika te isto tako mora na najvišoj razini izvesti proces razvoja proizvoda, s minimalnim vremenskim i financijskim gubicima. Proizvođači vozila su pod pritiskom zbog visokih emisijskih zahtjeva i to se odražava na dobavljače. U kontekstu razvoja spremnika za gorivo, to znači potrebu za neprekidnim ulaganjem sredstava u razvoj novih materijala, proizvodnih tehnologija i viših standarda kvalitete. Aktualne proizvodne tehnologije spremnika za gorivo mogu se podijeliti na dvije skupine:

- COEX (konvencionalno višeslojno oblikovanje puhanjem)
- NGFS (eng. *Next Generation Fuel Systems*)

Visoki standardi za smanjenje emisije ugljikovodika ostavljaju trag i na proizvodnji spremnika za gorivo, gdje je NGFS metoda razvijena upravo za minimiziranje emisije kroz stijenku spremnika. Emisija ugljikovodika se smanjuje na način da se sve komponente spremnika umeću tijekom puhanja u ovojnicu, dok se kod COEX metode komponente naknadno ugrađuju na proizvodnoj liniji. Naknadna ugradnja zahtijeva rezanje ovojnice te umetanje i varenje komponenti. NGFS metoda je mnogo logistički zahtjevnija i za njenu upotrebu potrebno je više resursa. U promatranom slučaju razvoja i proizvodnje spremnika za gorivo, proces razvoja proizvoda podijeljen je na 7 faza:

- Poslovna studija i zahtjevi klijenta
- Razvoj koncepta i studija izvedivosti
- Preliminarni dizajn
- Detaljni dizajn
- Validacija proizvoda i procesa
- Pilot
- 90 dana od početka proizvodnje

Za promatrani slučaj fokus je stavljen na drugu fazu u kojoj je najbitniji dio razvojnog procesa procjena troškova projekta te izrada konceptualnog modela proizvoda. Kvalitetna procjena troškova u drugoj fazi rezultira kompetitivnom ponudom prema kupcu te osigurava dobre temelje za daljnji razvoj proizvoda. Kako bi se osigurala od financijskih gubitaka tijekom projekta, tvrtka mora koristiti pouzdane metode za predviđanje troškova. Postoji mnogo načina kako se predviđaju i zapisuju troškovi projekta. Jedan od postojećih načina zapisivanja jest korištenje specijaliziranih Excel tablica koje služe za pregled cjelokupne financijske strane projekta te omogućuju odjelu prodaje da kreira valjanu ponudu prema kupcu. Tvrtka koristi tzv. ED&D tablice koje predstavljaju svojevrsni standard u zapisu resursa koje projekt zahtijeva za uspješno izvođenje. Tablica sadrži podatke kao što su: lokacije tvornica, godišnje stope cijena rada pojedinih odjela, opis projekta te tzv. mountain tablicu. Mountain tablica zapravo predstavlja mjesečni prikaz resursa projekta kroz detaljnu raspodjelu radnih sati po odjelima koji su uključeni u razvoj. Odjeli koji su uključeni u razvoj i čija podjela resursa čini pravi izazov su:

- *Program management*
- *Product engineering*
- *Advanced manufacturing*
- *CAD*
- *CAE*
- *Product quality*

Za svaki odjel potrebni sati su predviđeni od strane voditelja odjela ili odgovorne osobe u određenom projektu. Trenutno se metoda predviđanja oslanja na tri značajke: mišljenje stručnjaka, odobrenje odjela prodaje i predviđanje voditelja odjela. U procesu predviđanja resursa uloga voditelja projekta je najvažnija. Voditelj projekta ima zadaću uputiti i prikupiti sve procjene resursa od voditelja pojedinih odjela te ih evaluirati, korigirati i u suradnji s odjelom prodaje pripremiti za slanje kupcu.

Radi boljeg razumijevanja procesa predviđanja, provedena je analiza nekoliko ED&D tablica od projekata koji se nalaze u ranim fazama razvoja proizvoda. analiza je podijeljenja u tri dijela: udio odjela u ED&D tablici, distribucija resursa po stupnjevima projekta te analiza strukturne podjele poslova. 7 projekata je bilo dostupno za analizu te je uočeno kako postoje velike razlike u procjenama resursa za projekte sličnih karakteristika. Razlike su se kretale od 16 do preko 20%. Također raspodijela poslova tijekom faza razvoja proizvoda u različitim projektima nije bila slična. Analizom mape raspodijele poslovaa po odjelima tijekom projekta, utvrđeno je da je najviše sati potrebno utrošiti u prve tri faze razvoja. Neke od analiziranih tablica nisu pratile proces te je utvrđeno da ne postoji standard koji se slijedi prilikom alociranja sredstava po fazama projekta te da ne postoji određeni standard koji se mora pratiti pri procesu inicijalnog predviđanja potrebnih resursa u projektu. Nepostojanje striktnog standarda pri procjeni resursa za vrlo kompleksne proizvodne procese kao što je proizvodnja spremnika za gorivo, rezultira vrlo različitim procjenama čak i za projekte sličnih karakteristika. To navodi na zaključak da predviđanje troškova ovisi isključivo o iskustvu odgovorne osobe te je podložno subjektivnom utjecaju. Ovi zaključci motivirali su daljnje istraživanje o metodama predviđanja troškova te izradu alata koji bi smanjio utjecaj individualca na proces predviđanja.

Glavne metode predviđanja troškova razvoja proizvoda mogu se podijeliti na kvalitativne i kvantitativne. Kvalitativne metode se ogu dalje dijeliti na intuitivne i analogne. Kvalitativne metode se baziraju primarno na analizi proteklih projekata koji su slični onome za kojeg se vrši predviđanje. Glavni faktor pri dobroj kvalitativnoj procjeni jest valjana identifikacija sličnosti između prošlih i aktualnog projekta. U tom smislu, podaci o proteklim projektima mogu pružiti korisni uvid u to kako bi se razvoj novog proizvoda mogao odvijati. U promatranom slučaju korištene su upravo intuitivne kvalitativne metode. Intuitivne metode karakterizira oslanjanje na mišljenje stručnjaka i njegovo iskustvo u sličnim projektima. Glavna prednost ovakvih metoda je relativno brza procjena. Ovakav tip predviđanja izuzetno je podložan pristranosti i subjektivnom utjecaju individualca na konkretan iznos procjene. Greška procjene može se smanjiti korištenjem mišljenja više stručnjaka, no takvo predviđanje je i dalje pod utjecajem intuicije procjenitelja. Kvantitativne metode karakterizira detaljna analiza specifikacija proizvoda, njegovih značajki te vrsta proizvodnog procesa, za razliku od oslanjanja na znanje i iskustvo procjenitelja. Resursi su obično izračunati pomoću analitičkih funkcija koje se oslanjaju na varijable s najviše utjecaja na ishod projekta. Kvantitativne metode se dalje mogu podijeliti na parametarske i analitičke. Analitičke metode zahtijevaju detaljno znanje o proizvodnom procesu dok parametarske metode zahtijevaju uvid u faktore koji najviše utječu na troškove projekta. S obzirom na činjenicu da tvrtka ima bazu podataka koja sadrži dokumente i financijske planove proteklih projekata, donesen je zaključak da je parametarska metoda idealan izbor za daljnju upotrebu, s obzirom na mogućnost identifikacije utjecajnih faktora analizom podataka. Također, predviđanje se odvija u ranim fazama razvoja proizvoda te proces proizvodnje nije poznat u detalje, dok su do tada značajke proizvoda u velikoj mjeri definirane. Parametarska metoda je u suštini statistička metoda koja koristi razne alate za analizu podataka te formiranje valjanog predviđanja. Najčešći statistički alat koji se koristi je metoda regresije, dok se u posljednje vrijeme sve više koriste neizrazite metode. Za izradu alata odabrana je metoda višestruke linearne regresije, iako će biti napravljena usporedba s neizrazitim metodama u svrhu dobivanja boljeg uvida u tehnike predviđanja i njihove značajke.

Kako bi se napravili dobri temelji za izradu stabilnog alata, bilo je potrebno provesti nekoliko koraka prije samog početka izrade. Radi se prvenstveno o rudarenju podataka te o pronalaženju utjecajnih faktora. Pri rudarenju podataka korištena su dva software rješenja, SAP i ENOVIA PLM. SAP se u tvrtki koristi za financijsko praćenje projekta te sadrži podatke o potrebnim resursima za pojedine faze razvoja. PLM platforma se koristi za pohranu i dijeljenje svih dokumenata relevantnih za razvoj proizvoda. U njemu se nalaze svi nacrti, ugovori, prezentacije te rješenja bitna za izvođenje projekta. Za analizu odabrano je 13 projekata čije su faze razvoja proizvoda završene i koji se nalaze u fazama proizvodnje ili gotovog projekta. Odabrani projekti su iz iste skupine kupaca te ih karakteriziraju slični zahtjevi i procesi. Takav uzorak je odabran radi što bolje mogućnosti usporedbe podataka te analize utjecajnih faktora. Analiza utjecajnih faktora odrađena identificiranjem faktora koji utječu na svaki pojedini odjel te usporedbom radi prepoznavanja onih koji imaju utjecaj na sve odjele i najvažniji su u kontekstu raspodijele resursa. Pronalaženje utjecajnih faktora izvedeno je pomoću razgovora s odgovornim voditeljima timova te s analizom podataka. Nakon detaljne analize utvrđena je preliminarna lista utjecajnih faktora:

- Tip spremnika
- Tehnologija proizvodnje
- Kompleksnost geometrije
- Broj varijanti
- Broj dodatnih komponenti
- Kupac
- Količina prijenosa projekta
- Lokacija tvornice
- Broj alata
- Trajanje projekta

Nakon identifikacije utjecajnih faktora izvedena je analiza financijskog dijela projektnog uzorka gdje je napravljena usporedba realnih troškova projekta i predviđenih troškova projekta. Nakon prikupljanja svih podataka, tri projekta su odbačena zbog nepotpunih podataka i dokumenata koji nedostaju. Analizom je utvrđeno da se razlike između predviđenih i realnih troškova nalaze u rasponu od 4 do 31%, što potvrđuje činjenicu da nedostaje alat koji bi pružio podršku za točnije predviđanje troškova razvoja. Nakon analize troškova izrađena je matrica utjecajnih faktora povezana sa svim značajkama promatranih projekata. Matrica je tada iskorištena izračun korelacije između faktora. Nakon kreiranja matrica korelacija i validiranja značajnih statističkih vrijednosti, donesen je zaključak da na cjelokupni proces razvoja proizvoda najvećim dijelom djeluju 4 varijable: broj alata, broj varijanti, trajanje projekta i vrsta tehnologije. Utvrđene 4 varijable su nadalje iskorištene u primjenom metode višestruke linearne regresije. Upravo ta metoda je izabrana jer je analizom pokazano da varijable linearno ovise o broju sati utrošenom u razvoj proizvoda. Regresijom su dobivene jednadžbe koje će biti inkorporirane u alat i korištene za davanje konkretnih vrijednosti prilikom predviđanja.

Alat je izrađen korištenjem Microsoft Excel alata i Visual Basic for Application modula (VBA). Alat je zamišljen da se prilikom pokretanja pojavi korisničko sučelje programa u koji korisnik odabire parametre projekta koji su definirani matricom utjecajnih faktora. Program sadrži 5 kartica od kojih svaka čini zasebnu kategoriju utjecajnih faktora. Prilikom odabira svih parametara oni se zapisuju u tablicu te se pomoću formula dobivenih regresijom korisniku prikazuje procjena broja sati po svakom odjelu uključenom u razvoj proizvoda. Alat je zamišljen da zbog malog broja uzorka projekata dozvoljava blagu intervenciju korisnika te omogućuje korekciju dobivenog broja sati. Vrijednosti dobivene jednađbom čine 70% cjelokupnog fiksnog iznosa sati, dok se na drugih 30% iznosa može utjecati odabirom raznih parametara koji se ponašaju kao korektivni faktori. Alat stoga kreira baznu procjenu temeljenu na statističkoj analizi prethodnih projekata, no i dozvoljava korekciju ukoliko je to potrebno prema mišljenju korisnika. Osim što prikazuje troškove alat također omogućava korisniku slanje dokumenta kao privitka u e-mailu, spremanje tablice u PDF format te printanje. Također korisnik može u bilo kojem trenu promijeniti parametre i dobiti korigiranu procjenu. Na svakoj kartici u programu nalazi se i opcija za pomoć koja otvara prozor koji sadrži informacije o parametrima na toj kartici i kako se oni koriste pri predviđanju.

Budući da su neizrazite metode također dio parametarskih tehnika za predviđanje resursa, u radu je analizirana i mogućnost njihove eksploatacije. Najpoznatija tehnika za procjenu resursa ,a koja se ubraja u neizrazite metode, je zasigurno umjetna neuronska mreža. Umjetna neuronska mreža jest skup međusobno povezanih jednostavnih procesnih elemenata, jedinica ili čvorova, čija se funkcionalnost temelji na biološkom neuronu. Pri tome je obradbena moć mreže pohranjena u snazi veza između pojedinih neurona tj. težinama do kojih se dolazi postupkom prilagodbe odnosno učenjem iz skupa podataka za učenje. Neuronska mreža obrađuje podatke distribuiranim paralelnim radom svojih čvorova. Neuronske mreže su vrlo dobre u procjeni nelinearnih odnosa između utjecajnih faktora te pružaju mogućnost nalaženja obrazaca u velikom polju podataka. Za potrebe procjene troškova izrađena je neuronska mreža koja koristi 4 najvažnija utjecajna faktora dobivena prijašnjom statističkom analizom. Neuronska mreža izrađena je pomoću Matlab programskog paketa. Tri su najvažnija koraka pri kreiranju neuronske mreže: trening, validacija i testiranje. Također tri su glavne karakteristike koje čine arhitekturu neuronske mreže: broj skrivenih slojeva, broj neuron ate algoritam za učenje. Napravljena neuronska mreža koristila je tri sloja, jedan za unos podataka, jedan skriveni te jedan za izlaz podataka. Metodom pokušaja i pogrešaka odabran je broj od 10 neurona te pripadne aktivacijske funkcije za slojeve. Prilikom treninga neuronske mreže pažnja je bila usmjerena na izbjegavanje *over-fittinga* što bi onemogućavalo dobivanje dobrih izlaznih varijabli ukoliko se unese zahtjevan upit. Nakon treninga provedena je simulacija u kojoj su korišteni podaci o realnim troškovima tri projekta, te je ishod uspoređen s izrađenim alatom pomoću kojeg su također provedena predviđanja za iste projekte. Pri usporedbi predviđenih radnih sati, alat koji počiva na metodi regresije pokazao je neznatno veću točnost u odnosu na neuronske mreže. Postavlja se pitanje opravdanosti korištenja neuronskih mreža u svakidašnjem poslovanju tvrtke i u predviđanju resursa potrebnih za razvoj proizvoda. Ukoliko su faktori linearno zavisni, tada se mogu vrlo dobro opisati jednostavnim statističkim metodama koje su vrlo upotrebljive u svakodnevnom okruženju. Tada se ne javlja potreba za kreiranjem alata koji počiva na kompleksnim neizrazitim metodama. No ukoliko se omogući prikupljanje velikog broja projektnih podataka, tada su neuronske mreže idealne za pronalaženje obrazaca i veza između troška i utjecajnih faktora, što bi omogućilo još bolje razumijevanje u procese razvoja proizvoda i kako određeni faktori djeluju na ishod projekta.

1 INTRODUCTION

The need for valid resources estimation in automotive industry environment arises from its very competitive nature and competing for the market share. Automotive industry is characterized by cutting edge technology, advanced manufacturing methods, strong R&D capabilities and powerful logistics and planning on one side. On the other, it is characterized by tight schedules, constant changes and challenges in relationships between suppliers and OEMs. In Germany alone, automotive industry made revenue of 384 billion EUR in 2014. what is 20 percent of total German industry revenue. Furthermore, automotive industry is the corner stone of innovation and development in Europe. German OEMs have spent 17.6 billion EUR on R&D investments in 2014. alone and they are responsible for more than 60 percent of R&D growth in Europe [1]. According to the German Association of Automotive industry (VDA), the role of suppliers in the industry is increasing every year. That means suppliers needs to deal with increase in volume of production, create reliable supply chains and maintain or improve the product quality. Suppliers are thus competing for market share and they are trying to maintain existing business relationship with OEMs while establishing new ones.

For supplier companies it is important to have a good estimate of project costs so that awareness of resources expenditure can be created and used in strategic planning for the future business endeavors. Estimation is also important in the sense of offering a competitive quote without compromising quality of the product and being aware of the quantity of positive revenue for each project. The goal of the cost estimation and building the future strategy for the company would be strengthening the position on the market or ideally creating the breakthrough among the competitors. With accurate cost estimation it is possible to create reliable strategy for future company development and investments, based on the revenues from particular projects. Accurate cost estimation can contribute to creation of more competitive offers resulting in more project being awarded. Automotive companies often have to estimate the cost of a product development that contains significant amounts of new technology, and that requires considerable experience of previous projects, technology trends and new developments in other industry sectors. One of the main tasks for the automotive suppliers to remain competitive in the market, now that the market boom is slowly decreasing and high level uncertainty is present, is to drive smart efficiency improvements in all fields of operation, including the product development [2]. Motivation for the estimation tool development emerged after the comparison of the estimation data between the similar programs and realizing that there are unusually big differences in estimation caused primarily by lack of standardized methods. Present estimation relies on the responsible team leaders whose accuracy depends on their knowledge and experience with previous programs. The main goal with developing the tool was to eliminate, or to at least to minimize, influence of the individual on the program cost estimation. As numerous research has shown, the ability to lead and to successfully execute projects lies on manager's personality, knowledge and set of competencies that not everybody possesses in the same way [3][4]. While some managers are more experienced, others are new to the specific tasks that fuel tank development requires. Having all profiles of estimators in the company there is reasonable doubt that the cost estimation can improve on itself without any standardization or estimation tool introduction. Estimations given by the tool are useful for providing the insight in what shaped the costs in previous programs and it acts like lessons learned platform.

2 Early program phases and resources estimation analysis

2.1 The product and the development phases

Case study in the observed company was oriented towards their core business product and main source of income, plastic fuel tank (PFT) for the automotive industry. Company also produces selective catalytic reduction tanks (SCR), clear vision systems, camshafts and packaging. Being one of the top 20 automotive suppliers in Europe and top 100 in the world, company has to maintain its high standard of development and constant improvement in the field of new technologies and processes. The product itself requires continuous advancement because of the constantly rising requirements that come from regulatory bodies in the form of new standards and euro norms. Car manufacturers are being put through all time high hydrocarbon emission and sustainability requirements, which reflects heavily on the suppliers.

In the fuel tank production that means constant improvements on the materials, testing the different types of polymer layers and creating the cost effective solutions. Company tries to follow strict emission requirements by fulfilling the market needs on time. With the norms Euro V and Euro VI, it has been stated that nitrogen oxide in the exhaust gases needs to be minimized. That required from car manufacturers active introduction of aqueous urea solution that is sprayed into the exhaust system. Company followed that trend by introducing the SCR tanks into the program and has already proven as worthy competitor in that field. To follow market demand and increased amount of partial zero emission (PZEV) and hybrid vehicles, company is constantly improving the manufacturing process thus introducing the Next Generation Fuel Systems process. Hybrid and vehicles with start-stop systems have additional demands regarding the fuel tank geometry and that is installing the baffles that prevent the fuel slosh noise when vehicle switches to electric drive or turns off the engine. All these requirements are being dealt with in two main manufacturing methods:

- Multilayer co-extrusion conventional blow molding (COEX)
- Next Generation Fuel Systems (NGFS)

High hydrocarbon emission standards reflect on the manufacturing process of the fuel tank, requiring minimal intervention in the tank geometry after the molding procedure and eliminating undesirable cutting and welding of the tank shell. By using NGFS method, tank components are incorporated into the tank shell during the blow molding procedure. It means that the production down-line is less complex comparing to the COEX method, simply because it doesn't require cutting the shell, placing the component and then welding them again, at least not for most of the components. However, NGFS method is inherently more complex compared to COEX, and is used only if customer has specific needs and high emission requirements. NGFS method, as it will be seen in later chapters, influences the product development resulting in higher cost and need for resources. Manufacturing planning is important part of the product development and it needs special attention from logistics point of view and it has to be aligned with all other program phases, minor and major.

Fuel tank development process consists of 7 main phases which execution defines program outcome:

- Business case and voice of customer
- Conceptual design and feasibility
- Preliminary design
- Detailed design
- Product and process validation
- Pilot
- 90 day from SOP

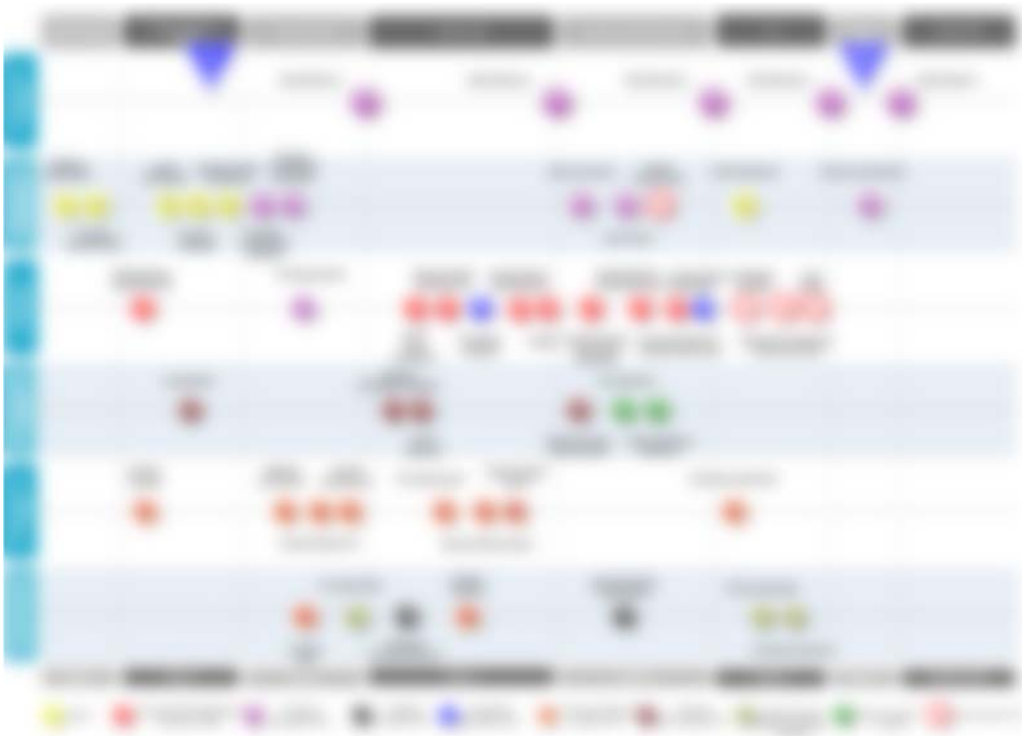


Figure 1 Product develpment map

For observed case main focus was on how to improve the estimation process and how to rightfully allocate project resources. Process of resources allocation itself is carried out in the second phase - Conceptual design and feasibility. Product concept, within second phase, is well defined process which besides engineering cost estimation (ECE) includes also: design concept, initial CAE simulations, BOM, design and process validation plan, quality check and purchase parts RFQ. Second phase should result in a sound quote to the customer and well planned product development phases, including all responsible departments. Overall responsibility for carrying out the quote relies on the sales department. Even though sales management doesn’t participate in ECE process itself, they do however affect the final outcome. Responsible who are involved in creation of the ECE are mostly department team leaders. Overall responsibility for ECE is on program manager, who through his expertise and guidance needs to facilitate estimation process in a way that will enable sales to create quotation on time. ECE challenge lies in the fact that it is done in early stages of the project, with little or no information available regarding the possible changes and outcomes during the project. Usually, the project length varies from 28 - 38 months, considering the regular fuel tank development. Inputs for the ECE are customer requirements consisting of product specifications, yearly production volume, simulations and design requirements. After getting the requirements the quotation team is assembled and quote preparation begins following the processes that are internally defined, specifically to fit the customer.

