UNIVERSITY OF NOVI SAD
FACULTY OF TECHNICAL SCIENCES IN NOVI SAD

Marija Milojević Jevrić

THE APPLICATION OF META-HEURISTICS TO OPTIMISE LOAD DISTRIBUTION IN MACHINE ELEMENTS AND ASSEMBLIES

DOCTORAL THESIS

Novi Sad, 2015

УНИВЕРЗИТЕТ У НОВОМ САДУ ● **ФАКУЛТЕТ ТЕХНИЧКИХ НАУКА**
21000 НОВИ САД, Трг Доситеја Обрадовића 6

КЉУЧНА ДОКУМЕНТАЦИЈСКА ИНФОРМАЦИЈА

Редни број, **РБР**:	
Идентификациони број, **ИБР**:	
Тип документације, **ТД**:	Монографска документација
Тип записа, **ТЗ**:	Текстуални штампани материјал
Врста рада, **ВР**:	Докторска дисертација
Аутор, **АУ**:	Марија Милојевић Јеврић
Ментор, **МН**:	Др Татјана Давидовић, Виши научни сарадник
Наслов рада, **НР**:	Примена метахеуристика на оптимизацију расподеле оптерећења код машинских елемената и склопова
Језик публикације, **ЈП**:	Енглески
Језик извода, **ЈИ**:	Срп. / Енг.
Земља публиковања, **ЗП**:	Србија
Уже географско подручје, **УГП**:	Војводина
Година, **ГО**:	2015
Издавач, **ИЗ**:	ауторски репринт
Место и адреса, **МА**:	Трг Доситеја Обрадовића 6, Нови Сад
Физички опис рада, **ФО**: (поглавља/страна/ цитата/табела/слика/графика/прилога)	(8 / xxviii+135 / 148 / 23 / 33 / 0 / 0)
Научна област, **НО**:	Математика, Информатика
Научна дисциплина, **НД**:	Оптимизација
Предметна одредница/Кључне речи, **ПО**:	Континуална оптимизација; Метахеуристике, Расподела оптерећења , Радни век, Лежаји, Зупчаници
УДК	
Чува се, **ЧУ**:	Библиотека факултета техничких наука у Новом Саду
Важна напомена, **ВН**:	
Извод, **ИЗ**:	Докторска дисертација се бави темама везаним за хеуристичке, метахеуристичке и предикционе методе. Извршена је метахеуристичка оптимизација проблема трансверзалног фактора расподеле оптерећења код цилиндричног зупчастог пара. Вишекритеријумски проблем геометрије планетарног преносника је решаван применом генетских алгоритама. Поређење неколико метахеуристичких метода је дато на решавању оптимизационог проблема радног века и динамичке носивости код котрљајних кугличних лежаја. Бајесове мреже, као предикциона метода, су уграђене у софтвер за подршку одлучивању у комплексним машинским постројењима.
Датум прихватања теме, **ДП**:	17. 04. 2014.
Датум одбране, **ДО**:	

Чланови комисије, **КО**:	Председник:	др Наташа Сладоје Матић, ванредни професор	
	Члан:	др Љиљана Теофанов, ванредни професор	
	Члан:	др Синиша Кузмановић, редовни професор	Потпис ментора
	Члан:	др Драган Урошевић, научни саветник	
	Члан	др Милан Дражић, ванредни професор	
	Члан, ментор:	др Татјана Давидовић, виши научни сарадник	

UNIVERSITY OF NOVI SAD ● **FACULTY OF TECHNICAL SCIENCES** 21000 NOVI SAD, Trg Dositeja Obradovića 6	

KEY WORDS DOCUMENTATION

Accession number, **ANO**:	
Identification number, **INO**:	
Document type, **DT**:	Monograph documentation
Type of record, **TR**:	Textual printed material
Contents code, **CC**:	Ph. D. thesis
Author, **AU**:	Marija Milojević Jevrić
Mentor, **MN**:	Tatjana Davidović, Ph.D., Researh Assosiate
Title, **TI**:	The application of meta-heuristics to optimise load distribution in machine elements and assemblies
Language of text, **LT**:	English
Language of abstract, **LA**:	eng. / srp..
Country of publication, **CP**:	The Republic of Serbia
Locality of publication, **LP**:	A. P. Vojvodina
Publication year, **PY**:	2015
Publisher, **PB**:	Author's reprint
Publication place, **PP**:	Trg Dositeja Obradovića 6, Novi Sad
Physical description, **PD**: (chapters/pages/ref./tables/pictures/graphs/appendixes)	(8 / xxviii+135 / 148 / 23 / 33 / 0 / 0)
Scientific field, **SF**:	Mathematics; Computer Science
Scientific discipline, **SD**:	Optimisation
Subject/Key words, **S/KW**:	Continual optimisation; Meta-heuristic; Load distribution; Working life; Bearings; Gears;
UC	
Holding data, **HD**:	Library of the Faculty of Technical Sciences in Novi Sad
Note, **N**:	
Abstract, **AB**:	In this Phd thesis, heuristic, meta-heuristic and predictive methods are considered. Meta-heuristics optimisation of the transverse load distribution factor of helical and spur gears is conducted. The multi-objective optimisation problem of the planetary gear train is done using Genetic Algorithm method. A comparative study of several meta-heuristic methods is given for solving the problem of dynamic load capacity and working life at radial ball bearings. Bayesian network, as a predictive method, is implemented in the software to support the decision making in complex mechanical plants.
Accepted by the Scientific Board on, **ASB**:	17. 04. 2014.
Defended on, **DE**:	

Defended Board, **DB**:	President:	Nataša Sladoje Matić, Ph.D., Associate Professor	
	Member:	Ljiljana Teofanov, Ph.D., Associate Professor	
	Member:	Siniša Kuzmanović, Ph.D., Full Professor	Menthor's sign
	Member:	Dragan Urošević, Ph.D., Research Professor	
	Member	Milan Dražić, Ph.D., Associate Professor	
	Member, Mentor:	Tatjana Davidović, Ph.D., Research Associate Professor	

Acknowledgments

First and foremost, I would like to express my gratitude and great appreciation to my supervisor, Dr. Tatjana Davidović, research associate professor, for the cooperation, patience, methodical approach and tactfulness, as well as for her enormous technical help and support which made me persevere in my scientific research. Dr Davidović's flexibility, genuine caring and concern, and faith in me during the dissertation process enabled me to attend to life while also earning my PhD. She has been motivating, encouraging, and enlightening. Her guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better supervisor for my PhD study. I will forever be thankful to my supervisor and she remains my best role model for a scientist and supervisor.

I also express my appreciation to the extinguished members of the committee:
Prof. Dr. Nataša Sladoje, the committee president
Prof. Dr. Ljiljana Teofanov
Prof. Dr. Siniša Kuzmanović
Prof. Dr. Milan Dražić and
Prof. Dr. Dragan Urošević,
for accepting to be part of the team and invest time in reading of my thesis. I would like to express them my gratitude for all their very useful comments and reviews. I feel honored to have my work examined by them and I would like to thank them for spending a day as opponents in my dissertation defence.

Special appreciation to professor Milan Dražić for his exhaustive and useful advice and guidance, and also to my supervisor Dr. Davidović for her patience and professional conduct through all the editing work.

I would like to thank to professor Siniša Kuzmanović, for his expertise and meticulousness, to professor Nataša Sladoje Matić for her great ideas and for encouraging me in highlighting my results. My gratitude also goes to professor Ljiljana Teofanov for her methodicalness and precision and to professor Dragan Urošević for his great expert comments and emphasising. Thank you all for helping me improving my PhD thesis.

I owe my gratitude to all the members of staff, all the professors at MI SASA for their support and trust. I am grateful to you for making me feel a part of your team.

I appreciate all the financial, technical and professional help and support received from my Home institution, Mathematical Institute of SASA (Serbian Academy of Science and Arts), III04406 project and the manager Dr Zoran Ognjanović.

Exeptional appreciation to Dr Silvia Ghilezan, for all the help and support and cofidence invested in me during my doctoral studies.

I am especially grateful to my parents, Novak and Dragana Milojević, and to the extended family members for their immense love and constant encouragement. In their honour, from the bottom of my heart, I dedicate this doctoral dissertation.

And finally, my deep gratitude goes to people who didn't want to be named, but without whose support this thesis would have not been possible, they know who they are!

Novi Sad, 2015.

R E Z I M E

Analiza dostupne literature u vezi sa projektovanjem mašinskih elemenata i sklopova ukazala je na pojavu mnogih optimizacionih problema, koji su najčešće kontinualnog tipa. Poznati optimizacioni problemi u mašinstvu su optimizacija u projektovanju kotrljajnih ležaja, optimizacija rotorskog sistema, optimizacija raspodele opterećenja, optimizacija menjača, optimizacija geometrije vratila, optimizacija dinamičke nosivosti i mnogi, mnogi drugi.

Da bi rešili ove probleme istraživači su primenjivali metodu konačnih elemenata, razne algebarske transformacije, Bajesove i veštačke neuronske mreže, optimizaciju rojem čestica, genetske i neke druge hibridne evolutivne algoritme. Modeli koji se javljaju u dostupnoj litaraturi uglavnom uzimaju u obzir jednu grupu uticajnih parametara na raspodelu opterećenja, dok svim ostalim parametrima dodele fiksne vrednosti radi pojednostavljenja. Izbor parametara koji će biti fiksirani moguće je izvršiti na razne načine pri čemu je bitno da se time ne gubi na opštosti modela.

Na primer, metode za proračun cilindričnih zupčanika sa pravim ili kosim zupcima za svrhe preliminarnog dizajna ili za standardizaciju se uglavnom svode na korišćenje linearne teorije elastičnosti (Navijeova jednačina) ili korišćenje Hercovih kontaktnih modela. Pored toga, kontaktni pritisak se, prilikom korišćenja metode konačnih elemenata, često izjednačava sa srednjim Hercovim pritiskom. Korišćenjem takvih i sličnih aproksimacija dobijaju se modeli čija simulacija može malo da odstupa od realnog stanja. Potrebno je simulirati raspodelu opterećenja u što realnijem obliku sa svim uticajnim parametrima što može biti vremenski i/ili memorijski zahtevno. Zbog toga su se aproksimativne metode, kao veštačke neuronske mreže ili genetski algoritmi, pokazale prikladnijima od klasičnih algebarskih metoda. Aproksimativne metode omogućavaju dobijanje rezultata zadovoljavajuće tačnosti za znatno kraće vreme rada.