2.2 ED&D charts and estimation

There are numerous specialized software services that can be used for gathering cost information and analysis. Companies today are mostly using custom made Excel sheets for several reasons. They can be personalized and made in a way that suits the company processes and development phases, it is easy to communicate sheets with the customer and Excel is fairly known so additional training is not required. One of used custom Excel sheets is known as Engineering Design and Development chart (ED&D). ED&D chart represents rather an overview of cost estimates than the tool for estimation itself. It contains general project information such as: yearly cost rates, plants and locations, project description and mountain chart. Mountain chart is essentially a monthly breakdown of individual project costs throughout the project duration. It serves a purpose of more detailed working hours allocation and workload distribution overview.

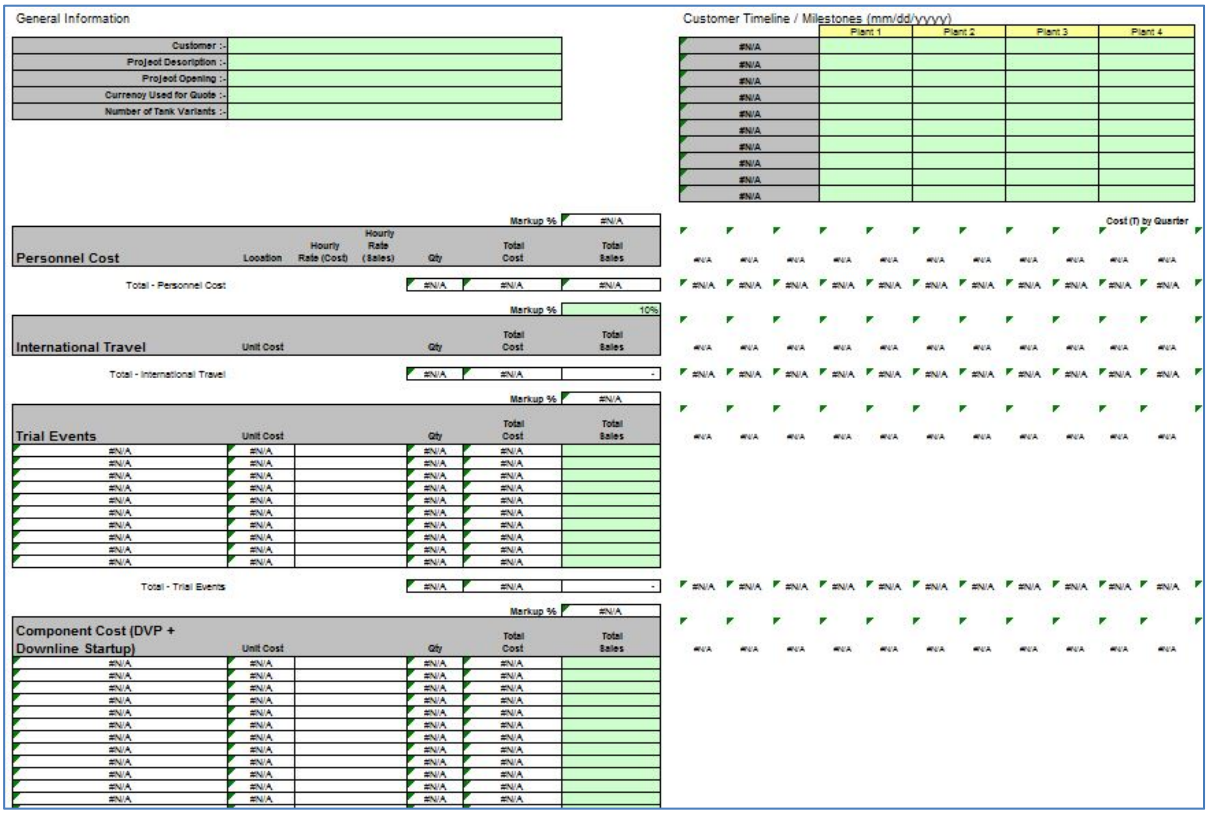


Figure 2 ED&D Chart - Summary sheet

In the cost estimation process main responsible role belongs to the program manager, whose task is to collect and articulate the estimations from team leaders and to create the input for ED&D chart. ED&D chart is used as a base for quote creation by sales department. Costs in ED&D chart are divided in several categories and overall it provides insight in cost breakdown. In a fuel tank development process main cost categories are:

- Personnel cost
- International travel
- Trial events
- Component cost (DVP + down-line startup)
- Product validation
- Startup costs

Main focus of the ECE is to estimate project cost as accurate as possible so the quote given to the customer would match the expected amount received for the development phase. Every cost category has a matching cost (cost for the company) and sales price depending on wanted markup. Cost prices are determined yearly and they are provided within the ED&D chart. Personnel cost represents a man-hour workload for project team members from every department involved in the development of the final product. Personnel cost shows how many hours is needed from each department for successful project development. Personnel costs are divided into categories:

- Program management
- Product engineering
- Advanced manufacturing
- CAD
- CAE
- Product quality

For each category man-hour load is estimated by its team leader/manager or responsible person. The load distribution between the departments will be shown in the next chapters. Example of the way that personnel cost is displayed in ED&D chart is given in the table below.

Table 1 Display of ED&D personnel cost notation

Role	Hourly rate - Cost	Hourly rate - Sales	Quantity	Total cost	Total sales	Markup
Program management	X_1	Y_1	Z_1	$X_1 * Z_1$	$Y_1 * Z_1$	$1 - \frac{\sum Y_n Z_n}{\sum X_n Z_n}$
Product engineering	X_2	Y_2	Z_2	$X_2 * Z_2$	$Y_2 * Z_2$	

Personnel cost determination is one of the biggest challenges in the cost estimation process because it relies on several expert opinions and tools built with different estimation methodology. Travel expenses associated with the personnel are also challenging to estimate. Having 32 factories in 16 countries sometimes requires of team members to be engaged in international travel. Costs for travelling can rise up significantly due to the unpredictable development of the project or specific customer requirements. That is the reason it needs to have its own budget line in the ED&D chart. Nevertheless, travel costs are difficult to estimate accurately due to the project contingencies. Usually, estimate is based on the planned number of plants for the product and based on the previous experiences with the certain customers.

2.3 ECE procedure

At this moment, populating ED&D charts relies on three interconnected features: program manager experience and intuition, department's estimations and sales approval. There is no defined order in how these features affect the cost estimation because the process itself, ideally, consists of mutual cooperation throughout the whole phase.

Figure 3 Features that affect cost estimation

2.3.1 Program manager role – Expert opinion

A program can be defined as a group of related projects, subprograms, and program activities managed in a coordinated way to obtain benefits not available from managing them individually. Programs may include elements of related work outside the scope of the discrete projects in the program. A project may or may not be part of a program but a program will always have projects [5]. Program management is the centralized coordinated management of a program to achieve the program's strategic benefits and objectives. In addition, it allows for the application of several broad management themes to help ensure the successful accomplishment of the program. Managing product development through program allows coordination of multiple projects by means of optimized or integrated cost, schedules, or effort; integrated or dependent deliverables across the program, delivery of incremental benefits, and optimization of staffing in the context of the overall program's needs. Projects may be interdependent because of the collective capability that is delivered, or they may share a common attribute such as client, customer, seller, technology, or resource [6].

PM is responsible for gathering the data from different departments and articulating it into reasonable forms that would be approved by sales, but more importantly, by customer. Apart from being responsible for initial cost estimation, PM must decide and "balance" program issues between the performance, schedule, and the cost of the program as variables change [7]. Along with good understanding of the product and its development phases, it is an absolute necessity to have good communication with other departments in order to get more accurate and in time data. After data and working hours are gathered, PM needs to evaluate and, if necessary, adjust the estimation according to the project dynamics, customer requirements and the project development up to that phase. Another important task for PM is to make sure that all estimations come in time and that there are no delays that can affect further program development. Based on ED&D charts available for analysis, it was possible to define a few different program types.

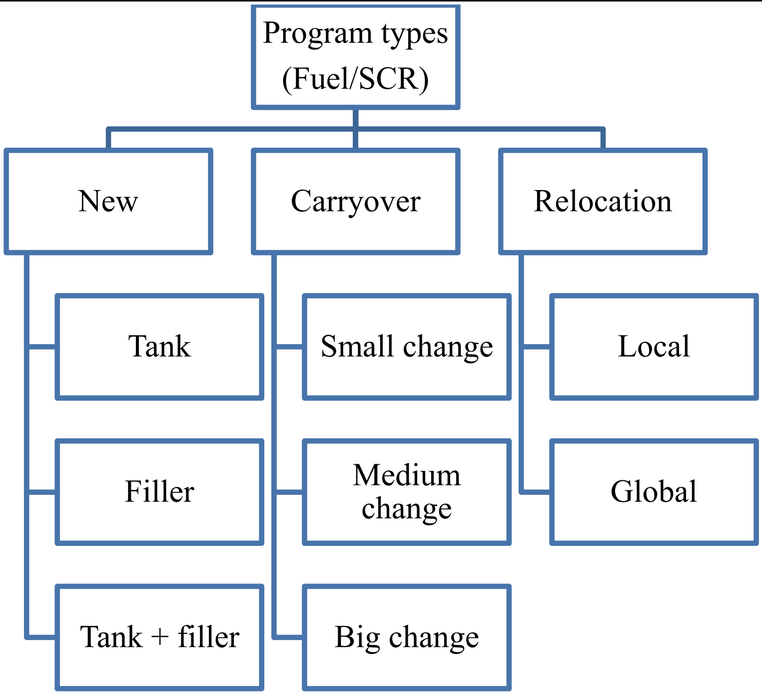


Figure 4 Program types classification

Based on the program types (from available ED&D charts) PM has to be able to evaluate the gathered data and if needed make the adjustments in coordination with the person responsible. Program types are the same for development of fuel or SCR tanks. “New” represents category of programs that are newly developed with small to none amount of carryover from the previous programs. Categories that affect program dynamics and time span are dependable on what the product is. Products can be divided in two categories: tanks and filler pipes (PFT/SCR). Sometimes, the program requires development of filler pipe alone, without the fuel tank, while some programs require both filler pipe and tank development. As it would be shown in later chapters, nearly third of all the programs have both, tank and filler development. Another program type is “Carryover”, what essentially means the amount of carryover is usually large and there are change requests for already existing tanks/filler pipes. Depending on the change magnitude or amount of carryover (small, medium, big), program time span is affected proportionally. Third program type refers to “Relocation of production”. Having over 30 factories around the world and with constant expansion, it can happen that it is more feasible to change the location of the production line for the purpose of optimization of the process, minimizing the transport costs or filling the production gaps. That however, affects program from the logistics standpoint and it is important to take into account multiple factors that influence decision making process, such as: overproduction for compensation, amount of time out of work, tools adjustment, (down)line optimization etc. PM role in estimation process is to have good understanding of all possible outcomes for certain types of programs and enough experience to deal with possible issues accordingly.

2.3.1.1 PM and communication

Within the early program development itself, PM has to create balance between engineering and production requirements, since they often differ regarding the product complexity and exploitation options. In order to achieve program success, cross-functional cooperation has to be established through manager encouraged intra communication. Since the estimation process is early in the development phase, having open communication channels needs to be established as one of the main priorities. The goal is to have an interconnected program team and there are various ways to achieve that in a short time. There are significant differences in the use of different methods of communication between low and high cooperation groups, with strong cross-functional cooperation related to enhanced informal communication [8]. High accessibility of team members encourages discussion about program issues and eases finding the solution with mutual agreement. The term “accessibility” in the program context represents an individual’s perception of liberty or ability to approach or communicate with another project member. Creating an atmosphere of an accessible team members lies solely on PM and requires additional effort added to the existing dedication to the technical part of the program [9]. ECE process is susceptible to many changes in its buildup. It is of crucial importance that changes are communicated to responsible members because of the possible customer inquiries about the amount of working hours defined. Communication channel can be broken down according to the three interconnected features that define ECE process as shown in the upper figure:

- Sales
- Program manager
- Responsible departments

Each of this roles, having their own tasks in the ECE, can act independent in their estimations. However, end figures needs to be confirmed by all three sides with mutual agreement and understanding of carried out decisions. Due to the high competitiveness and challenging timelines in automotive industry, there is usually no time to create program communication plan for every program that comes along. The real long term solution would be establishing communication plan as a standard within the company, with clearly stated preferred media and methods for every phase and situation within program development. Or as an alternative, if a company doesn't have the communication plan as a standard, PM should consider understanding basics of the communication process in a program, as it surely affects the cost estimation among other activities. There are five important factors to consider when establishing clear communication channels [10]:

- Who is involved in the communications process – stakeholders such as: project team members, customer management and staff, external stakeholders
- What is being communicated – what kind of information needs communication
- When the information is communicated – daily, weekly, monthly, as needed or as identified
- How the information is distributed – in a meeting, a memorandum, e-mail, presentation, etc.
- Who will provide the information being communicated

By considering these factors PM has a clearer vision of what needs to be done in regard to having a strong foundation for effortless communication exchange between the team members. Due to the fact that cost estimation process also demands clear external communication (customer), it is even more important that the whole team is synced and able to answer the questions.

2.3.2 The department's estimations

Responsible team members are the ones who are requested to make an estimation of working hours for their departments/functions. Usually, the team leaders/managers are responsible. Existing method can be divided in three steps and is based on the request for estimation from PM/Customer/TPL to the other team members, making the estimation and sending it to the person who requested it. As shown in figure 4, it's a three step process and every step consists of its own requirements and problems.

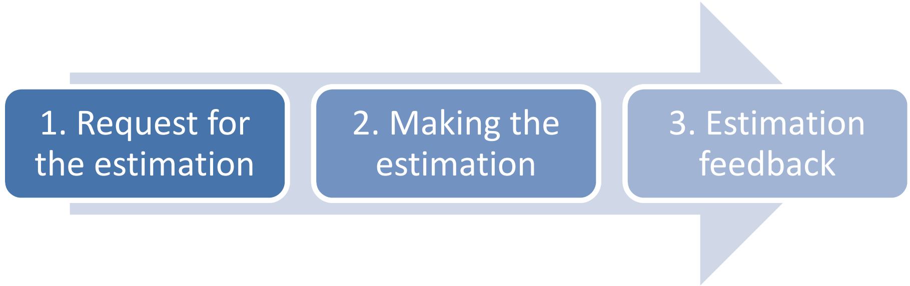


Figure 5 Existing estimation steps

Every step requires some rules to be followed in order for it to be effective and transparent. Currently, there is no standardized procedure for communication when requesting the estimation. Furthermore, there are no standardized forms/documents that would be used for that purpose. Main communication tool is email, but quotes are sometimes requested in informal manner over the phone or even at the kickoff meetings.

1. Request for the estimation - Person sending the request (PM/TPL/C) needs to collect the preliminary data in order to create sufficient input so that the request receiver is able to make accurate and fast estimation. It usually means collecting as much data as possible regarding the product specifications, design, quantity and timing. It can be very challenging to collect all the data needed, especially from the customer. Because of that, this step requires clear communication with customer, well defined program scope and capability to “ask the right questions”. Skills that are, arguably, dependable on the professional experience and the person itself. It is advised to achieve stable and trustful communication with the customer in order to get the needed information. Depending on when the request for estimation is sent, the receiver has limited amount of time for an estimation process.

2. Making the estimation - Person responsible for making the estimation receives the request from PM/TPL/C. Upon receiving of the request, the goal is to estimate the hours needed for ED&D chart. Requests are mainly informal and sometimes without sufficient data for accurate estimation. Some departments have tools (Excel sheets) to help them with the cost estimations while some are using numbers based on the experience in programs so far. After initial estimation is given by the tool, responsible expert can modify it according to the specific program needs and with the cost awareness. In order

to improve future cost estimation in specific departments it is important to have estimated cost evaluation for every phase of the project. For example, it is estimated that 1000 working hours will be needed for product engineers to reach the first milestone. After the first milestone was reached, it would be advisable to take a look back on the program, realize how many hours were spent in a real work and compare with the estimation. Additionally, if there is a significant difference, comments on what caused it and how did it affect the program development, should be made.

3. Estimation feedback - Main criteria for the accurate estimation of hours are valid inputs. Depending on the department there is a need for different inputs, but some of them are the same. There is a big issue in articulating the inputs that are being requested, therefore, standardization of input classification would be a possible solution. Inputs should be delivered via form specified for each department. Ideally, there would be an online platform for every project (wiki) where all the product specification, that are known up to date, would be available to all team members. Forms should be created in cooperation with departments so they can specify what is the information that they need. Caution is needed not to overcomplicate the forms because some of the information is simply not available at the early stages of the project. Some of the most required inputs are: ruff timing, prototype availability, quantity, BOM, product dimensions and specs, design, specification of components and used technology. Having an online platform would enable execution of unified cost estimation between the departments, where all the responsible experts would put the estimations for their own departments, visible to themselves and the PM. It would centralize communication and in that way enable PM to have an easy overview of all the department’s estimations in one place. It would eliminate the need for fishing emails that contain the relevant information, thus minimizing the time needed for making the final estimation.

Other than valid inputs, timing of the request is also important. Very often customer request comes on a short notice and timing for the cost estimation is not ideal but it needs to be carried out in the best way possible. If there is a time for estimation then it’s the PM responsibility to send the input form early enough to the departments. When sending it back, estimation data should be completed, if there are some unknowns it is necessary to contact the PM/TPL/Customer for clarification.

2.3.3 Sales department approval

Usually estimated costs are recognized to be too high by sales department. It is up to sales to make a quotation competitive and appealing to customer so occasional intervention is needed and understandable. When changing the estimation, sales should make a contact/report to affected departments about taken actions so that they can be up to date with current ED&D. The reason is that the customer can ask for clarification on the budget item and why do the costs in a current program differ from the costs in previous ones. Head of department should be able to answer such inquiries with confidence and good arguments. The goal is to charge as much engineering costs as possible in the development phase of the Program and to put less to the overhead costs. Information flow between sales and other departments is usually not well established, mostly because e-mail communication is time consuming on already tight schedule. Program changes should be reported across all functions in the program team, if not for the business matter than for the social one. It elevates the team spirit, makes the working environment more motivating and creates the sense of working together towards the successful completion of the project [11]. Surely, there is no need in reporting every minor change, but the ones that are relevant to the department, especially in the quotation phase, helps to build foundation upon which team will work on in the further program development. Again, simple solution for this would be to establish the web based news board that is easy to edit and to update. If it's used properly even changes through program could be evaluated after completion of each phase. After sales department and other responsible had come to an agreement, quote is sent out to the customer and father program development is expected.

2.4 ED&D chart initial analysis

To gain better understanding of the estimation process and the role of every department, research on ED&D charts was carried out. Research included 7 different charts for different programs, focusing on identifying similarities and differences in the estimation values based on program types. Available charts were from specific business unit dealing with certain customers, so it cannot be accounted for the other business units, even though findings could be easily projected since in all business units programs are carried out following the same procedures and processes. Available ED&D charts were from the programs that are in their first stages of development, hence the latest estimation tools and knowledge were used. Analysis was divided in three major steps:

- Department share in ED&D
- Milestone resources distribution
- WBS analysis

Table 2 Initial analysis program types

Program	Type	Product
P1	New	SCR
P2	Carryover	PFT
P3	New	SCR
P4	New	PFT
P5	Carryover	PFT
P6	Relocation	PFT
P7	New	PFT

2.4.1 Department share in ED&D

Personnel costs are divided in six different categories (departments). Every program has a different distribution of department share in overall cost amount, depending on various product and manufacturing parameters. It is important that similarities in programs are identified, or differences for that matter, connections are created and shares compared regarding the latter. As expected, department shares between the programs were very different because of the different program types. However, there are also significant differences between similar programs. To some extent differences can be explained through product variations, manufacturing locations, customer complexity, etc. Absolute values of working hours are irrelevant in conducted analysis because of the various program lengths and product variants. Focus is, instead, on ratios that should be similar. Usually, in a regular program (new), PE takes most of the percentage/hours required in the chart. Compared to the WBS job distribution (see chapter 2.4.3), that is expected. In new programs, distribution of hours is somewhat logical and can be explained in tank design variations, program complexity, scope and time span. Figure below shows percentage of share within the program cost for every department, over 7 analyzed charts.

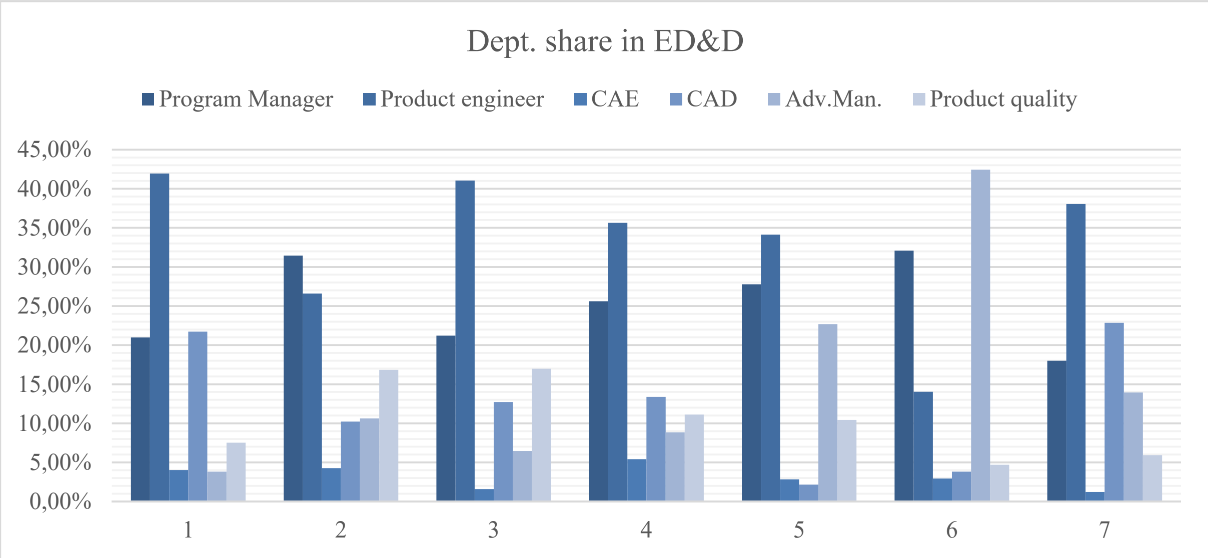


Figure 6 Department's share in total cost of a program

In the program 6 one can see that Adv. manufacturing is very dominant along with program management. The reason is that changing the manufacturing location requires a lot of logistics and management involvement, and if the product has already passed the DVP (design verification process) it doesn't require much more CAD/CAE hours. PM working hours are influenced by nearly the same factors as PE working hours so at most they should be around equal, if regular program is monitored. It is evident that differences between program estimations are substantial. For some it is expected because of different types of programs, but programs P4 and P7 should have similar distribution between roles. Every program has different requirements and absolute values but for programs that belong to the same customer and product complexity category it is expected that ratios between roles are similar. Comparing the programs P1 and P3, that are both new and for SCR tank, there is an evident difference in CAD and Adv. manufacturing working hours. Also CAE is fairly low in the program P3. Along with many similarities between programs, the biggest distinction is that they are managed by different PM's and conclusion has been made that they have the main influence on end estimation values. Within the same department PM's can have informal agreement on values distribution based on the experience with the customers so far. However, that by no means accounts for a standard in the company because it lacks definition and standardization. In the later chapters description it is possible to see that the company has very well defined work breakdown structure (WBS) and that leads to possibility to extract ratios between the departments and to determine their involvement in the product development. The idea of having a WBS defined in detail is for program teams to follow it and to deliver high standard products throughout business units. Having that in mind, a discussion can be made on how it is possible that there are substantial differences across similar program estimations. Certain amount of flexibility is desired when it comes to following the WBS simply because not every customer has the same internal milestones and product development requirements. For instance, some customers may require extensive CAE simulations being carried out, which would require notable amount of hours more compared to the programs for less demanding customers and would result in different ratios between the departments. Similar examples could be given for other development requests and dealing with that requires flexibility in program planning. That is the reason why there are several business units each assigned to a group of specific customers based on their strategic importance, similar requirements and location. One could argue that apart from department ratios in ED&D chart, milestone distribution is also important in order to see how resources are allocated throughout the whole project.