U ovom radu su opisane neke optimizacione metode koje su primenjivane u rešavanju raznih teških optimizacionih problema kod mašinskih elemenata i sklopova. Pre svega su prikazane egzaktne optimizacione metode, posebno one zasnovane na Njutnovom algoritmu i njegovim modifikacijama. Aproksimativne metode nalaze rešenje koje je blizu tačnom rešenju, dok se simulacione metode koriste da bi se izvršila statička i dinamička analiza različitih tipova objekata. Prediktivne metode kao ulaz koriste empirijske podatke u poznatim uslovima rada sistema, a na izlazu daju predikciju ponašanja sistema za nepoznate uslove rada. U tu svrhu, u ovom istraživanju su korišćene Bajesove mreže. Heurističke metode su efikasne zato što koriste neka *a priori* znanja o problemu koji se rešava i na taj način generišu prilično kvalitetna rešenja. Iz tog razloga se kod metaheuristika koriste za generisanje početnih rešenja. Metaheurističke metode su veoma efikasne u traženju rešenja optimizacionih problema koje je blizu optimalnom za znatno kraće vreme rada i uz manju upotrebu procesorke memorije u odnosu na egzaktne metode. Postoje razne

vrste metaheurističkih metoda, a među najviše korišćenim u mašinstvu su genetski algoritmi kao i hibridi zasnovani na genetskim algoritmima. Genetski algoritmi su zasnovani na populaciji rešenja, koja pokušava da pređe iz jedne generacije slučajno generisanih dopustivih rešenja u drugu generaciju primenom genetskih operatora. Genetski operatori predstavljaju mehanizme pretraživanja u cilju pronalaženja optimalnog hromozoma u jednom regionu. Hibridni genetski algoritmi uključuju i neke dodatne optimizacione metode da bi se efikasnije dobio očekivani rezultat. Za poređenje rezultata dobijenih genetskim algoritmima, u ovom istraživanju su korišćene heurističke metode lokalnog pretraživanja patern search i active set u multistart okruženju.

Uopšteno govoreći, doprinosi u ovom doktorskom radu se sastoje u primeni optimizacionih i predikcionih metoda na rešavanje problema izabranih mašinskih elemenata i sklopova. Samim tim, doprinosi pripadaju dvema oblastima istraživanja, matematika i mašinstvo, što daje disertaciji multidisciplinarni karakter. Doprinosi u oblasti matematike, pre svega, uključuju razvoj efikasnih matematičkih modela razmatranih mašinskih sistema. Tako razvijeni modeli omogućuju jednostavnu primenu metoda optimizacije kao i alata veštačke inteligencije na rešavanje odgovarajućih problema. Razvijeni modeli su omogućili da se, za pojedine mašinske elemente i sklopove, identifikuju parametri koji imaju presudan uticaj na povećanje njihove efikasnosti. Softverska implementacija predloženih metoda predstavlja ključne elemente u razvoju sistema za podršku odlučivanju koji u poslednje vreme postaje nezamenljiv alat inženjerima u procesu projektovanja. Doprinos u mašinstvu je primena ove metodologije na konkretne primere problema raspodele opterećenja kod zupčanika i radnog veka kod ležaja. Postojeći rezultati iz novije literature su poboljšani primenom ove metodologije i pružaju preporuke za buduće projektovanje efikasnijih mašinskih elemenata i sklopova. U nastavku su ukratko opisani primeri razmatrani u ovoj disertaciji i ukazano je na publikacije autora u kojima su objavljeni glavni rezultati.

Jedan od primera razmatranih u ovom radu je optimizacija raspodele opterećenja spregnutog cilindričnog zupčastog para. Istraživanje je bazirano na ISO standardima [ISO 6336-1], [ISO 6336-2], [ISO 6336-3], [ISO 1328], [ISO 53], kao i na postojećim metodama optimizacije [Sánchez 2008], [Schneider 2006], [Szabó 2005]. Korišćene su heurističke, metaheurističke metode uglavnom bazirane na genetskim algoritmima kao i hibridi.

Razmatrani su matematički modeli spregnutih cilindričnih zupčastih parova sa pravim i kosim zupcima utvrđeni u skladu sa ISO stadardima. Oni zavise od nekoliko desetina geometrijsko-konstrukcionih parametara. Izbor najuticajnih predstavlja značajnu istraživačku temu u kojoj metaheuristike mogu mnogo da pomognu. Metaheurističke metode je moguće primeniti na rešavanje optimizacionog problema kada imamo egzaktan matematički model, predstavljen jednačinama ili sistemima jednačina.

Doprinos u mašinstvu uključuje korišćenje razvijenog hibridnog algoritma za rešavanje optimizacionog problema transverzalnog parametra raspodele opterećenja kod cilindičnih zupčastih prenosnika. Hibridizacija je ostvarena obogaćivanjem genetskog algoritma funkcijom lokalnog pretraživanja. Pretpostavljen je idealni slučaj u kome je raspodela opterećenja tokom vremena ravnomerna kako bi se došlo do zaključka koji su to najuticajniji parametri. Korišćenjem ove metode, izveden je zaključak da ugao nagiba zupca kao i koeficijenti pomeranja profila alata najviše utiču na promenu vrednosti paramentra raspodele opterećenja. U okviru simulacije rada matematičkog modela opterećenja spregnutog cilindričnog zupčastog para, menjane su vrednosti uticajnih parametara, da bi se izvela odstupanja vrednosti parametra raspodele opterećenja od pretpostavljenih idealnih vrednosti.

U procesu evaluacije je pokazano da je metoda, predložena u ovom radu, efikasnija od metoda koje su uzete za poređenje prema unapred zadatim kriterijumima. Ovo pripada doprinosu u mašinskom inženjerstvu. Konačni rezultat predložene metode je povećanje kvaliteta raspodele opterećenja spregnutih cilindričnih zupčastih parova.
Rezultati u ovoj tezi su poboljšanja rezultata dobijenih u radovima [Sánchez 2013], [Pedrero 1996], [Pedrero 2011], [Sánchez 2013], [Simon 1988], [Zhang 2010], [Zhang 1999]. Poboljšanja u odnosu na prethodno uspostavljene rezultate se sastoje u sledećem:

- Manje pojednostavljenja nego u modelima [Pedrero 2011], [Pedrero 2010], [Sánchez 2013].

- Raspodela optrećenja je optimizovana korišćenjem genetskih algoritama, za razliku od metoda korišćenih u literaturi. Uticajni parametri su optimizovani tako da je dobijena raspodela opterećenja je približnija ravnomernoj nego u radovima [Simon 1988], [Zhang 2010], [Zhang 1999], što dovodi do njihovog poboljšanja.

Ova poboljšanja pripadaju doprinosima u mašinskom inženjerstvu. Razvijena je MATLAB biblioteka za proračun transverzalnog faktora raspodele opterećenja kod cilindričnih zupčanika koja je javno dostupna a predložena u radu [Milojević 2013]. Dobijeni rezultati su opisani u [Milojević 2013].

Drugi primer primene genetskih algoritama, razmatran u okviru ovog istraživanja, je rešavanje višekriterijumskog nelinearnog optimizacionog problema planetarnih prenosnika. Pri optimizaciji geometrije planetarnog prenosnika, jako je važno uzeti u obzir ograničenja u obliku osnog rastojanja između centralnog zupčanika i planetarnog zupčanika u cilju dobijanja što većeg stepena iskorišćenja. U toku simulacije rada predloženog modela, uzete su u obzir promene vrednosti nekih uticajnih parametara u odnosu na preporučene u fazi projektovanja. Promenom parametara potvrđuje se osetljivost stepena iskorišćenja što je i prikazano na odgovarajućem grafiku zavisnosti (Pareto skup tačaka). Rezultati mogu biti vrlo korisni inženjerima koji se bave projektovanjem planetarnih prenosnika sa raznim prenosnim odnosima. Dobijeni rezultati pripadaju doprinosima u mašinskom inženjerstvu i oni su publikovani u [Rosić 2011a, Rosić 2011b] i jasno ukazuju na važnost formulisanja problema u obliku višekriterijumske optimizacije.

Genetski algoritmi su veoma pogodna tehnika u situacijama kada je potrebno rešiti problem kontinualne optimizacije. Međutim, uvek je dobro imati i uporedne rezultate dobijene nekim drugim metaheurističkim metodama. Kako su, za problem optimizacije dinamičke nosivosti kod kugličnih ležaja najčešće primenjivani genetski algoritmi, u ovom istraživanju je urađena komparacija rezultata dobijenih genetskih algoritmom i rezultata dobijenih drugim metodama. U tom smislu, treći primer razmatran u ovoj tezi je razvijanje nove metode za optimizaciju dinamičke nosivosti i radnog veka kod kotrljajnih kugličnih ležaja u funkciji 10 različitih parametara, korišćenjem tri metaheurističke metode. Razvijena metoda pripada doprinosu u mašinskom inženjerstvu. U tu svrhu, za rešavanje višekriterijumskog optimizacionog problema dinamičke nosivosti i radnog veka kugličnog ležaja, pored genetskih algoritama su korišćene još dve metaheurističke metode (patern search i active set). Dobijeni rezultati su pokazali da patern search i active set postižu bolje rezultate od genetskog algoritma i to za kraće vreme izvršavanja što je dalo smisao i opravdalo upotrebu novih metoda. Optimizacija primenom patern search i active set daje veći dinamički kapacitet i duži radni vek za isti matematički model i to za kraće vreme sopstvenog izršavanja, i uz korišćenje manje memorijskog prostora. U dosadašnjim istraživanjima patern search i active set nisu primenjivani na posmatrani problem, što predstavlja još jedan doprinos ove doktorske disertacije u matematici. Na osnovu dobijenih rezultata, zaključeno je da povećanje broja kuglica i pogodna promena unutrašnje geometrije ležaja dovodi do blagog povećanja dinamičkog kapaciteta i radnog veka ležaja u odnosu na standardne kataloške vrednosti. Rad [Milojević 2014] sadrži kratak opis dobijenih rezultata i predstavlja doprinos u mašinskom inženjerstvu. Dobijeni rezultati u ovom istraživanju predstavljaju poboljšanja u odnosu na rezultate prikazane u sledećim radovima [Chakraborty 2003], [Costin 2010], [Gupta 2007], [Mendi 2010a], [Rao 2007], [Waghole 2014], [Shigley 1989].

U slučajevima kada matematički model nije poznat ili je suviše komplikovan da bi se mogao predstaviti korišćenjem sistema jednačina, mogu se primeniti neke druge tehnike kao što su Bajesove ili Neuralne mreže. Kao ilustracija, Bajesove mreže su primenjene na dva mašinska sistema u cilju prevencije preopterećenja i potencijalnih otkaza [Milojević 2012], [Glišović 2013].