2.4.2 Milestone resources distribution

Next step in analysis was to extract data from mountain chart for every milestone in a program and to compare working hours that were spent. Mountain chart represents cumulative estimated cost monthly breakdown. That way it is possible to see how many months are needed to achieve certain milestone and how many hours for every department needs to be utilized. Also, it is possible to see amount of working hours spent every month of the project up to start of production (SOP). Overall program development is, as mentioned at the beginning of the chapter, internally divided in 7 phases (gates), first being business case and the last being the SOP. All of the internal company phases needs to be aligned with the customer milestones. In order to achieve that, mountain chart is made regarding to the customer milestones and working hours are allocated respectfully. From the WBS it is not possible to extract percentages or amount of working hours that needs to be allocated in specific phases, consequently, allocation relies on the company program experience and the development requirements. What can be seen from WBS, is in which phases are certain departments needed and in which are not.

ED&D Mountain Chart																	
Customer :-																	
Project Description :-																	
Project Opening :-																	
SOP :-	#N/A																
Year	#VALUE!		Plant 1 Plant 2														
	Location	Annual Allocation (%)	Annual Allocation (hours)	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total	Total Cost
PERSONNEL COST																	
Product Engineer				0	0	0	0	0	0	0	0	0	0	0	0	0	-
Program Manager				0	0	0	0	0	0	0	0	0	0	0	0	0	-
				0	0	0	0	0	0	0	0	0	0	0	0	0	-
CAD				0	0	0	0	0	0	0	0	0	0	0	0	0	-
Prototype Production				0	0	0	0	0	0	0	0	0	0	0	0	0	-
Product Engineer				0	0	0	0	0	0	0	0	0	0	0	0	0	-
Program Manager				0	0	0	0	0	0	0	0	0	0	0	0	0	-
Product Quality				0	0	0	0	0	0	0	0	0	0	0	0	0	-
Product Validation				0	0	0	0	0	0	0	0	0	0	0	0	0	-
				0	0	0	0	0	0	0	0	0	0	0	0	0	-
				0	0	0	0	0	0	0	0	0	0	0	0	0	-
				0	0	0	0	0	0	0	0	0	0	0	0	0	-
				0	0	0	0	0	0	0	0	0	0	0	0	0	-
				0	0	0	0	0	0	0	0	0	0	0	0	0	-
				0	0	0	0	0	0	0	0	0	0	0	0	0	-
				0	0	0	0	0	0	0	0	0	0	0	0	0	-
				0	0	0	0	0	0	0	0	0	0	0	0	0	-
				0	0	0	0	0	0	0	0	0	0	0	0	0	-
				0	0	0	0	0	0	0	0	0	0	0	0	0	-
				0	0	0	0	0	0	0	0	0	0	0	0	0	-
Monthly Total - PERSONNEL COST				0	0	0	0	0	0	0	0	0	0	0	0	0	-
INTERNATIONAL TRAVEL																	
Monthly Total - INTERNATIONAL TRAVEL				0	0	0	0	0	0	0	0	0	0	0	0	0	-
TRIAL EVENTS																	
Monthly Total - TRIAL EVENTS				0	0	0	0	0	0	0	0	0	0	0	0	0	-
COMPONENT COST (DVP + DOWNLINE STARTUP)																	
Monthly Total - COMPONENT COST (DVP + DOWNLINE STARTUP)				0	0	0	0	0	0	0	0	0	0	0	0	0	-
PRODUCT VALIDATION	From	To	Apply Cost	#VALUE!	#VALUE!	#V E!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!		
Monthly Total - PRODUCT VALIDATION				0	0	0	0	0	0	0	0	0	0	0	0	0	-

Figure 7 Mountain chart in ED&D

There is no use in analyzing mountain chart for more detailed periods than milestone completion, because it's generated automatically from annually allocated percentages. Milestone distribution was done for the programs P1, P3 and P4 because the other programs were not comparable regarding the program groups or mountain chart lacked the data for comparison. Regardless of product being for PFT or SCR, distribution of resources should be very similar given the fact they follow the same product development layout. As shown in fig. 7, working hours to achieve milestones are substantially different among programs. It shows how many percent of total working hours within a program is estimated to be spent on individual customer milestones. Big difference can be seen immediately from the milestone 1. Programs P3 and P4 differ in 16% of allocated hours, while the difference between P1 and P4 can be acceptable. Trend continues in the milestone 2 also, where total difference between P3 and P4 is more than 18%. P4 also has one additional customer milestone when compared to the other two (5).

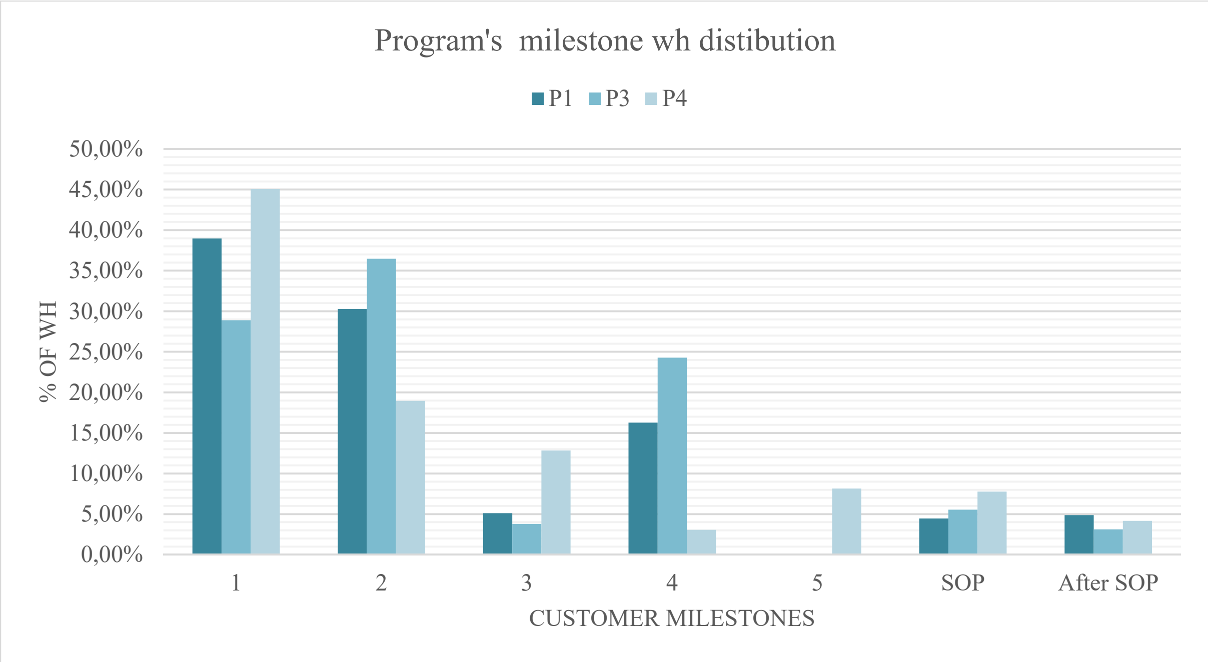


Figure 8 Distribution of wh for specific milestones

It is expected for programs early stages to have the most hours allocated. As defined in WBS, it is visible that nearly every aspect of the program is involved in the first period, specifically to the gate 4. One of the problems in ED&D charts, that shows lack of standardization, is that milestones order is not well specified for all programs. Some of them (e.g. P4) have the milestone 4 prior to milestone 3. It could affect the distribution in the charts, but nevertheless, total cost remains the same. Also, there is no agreement on whether "After SOP" phase should be calculated in the ED&D or not and from the charts overview 3 out of 7 didn't have the After SOP milestone. There is a room for improvement in achieving mutual agreement within PM department to use the same nomenclature and timetable layout for the given customer. It will not make a great effect to the cost estimation in general but it would be easier for future comparisons and evaluations.

In the figures below it is shown how working hours are distributed for every department regarding the customer milestones. The charts show the percentage of total department hours spent in the specific milestone. PM and PE are the most demanding categories and require most of the program funds, with PE being more dominant. It is understandable that the distribution of working hours is leaned heavily towards the first three phases. All of the major engineering development is being done in the first three phases, starting with initial simulations and conceptual design and ending with the design freeze and finalized product specifications.

It is evident that CAE is needed only in the first milestone, rarely the second. The reason is that the numerical validation of the tank designs are conducted prior to the detailed tank development, to make sure that the structural integrity and tank properties are aligned with the customer requirements. CAD is also needed mostly through the first three phases because those are the conceptual design development as well as the product design finalization phases. In real program development, CAD is involved usually throughout the whole program, but only if there are requests for changes. In ED&D charts changes are not considered, so CAD hours are seen only for the initial tank development. Advanced manufacturing support in logistics is there throughout the whole program, but mostly in the beginning and prior to SOP. Similar distribution can be seen for PQ also. In these detailed charts it is clearly visible how differences affect particular departments. For instance, PM and PE in P1 are similar for the first two milestones, in P3 they are around 10% higher in milestone 2 while in P3 they are more than 20% higher in milestone 1. Such differences in distribution of hours gives the customer a basis to argue that it was different in previous programs and they can rightfully ask question about the issue. If the distribution would be standardized in the form of the formal agreement for the departments, those issues could be easily avoided.

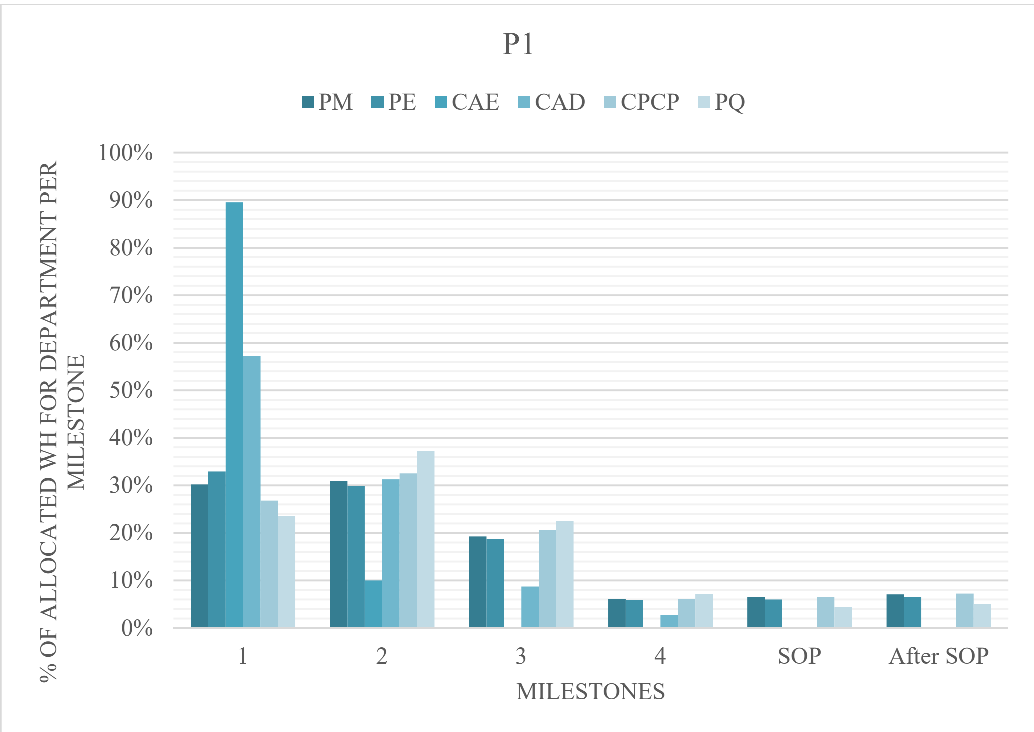


Figure 9 Distribution of dept. wh through milestones for P1

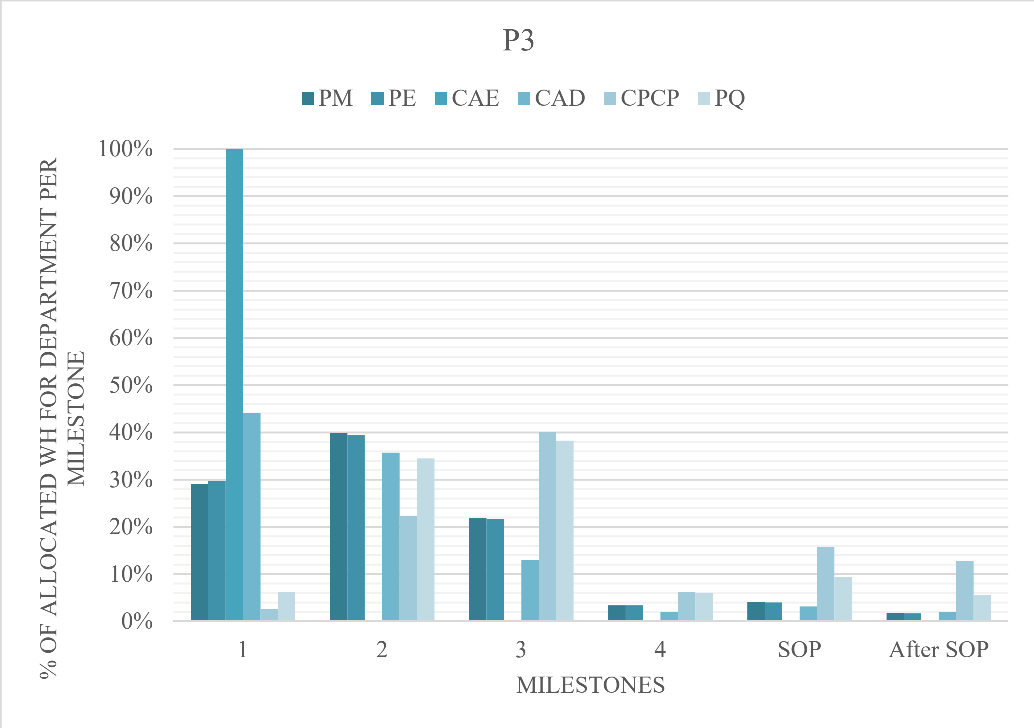


Figure 10 Distribution of dept. wh through milestones for P3

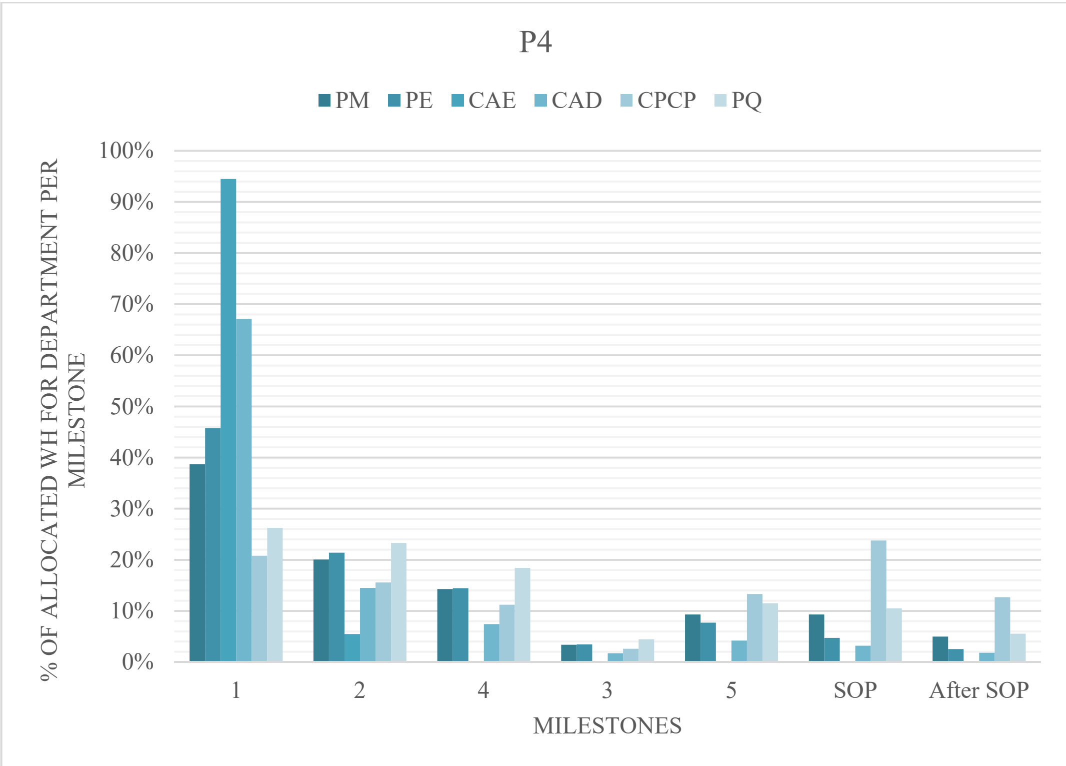


Figure 11 Distribution of dept. wh through milestones for P4

2.4.3 WBS

The program and the final product it delivers are inseparable. Details of what is needed for development, testing and delivery is determined by the product requirements, its architecture, and design. Studies have shown that the relationships and interactions between the architecture of products, their development projects, and the organizational teams involved, should be aligned in order for a company to become successful [12]. Since the development of a fuel tank includes the design process and other iterative phases, which differ from workflow processes, company had to decompose program and the product hierarchically, what is better known as a Work Breakdown Structure.

Among companies that provide complex product development services WBS is a widely used planning method. A WBS is a hierarchical tree structure decomposition of a project/program into smaller components down to the level of work elements—the leaves of the hierarchy. The role of the WBS is to reflect the total scope of work involved in the program. Capturing the total work and efforts required for the project is most important, as WBS is the source for project cost estimations, schedule planning, and risk mitigation [13]. WBS created for the PFT/SCR development consists of 341 tasks, grouped in 49 deliverables which are further assigned to 7 program phases.

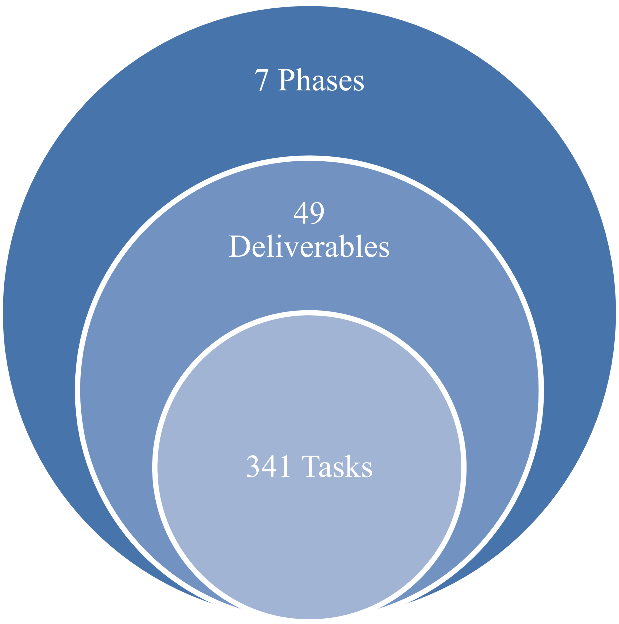


Figure 12 Display of the WBS sub-categories

Each of the tasks are assigned to a specific role in the program team which ultimately belongs to one of the departments. Deliverables are assigned to the specific departments that are responsible for carrying them out within the certain phase of the program. By using the company's WBS it was possible to extract the quantity of how many times each assignment is given to the specific role. The goal was to see how many times is each department involved as a responsible for the deliverable. This method, as the one that will be described later, has a significant drawback in inability to assess time span of every deliverable. It may seem that some departments are under great load based on the quantity of responsibilities, but some of the deliverables take more time than others and are more complex. Nevertheless, this shows a rough picture of the department responsibility distribution throughout the program.

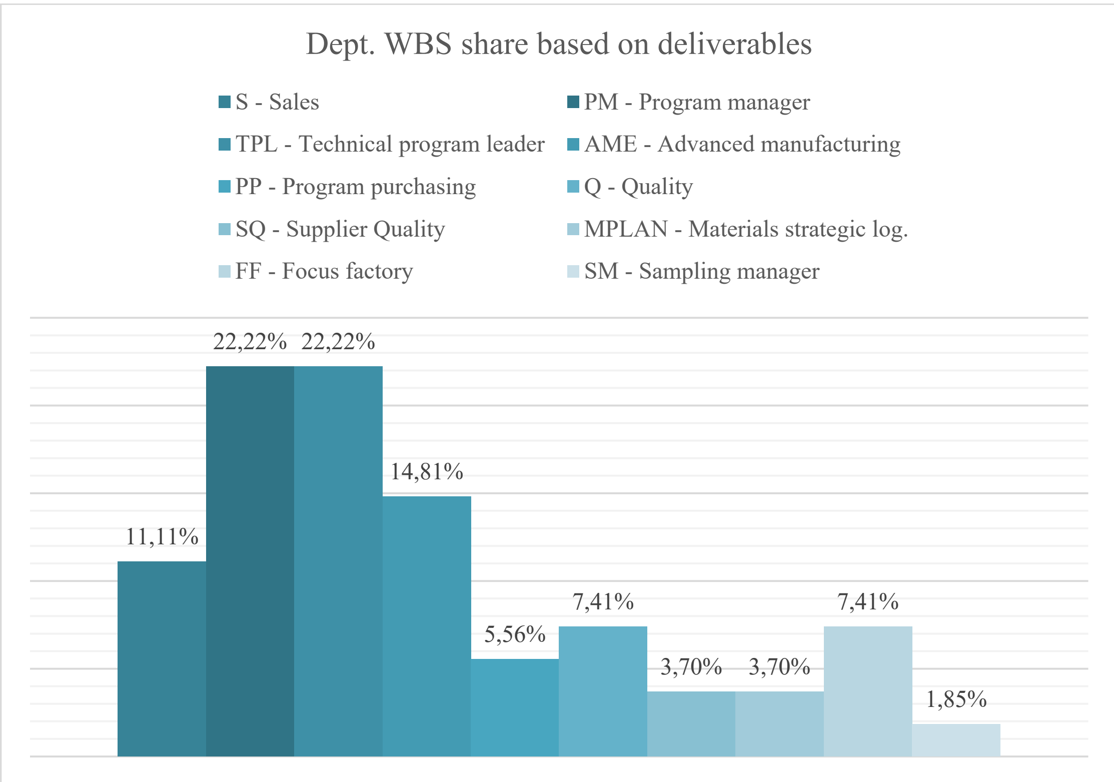


Figure 13 Responsibility distribution based on WBS deliverables

The end result looks nothing like programs estimation that were mentioned before. Distribution of responsibility is logical but that cannot in any way resemble the working hours needed to complete the deliverable. As those 49 deliverables in WBS are divided into 341 tasks that are also assigned to responsible roles, it was possible to extract that data, in the same manner as for deliverables, and to see how many times specific role was mentioned as responsible for the task. For that amount of tasks it is expected that results should get closer to the real looking program estimations.

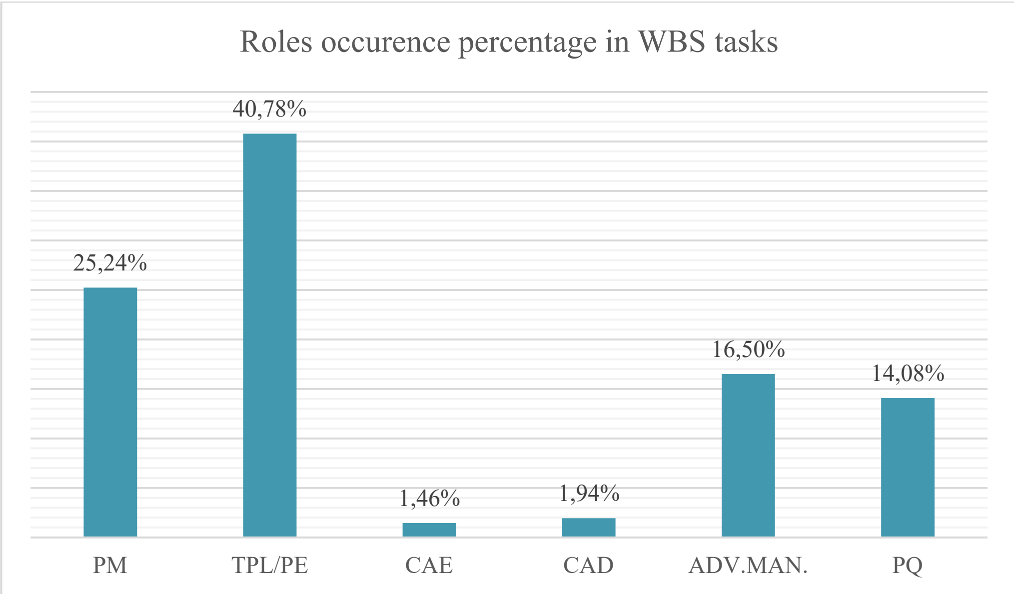


Figure 14 Responsibility distribution based on WBS tasks

From the figure above it can be seen that using the data from 341 tasks in WBS resulted in more realistic display of the workload in the program. As already mentioned, these percentages are based solely on the number of occurrences in the WBS, not taking into account task complexity and duration. While PM and PE percentages are realistic compared to the real program, CAD and CAE percentages are fairly low, because they have the few tasks. However, in real environment those few tasks takes a lot of time and resources to accomplish so the real number would be higher. By using the history data from past programs and checking what are the real shares of CAD and CAE (and other departments) in the program, this chart could be modified and used as a base for a workload distribution within the program development.