U prvom primeru, Bajesove mreže su primenjene na procenu pouzdanosti mašinskog postrojenja za farbanje metalnih poluproizvoda. Predstavljena je funkcionalnost mašinskog postrojenja i razvijena je metoda za predikciju kvaliteta proizvoda na kraju procesa ukoliko dođe do otkaza nekog od podsistema. Ispitano je nekoliko različitih scenarija: verovatnoća otkaza nekog podsistema ako je poznato stanje kvaliteta proizvoda, verovatnoća stanja kvaliteta proizvoda ukoliko je došlo do otkaza jednog podsistema ili nekoliko podsistema istovremeno. Urađena je komparativna analiza između iskustvenih (istorijskih) verovatnoća i verovatnoća dobijenih korišćenjem Bajesovih mreža. Dobijeni rezultati (predstavljeni i u [Milojević 2012]), koji su doprinos u mašinskom inženjerstvu, pokazuju da se te dve verovatnoće međusobno razlikuju za manje od 3% što potvrđuje da Bajesove mreže dobro modeliraju realno stanje sistema. Mera kvaliteta predikcije Bajesovim mrežama se ogleda u razlici između istorijskih verovatnoća i verovatnoća dobijenih modelovanjem. Ukoliko je ta razlika manja od 3%, kao što je slučaj u radu [Milojević 2012], smatra se da je predikcija veoma zadovoljavajuća [Darwiche 2009].

Drugi primer primene Bajesovih mreža odnosi se na mašinsko postrojenje za filtraciju transformatorskog ulja. Razvoj Bajesove mreže je uključio tri koraka: identifikaciju važnih promenljivih zajedno sa njihovim sopstvenim verovatnoćama; veze između promenljivih i njihove prikaze u grafičkoj strukturi; procenu pouzdanosti rezultujućih verovatnoća dobijenih korišćenjem Bajesovih mreža. Rezultujuća mreža je ugrađena u softver za podršku odlučivanju tokom proizvodnih procesa u situacijama kada postoji nedostatak informacija o stanju sistema i/ili proizvoda što je opisano u radu [Glišović 2013]. Ovaj rezultat je praktičan doprinos u mašinskom inženjerstvu.

U modernim poslovnim okruženjima kada su investitori izloženi čestim i stalnim promenama, teorije pouzdanosti i održavanja dobijaju značajnu ulogu u upravljanju poslovnim procesima. Bajesovi modeli mogu efikasno da se primene na probleme predikcije i prevencije i na taj način utiču na smanjenje troškova održavanja. Studije sprovedene u [Milojević 2012], [Glišović 2013] pokazuju visoku pouzdanost Bajesovih modela kroz simulacione eksperimente.

Razmatrani primeri primene metaheurističkih i predikcionih metoda na nekoliko reprezentativnih optimizacionih problema u mašinstvu su potvrdili njihovu efikasnost. Samim tim, otvorena je mogućnost primene tih i drugih metoda na slične probleme, tj. predložena je nova metodologija za rešavanje teških optimizacionih problema u mašinstvu.

U mašinstvu postoji niz problema koji zahtevaju primenu predikcije ili optimizacije velikog broja konstrukcionih parametara u cilju omogućavanja donošenja temeljne odluke. Višekriterijumska optimizacija je takođe često potrebna u mašinstvu usled pojave suprotstavljenih kriterijuma cene koštanja materijala i kvaliteta materijala, ili veličine mašinskog dela i njegove dinamičke izdržljivosti itd. Ova doktorska teza se upravo bavi integracijom matematičkih modela pojedinih mašinskih elemenata i metoda veštačke inteligencije što otvara vrata za buduća istraživanja u ovoj sferi. Proračuni pri konstrukciji mašinskih elemenata i sklopova u skladu sa ISO standardima zahtevaju dosta vremena čak i u slučaju kada su svi podaci unapred poznati. Ukoliko je potrebno vršiti dalja istraživanja u cilju poboljšanja rada mašinskih elemenata, klasični proračuni nisu efikasni. Radovi [Milojević 2013], [Rosić 2011a] i [Rosić 2011b] se bave modelovanjem funkcionalnosti mašinskog sklopa matematičkim modelom na koji se zatim primenjuje genetski algoritam. Cilj optimizacije je poboljšanje konstukcionih parametara radi ostvarivanja ravnomerne raspodele opterećenja, odnosno visokog stepena iskorišćenja. U svakom slučaju, ova doktorska disertacija predstavlja sintezu matematike, mašinstva i kontrole kvaliteta industrijskog proizvoda. U ovoj disertaciji, naglasak je bilo na kvalitetu prozivoda (bez obzira na povećanje cene). Međutim, isti principi važe i u situacijama kada je primarna minimalna cena koštanja proizvoda, ali zahtevaju modifikacije matematičkih modela u skladu sa tim zahtevima. U radu [Milojević 2014], predstavljena je optimizacija dinamičke nosivosti kugličnih kotrljajnih ležajeva primenom tri različite metaheurističke metode. Postignuti rezultati ovog istraživanja su doveli do poboljšanja od čak 30.3% u odnosu na dosadašnji najbolji poznat rezultat u literaturi, što opravdava dalja istraživanja primene metaheurističkih metoda u mašinstvu. Upotreba predikcije u mašinskoj industriji je u današnje vreme sve više potrebna radi donošenja brzih i ispravnih odluka, kao i sprečavanja mogućeg gubljenja novca. Radovi [Milojević 2012] i [Glišović 2013] se bave povezivanjem predikcionih metoda sa postojećim mašinskim postrojenjima u fabrikama. Pristup zastupljen u navedenim radovima se pojavljuje po prvi put u literaturi i pravi konekciju između iskustvenih podataka iz mašinskog fabričkog postrojenja i predikcione mreže. Kako su dobijeni predikcioni rezultati izuzetno precizni na testiranom skupu podataka, očekuje se da oni daju adekvatne odgovore i u nepoznatim okolnostima. Time se potvrđuje značaj upotrebe ovakvih i sličnih matematičkih aparata u mašinskoj industriji.