On the figures below it is possible to see the distribution of workload through the program phases based on the WBS. That map of responsibilities shows for instance, that the sales department has the overall responsibility in the quotation phase while PM and PE have more responsible tasks in the early product development period. This map of responsibilities is created by using the 49 deliverables assignments. If one department has an overall responsibility, it doesn't mean that others don't have responsibility at all. When looking at the detailed tasks within the deliverables, it is obvious that other departments are heavily involved in the certain phases of the development. Also, this map doesn't represent the overall program responsibility, which relies mostly on PM, yet it show responsibility for particular deliverables. One could argue that this has nothing to do with how real programs are carried out. That would be understandable especially if the WBS is of newer date. That is the reason why it has to be constantly updated, improved and reintroduced on account of the lessons learned through the program development experience. Workload distribution shown in the figures could be used a base for the standard in mountain chart hours allocation. Later on when there are sufficient data from past programs, it can be used for WBS adjustment and it could lead to better and more accurate mountain chart filling. These techniques for workload distribution allow for easier management of multiple programs carried out by the same team. Even if at the moment company doesn't have such a workload for engaging the same team in multiple programs, it will definitely happen as every company has a mission to grow on the market. That way it is possible to evaluate at which stage is the current program, what is the current workload among the departments and what are the possibilities in filling the workload holes that certain department is experiencing. Needless to say, the departments are rarely without workload, but it can always be better managed and used more effectively when having a clear overview of specific project phases and responsibility assignments.

The cost estimation procedure can be heavily influenced by the workload distribution among the departments. For example, when PM receives all the estimations and puts them together, short analysis can be done and conclusion can be made whether there are some illogical "spikes" of workload in regards to the WBS. If so, another iteration of the estimation process is needed. There are some authors that offer the idea of devising the whole cost estimation procedure based entirely on the WBS [14], but that model would be too rigid for companies that offer product whose requirements differ greatly between various programs. There are several other methods that are better suited for cost estimation, especially when automotive environment is taken into account.

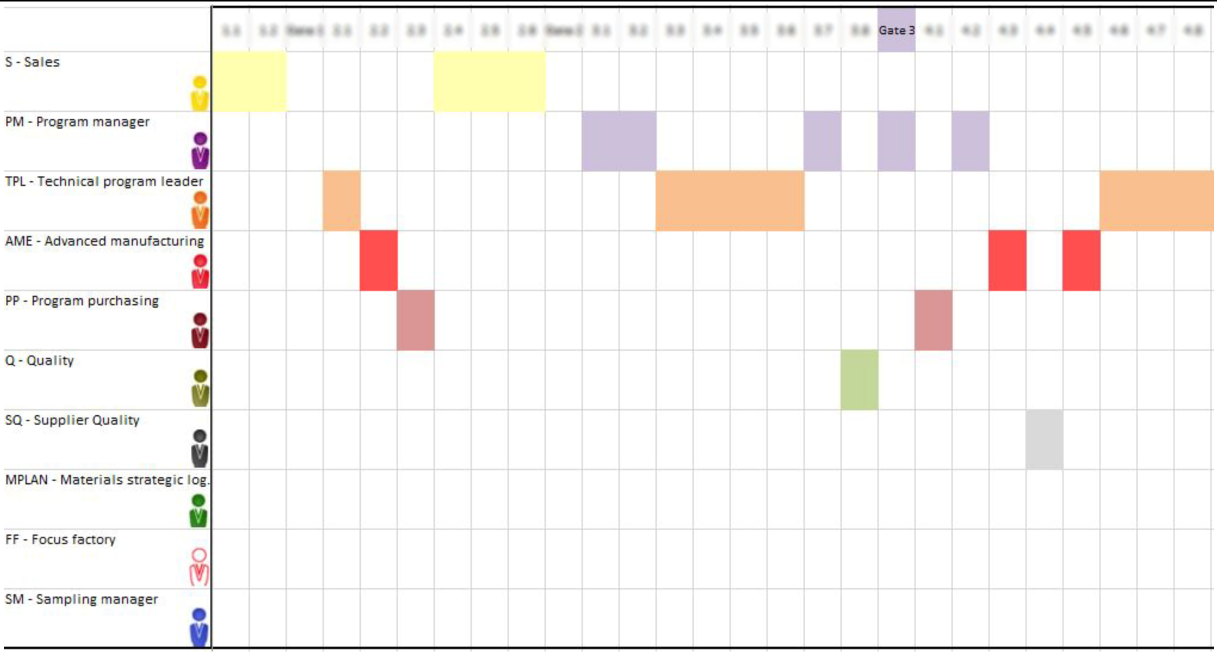


Figure 15 Workload distribution map part 1

Figure 16 Workload distribution map part 2

3 Cost estimation methods

The accuracy of the cost estimation process is the „make or break” of the project success. One of the greatest challenges for a PM is to successfully deliver on all aspects of a program, both according to the customer specifications and within the specified budget. It can often happen that either one aspect or the other can be accomplished, but not necessarily both. When it comes to controlling costs one of the most important things is the first step to make appropriate estimations at the outset of a program. Ability to estimate and control costs is largely a matter of adhering to established guidelines, often by learning from previous programs and promptly reacting to current circumstances.

Cost estimating is the predictive process used to quantify, cost and price the resources required by the scope of an investment option, activity, or project [15].

Good cost estimation has a direct bearing on the performance and effectiveness of a business enterprise because overestimation can result in loss of business and goodwill in the market, whereas underestimation may lead toward financial losses to the enterprise [16]. Because of the sensitivity and crucial role of cost estimation in an organization, it has been a focal point for design and operational strategies and certainly represents a key agenda for managerial policies and business decisions. Over the years, cost estimation methods have gone through evolution from simple rules of thumb to relatively complex models. This evolution has been primarily motivated by the growing complexity of products and processes, and the ever-increasing pressure on performance improvement motivated by the fierce competition.

Cost estimation is the most important preliminary process in any development project and is crucial to ensure the successful completion of a project. Estimating a product development cost is an example of a knowledge-intensive engineering task [17]; that is, it depends on the knowledge of the responsible expert. In fact, estimators require more than several years in the specific environment to develop the expertise to conduct the cost estimation process. The main problem is that the engineer's/manager's expertise is often not documented or authenticated. Hence, this expertise and knowledge affect the estimation process which is prone to subjectivity. Accuracy and comprehensiveness in cost estimation are fairly delicate issues. They depend on, and can be easily affected with many different parameters. In addition, each parameter has to be properly addressed in order achieve and maintain an acceptable level of accuracy and reliability during the process. There is no general consensus that placed the cost estimation techniques and methods into well-defined categories. There are numerous authors that dealt with the matter providing an overview of the methods in their own different categories. Also, several project management textbooks have chapters dedicated to the very subject of cost estimating [5] [7] [11] [27].

Table 3 Cost estimation classification matrix for process industries [27]

	Primary characteristic	**Secondary characteristic**		
Estimate Class	**Level of project definition** Expressed as % of complete definition	**End usage** Typical purpose of estimate	**Methodology** Typical estimation method	**Preparation effort** Typical degree of effort relative to least cost index of 1
Class 5	0% to 2%	Concept screening	Parametric models, judgment or analogy	1
Class 4	1% to 15%	Study or feasibility	Equipment factored or parametric models	2 to 4
Class 3	10% to 40%	Budget authorization or control	Semi detailed unit costs with assembly level line items	3 to 10
Class 2	30% to 70%	Control or Bid/Tender	Detailed unit cost with forced take off	4 to 20
Class 1	50% to 100%	Check estimate or Bid/Tender	Detailed unit cost with detailed take off	5 to 100

A number of authors have attempted to categorize estimation methods using certain criteria. Among many different types of classifications, criteria can be based on the type of industry as shown in the table above, where estimation methods are divided in five classes [27]. Some divided the methods into traditional detailed breakdown, simplified-breakdown, group-technology-based, regression-based, and activity-based [18], while some authors classified the very same methods into intuitive, analogical, parametric, and analytical [19]. Also there are authors that grouped methods in intuitive, parametric, variant-based, and generative cost estimating approaches without defining them clearly [20]. There are also authors that provided the comprehensive review of the existing methods and classifications and put them into an extensive hierarchical classification [16]. The latter hierarchical classification is made for the product cost estimation, but since the cost drivers and influence factors are the same as for the program cost estimation, they can be used in analogy. In the figure below is shown how cost estimation methods can be categorized. Some could say that the methods being used depend on the phase of the product development and on the quantity of information known about the product so far.

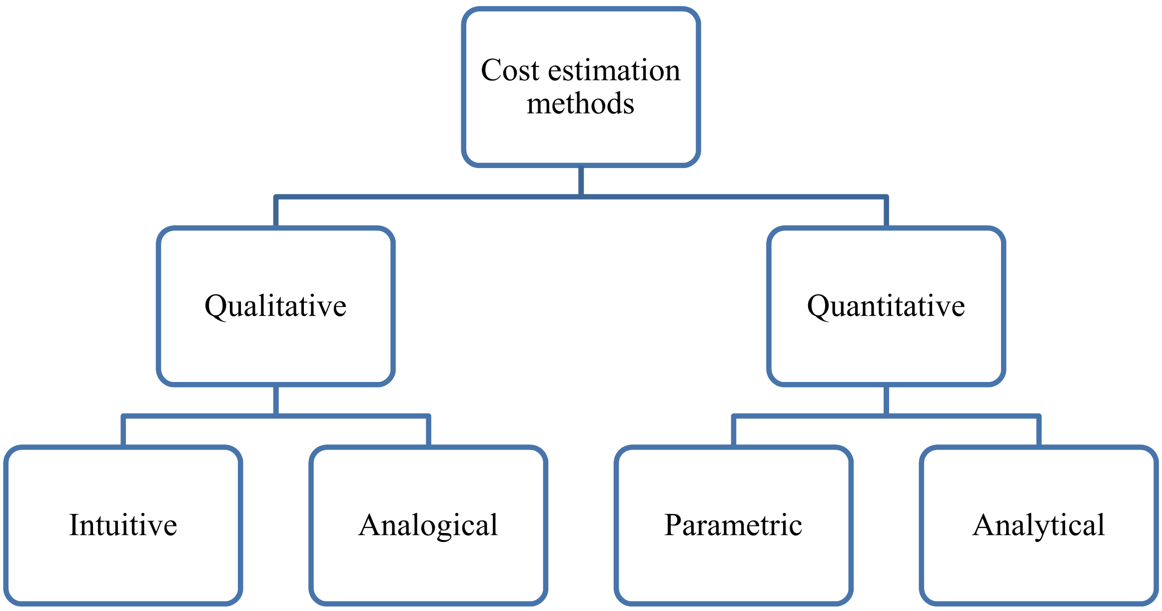


Figure 17 Cost estimation methods breakdown

3.1 Qualitative methods

3.1.1 *Intuitive*

Intuitive estimation methods, also known as expert judgment methods, are based on the human expert's prior project knowledge and experience [16]. A major benefit of these methods is moderate time consumption. They are usually applied in early project stages. Because the judgment depends on individual opinion, it is subjective and is void of mathematical estimating relationships. If not carefully treated expert judgment is prone to bias hence liable to wrong estimates. A major downside to the intuitive methods is that the estimations are very susceptible to a wide variety of subjective factors. Because of that, the obtained estimations face problems regarding accuracy and repeatability. These issues with accuracy can be reduced to a certain extent by using methods that are considering estimations from more than one estimator. Even experienced PM's often consult with their peers in order to fine-tune what they believe is a reasonable estimate. The expert judgment technique involves consulting one or more experts to validate the estimate of the proposed project against the experience and understanding of the experts, who will consider the details of project complexities and characteristics in tempering the estimate or concurring with it. Using an expert judgment opinion is somewhat akin to identifying and using the results of a parametric technique of personal nature, which is based on intuition, experience, and not-yet-articulated indices [21].

There have been some endeavors to reduce subjectivity through recursive refinement of estimates by a number of subject matter experts who reach consensus anonymously. This process is called the Delphi method. The whole idea behind this method is to minimize individual bias. Although bias can be minimized some problems still persist. It is difficult to provide accurate estimates using expert judgment because this method bases its calculations on basic estimates which are not adjusted according to different peculiarities of planned project deployments [22]. Because of various opinions that may produce different outputs given the same inputs, this method is highly uncertain and in terms of confidence rating it rates low. Expert‘s opinions are subjective hence estimates developed on them sometimes lack rationality and detailed analysis. Nonetheless, until these unspoken extrapolation techniques are formalized, expert judgment will remain one of the more reliable sources for checking the realism of the estimates.

Having in mind the descriptions of the company procedures shown in a previous chapter, conclusion can be made that it is exactly the intuitive cost estimation methods that is used. Even though some departments have the tools which resemble the parametric estimation, the data used in a tool is not based on history data, rather it's based on the expert intuition and that fundamentally represents the expert opinion estimation method.

3.1.2 Analogical

Analogous techniques are the simplest forms of estimating that still deliver reliable estimations. Analogous estimating refers to the estimating process where, in the PM‘s opinion, there is significant similarity between the proposed project and those projects that had already been carried out in the past. The analogy method uses the cost of a system that can be identified with the new one, adjusts for different values, and estimates the cost of the new system. This technique identifies a currently comparable system similar in design and/or operation to the proposed system [23]. They are also considered as conditionally reliable methods. The reason being that the relations between similarities are usually estimated by an expert. The main positive sides of these methods are the ability to achieve estimation quickly and the process being transparent. These methods strongly rely on the databases of past projects, and become unreliable if proper connection of similar characteristics can‘t be obtained [24]. Analogous methods tend to be less complex, easier to use, and more inexact compared to the parametric ones. These early estimates are used to formulate various options and to determine the relative viability of a project in the process of screening alternate projects. The base purpose of analogous methods is to develop order of magnitude estimates based on scarce information. Because of that the PM usually has to make several assumptions about some of the project's environmental or functional characteristics. Some of those characteristics are design properties, manufacturing process, implementation techniques, and resources availability. Here again, the PM‘s experience will be the deciding factor in judging the proposed project to be similar to those in the database or those that have formed the basis of the customized model [21].

The estimation task in analogous method is created by two sub-tasks: transfer of relative costs and the estimate of the new costs. When estimations are ready, they undergo a review procedure. This procedure includes the review of the estimated costs and determination of these costs rationality. One of the difficulties that poses themselves in such methodologies is that one can hardly find systems that are exactly the same. Hence, as mentioned before, estimates based on analogous method are prone to a high degree of subjectivity when determining the similar and dissimilar technical parameters of the analogous system. Because it is subjective, the level of uncertainty in estimates is very high. Being a judgment process, successful utilization of analogy method requires considerable expertise and experience [22].

3.2 Quantitative methods

3.2.1 Analytical

Analytical estimation methods can be used when both product data and manufacturing technology are specified in detail. They are usually applicable in the later stages of the product life cycle. This approach requires a detailed breakdown of the complete process into elementary tasks [16]. The estimation is made on decomposing a product into elementary units and operations. For every task relations between inputs and corresponding outputs are analytically determined. Also, activities that represent different resources consumed in the production cycle have to be taken into account. Final cost is determined as a summation of units, operations and activities. These methods are usually rigid and relations between parameters are not easily modified. They do not have adaptation ability. The gained results gives the most accurate estimations given that the data is reliable. The major downsides of these methods are time consumption and limited applicability in the early project stages [24]. Condition to use these methods in earlier stages of the project would be very well known product design specifications and manufacturing line properties. Since neither of those are available at the early stages of fuel tank development, these methods cannot be used effectively. It could be used with the introduction of the history data but then it would be a hybrid with the parametric cost estimation.

3.2.2 Parametric

Parametric estimating is a technique that develops cost estimates based upon the examination and validation of the relationships which exist between a project's technical, programmatic, and cost characteristics as well as the resources consumed during its development, manufacture, maintenance, and/or modification [25].

Parametric methods could be classified as statistical because they are derived by applying the statistical methodologies and by expressing cost as a function of its constituent variables. These techniques could be effective in those situations where the parameters, sometimes known as cost drivers, could be easily identified [16]. The goal is to develop cost estimating relationships (CERs) which are then used to estimate cost based on one or more product performance or design characteristics (e.g., dimensions, variants, weight, type). The parametric method is most commonly performed in the initial phases of product development, when the product is still in the conceptual design phase. Although during this phase the design is not yet very well specified, but the core properties are available and usually good enough to create an estimation. Similar to the other conceptual estimating methods, parametric estimating is reliant on the collection and analysis of previous project cost data in order to develop the CERs. An underlying assumption of parametric estimating is that the historical framework on which the parametric model is based is applicable to the new project (i.e., the technology has not radically changed) [26].

Even though statistical approach in parametric estimation is mostly done by regression methods, there has been an introduction of fuzzy methods in recent years, so data analysis within parametric estimation could be divided into:

- Regression methods
- Fuzzy methods

Both of those methods typically involve several independent variables or cost drivers. Difference is in how to approach the history data and what algorithm to use in data analysis. Tipically regression methods are easier to use and they can operate on smaller quantity of data if the quality is sufficient. There are several types of regression methods ranging from single to multiple linear and non-linear ones. Regarding the fuzzy methods there are numerous algorithms and techniques available for use, but the most widespread method is artificial neural networks (ANN). The inherent advantage of ANN is that they can effectively extrapolate and generalize. Furthermore, they can model cases where the functional relationship between case data is hidden or cannot be expressed in polynomial form, because an input–output mapping is allowed without understanding the functional relationship between influence factors [26]. However, ANN usually requires a large set of training cases what makes it a demanding method for the commercial use. Although, testing the ANN capabilities for the observed case, can provide useful insight on what is needed in order to effectively introduce fuzzy methods for cost estimation, as will be seen in later chapters.

4 Base for the tool development

It is evident from the ED&D charts data that cost estimation ratios are not aligned even between similar programs. Reasons for that are possibly lack of standardized procedures in internal communication, shortage of tools for cost estimation and error in expert opinion. Expert opinion is very valuable tool built by years of experience in conducting similar projects and tasks. However, different programs are led by different managers, and not all of them have the same experience and relevant knowledge. From that, conclusion can be drawn, if standardization in cost estimation wants to be achieved, then sole usage of expert opinion needs to be lowered or at least used alongside tools that were developed by experts. That way the influence of an individual is limited, but still allowing more experienced managers to make changes if needed.

Acknowledging the need for more accurate estimation and improvement over the present methods, it is crucial to create a solid foundation for implementation of the new method in form of an estimation tool. The tool's target users would be in the program management department, what is expected with them being responsible for overall estimations in the program. Tool creation is, among other reasons, motivated by great difference in cost estimation values between comparable programs, as seen in previous chapters. Tool’s main goal would be to create a better insight in program cost breakdown and to enable PM to discuss the costs with other departments with more confidence and to conclude the whole process in shorter period of time. In order to create such complex tool there are some key steps that needs to be followed.

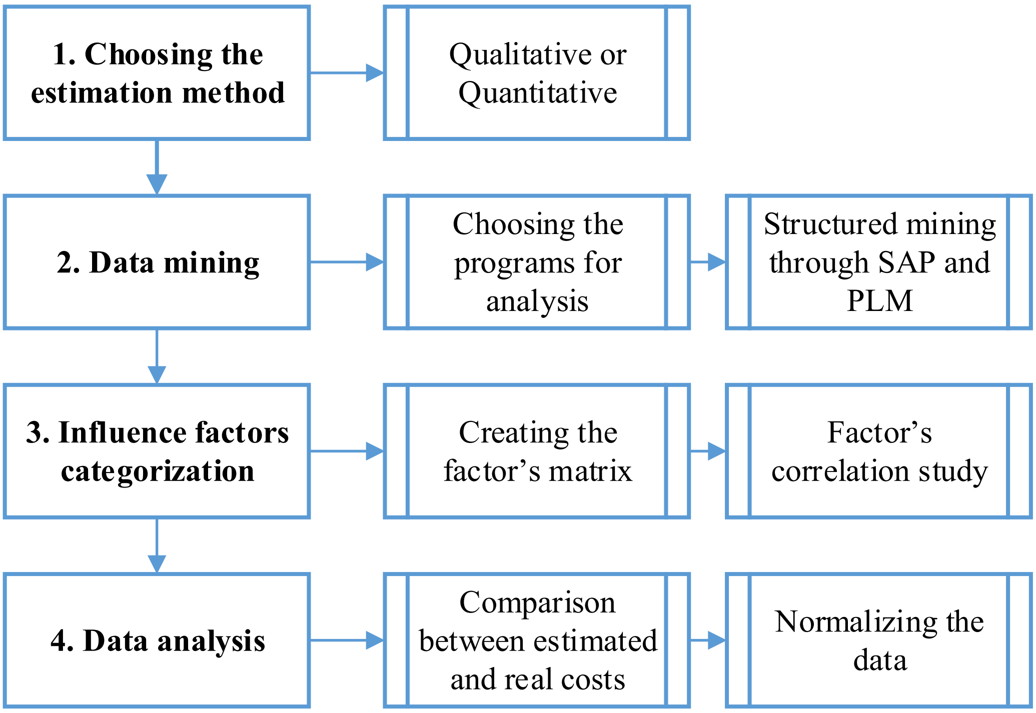


Figure 18 Key steps for the tool predevelopment

4.1 Choosing the estimation method

With base estimation methods being reviewed, one can set a number of elimination criteria in order to choose the right method that will suite the company needs and, most importantly, that is able to create a valid output based on the available data in specific time of the program. The criteria for the automotive industry, in the advanced technologies manufacturing environment, can be easily set based on the company's core product type. Particular case observed in this case study is the production of fuel tanks. As studies have shown, there are numerous estimation methods but also there is always clear distinction of why, when and how they can be used. Also, some of the most respected high technology companies (NASA & ESA) have devised, based on their experience, simple graphic for choosing the estimation method based on the actual program life cycle [23].

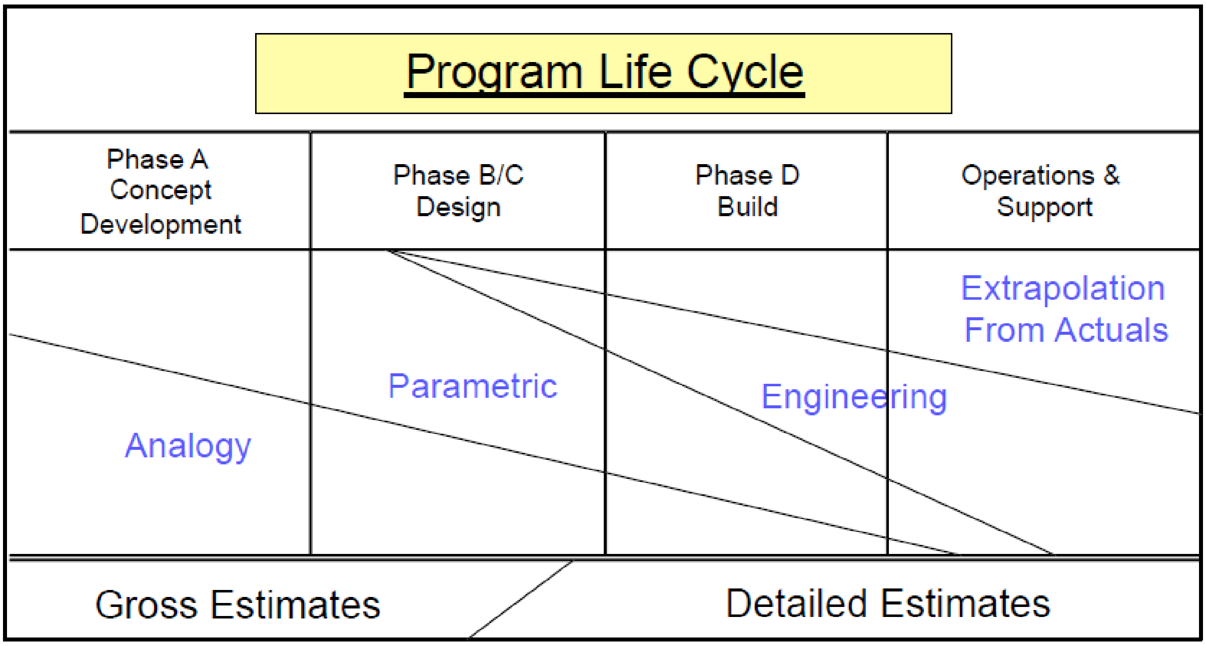


Figure 19 NASA's cost estimation method map [23]

Since the observed company has already developed reliable product development processes and has gathered great experience in the field of blow molding, it is possible to decide between qualitative and quantitative methods of the ECE.