A B S T R A C T

An analysis of the available literature related to the design of mechanical elements and assemblies indicated the existence of many optimisation problems, mostly continuous type. Known optimisation problems in the mechanical engineering include optimisation in design of rolling bearings, rotor system optimisation, load distribution, gears optimisation, shaft geometry optimisation, dynamic load optimisation, and many others.

To derive these problems researchers have applied finite elements methods, various algebraic transformation, Bayesian and Artificial Neural Network (ANN), Particle Swarm Optimisation, Genetic and some other hybrid Evolutionary Algorithms. Models that are appearing in available literature usually take into consideration one group of influential parameters for load distribution, while all other parameters are set to the fixed values, for simplification. The selection of parameters to be fixed should enable good presentation of model without loos of generality.

For example, calculation methods for helical and spur cylindrical gears for purposes of preliminary design or for standardization are usually reduced to linear elasticity theory (Navier equations) or to implementation of Hertz contact models. Furthermore, by using the finite element models, contact pressure is often taken to be equal to an average Hertz pressure. Such and similar approximations can lead to models which simulation may not agree with experimental results. It is necessary to simulate load distribution in the realistic form with all influential parameters, which can be time and / or memory demanding. Because of that, approximate methods, like ANN or Genetic Algorithms (GA), are shown to be more appropriate than classical algebraic methods. Approximation methods allow getting results of good quality within considerably shorter execution time.

In this paper, some optimisation methods that are applicable for solving hard optimisation problems of machine elements and assemblies, are described. First of all, exact optimisation methods are explained, especially those based on Newton's algorithm and its modifications. Approximation methods find solution that is close to exact one, while simulation methods are used in order to perform static and dynamic analysis of different type of objects. Prediction methods use empirical data as an input in known condition of a working system and, as an output, they give prediction of the system's behavior in case of unknown working conditions. To this purpose, in this research Bayesian networks are used. Heuristic methods are efficient because they use some *a priori* knowledge about the problem they are solving, and based on that knowledge they generate high quality solutions. Because of that, they are used in meta-heuristics to generate initial solutions. Comparing to exact methods, meta-heuristic methods are very efficient to find solutions for optimisation problems that are close to optimal for shorter working time and using lees memory

than the exact methods. There are many types of meta-heuristic methods, however the most frequently used in mechanical engineering are GAs as well as hybrid algorithms based on GAs. GAs are based on randomly generated population of solutions and they try o enhance the solutions from one generation into solutions of the next generation by applying genetic's operators. The operators represent searching mechanisms in order to find optimal chromosome in one region. Hybrid GAs include some additional optimisation methods in order to reach expected result more efficiently. In order to evaluate results obtained by GAs, some other local search based heuristic methods, such as pattern search and active set within multi-start framework, are used in this thesis.

Generally speaking, contributions in this thesis are addressed to application of optimisation and predictive methods to solve problems related to mechanical elements and assemblies. Hence, achieved contributions belong to two fields of research i.e. mathematics and mechanical engineering, which gives multidisciplinary dimension to this thesis. First of all, contributions in mathematics include the development of effective mathematical models for the considered mechanical elements. The models developed in such a way are enabling easy application of optimisation methods, as well as artificial intelligence tools to solve the considered problems. The developed models have allowed to identify some parameters, for particular mechanical elements and assemblies that have crucial impact to increasing their effectiveness. Software implementation of the proposed methods is the key element for the development of decision support systems that lately exist as permanent tool used by engineers in the design process. Contribution in mechanical engineering is application this methodology to the concrete examples of distribution problem at gears and life of a bearing. Existing, more recent, results are improved by applying this methodology and offer recommendations for future modeling more efficient mechanical elements and assemblies. In the following, we briefly describe the considered examples in this thesis and refer to author's publications where main results are published.

The first considered example in this thesis is optimisation of load distribution problem at helical and spur gears. Research is based on ISO standards [ISO 6336-1], [ISO 6336-2], [ISO 6336-3], [ISO 1328], [ISO 53], as well as on the existing optimisation methods [Sánchez 2008], [Schneider 2006], [Szabó 2005]. Heuristic and meta-heuristic methods mostly based on GAs are used, as well as hybrid algorithms.

Mathematical models of helical and spur cylindrical gears are considered according to ISO standards. These models are depending on a few dozens of geometrical and design parameters. Selection of the most influential parameters represents an important research challenge where meta-heuristic methods may help a lot. Meta-heuristic methods are usually applicable to optimisation problems represented by the known exact mathematical model, expressed by equations or system of equations.

Contribution in mechanical engineering includes the usage of developed hybrid algorithm for solving optimisation problem of transversal load distribution factor for helical and spur gears. Hybridisation is achieved with enhancement of GA with a local search function. In order to conclude which parameters are the most influential, the ideal case is assumed, which has uniform load distribution during the considered period of time. By using this method, it turned out that helix angle and profile shift coefficients of the pinion and the wheel mostly affect the changing of transversal load distribution factor. Within simulation of mathematical model for load distribution at helical and spur gear pair the influential factors are changed over the time, in order to derive the deviations of the obtained values for the distributional load from the assumed ideal values.