Qualitative cost estimation methods are primarily based on analysis of previous products that were manufactured and their comparison in between. Cornerstones of successfully executed qualitative methods are valid identification of similarities between the products and the ability to extrapolate them within the new product development. In that sense, history data consisting of design and manufacturing process with known cost details provides useful insight in how new program development might behave. At the moment, qualitative techniques are utilized in the company's ECE process, precisely, intuitive methods.

Quantitative methods, on the other hand, are based on a detailed analysis of a product design, its features, and corresponding manufacturing processes instead of simply relying on the past data or knowledge of an estimator. Therefore, costs are either calculated using an analytical function of certain variables representing various product parameters or as the sum of elementary units representing different resources consumed during a whole production cycle of a given product. These techniques are known to provide more accurate results but their use is usually restricted to the final phases in the design cycle due to the requirement of a detailed product design, unless there is a considerable experience with the previous similar projects. For a tank development cost estimation, quantitative methods are better choice for several reasons:

- There is a considerable amount of similar programs made with the same technologies available for analysis
- Product materials and base components are usually the same kind
- There is a number of relevant parameters available prior to the design freeze
- If applicable, quantitative methods are more accurate than the existing intuitive ones

Quantitative methods are divided in two categories described in the prior segment. Analytical methods require knowledge about the product and the manufacturing process in detail, while parametric methods require knowledge about program factors that are critical for its performance. Since the company has a database for cost distribution tracking throughout the program development phase, assumption was made that it is possible to identify and extract the main cost drivers from the data sets. On that assumption parametric method has proven to be more feasible for use then analytical. The reason is also that in quotation phase of the project, fine details of manufacturing process are not fully known and are susceptible to sudden changes. To successfully carry out parametric method one needs to choose wisely between regression and fuzzy methods. Even though fuzzy methods have proven their efficiency, especially in production cost estimation, they usually need large datasets with high quality data in order to operate effectively and to create good estimations. If the dataset is linearly dependent than fuzzy methods could also be used successfully. For the observed case choice is on the regression methods which have proven their efficiency and have been used for forecasting for long period of time. Main reason to choose regression over fuzzy methods, as a primary tool calculation method, is their simplicity and ability to be very easily updated.

After choosing the parametric method and analysis technique in further tool development, two major steps needed to be carried out. One being history data mining and other, identification of program influence factors.

4.2 Data mining

Data mining is process of sorting through the data and finding the usable patterns and relationships in between. The overall goal of the data mining process is to extract information from a particular data set and transform it into an understandable structure for further use [28]. Even though number of available programs that meet the analysis criteria is not large enough to require usage of big data algorithms, methodology is still the same. For purpose of gathering the real working hours from the database, it is very important to have clear vision of needed data and to have predefined categories under which data would be classified after extraction. Furthermore, it is important to look for emerging patterns in the data through the perspective of influence factors, so it could be simpler to create valid prediction models. For data mining to be efficient and not time consuming, it is advisable to ask the right question while looking for needed data:

- What data type is needed?
- How to group the data?
- Is this data needed for further analysis?
- What is the correlation between data?
- In which category to put the data?
- Is this data subject to error?
- Is the quality of data sufficient for further use?

Because the potential big quantity of data, it is necessary to establish a good foundation for clustering before analysis and mining itself. At first, without identified patterns, clustering based on responsible roles should be sufficient to keep data well organized and monitored. All the data, initial ED&D and real working hours, should be clustered in exact same categories so that it is more comparable and easier to manipulate. First, ED&D data should be used to create categories because it is very well defined and filled with clear distinction between roles and responsibilities. Based on the created categories real working hours can be added for comparison and that way, it is easy to notice potential irregularities.

4.2.1 Choosing the programs and analysis methodology

To make the programs more comparable it is essential to choose the ones with similar goals, tasks and objectives. Hence, only programs from the same customer group are chosen for analysis. Product that is developed within these programs is fuel tank only. SCR tanks and filler pipes would be analyzed separately, depending on the outcome of the current research. Reason why only specific customer group programs were analyzed is that milestone and resources distribution among them is roughly the same, thus making it comparable. Furthermore, they are carried on within the same department making it easier to track and to connect the responsible roles with the outcome. It is expected from working hours distribution to be proportional between all programs. Criteria to pick particular program is:

- Specific customer
- New program (minimum carryover)
- Fuel tanks
- With minimal changes during the program
- Programs that are finished / in ending phase
- With available data for analysis

The premise of the research is to find usable data in the form of working hours spent on certain projects by individual roles within departments. Purpose of the mined data is to compare it with estimated working hours for the same projects and through the perspective of influence factors, make certain viable assumptions that would enable estimation tool development based on chosen method.

The goal is to get the raw data entered by program team members. Ideally, the data would consist only of working hours for the raw program development, excluding additional hours required by unexpected and initially unpredictable events. Program duration for observed customers is 22 to 37 months. In that period of time it is inevitable that some changes in the product development are going to occur. Engineering change management has its own process procedures that are followed in case it is needed. The reason why engineering changes costs (EnC) are not trying to be predicted by managers is that they can be charged additionally, depending on what the request is. It would also be nearly impossible to predict EnC simply because the product is almost always completely newly developed and even though observed programs are specifically for fuel tanks, requirements are very different for every project.

Potential challenge in data mining process could be identifying working hours spent on EnC requests and dealing with them. In ED&D charts, only personnel working hours for ideal outcomes is estimated. Data collected from ED&D is then not comparable to the real program data because of the EnC occurrence within the program. There are two possible ways to eliminate the extra hours added because of the EnC. First is to try and find all EnC request forms that are filled and to connect them to working hours that were spent on that EC. EC forms can contain the number of additional working hours requested. Second way would be to use EnC forms and see the cost of the change and calculate the working hours needed for its completion. Additional working hours used by EnC should be subtracted from the full amount, but with the great care for which responsible role is owner of that additional hours. If it would be difficult to find which roles are responsible for particular EnC, therefore, it is necessary to consult with the PM/TPL responsible for that program.

There are numerous potential problems within the systems that are used for data storage. Program itself is very dynamic process which outcome, among other, depends on the team member's expertise, willingness and motivation. Usually, the deadlines are very challenging and that requires of the program team to be fully focused on the current task. That is the reason why sometimes tracking the working hours on specific tasks is not in the main focus. Consequences of that approach are that working hours are not entered regularly to the database, what can result in inaccurate data unusable for analysis. Worst case scenario is that data is not entered at all or it is entered just to have something in the system.

Another type of problem that can arise is regarding the role responsibilities within the program team. In the beginning roles are clearly defined and working hours are estimated accordingly. Through the program development it can happen that one task is being done by a role that is not responsible for it, so in real data it manifests as increase in working hours that is not estimated or that should be assigned by other role. In order to eliminate this it is important to group the data by general departments, not individual roles. Another reason why data should not be grouped by individuals is that program time span is long enough for team members to change positions or simply leave the team.

Data validity depends on many factors that sometimes cannot be anticipated. If there are some datasets that look incomparable to other programs or estimations, they should be reviewed by an expert who will give opinion on reason why it turned out that way and how to adjust it, if possible. It could be helpful to find a program that went with minimum EnC and that was conducted under supervision of experienced manager, so it can be used as a benchmark for other datasets. One thing that could compromise the research is if validity, consistency and amount of data are significantly different between compared programs. Understandably, if estimations of the programs follow the differences than it would still be possible to compare them to real cost, but comparison of programs in between would not be possible. To achieve desired outcome, and that is to have well defined comparable data categories, every possible error needs to be well examined and resolved. The reason is that the number of programs for analysis is 13, and that number is not sufficient to the extent that possible errors would minimize with introduction of greater quantity of data.

4.2.2 Databases and mining outcomes

There are two most important software platforms used for storing the data and increasing the operational efficiency, SAP and ENOVIA PLM from Dassault Systemes. SAP's software is primarily used for keeping track of finances of the programs and company itself. Within programs it is used for resources hours allocation and for monitoring all the costs througout the program for every specific department, task and phase. Cost tracking provides an overview of program financial development and enables strategic planning of the resources reallocation if needed. Ideally, program gets identification number in the software prior to the development start and every major phase and request for change has its own identification connected to the main one. That way it has a well structured overview of the expenses and transparent development. ENOVIA PLM (Product Lifecycle Management) is another important software platform that was used for data mining. While SAP is used to retrieve information on program finances and to determine resources allocation, PLM is used to get the program documentation, product specifications and other data relevant for finding the cost drivers. Idea behind PLM is to store program documents divided into several subcategories ranging from quotation phase and calculation to the detailed design and product completion. Data is available at all times for the program team members and it allows for steady product evolution.

There were several issues that were encountered while mining from SAP database. The main financial data needed from the database were development costs and engineering changes costs. Some of the programs had costs other than development included in that very phase thus resulting in unreliable figures that couldn't be used without identifying those extra costs and reconfiguring the values. Some programs were very well managed from the financial side and all of the EnCs were added to the separate code with descriptions and costs, while others had changes costs added to the development code and again resulting in data that needed to be corrected prior to further use. One of the thirteen programs had to be discarded because it was very difficult to identify amount of costs that is assigned to development versus the amount that is assigned for changes and other program phases. PLM database research was used to gather the knowledge about the influence factors that are main cost drivers for the program and the outcome is described in the next chapter.

4.3 Influence factors

In order to make a representative tool for cost estimation, identification on program influence factors is crucial. Influence factors represent program and product specifications whose complexity and properties are directly related with the program outcome.

Every department has specific influence factors that affect the outcome on the local level within the sub team. But there are capital influence factors that are the same for every team and that are important for program in general and to the greater extent. Influence factors (IF) importance grows with quantity of tasks they are relevant to.

4.3.1 Influence factor extraction

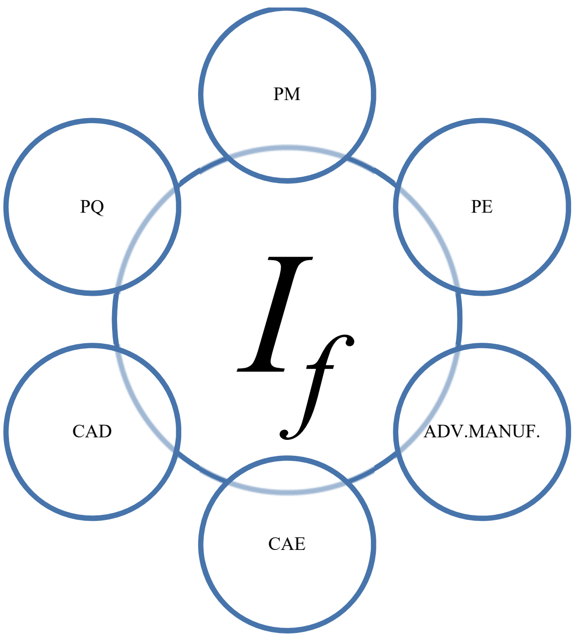


Figure 20 Influence factors - departments map

Every department has its own course of action and set of tasks that need to be fulfilled in order to meet the deadlines and in the end have a successful project. Every department also needs to estimate working hours load before starting the project so a valid quote can be delivered to customer. Working hours are mostly determined by product specifications, manufacturing location and the customer complexity. Most of the departments have excel tools where project/product specifications can be used as an input, and as an output it's expected to get feasible number of working hours. Understandably, not all of the specifications are of the same importance to every department. The goal is to extract the list of specifications used in every department and to compare them in between. Focus should be on the specifications that are shown to be important for every department and hence to the program in general. Furthermore, those specifications will be used as IFs and by that, history data would be easier to interpret. One of the main issues is where to draw the line with IF extraction. If there are too many IF, their impact is made relative and another corrective factor needs introduction. On the other side, it can also happen that number of overlapping specifications is low, in that case there should be a second step evaluation in which the rest of IF would be chosen. IF identification was conducted by using:

- Interviews with responsible team leaders
- Data analysis

Interviews were of the semi structured non bias type with the same set of questions for every responsible team leader. By using the same question and structure of the interview it was possible to find out what are the core factors needed for every department, how ECE works from their point of view and what can be improved on the matter. Analyzing the interview answers it was already possible to make a prediction on what IFs are going to be relevant for the program development. CAD and CAE departments working hours are mostly dependent on the fuel tank geometry and components positioning while PE and PM working hours are depending more on number of components that are built into the tank, BOM and overall complexity of the tank. Advanced manufacturing and PQ are, on the other hand, depending more on the plant location, used manufacturing technology and timing of the whole process. Some of the general preliminary IFs are:

- Tank type
- Technology (NGFS/Conventional blow molding)

- Geometry complexity
- Number of variants
- Number of additional components
- Customer
- Carryover amount
- Plant location
- Number of tools
- Program duration

Every separate tool used in ECE has a lot more parameters for input, but these should be the ones that are related to all departments. Purpose of using them is to make a new perspective on cost estimation and to use them as correction factors in the estimation tools. Also, IFs can be used in data mining to make a correlation between real working hours and the events that influence them directly. For example, if two projects are similar and one has been planned for factory on different location, through the IF related to plant location, it is possible to see what is the outcome regarding the working hours and to compare them with other IF outcomes. Result would be a clear conclusion on how project location affects working hours in real environment and possibility to adjust future estimations accordingly.

4.3.2 PLM documents

Through PLM database it was possible to extract documents that have information about almost all of the IFs mentioned above. For ones that didn't have information it was needed to contact the person responsible for the program and get the information from them. 13 programs in total were available for analysis and most of them had nearly all of the documents needed to identify IFs. Two programs had to be discarded because data was simply insufficient to make a conclusion about the cost drivers. Discarded programs were usually finished some time ago and the responsible person was not available for questions or if it is, very specific information that is needed simply is not remembered in detail. The real challenge in extracting the data from PLM was that it was not structured between the programs. Base folders are uniform and standardized but the content placement is not defined. That results in various data being sorted in wrong categories and making it very difficult to find and identify. Also, nomenclature of the documents is not agreed upon, hence the same documents were called differently among various programs. The same as with the existing cost estimation method, data management and sorting is also very dependent on individual's knowledge and experience. Some programs had very well managed documents that were transparent and easily identified, while others had missing documents and no obvious logic when it comes to sorting them. What could help in that particular case is to create a standard, or rule, of what documents are needed in PLM, in which categories and under what specific names. After all the data was gathered and IFs identified, next step was to create more thorough analysis.

4.4 Data analysis

Upon retrieving all the needed documents and cost values, next step was to analyze, evaluate and compare the data to create the foundation for further tool development. After discarding three programs for insufficient or low quality data ten other programs were left available for analysis. First it was needed to establish the existing estimation error within analyzed programs. Two major values were needed to proceed with the analysis:

- Estimated cost at the beginning of the program
- Real program cost

Estimated cost at the beginning of the programs was available from ED&D charts that were stored in the PLM database, or from project approval forms which are used by the controlling department prior to program start. Real program costs were retrieved from the SAP's system. There were several issues to consider while retrieving the real costs. In the SAP's categories under which costs are listed are not the same as in ED&D chart. For example, there are no PM and PE categories for themselves, instead, they are grouped into one category named project engineering. Furthermore, it was needed to determine how much of the real cost belongs to the PM and how much to the PE. There were several programs that had document consisting of program tracking within ED&D categories and that was used to determine percentage attributed to the each department. For the programs that didn't have that documents, percentages were used from the WBS distribution where it has been concluded that from overall project engineering cost PM has share of 38,4% (table below).

Table 4 Percentage of PM hours in Project engineering

Program	% of Project engineering in total working hours	PM % (WBS 38,4%)	Comment
1	64	36	
2	58	34	More CAD/CAE hours for filler
3	79	50	Design carryover -low CAD/CAE
4	58	35	More CAD/CAE hours for filler
6	61	40	
8	64	40	

After the data adjustment and alignment with the ED&D categories, it was needed to identify if engineering changes and other costs were entered into development phase and to correct/subtract the data with corresponding values. For every engineering change and budget intervention there is a matching project approval form that has detailed budget overview, where it is visible what are the changes to the cost and for what is the reason for them. Problem was that some of the programs had only few project approval forms available in PLM, and some of them were missing. Fortunately, extrapolation of values from those that were available has proven to be sufficient for data correction. After having the data prepared and categorized, the next step was to analyze the differences between estimated and real values. However, prior to analysis the next step in the process of developing a parametric model is to normalize the data. Normalizing the data refers to making the necessary adjustments to the base cost data to account for the differences between the actual basis of the data for all programs, and a desired standard basis of data to be used for the parametric model [26]. When normalizing the data the most important thing is not to affect it in a way that it is not reliable anymore.

4.4.1 Estimated versus real cost

SAP's software has all the cost entered in Euros, and for the analysis that is not relevant because of the labor price changing every year. Also, price of man-hour is not the same for every department. One of the most important things was to convert the cost to working hours and to normalize the data. In order to do that correctly there were several things to consider. First task was to identify correctly the timeframe for every program. Some of the programs started their development phases from 2007. and some up to 2012. Second, what was of great importance, was to create a base of all hourly rates for every department and for every year from 2007. After having those done correctly the next step was to identify when the data was entered into database and finally to convert the cost from Euros to working hours by using simple division. After all of the working hours were available it was easier to proceed to the next step, which was to finally compare the differences between the estimated and real working hours. The table below represents an absolute difference of total program hours in percentage. If there is a minus that means that program has spent less hours than it was initially planned. It is obvious that differences range between four and thirty one percent. Even though some programs have fairly low absolute differences there are still great ones between the departments. For example, program number 6 has lowest absolute difference but has one of the biggest in estimated and real advanced manufacturing and CAE. One could argue that it is not important how great department differences are if total value is within desirable limits, however, that is not correct due to the fact that estimation is wrong in itself. What happens is that some departments are overestimated and some are underestimated and final absolute value positions itself under 5% mark. What is important is to create valid estimation between the departments that will result in sound absolute value, thus enabling future reliance on that program while creating the new estimations.

Table 5 Absolute difference of total program wh's (normalized)

Program	Estimated	Real	Δ(%)
1	0,357	0,337	-6
2	0,390	0,317	-19
3	0,118	0,150	31
4	1,000	0,900	-10
5	0,436	0,338	-22
6	0,322	0,310	-4
7	0,564	0,496	-12
8	0,442	0,488	10
9	0,825	0,760	-9
10	0,791	0,581	-27

From the table above, that shows normalized data, it is obvious that most of the programs have spent less amount of working hours than estimated in the beginning. One of the reasons could be that PMs are basing the estimation on their experience, and it tells them that there are numerous changes and difficulties occurring throughout the program. Because of that it is understandable why they sometimes overestimate the program workload, even though they know additional changes are charged separately from the development costs. These effects also support the main premise of the estimation tool, and that is to minimize influence of an individual and to base it on relevant history data.

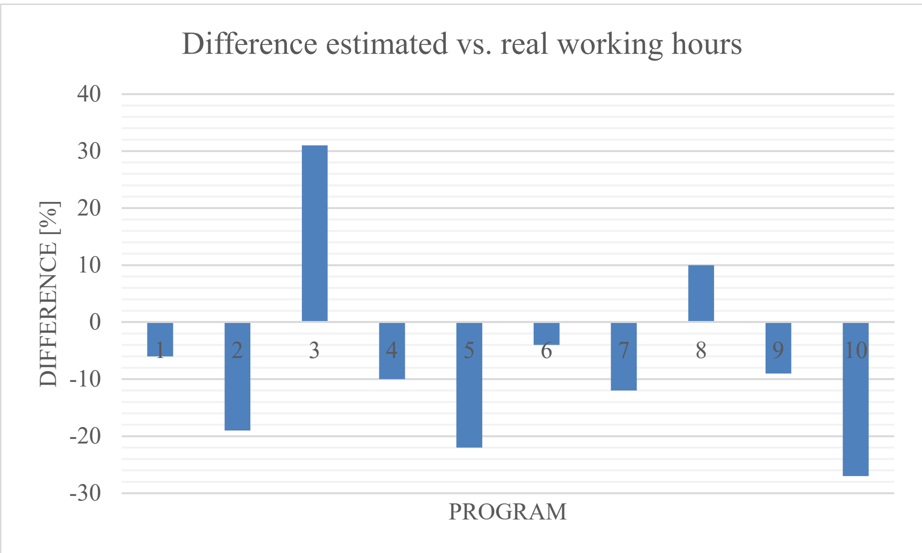


Figure 21 Difference est. vs real wh's for analyzed programs

4.4.2 IF matrix

By completing the working hours data correction, comparisons and analysis, it was possible to create the influence factor matrix for the programs. Using the PLM data, identifying relevant information and talking with responsible program members, enabled the creation of such matrix. Matrix provides insight to the program's properties and creates the possibility to connect relevant cost drivers to the cost itself. Factors used for the matrix were based on the preliminary ones given in the previous chapter. Preliminary factors were listed only from interviews with the responsible team members. However, through data analysis, detailed budget reviews and interviews, number of the most important cost drivers that are the same all across the departments has been narrowed down to those in the table below. Although, the process that helped the most in deciding which factors are relevant is examination of the existing tools for estimation within the departments.

Table 6 IF's matrix

Program	No. of variants	No. of tools	Technology	No. of plants	Shape	Fillers	Duration (M)
1	2	1	BM	2	Single	No	35
2	2	1	BM	1	Single	Yes	32
3	2	1	BM	1	Saddle	No	23
4	7	3	NGFS	2	Saddle/Single	Yes	37
5	3	1	NGFS	1	L-shape	No	28
6	2	1	BM	1	L-shape	No	26
7	6	2	BM	1	Saddle/Single	No	27
8	4	2	BM	1	Single	No	31
9	3	1	BM	1	Single	Yes	26
10	3	1	BM	2	Single	No	22

Number of tank variants has proven to be one of the most important program influence factors when it comes to the program costs. One variant represents a type of fuel tank created from a single tool (mold). One tool (mold) can be used to produce a number of variants and all variants are of the same shape but they differ in additional components that are built into the tank shell on the manufacturing down-line. Usually the variants that are made from one tool are for diesel and petrol fuel, but also due to the different emission requirements those tank shells can be additionally modified depending on the customer requirements. Most of the time there are two to three variants per one tool.

Number of tools essentially represents number of molds that are used to produce the tank shell. PFT development is arguably most challenging up to the point where tool is finished. Number of tools can be perceived as a number of projects within the program thus making this factor directly correlated with the working hours spent in total. Motivation for more than one tool is customer’s need for various tank shell shapes. Sometimes, within the same program, customer requires tanks for different car models or different drivetrains such as all-wheel/rear wheel drive or front wheel drive and that requires different shapes of the tank.

Manufacturing technology has a great impact on program development cost. In the introduction it was mentioned that NGFS procedure is more complex than the conventional blow molding. Consequentially, that requires more planning, engineering and logistics throughout the program. Understandably, if the manufacturing process is more complex there are greater risks involved in the procedure and it involves more attention from nearly every department included in the development.