It is shown, in the process of evaluation, that method proposed in this thesis is more efficient than methods taken for comparison, according to some predefined criteria. This belongs to contribution in mechanical engineering. Final result of the proposed method is increasing the quality of load distribution at helical and spur cylindrical gear pairs. These results are presenting improvements of the results obtained in papers [Sánchez 2013], [Pedrero 1996], [Pedrero 2011], [Sánchez 2013], [Simon 1988], [Zhang 2010], [Zhang 1999]. There are the following improvements of previous results:

- Less simplification than in models [Pedrero 2011], [Pedrero 2010], [Sánchez 2013].
- Load distribution is optimised by using GA instead of other methods available in the literature. Influential parameters are optimised and obtained load distribution is more uniform than the load distribution obtained in [Simon 1988], [Zhang 2010], [Zhang 1999], which improves these results.

These improvements belong to contributions in mechanical engineering. Publicly available MatLab library for computation of transversal load distribution factor optimisation at helical and spur gears, based on proposed framework [Milojević 2013] is implemented. The obtained results are described in [Milojević 2013].

The second example of GAs application, conducted within this research, is solving the multi-objective nonlinear optimisation problem of planetary gear trains. To perform the geometry optimisation of planetary gear trains, it is very important to take into consideration the limitations in the form of an axial distance between the sun and planetary gear in purpose of obtaining higher degree of efficiency. During the simulation of the proposed working model, the changing of the most influential parameters, with respect to the ones derived in the design phase, are taken in consideration. The changing of the project parameters reveals the sensitivity of the degree of efficiency as it is shown on the dependance graphics (Pareto set of points). The results are very useful for the designers dealing with planetary gear trains with various transmission ratio. Obtained results belong to contribution in mechanical engineering and they are published in [Rosić 2011a], [Rosić 2011b] and clearly indicates the importance of problem formulation in the form of multi-criteria optimisation.

GAs are very suitable technique in cases when it is necessary to solve continuous optimisation problems. However, it is always good to have comparative results obtained by some other meta-heuristic methods. For the dynamic load optimisation problem, in the recent literature GAs are among the most deployed methods for solving maximisation problem of dynamical load at ball bearings. To examine the

efficiency of GA, In this research a comparison of the results obtained by GA and other methods is conducted. In that sense, third example, considered in this thesis, is related to the development of a new method for optimisation of dynamical load ratings and rating life at radial ball bearings, as a function of 10 different parameters by using three meta-heuristic methods. The developed method belongs to contribution in mechanical engineering. For that purpose, beside GA, two other meta-heuristic methods (patterns search and active set) are used for solving multi-criteria optimisation problem of dynamic load capacity and service life of ball bearings. Obtained results showed that pattern search and active set are giving better results than GA for shorter computing time which justify employment of the two additional methods. Optimisation by pattern search and active set is providing better results with regards to dynamic capacity and working life for shorter computing time and less memory consuming. From the-state-of-the-art, pattern search and active set are not applied to the observed problem, which gives another contribution of this PhD thesis in mathematics. Based on the obtained results, it appears that increase in the number of balls and adequate improvement in the terms of internal geometry lead to slight increase in the dynamic capacity and bearing life within comparison to standard catalog values. The paper [Milojević 2014] contains short description of obtained results and represents contribution to mechanical engineering. The obtained results are improving the following results [Chakraborty 2003], [Costin 2010], [Gupta 2007], [Mendi 2010a], [Rao 2007], [Waghole 2014], [Shigley 1989].

In the cases when mathematical model is not known or it is too complicated to be expressed by using a system of equations, other techniques like Bayesian or Neural Networks can be applied. As an illustration, Bayesian Networks (BN) are applied to two mechanical systems in order to prevent the overload or potential failures [Milojević 2012], [Glišović 2013].

As the first example, BNs are applied to reliability assessment of the mechanical assembly for painting of the metal semi-products. A functionality of the mechanical assembly is presented and the method for prediction of product quality at the end of the process in the case of a failure of one of the subsystems is developed. Several different scenarios are tested: probability of system failure if product quality state is known, probability of the product quality conditions if one subsystem fails and/or several subsystems fail simultaneously. Comparative analysis between experienced (historical) probabilities and probabilities obtained by using BNs is performed. The obtained results (presented in [Milojević 2012]), that are the contribution to mechanical engineering, show that these two probabilities differ in less than 3%, which confirms accurate modelling of realistic system conditions. A measure of quality of BNs prediction is reflected in the difference between the historical probabilities and the probabilities obtained by modeling. If the ratio is less than 3%, which is proofed in [Milojević 2012], it is considered that the prediction is very satisfactory [Darwiche 2009].

The second example of the Bayesian network application refers to machine assembly for the transformer oil filtration. The development of the Bayesian Network was performed within three steps: identification of important variables together with their state probability; relations between variables and their representation in graphic structure; assessment of resulting probabilities obtained by using Bayesian Networks. Resulting network is incorporated into decision support system software during manufacturing process in cases of missing information about state of a system and/or products, as it is explained in [Glišović 2013]. This result is practical contribution to mechanical engineering.

In modern business environments, when investors are exposed to frequent and constant changes, theory of reliability and maintenance has gain a new role in business processes management. Bayesian models can be effectively applied to solve problems of the prediction and prevention in order to reduce the costs of maintenance. Case studies in [Milojević 2012], [Glišović 2013] show high reliability of Bayesian models through simulation experiments.