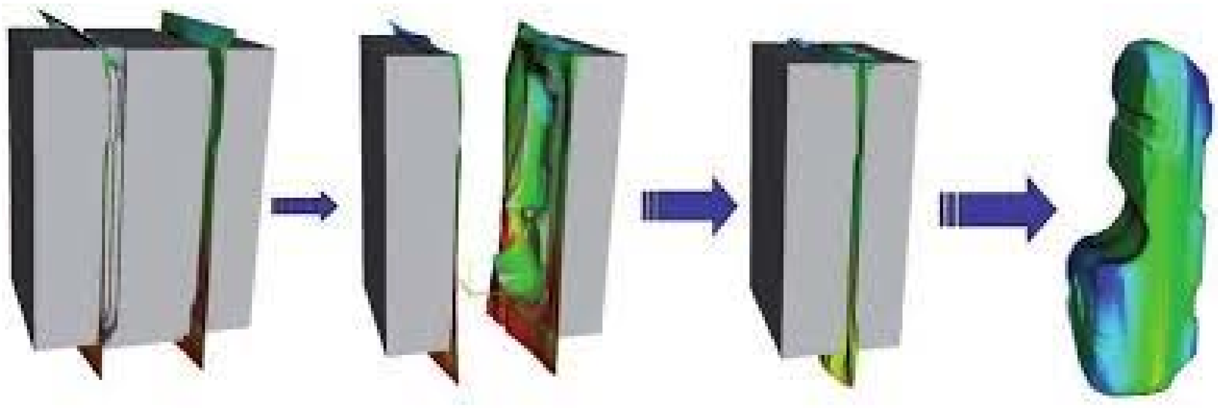


Figure 22 NGFS process

Number of plants mainly affects PM and advanced manufacturing. Having over 30 factories all over the world, and expanding, means not only engineering challenges but also social ones. That is the reason why there are local teams handling actual problems and being on disposal when needed. Some plants are fairly new and unexperienced and there could be potentially challenging situations while implementing new technologies and new production lines. That is why advanced manufacturing costs are affected primarily by technology and plants factors.

Shape of the tank shell is directly related to the tool that forms it. As mentioned, different shapes are motivated with different customer requirements and they primarily affect product engineering department according to their estimation tool. For this estimation tool shape is mostly put as a factor to have it stored for future analysis of correlation with the program cost. There are three main shapes of the tank shell:

- Single chamber - used mostly for the front wheel drive vehicles (Fig. 21)
- Saddle tank - used for rear wheel or four wheel drive cars (Fig. 22)
- L-shaped - could be used for all drivetrains depending on the position (Fig. 23)

Saddle tank has the most complex geometry that allows driveshaft to reach the rear differential. Also, it requires a heat shield to protect the tank from the heat sources and rotating parts.

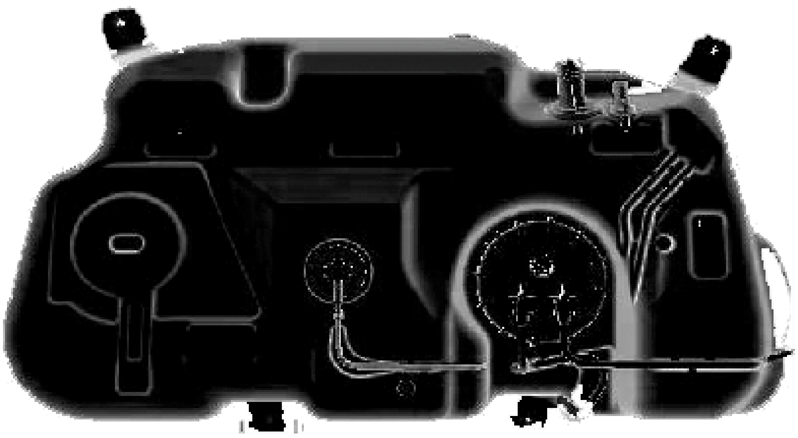


Figure 23 Single chamber fuel tank

Figure 24 Saddle fuel tank

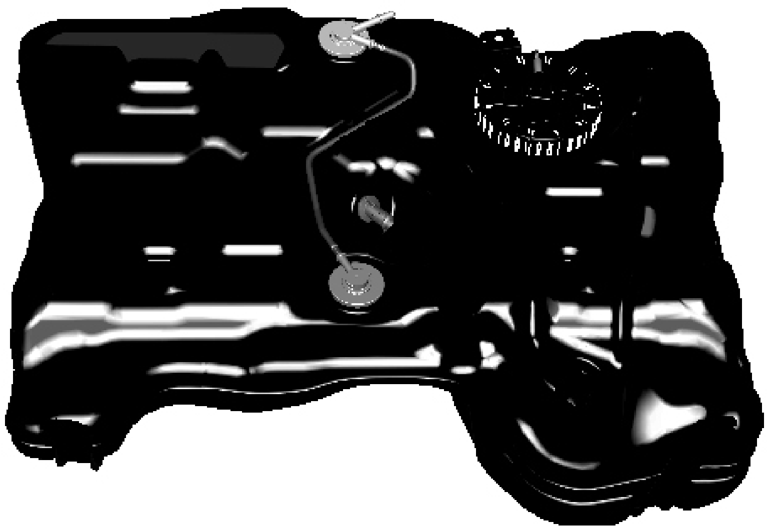


Figure 25 L-shaped fuel tank

Tank fillers are not the part of every program, their development depends on customer's request. However, if they are requested then the separate estimation is needed and subproject needs to be carried out. In this case, having a filler affects the overall tank development and when treated under the same program, it requires additional hours.

Figure 26 Tank filler

Duration of the program is the most important factor correlated with the working hours. As it can be seen programs duration ranges from 22 to 37 months and that reflects on the amount of working hours needed. Programs that have more tools, therefore more complex, are more likely to be finished in a longer time span. How each of these factors correlates with working hours will be shown in the next chapter.

5 The tool algorithm and development

Completing the data gathering and analysis enabled the further development of the estimation tool. For tool to be reliable the most important thing is to have history data used in a way that would enable confident and reliable estimation. Having data sorted and grouped into categories that can be used by statistical methods provides a solid ground for estimation. As much as quantity of data is important, quality seems to be of greater value for precise estimations. The idea behind the tool is, as already mentioned, to minimize the individual influence. Ideally, the tool estimation would rely solely on history data and there would be no intervention needed by responsible PM. It could be achieved only if there would be enough data to create the pattern for every possible program/product type that company offers. For example, several programs with NGFS with one, two and three molds, several programs with more than three plants, etc. It is obviously not the case since the company has evolved through time and is constantly improving new technologies and data storage platforms. In this particular case one could not rely only on history data because the sample simply is not big enough to draw conclusions for every possible program. That is why the estimation tool is built as a hybrid between parametric cost estimation using statistical methods and the expert opinion. The most important influence factors were used as a base for estimation using statistical methods. Values calculated using them are divided into two parts. One is non-manageable and holds 70% of the estimated value. Other part, the 30% of the value, can be affected by the estimator through usage of other influence factors as multipliers. Multipliers are derived from previous experience with the program and were being somewhat successfully used in existing estimation tools. This approach permits expert intervention but still allowing majority of the estimation being drawn from solid relevant history data. The tool will use regression models to estimate the data for PM, PE and CAD departments. For other departments the data was not sufficient and when performing correlation study there was no statistical significance of given influence factors. The main reason is that the data for those departments is not stable and could not be described linearly. For these departments estimation will rely on intuitive methods and it will also be incorporated into the tool.

5.1 The tool analytics

Parametric estimation is relying on usable historical framework of previous programs and using the statistical methods to devise equations that would describe a behavior of potential program based on the chosen influence factors. There are many diverse techniques that can be deployed for in depth data analysis as mentioned in previous chapters. Typically, data analysis consists of performing regression analysis of costs versus selected influence factors. As already mentioned, there are several fuzzy methods that could be used, the most popular being artificial neural networks, and their possible application will be described in later chapters.

5.1.1 Multiple linear regression

For cost estimating purposes, regression techniques are used to examine the contribution of different variables to the cost. One starts with a list of variables possibly influencing the cost. Based on statistical tests, significant variables are selected from the list and combined into cost-estimation relationships [29]. One of the best strengths of regression techniques is the ability to interpret the relationship between the cost and the influence factors considered. Observed influence factors depent linearly on needed working hours in the programs. that is the reason why linear regression is chosen over non-linear one. Furthermore, there are two types of regression, simple and multiple. In simple linear regression only one predictor variable (IF) is being considered. When there are more than one predictor variables, such as in observed case, one has to use a multiple linear regression model. Multiple regression model is just an extension of the simple model where parameter (i.e. slope) estimates are now included for each predictor variable in the model. These coefficient values for each predictor are the slope estimates [31]. Multiple linear regression theoretical equation is as follows:

$$Y_i = \beta_0 + \beta_1 X_{1,i} + \cdots + \beta_k X_{k,i} + \varepsilon_i \qquad (1)$$

Where Y_i, $X_{1,i}, \ldots, X_{k,i}$ represent the *i*th observations of each of the variables *Y* as the dependent variable and *X* as an independent variable. $\beta_0, \beta_1, \ldots, \beta_k$ are fixed and unknown parameters while ε_i is a random variable that in practice represents estimated error.

The main objective of regression analysis is to create a CER in the form of an equation that would represent a relationship between a dependent variable to one or more independent variables, or in this case influence factors. The dependent variable is called like that because it responds to changes in the independent variable where:

- The value of the dependent variable, designated by the symbol *Y*, is calculated
- The values of the independent variables designated by the symbol *X*, is known
- The resulting relationship between *Y* and *X* can be described mathematically

When working with CERs, the dependent *Y* variable typically represents cost, while the independent *X* represents the parameters (IFs) of the program being estimated. Influence factors should be chosen because there is a strong correlation between these variables and cost (*Y*) and because there are sound principles for the relationship being investigated [23].
To evaluate the regression model, two measures are of main importance. The adjusted R^2 value expresses the percentage variability in the cost that can be explained by the influence factors included in the model. The p-values give an idea of the significance of the individual influence factors.

5.1.2 Cost equation

By the analysis method being multiple linear regression, there are some steps and rules in statistics that have to be respected in order to get constructive and usable results. First, correlation matrix had to be created to evaluate the correlation between real cost of the program and influence factors. With the trial and error method there were three most important influence factors discovered as continouous predictors and one influence factor as a categorical predictor. Using the Pearson correlation method correlation matrix was created as in figure below.

Correlation: Wh; Tool; Mnths; Var; Ngfs

```
          Wh   Tool  Mnths    Var
Tool   0,972
       0,000

Mnths  0,698  0,553
       0,025  0,098

Var    0,908  0,930  0,369
       0,000  0,000  0,294

Ngfs   0,751  0,452  0,405  0,475
       0,04  0,189  0,246  0,166
```

Figure 27 Correlation matrix

Correlation matrix shows p-values among influence factors and working hours (Wh). The p-value is a probability that measures the evidence against the null hypothesis. Lower probabilities provide stronger evidence against the null hypothesis. In other words, p-values are used to determine whether the results are statistically significant. Common p-value is compared with significance level α which is usually 0.05. The p-value is a probability that measures the evidence against the null hypothesis. If the p-value is less than or equal to α than the null hypotesis is rejected (H_0), otherwise null hypothesis is accepted, or rather failed to reject. Upper number in the correlation matrix represents Pearson correlation coefficient and the lower one represents p-values. As can be seen from the figure, correlation between working hour and all the other influence factors has p-values under 0.05, that means that the null hypothesis, which states that there is no significant linear correlation, is rejected. Having p-values under α shows statistical significance of influence factors and allows for the next step in analysis which is performing multiple linear regression.

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
186,279	99,81%	99,65%	98,40%

Figure 28 R squared values

As mentioned earlier, one of main indications of valid regression model is the R^2. R^2 is the percentage of response variable variation that is explained by its relationship with one or more predictor variables. Usually, the higher the R^2, the better the model fits your data. R^2 is always between 0 and 100%. In the performed analysis R^2 equals 99,81% which means that almost all variations in the data can be described with these factors for given working hours. Also, R^2 adjusted and R^2 predicted, confirm that the model is valid and not over fitted.

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-3720	460	-8,09	0,000	
Tool	2312	312	7,41	0,001	12,35
Mnths	177,2	18,1	9,77	0,000	2,08
Var	527	114	4,63	0,006	10,59
Ngfs					
1	491	122	2,57	0,039	1,45

Figure 29 Multiple regression matrix

Again, the T-value and the corresponding p-value shows great statistical significance of the data and the validity of the regression outcome. The normal probability plot shows that data distribution is fairly normal, even though one dataset may be an outlier. However that doesn't affect the validity. The normality of data vould increase with more program samples with the sae quality.

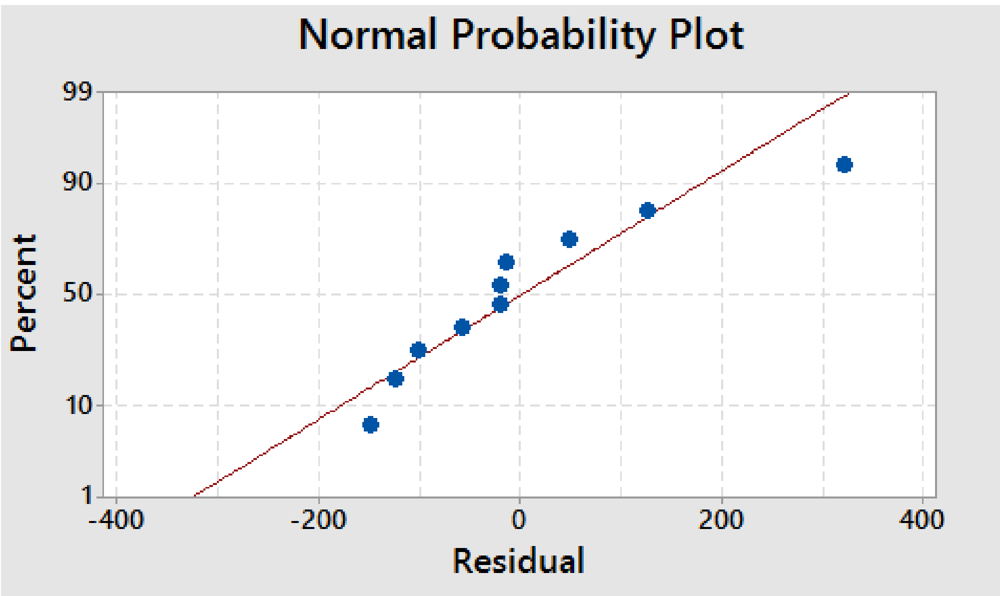


Figure 30 Normal probability plot of data

In the versus fits graph it is visible how samples behave around the fitted line and what are the residuals for given fitted value. It is shown how samples are behaving in a random distribution without emerging patterns indicating the desired outcome.

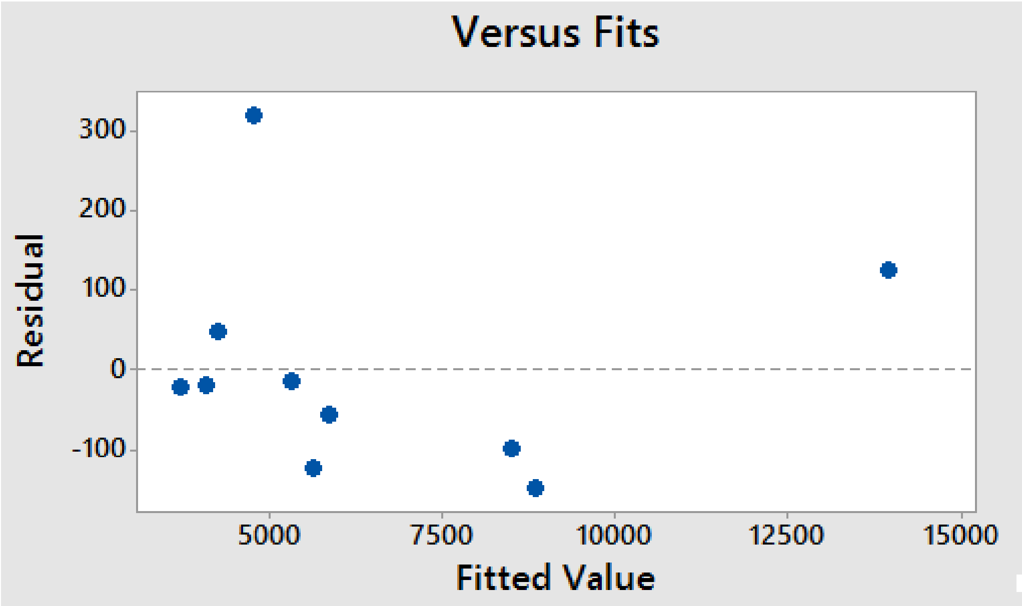


Figure 31 Residuals vs. fitted data value

There are three continuous variables which means that they are affecting the cost in a uniform way with constat values of their multipliers. One categorized variable is related to which technology is used in manufacturing. Categorized variables can have a binary state, having values 0 or 1. If the value is 0 that means that, in this case, technology used in manufacturing is conventional blow molding, and if the value is 1 then the technology used is NGFS. Categorized variables operate in a way that each state represents different equation that can be used to describe the behavior of the observed system. There are two equations that are able to describe amount of working hours for PE and PM departments, depending on what technology is used.

```
Regression Equation

Ngfs
0     Wh = -3720 + 2312 Tool + 177,2 Mnths + 527 Var

1     Wh = -3228 + 2312 Tool + 177,2 Mnths + 527 Var
```

Figure 32 Regression equations used in the tool

These equations will be incorporated into the tool to provide the base cost estimation which will be adjusted by other factors.

5.2 The tool description

The tool was developed with the Microsoft Excel and Visual Basic for Applications (VBA). It consists of two main parts, one being the summary sheet and other being input user form. Input user form provides an opportunity for estimator to enter all the known program parameters and to receive an estimation in the form of working hours needed for most of the departments. The tool estimates working hours for PM, PE, Adv. Man., CAD and CAE departments. After thorough research it was concluded that PQ hours depend on the specific factory and that every factory already has an estimation tool that is used in the quotation phase. That tool is used by the person responsible for manufacturing supervision. Ideally it would be possible to include all the factories into this estimation tool but the available data was heavily insufficient to do so. There are several input types of user form controls that the tool relies on:

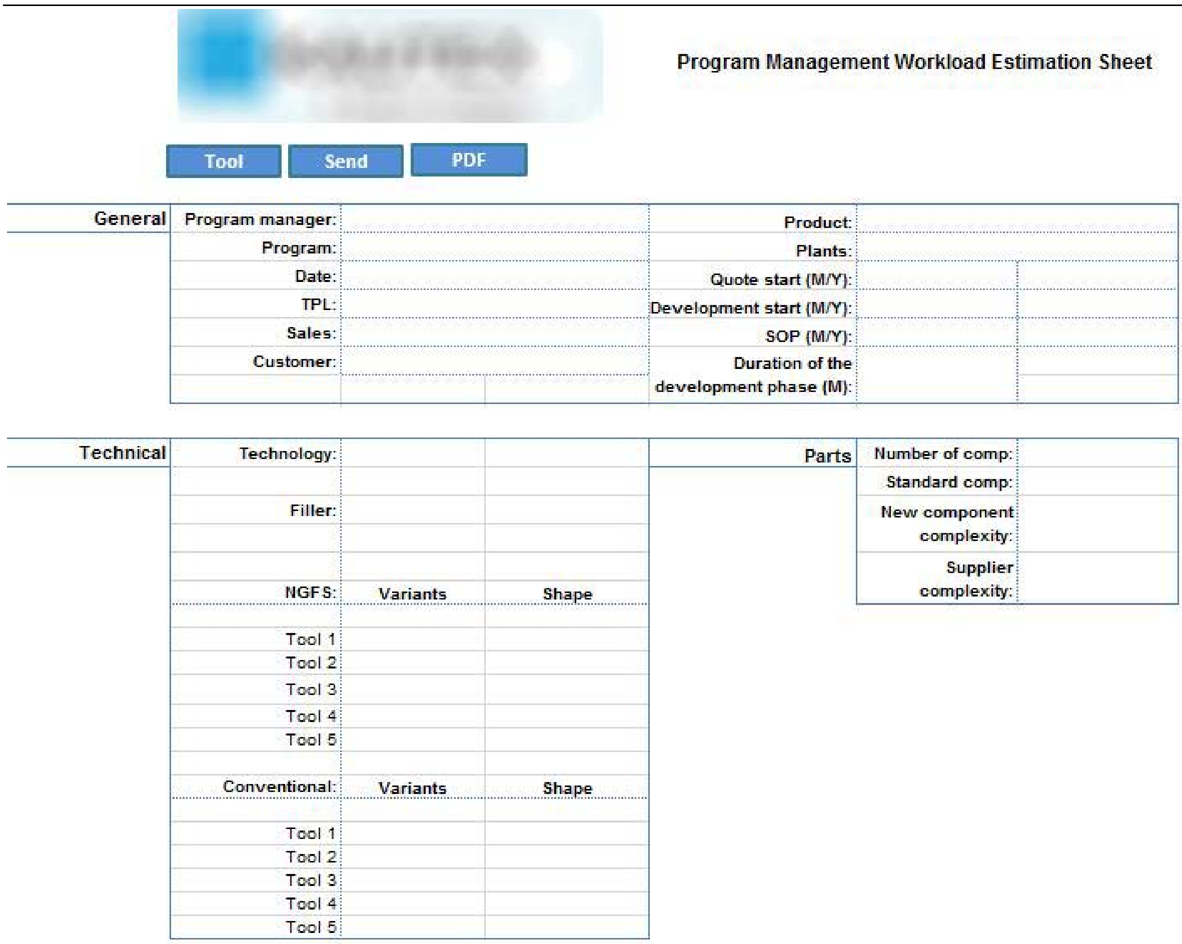
- Choose quantity / Drop down (1)
- Choose type / Drop down (2)
- Calculation (3)
- Slider (4)
- Written input (5)
- Checkbox (6)

All of these controls are used as an input tools to enter the data into the background worksheet which afterwards uses them for the estimation calculation. Numbers assigned to input types are used to describe the connection between factors and controls used for input.

Table 7 Relationship between IF's and input type

General tab		**Technical tab**		**Parts tab**	
Factor	Input type	Factor	Input type	Factor	Input type
Customer	2	Technology	2	#Components	1
#Plants	1	#Tools	1	#Standard components	1
Locations	5	#Variants	1	Supplier	4
Duration	3	Shape	2	New components complexity	4
Linetype/Add.	2	Filler	6	Carryover amount	4
Team exp.	4				

Summary sheet is used to store the data entered via user form. Also, it is used as a primary source of information when sharing the estimations, while user form serves the purpose only to populate the summary sheet more easily. By using the command buttons on the sheet, it can be printed out into two pages ISO A4 format, sent as an attachment in the e-mail or it can restart the user form if needed. Summary sheet is shown in the figures below.



Program Management Workload Estimation Sheet

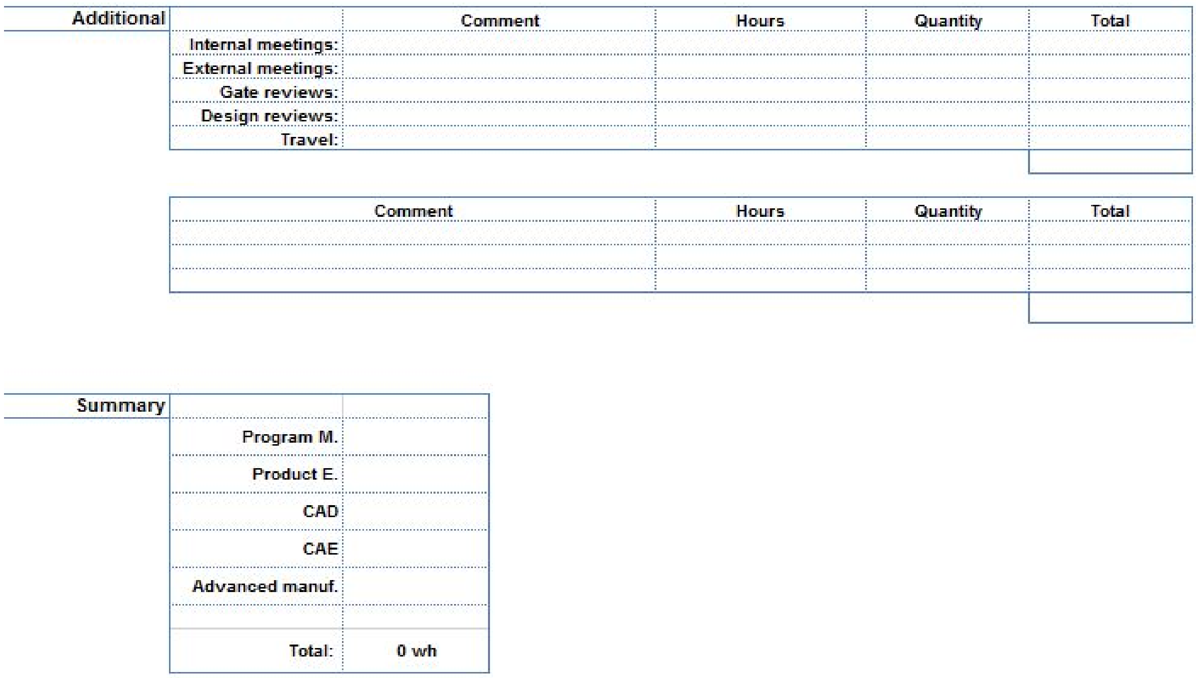
Tool | Send | PDF

General			
Program manager:		Product:	
Program:		Plants:	
Date:		Quote start (M/Y):	
TPL:		Development start (M/Y):	
Sales:		SOP (M/Y):	
Customer:		Duration of the development phase (M):	

Technical		
Technology:		
Filler:		
NGFS:	Variants	Shape
Tool 1		
Tool 2		
Tool 3		
Tool 4		
Tool 5		
Conventional:	Variants	Shape
Tool 1		
Tool 2		
Tool 3		
Tool 4		
Tool 5		

Parts	
Number of comp:	
Standard comp:	
New component complexity:	
Supplier complexity:	

Figure 33 Summary sheet part 1



Additional	Comment	Hours	Quantity	Total
Internal meetings:				
External meetings:				
Gate reviews:				
Design reviews:				
Travel:				

Comment	Hours	Quantity	Total

Summary	
Program M.	
Product E.	
CAD	
CAE	
Advanced manuf.	
Total:	0 wh

Figure 34 Summary sheet part 2

5.2.1 General tab

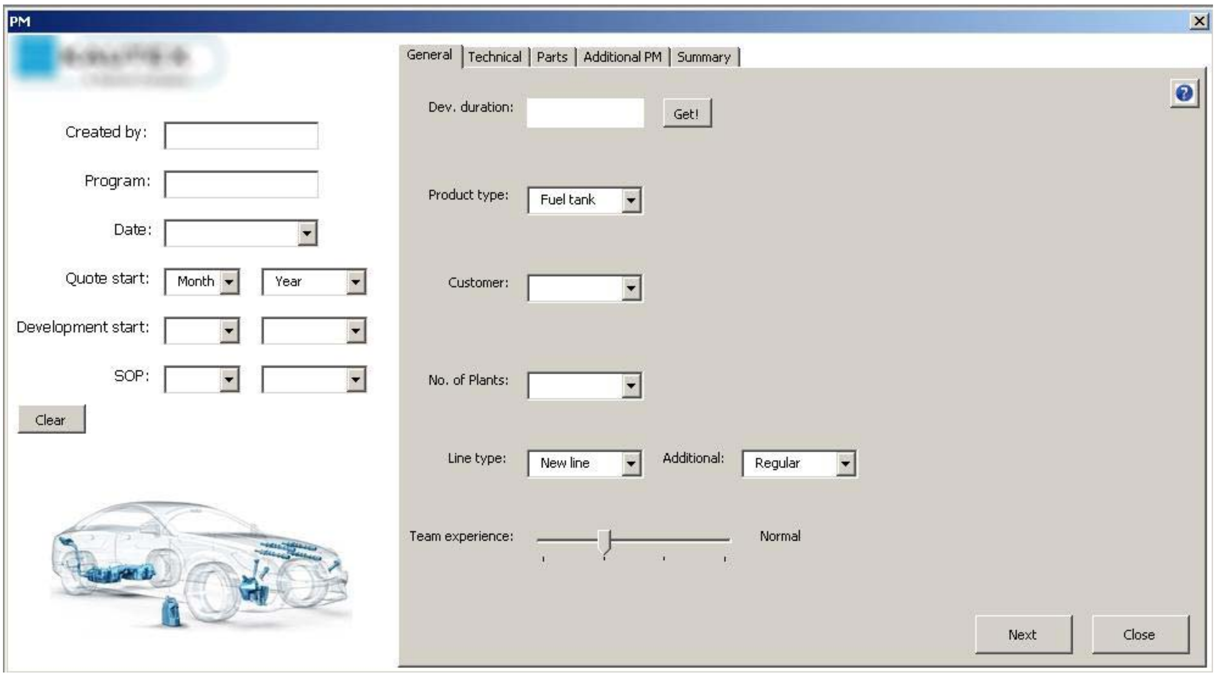


Figure 35 The tool interface - General tab

When opening the tool, the user form is the first thing that shows up and the first step of the estimation would be to enter the program data to the *General* tab. *General* tab requires input of the main program characteristics. Product type in this tool is limited only on fuel tanks, with possibilities to add SCR tanks upon further development of the tool. Customer group holds a list of all the company's customers and the most important OEM's. Number of plants is limited to 5, even though most of the time in real programs number of plants is two or under. By choosing the number of plants matching number of textboxes shows up enabling the estimator to add the names of the plants using the standard company code. Line type offer a choice between new line and copied line, because it heavily affects the advanced manufacturing hours. Also, additional list offers a choice between: regular, new plant, transfer BMM, expansion. All of these additional factors are used to modify advanced manufacturing hours depending on whether the program is regular, new plant is required, there has to be transfer of the blow molding machine or the plant needs additional expansion in order to accommodate the program needs. Team experience is something that is usually not very well known at the beginning of the program, but can be used as a correction factor if needed. Development duration is retrieved by using the command button and it is calculated based on the time data entered for the development start and the SOP.

5.2.2 Technical tab

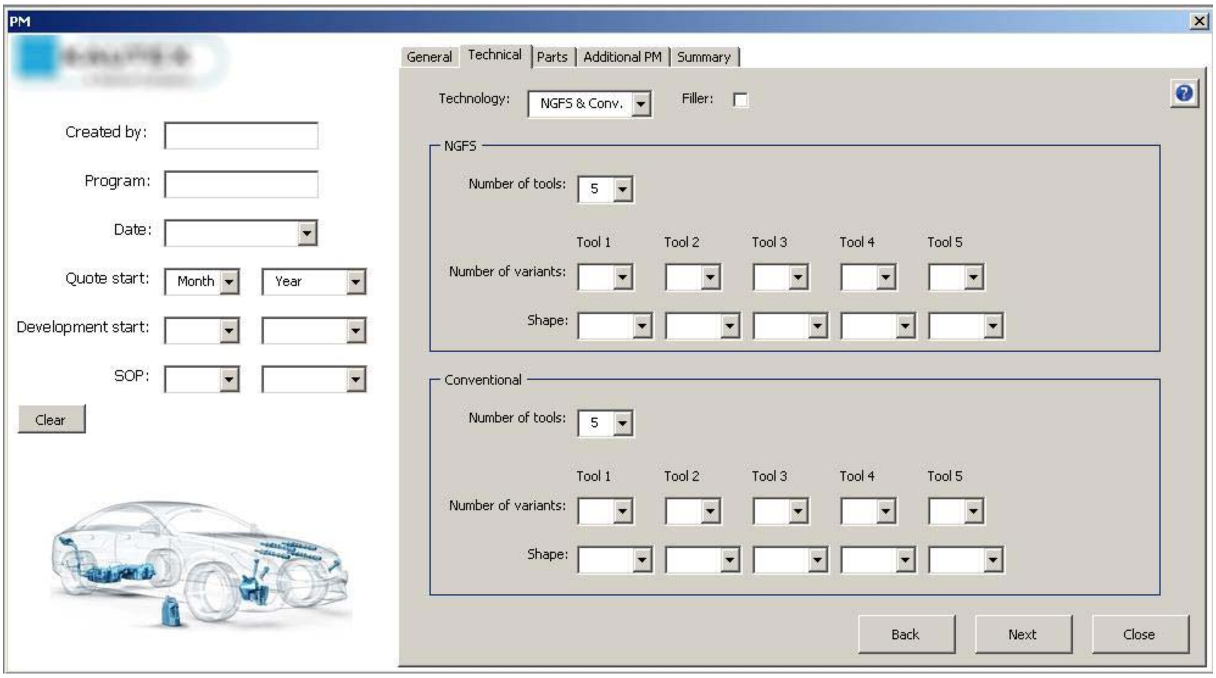


Figure 36 The tool interface - Technical tab

The next tab for factor input is *Technical.* Its purpose is to gather the information on manufacturing technology needed for the program. Based on chosen technology the form shows NGFS, conventional or both frames. Next step is to choose number of tools that enables matching number of dropdown menus where it is possible to choose number of variants per tool and to assign a specific shape to those variants. As shown in the previous chapter, right next to the duration of the program, these factors are the most influential cost drivers. Maximum number of tools allowed is five per technology and maximum number of variants if five per tool. Offered shapes are: single chamber, saddle, L-shape and PHEV. Filler presence in a program can be confirmed with a checkbox and that mostly affects PE hours by increasing the workload with the multiplying factor derived from available existing estimation tools.

5.2.3 Parts tab

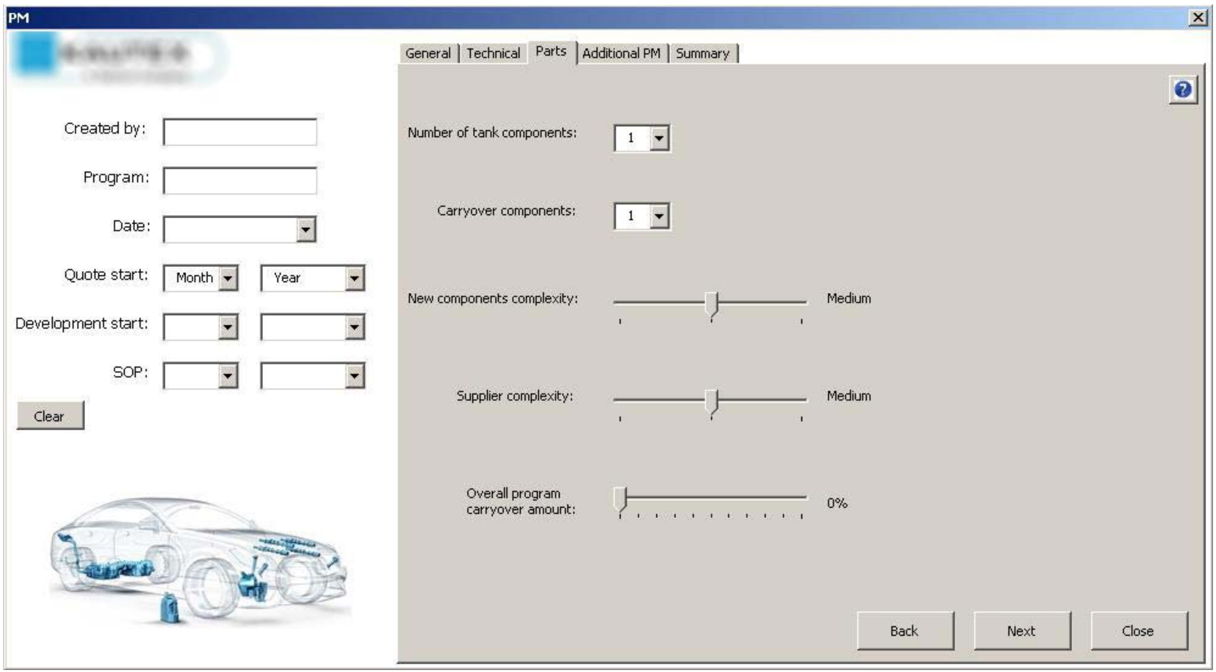


Figure 37 The tool interface - Parts tab

In the *Parts* tab estimator has to choose from the dropdown menus the number of overall tank components as well as the number of carryover components. Carryover components represent how many components, within the chosen overall quantity, were used in previous programs. This tab is mostly used for the additional expert adjustment as described in the previous chapter. Here the estimator can affect the 30% of the costs with the multiplying factors such as new components complexity and supplier complexity. Overall program carryover amount offers choice of percentage from 0 to 100%. It is used as a correction factor for PE, CAD and CAE hours respectively. If the program carryover amount is high, that certainly lowers the need for the working hours compared to the program with no carryover at all. Some of the cases for having the high carryover, among many, could be usage of previously made tool or the existing product modification according to new customer requirements. With carryover percentage being higher, engineering working hours are getting lower.

5.2.4 Additional PM tab

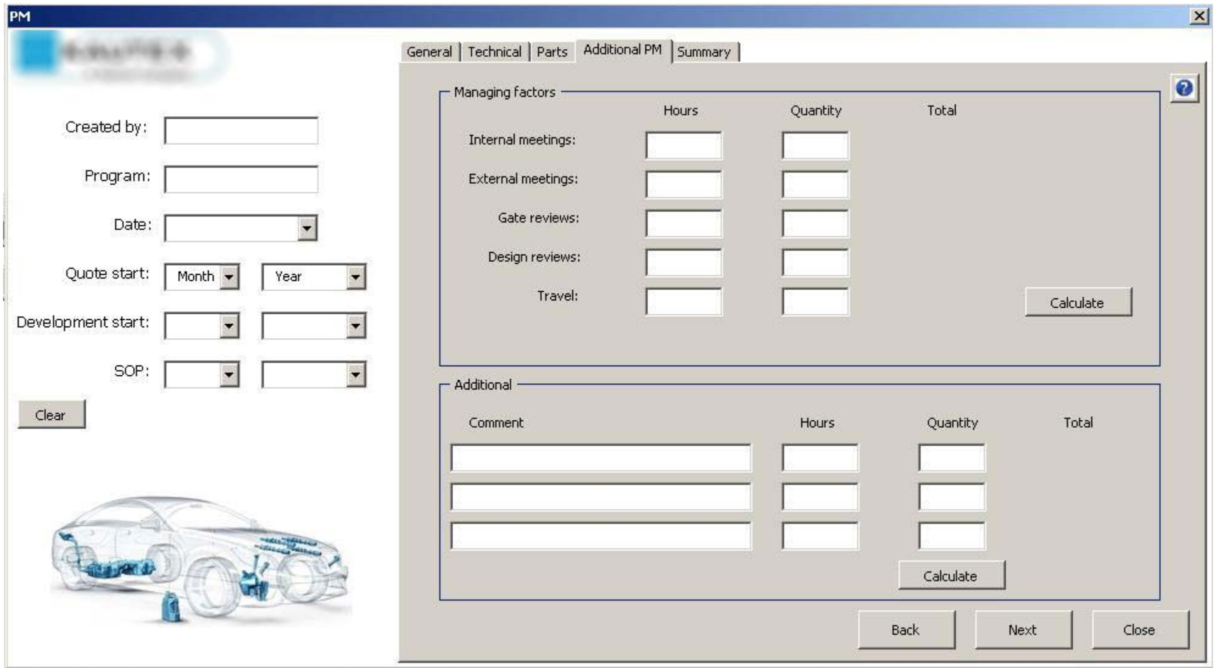


Figure 38 The tool interface - Additional PM tab

Additional PM tab serves the purpose of creating the additional inputs for PM department hours. PM's role in managing the program development certainly includes all of the managing factors stated in the tool thus their effect on working hours is already estimated within the equation. However, these additional hours enable PM to deal with specific programs that by personal opinion require more attention in some of these fields. This way it is possible to create an estimation for programs that are not generic and for programs that require more dedication in order to achieve desired outcome. *Additional* frame allows for totally customized input of factors that are considered to increase the PM workload. By using the calculate command button user gets an overview of additional hours quantity per every factor used.

5.2.5 Summary tab

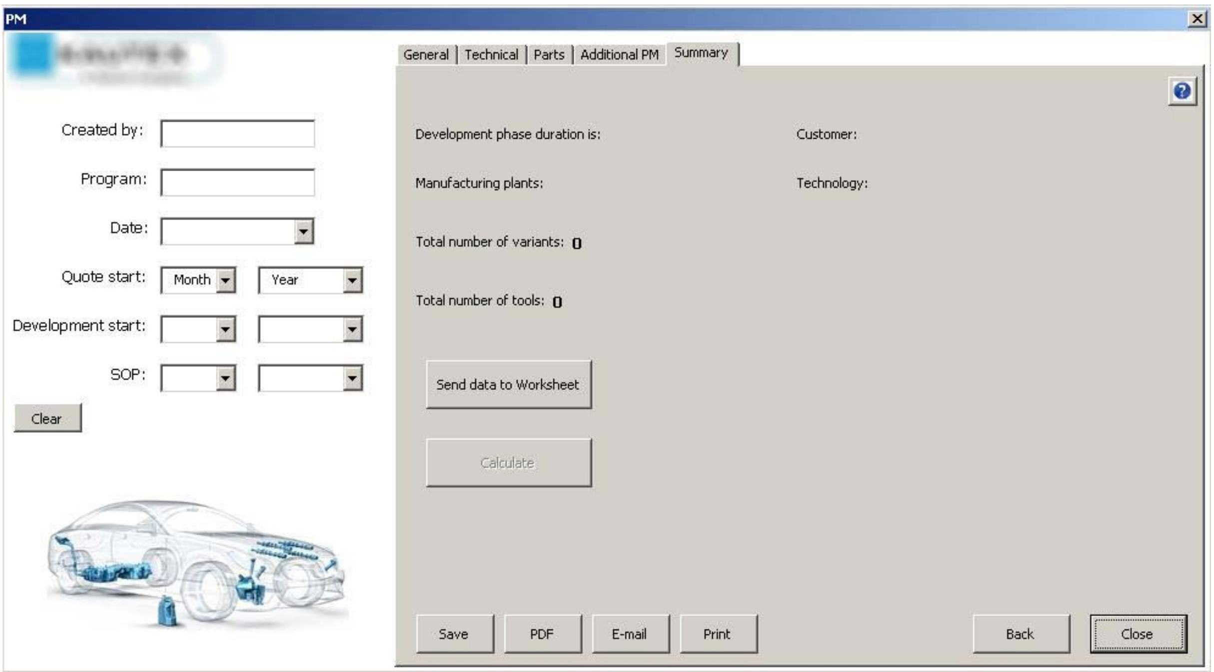


Figure 39 The tool interface - Summary tab

Summary tab provides an overview of the most important data entered into the user form. After completing the input procedure, estimator has to *Send data to Worksheet* which enables the *Calculate* command button. Upon choosing the calculate button the estimations are shown in the center of the user form. Those estimations are at the same time stored, as all other entered variables, into the worksheet. Command buttons in the lower part of the tab allows to the estimator to save the file, to save it as PDF with creation date automatically added to the name, to e-mail the file which creates new e-mail template with the file attached and to print the sheet which opens the print dialogue. User can return to every tab to change the inputs whenever needed. After changing the input, procedure is the same, first it is needed to send data to worksheet and then to calculate the values.

Every tab has the help button in upper right corner. Help brings the pop-up window that has the information on the factors that are on that particular tab and it shows the way how that factors are used to calculate the data. Help is created to provide additional information to the user and to enable better insight in how the tool works. An example of a help pop-up is given in a figure below.

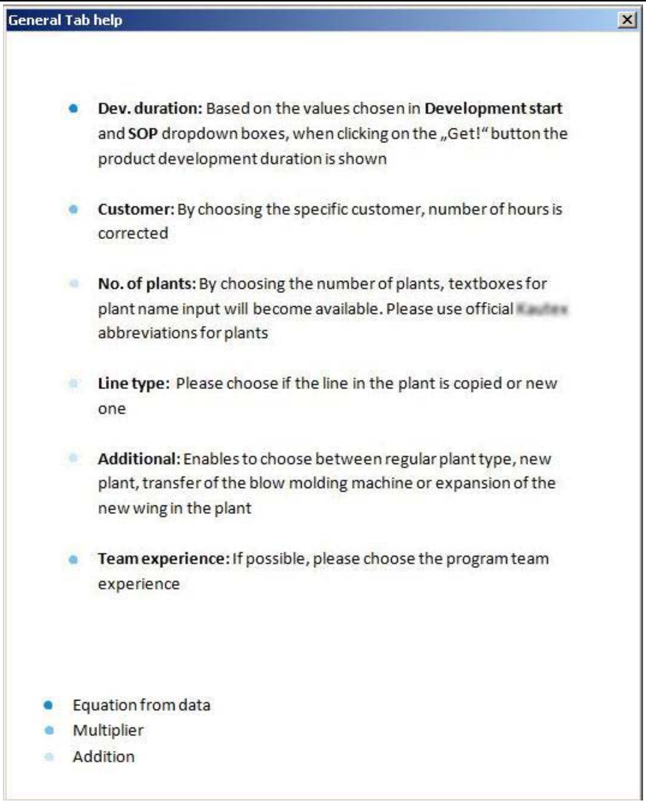


Figure 40 The tool interface - General tab help

There are three ways the factors are used:

- Equation from data
- Multiplier
- Addition

Having all these operations allows the tool to be a hybrid in estimation procedure as described in the previous chapter. In the table below it is shown how certain factors are used in the estimation procedure. The tool essentially uses regression methods to create the base estimation which can be influenced to the certain extent with the expert opinion. There are also fuzzy methods that can be used to create the base estimation. Their applicability on product development cost estimation is examined in the next chapter.

Table 8 Connection between IF's and estimation tool

Equation	Multiplier	Addition
Duration	Customer	#Plants
Technology	Team exp.	Line type
Carryover amount	Filler	Additional
	#Components	Additional PM
	New comp. Complexity	
	Supplier complexity	

5.2.6 Further development of the tool

The tool created for estimation relies on the data from 10 previous programs. Accuracy of the tool is limited by the number of programs introduced in its base equations. Understandably, by introducing more history data the tool would be able to predict more accurately and with greater scope. However, it would be too complex to go through the data mining procedure every time when one wants to improve the tool with new dataset. Base data upon which the equations rely is relatively simple. All that is needed is 4 most important variables from the program and the accurate cost tracking through the development. ED&D chart already has information on all 4 variables, it only lacks the ability to track the data. Although, there were some custom made tracking sheets within ED&D charts in the analyzed datasets from the PLM. The proposal would be to create the simple tracking sheet within the ED&D, which would enable PM to enter real development cost (without additional cost) for every month or quarter of the program. There is no need for additional databases when ED&D is stored in the PLM thus available to every team member. This way it would be easier to include the program data to the tool immediately upon the program development phase completion. There would be no need in identifying the changes cost and the other additional financial rotations when collecting the data. This way the tool would be constantly improved by introduction of the new data. Also, having the program closely tracked enables PM to draw conclusions on the phenomena occurring during the development and it could also act as a lessons learned platform.

6 Fuzzy forecasting methods

As described in chapter 3, there are numerous cost estimation methods with fuzzy methods being one of them. Fuzzy methods in product development cost estimation are relatively new and arguably unexplored. Most of the studies related to the use of fuzzy methods for cost estimation are oriented towards software cost estimation. There are several fuzzy methods that could be adapted for use in the cost estimation: artificial neural networks (ANNs), case based reasoning, genetic algorithms, rule indication and various clustering techniques [32]. However, most of the available literature is oriented towards use of the ANNs. One of the reasons being that ANNs are relatively easy to use without the need for deep theoretical understanding of the matter. ANNs are generally considered to be prospective in the cost estimation field because they can be used to model non-linear behavior with no apparent connection between the cost drivers and the outcome. They have the ability to learn from past incomplete and hazardous data as well as to generalize solutions for future practices. Also, ANNs can be easily modified for various types of projects and used for validation of parametric regression models since the same cost drivers can be used in both simultaneously.

6.1 Artificial neural networks

Artificial neural networks are distributed information processing systems composed of many simple computational elements interacting across weighted connections. ANNs can identify and learn correlated patterns emerging between input datasets and corresponding target values. After training is performed, ANNs can be used to predict the outcome of new independent input data. ANNs seek to simulate the human brain structure, human thinking and human learning in a machine. Because of that, they are ideally suited for the modeling of complex and non linear datasets. An ANN consists of many single processors, which interact through a dense web of interconnections. A neuron or processing element has primarily two tasks. First, it computes output sent to another neuron which determines its output value by applying a transfer function. Second, it updates the local memory, i.e. weights and other data types called data variables. These processing elements are organized into two separate layers. The first layer is known as the input layer and the last layer is referred to as the output layer. The one or more inner layers are known as hidden layers. The input neurons receive values from outside the neural networks' environment, whereas the output neurons send the values to this outside environment [33][34].

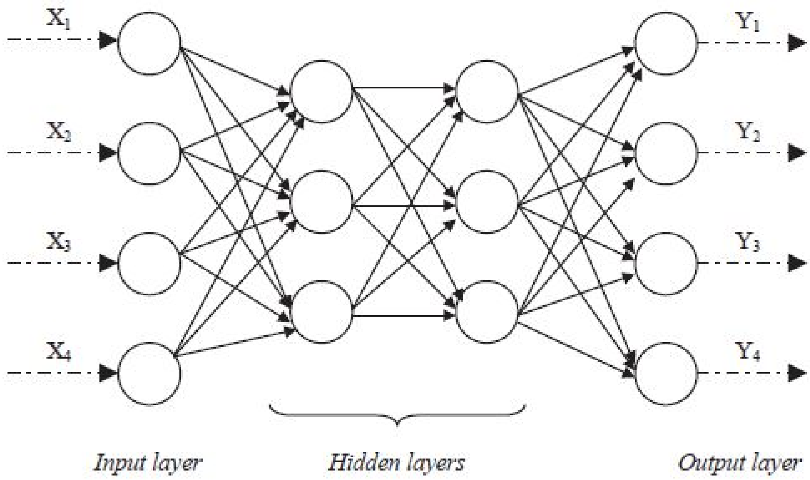


Figure 41 ANN architecture with two hidden layers [34]

6.2 Project engineering estimation with ANNs

Influence factors correlation matrix represents a solid base for application of parametric estimation methods. The same matrix used for the tool was also used to set up the ANN. Many specialized software solutions are available for modeling the fuzzy methods. For this purpose used software was Matlab and its Neural Network toolbox. There are three main operations important for valid ANN creation:

- Training
- Validation
- Testing

When considering cost estimation, the different input signals are weighed and combined into a final cost model. By feeding program data with a known cost to the network, the network is trained to estimate the cost. Network training implies that the different interconnection weights will be adapted every time new data is fed to the network. Adaptation of the weights is done based on a simple punishment/reward principle. If the interconnection performed as expected during estimation of the previous dataset, this variable is rewarded by increasing the weight of its interconnection to the output. If the interconnection performed badly, the weights are decreased in the next iteration step. By minimizing the squared error between the estimate and the desired output (in this case the cost), the network is trained and as more factors are fed, the learning effect increases. However, one should pay attention on when to stop the network training. When new factors, not included in the training set, are fed to the network, inaccurate estimates can be generated. Undesirable effect would be over-fitting the network, what would result in failed generalization upon feeding the new different data. One way to detect when to stop the training is when the error of testing set starts to increase, that means over-fitting is usually avoided [29].

ANN architecture is typically characterized by the number of hidden layers, the number of neurons and the learning algorithm. While there are many possibilities in ANN layer creation, for the purpose of this parametric estimation only one hidden layer was chosen. It was proven that regardless of the activation function and the dimension of the input space, a neural network with one hidden layer can approximate any continuous function arbitrarily well [29]. There are several activation functions used in ANN layers. For the purpose of parametric estimation ANN was used as a function fitting tool. Since statistically significant linearity of data was already positively evaluated in regression analysis, using the linear transfer function in the output layer was a logical choice (Fig. 40). Hidden layer used the tan-sigmoid transfer function that is useful for pattern recognition (Fig. 39).

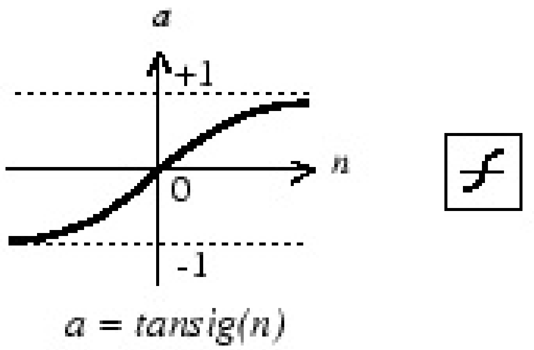


Figure 42 Tan - sigmoid transfer function [36]

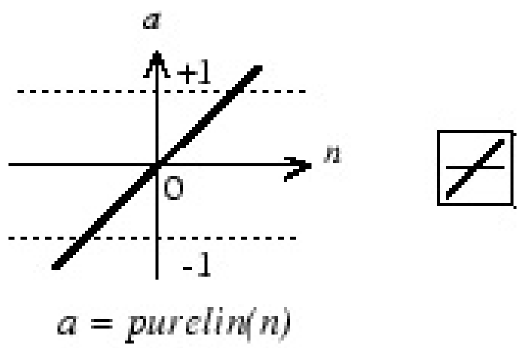


Figure 43 Linear transfer function [36]

Matlab offers three different training algorithms by the default. Training algorithm chosen for this particular case was Bayesian Regularization. This algorithm typically takes more time, but can result in good generalization for difficult, small or noisy datasets such as the one available. Training stops according to adaptive weight minimization (regularization). Number of neurons was chosen by trial and error and the best results were obtained with the 10 neurons in the hidden layer. Resulting neural network architecture is shown in a figure below.

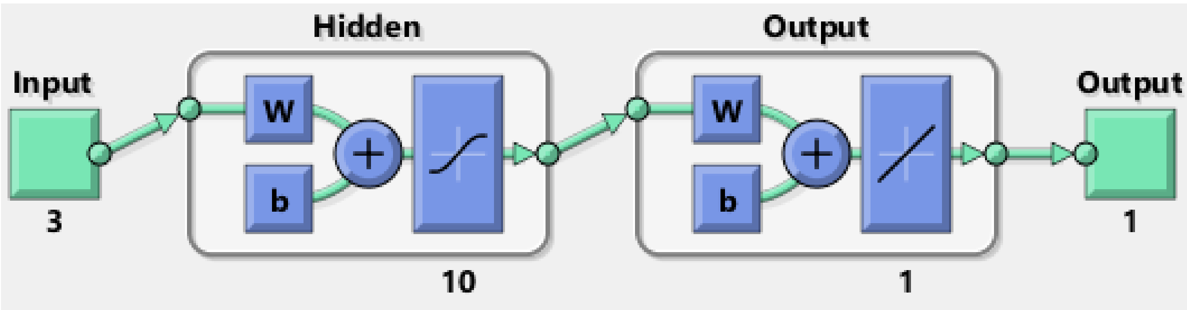


Figure 44 ANN architecture

Training of the ANN was performed by using the same inputs as for the regression in the estimating tool. Input variables were influence factors and target dataset were working hours. Training was performed while slight increase in error was noticed. *R* value shows very high statistical viability of the data that was fed to the network.

net | Input

3x10 double

	1	2	3	4	5	6	7	8	9	10
1	23	22	32	26	26	35	28	27	31	37
2	1	1	1	1	1	1	1	2	2	3
3	2	3	2	2	3	2	3	6	4	7

Figure 45 Input dataset

net | Input | Target

1x10 double

	1	2	3	4	5	6	7	8	9	10
1	3700	4050	5300	4300	5100	5789	5500	8700	8403	14077

Figure 46 Target dataset

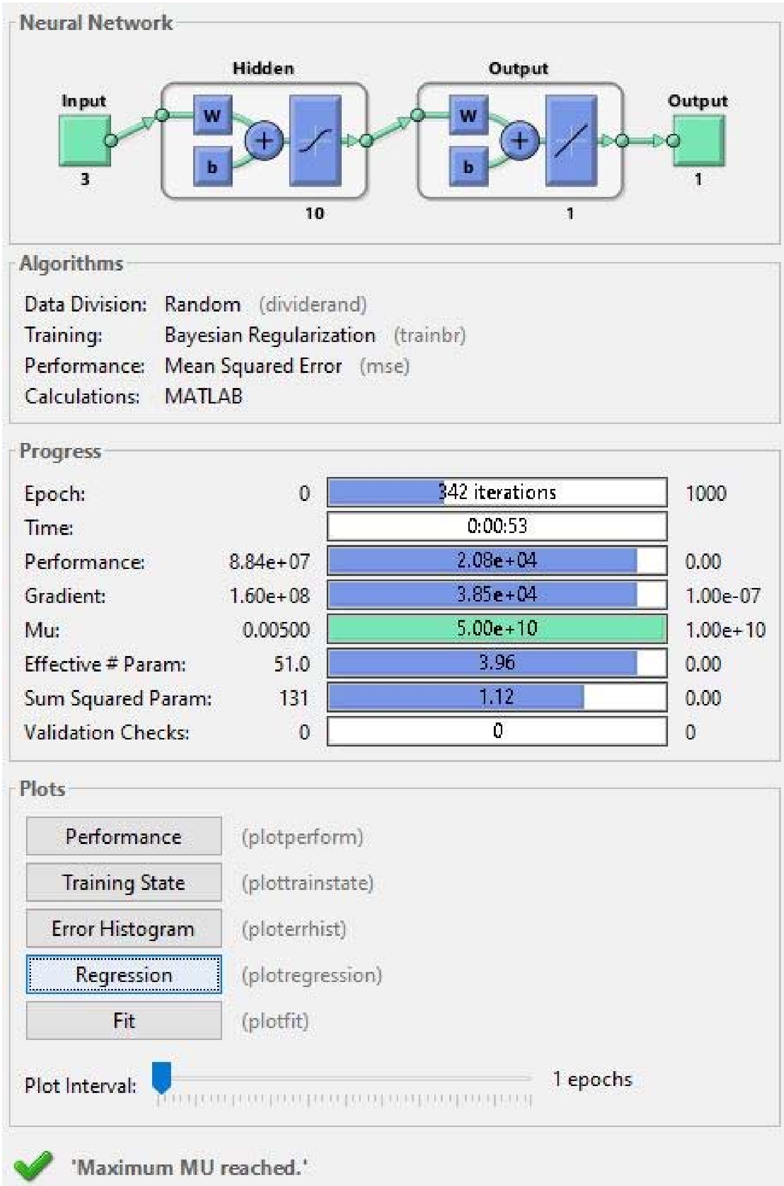
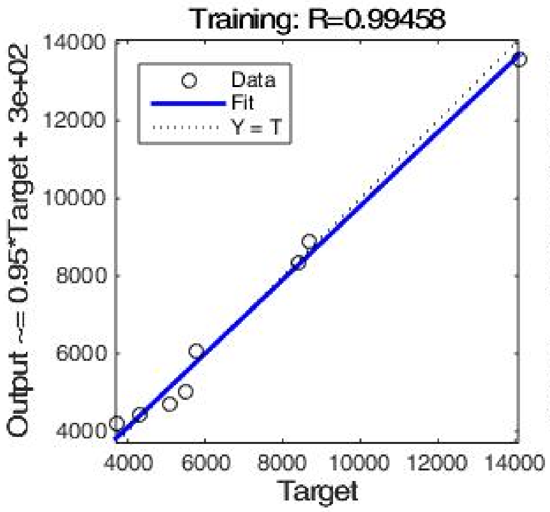


Figure 47 ANN training interface



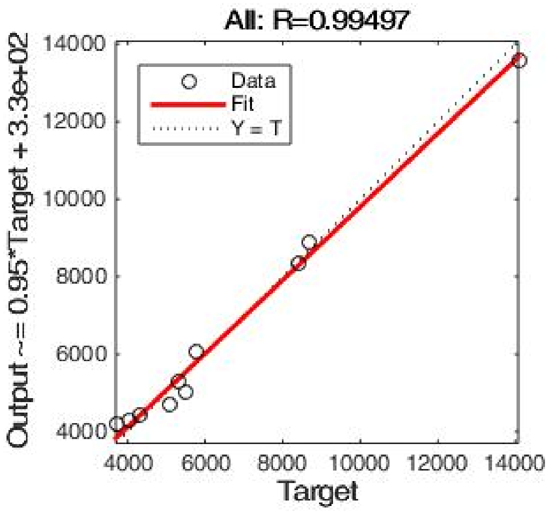


Figure 48 Neural network training performance

6.3 Comparison between ANN and the estimation tool

After the ANN was successfully trained, estimation simulations could be performed. Simulations were performed only for project engineering category (PM+PE), because of the data reliability. There were three real programs data values available for simulation. The data was fed into the network through the input layer as the sample variable. From the output layer it was expected to receive estimations based on the trained network. The same data was also used in regression equations that are used in the tool. For evaluating and comparing regression techniques and neural networks, measure of performance was the absolute deviation from the real cost (Table 9).

Table 9 Absolute deviation from the real cost

Absolute error %	
ANN	Regression
3,45%	3,06%
7,29%	5,79%
-11,68%	-8,23%

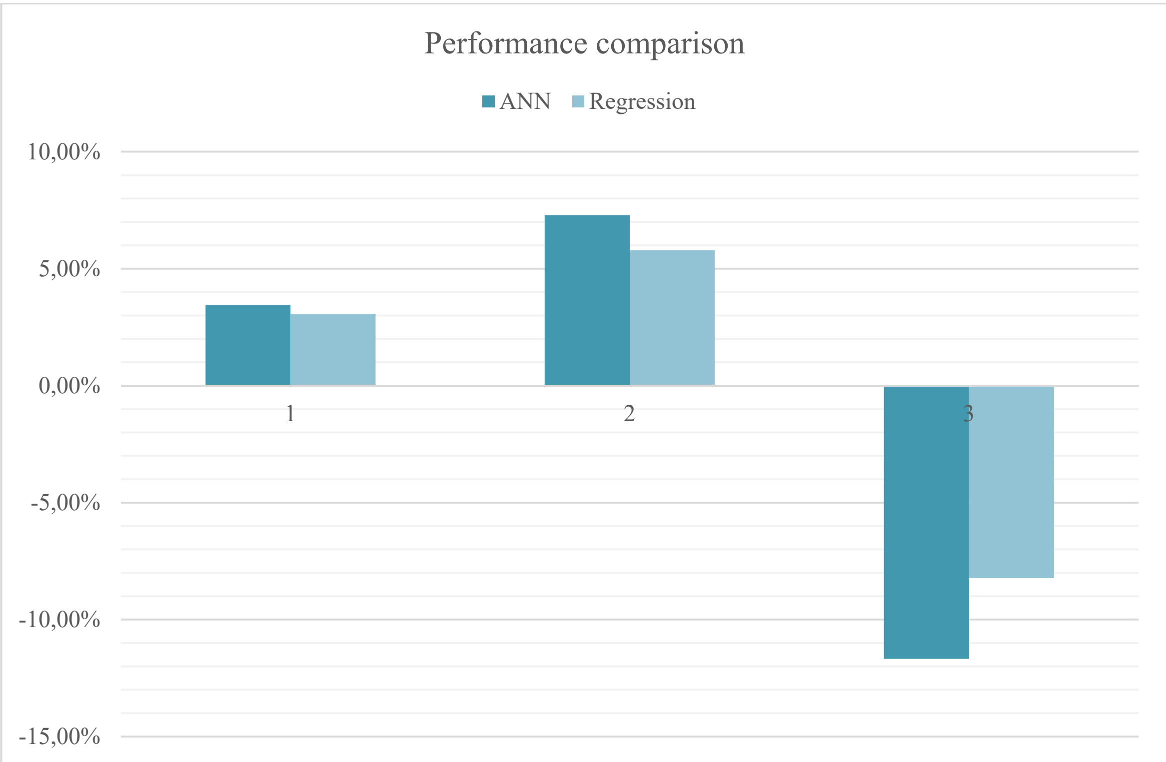


Figure 49 Cost estimation comparison ANN vs. Regression

From the figure above it is visible that the parametric regression model is more efficient in these particular cases, but not significantly. The nature of tested data was that the programs total working hours were getting higher from program 1 to program 3. Data that was used to create ANN and the regression equation holds more programs with lower amount of working hours. That is the reason why the error is getting higher, there is a lack of data to describe more complex and bigger programs. ANN error gets higher when there is lower amount of specific data that could be used in training, in this case for more demanding programs. ANN creates higher error because it is not suitable for data extrapolation, however if there was more data to be fed, accuracy would consequentially be increased.

ANN has proven to be efficient even when training is performed with small amount of data. The reason for that is the factors are linearly dependent and that doesn't require high level of network adjustments. If the factors were to be non-linearly dependent on the cost, ANN performance would be significantly lower. Over the time the list of influence factors could be expanded and reintroduced to the network architecture and modified to accommodate possible deviations from the linear dataset. True power of ANNs lies in interpretation of large databases with factors that are not emerged in patterns when used as input. Besides for cost estimation, it is a diverse and powerful tool that can be used for risk forecasting and various project activities planning. However for the purpose of the cost estimation, everyday use in a dynamic company environment would require a lot of resources to create ANN that could be used uniformly, without having to tweak the properties prior to every estimation. If the influence factors are, like in the observed case, linearly dependent, than usage of ANN would be uncalled for since regression methods are more than sufficient to describe product development outcome.

7 Conclusion

Initial analysis of cost estimation methods has shown great differences between similar programs thus confirming the hypothesis that some form of tool based on history data should be introduced into the estimation process. As the fuel tank production is very specific and requires detailed product properties in early stages of the program, it was possible to choose the parametric estimation method. Parametric method relies heavily on the quality of history data. 13 past programs were chosen for analysis and after data mining and thorough examination, 3 of them were discarded either because of the missing data or because of the low quality of data. In order for parametric method to be viable, main cost drivers were identified and the influence factor matrix was created. After data normalization and analysis, it was possible to perform statistical analysis and to identify the factors with the biggest influence on the program cost. Those factors are number of tools, number of variants, duration and manufacturing technology. Correlation matrix confirms statistical significance of the factors thus enabling the multiple linear regression to be performed. With data being normalized, fuzzy methods could later also take place for the estimation evaluation. After performing regression analysis and retrieving equations that describe program cost, the tool development could take place. Tool vas developed with Microsoft Excel and VBA and was created to enable influence of an expert on the estimated data. Expert influence is minimized but still allowing for intervention if needed. The main goal for the tool is to provide support for program management department when program enters the quotation phase. Apart from using the regression techniques, possibility for introduction of fuzzy methods was also examined. Using the same dataset as for the tool creation, ANN was set up and evaluated. In performance comparison between the regression equation and the ANN, there were small differences in favor of regression methods. Thus proving that ANNs could be used for valid cost estimation even with small datasets, if the factors are linearly dependent. However, introducing fuzzy methods instead of multiple regression tools would require much more effort in order to create the tool that could be used on everyday basis. Although, introducing more complex algorithms to the system where influence factors are completely linearly dependent on program cost, should be discarded because of the unjustified increase in the complexity of the tool. If the larger database of the projects would prove to have some non-linear dependencies among influence factors and program cost, then fuzzy methods would be certainly an interesting choice and with their use beneficial conclusions could be made.

8 LITERATURE

[1] Di Bitonto, S., Trost, R.: The Automotive Industry in Germany, Germany Trade and invest, 2015

[2] Global automotive supplier study, Roland Berger Strategy Consultants / Lazard, 2014

[3] Montequin, VR., Nieto, AG.: Managerial style profiles of successful project managers: a survey, Procedia computer science Vol. 64, 2015, pp. 55-62

[4] Miterev, M., Engwall, M.: Exploring program management competences for various program types, International Journal of Project Management, 2015

[5] Project Management Institute: A guide to the project management body of knowledge - Fifth edition, PMI Inc., 2013

[6] Project Management Institute: The standard for program management, PMI Inc., 2006

[7] Mislick, G. K., Nussbaum, D.: Cost estimation methods and tools, John Wiley & Sons, 2015

[8] Pinto, M. B., Pinto J. K.: Project team communication and cross functional cooperation in new program development, Journal or product innovation, 1990

[9] Zulch, B.: Leadership communication in project management, Social and Behavioral Sciences, 2014, pp 172-181

[10] Project communication handbook - Second edition, Office of project management process improvement, 2007

[11] Kerzner, H.,: Project management: A systems approach to planning, scheduling and controlling 7th ed., John Wiley & Sons, 2001

[12] Eppinger, S., Salminen, V.: Patterns of Product Development Interactions, Proc. International Conference On Engineering Design, ICED01, Glasgow, 2001, pp. 283-290

[13] Sharon, A., Dori, D.: A Model-Based Approach for Planning Work Breakdown Structures of Complex Systems Projects, Proc. of the 14th IFAC Symposium on Information Control Problems in Manufacturing, Bucharest, 2012

[14] Sequeira, S., Lopes, E.: Simple Method Proposal for Cost Estimation from Work Breakdown Structure, International Conference on Project MANagement, Procedia Computer Science 64, 2005, pp. 537 – 544

[15] Hollman, J.K.,: Total cost management framework – 1st Edition, AACE International, 2006

[16] Niazi, A., Dai, J.S.: Product Cost Estimation: Technique Classification and Methodology Review, Journal of manufacturing science and engineering, 2006

[17] Elfaki, A.O., Alatawi, S., Abushandi, E.: Using Intelligent Techniques in Construction Project Cost Estimation: 10-Year Survey, Advances in Civil engineering, 2014

[18] Zhang, Y. F., Fuh, J. Y. H., Chan, W. T.: Feature-Based Cost Estimation for Packaging Products Using Neural Networks, Computers in Industry Vol. 32, 1996, pp. 95–113.

[19] Ben-Arieh, D., Qian, L.: Activity-Based Cost Management for Design and Development Stage, International Journal of Prod. Econ. Vol. 83, 2003, pp. 169–183

[20] Shehab, E. M., Abdalla, H. S.: Manufacturing Cost Modeling for Concurrent Product Development, Rob. Comput.-Integr. Manufact. Vol 17, 2001, pp. 341–353

[21] Parviz, F.R.: Project estimating and cost management, Management Concepts Inc., Virginia, 2002

[22] Sibanda, K.: A review of generic cost estimation techniques for information technology projects, International Journal of IT and Business Management Vol. 18, 2013

[23] Nasa Cost Estimating Handbook V4.0, National Aeronautics and Space Administration, Washington D.C., 2015

[24] Florjanic, B., Kuzman, K.: Estimation of Time for Manufacturing of Injection Moulds Using Artificial Neural Networks-based Model, University of Ljubljana, Faculty of Mechanical Engineering, Ljubljana, 2012

[25] Parametric Estimating Handbook - Fourth Edition, International Society of Parametric Analysts, Vienna, 2008

[26] Dysert, L.R.: An introduction to Parametric Estimating, AACE International transactions, 2008

[27] Amos, S.: Skills and knowledge of cost engineering - Fifth edition, AACE International, 2007

[28] Hand, D., Heikki, M., Padhraic, S.: Principles of data mining, The MIT Press, Massachusets, 2001

[29] Verlinden, B., Duflou, J.R.: Cost estimation for sheet metal parts using multiple regression and artificial neural networks, Production Economics Vol. 111, 2008, pp. 484-492

[30] Caputo, A.C., Pacifico, M.P.: Parametric and neural methods for cost estimation of process vessels, Production Economics Vol. 112, 2008, pp. 934-954

[31] Makridakis, S.G., Wheelwright, S.C., Hyndman, R.J.: Forecasting: methods and applications - 3rd edition, John Wiley & Sons, 1998

[32] Grewal, A.S., Gupta, V.: Cost Estimation with Neural Network and Affect of Different Input parameters on Cost Estimation, International Journal of Emerging Technologies in Computational and Applied Sciences, 2013, pp. 312-316

[33] Sharma, A., Chopra, A.: Artificial Neural Networks: Applications In Management, Journal of Business and Management Vol. 12, 2013, pp. 32-40

[34] Cavalieri, S., Maccarrone, P., Pinto, R.: Parametric vs. neural network models for the estimation of production costs: A case study in the automotive industry, International Journal of Production Economics Vol. 91, 2003, pp. 165-177

[35] Elsawy, I., Hosny, H., Razek, M.A.: A Neural Network Model for Construction Projects Overhead Cost Estimating, International Journal of Computer Science Issues Vol. 8, 2011

[36] Matlab help documentation, MathWorks

APPENDIX

I. CD-R disc