The considered examples of meta-heuristic and prediction methods application to several typical optimisation problems in mechanical engineering confirmed their effectiveness. Therefore, there is an open possibility to apply these and other methods to similar problems. In other words, a new methodology for solving hard optimisation problems in manufacturing engineering is proposed.

Mechanical engineering has huge number of construction tasks that require deployment of prediction or optimisation techniques to enable easier decision making. Deployment of multi-objective optimisation is often needed in mechanical engineering due to appearing of conflicting criteria, such as price and quality ratio or mechanical element size and dynamical capacity ratio, etc. In this thesis, the integration between mathematical models of mechanical elements and artificial intelligence methods is done which opens a new research area. The calculations in mechanical elements and assemblies with respect to ISO standards require plenty of constructors time even if all the data is known in advance. If the further improvements are needed, classical calculations are not very efficient. The work presented in [Milojević 2013], [Rosić 2011a] and [Rosić 2011b] is referring to modeling functionality of mechanical assembly with mathematical models, which present basis for the application of GA. Optimisation goal in this research is improvement of the design parameters for load distribution or efficiency. In any case, this doctoral thesis is the synthesis of mathematics, mechanical engineering and quality control of the industrial products. The emphasis in this thesis was on the quality of the product (regardless the potential increase in the cost). However, the same principles can be applied in the situations when the primary objective is the minimal cost of the manufacturing, but the mathematical models would require certain adjustments in accordance with these requirements. A work presented in [Milojević 2014] is dealing with the optimisation of the dynamical capacity of rolling bearings with 3 different meta-heuristics. Achieved results yield more than 30.3% improvement with respect to the current best known result in the relevant literature, which justifies further research about possible deployment of meta-heuristics in mechanical engineering. The prediction in mechanical industry is needed nowadays for fast decision making and preventing unnecessary costs due to damages. Research results presented in [Milojević 2012] and [Glišović 2013] are dealing with the connection of predictive methods and fault detection of mechanical systems in the existing factories. The approach proposed in the mentioned works presents the connection between the historical data from factories and prediction networks. As achieved prediction results are shown to be highly precise, it is expected that the network will give accurate responses even in the unknown conditions. Following this, a deployment of these and similar mathematical apparatus in mechanical industry is completely justify.

N O M E N C L A T U R E

The used labels, abbreviations, and names of variables are listed here. The measurement units (when applicable) are given in square brackets.

a - centre distance of gears [mm]
α - radial contact angle of bearings [°]
α_n - normal pressure angle of gears [°]
α_P - pressure angle of the basic rack for cylindrical gears [°]
α_{Pn} - normal pressure angle of the basic rack for cylindrical gears [°]
α_t - transverse pressure angle of gears [°]
α_w - operating (working) pressure angle of gears [°]
α_{wt} - pressure angle at the pitch cylinder [°]
b - face width of gears [mm]
b_s - central web thickness [mm]
β - an unknown constant, it is the factor which determines the upper bound of the rolling element diameter of bearings; helix angle of gears [°]
b_m - rating factor for contemporary, commonly used, high quality hardener bearing steel in accordance with good manufacturing practices, the value of which varies with bearing type and design
c_1 - tip clearance of pinion [mm]
c_2 - tip clearance of wheel [mm]
$c_{\gamma\alpha}$ - mean value of mesh stiffness per unit face width (used for K_v, $K_{H\alpha}$, $K_{F\alpha}$) $[\frac{N}{mm\mu m}]$
C - single stiffness $[\frac{N}{mm\mu m}]$
C_b - basic rack factor (same rack for pinion and wheel)
C_d - dynamic load capacity [N]
C_M - correction factor
C_{or} - basic static radial load rating of bearings [N]
C_r - gear blank factor
C_s - static load capacity [N]
C_{th} - theoretical single stiffness $[\frac{N}{mm\mu m}]$
d - bearing bore diameter [mm]
d_1 - reference diameter of pinion [mm]
d_2 - reference diameter of wheel [mm]
$d_{(a)}$ - reference diameter of the sun gear [mm]
d_{a1} - tip diameter of the pinion [mm]
d_{a2} - tip diameter of the wheel [mm]
d_b - base diameter of gears [mm]
d_{b1} - base diameter of the pinion [mm]
d_{b2} - base diameter of the wheel [mm]
d_{f1} - root diameter of the pinion [mm]

d_{f2} - root diameter of the wheel [mm]
d_i - the inner raceway diameter at the grooves of bearings [mm]
d_o - the outer raceway diameter at the grooves of bearings [mm]
D - bearing outer diameter [mm]
D_b - ball diameter of bearings [mm]
D_m - pitch diameter of bearings [mm]
D_{out} - outer diameter of planetary gear train [mm]
e - parameter for mobility conditions of bearings
E - modulus of elasticity, the Young modulus [$\frac{N}{mm^2}$]
ε - parameter for outer ring strength consideration of bearings
ε_α - transverse contact ratio of gears
ε_γ - total contact ratio of gears
f_c - factor which depends on the geometry of the bearing components, the accuracy which the various components are made, and the material
f_i - inner raceway curvature coefficient
f_o - outer raceway curvature coefficient
f_{pb} - transverse base pitch deviation (the values of f_{pt} may be used for calculations in accordance with ISO 6336, using tolerances complying with ISO 1328-1) [μm]
F_a - bearing axial load (axial component of actual bearing load) [N]
F_r - bearing radial load (radial component of actual bearing load) [N]
F_t - transverse tangential force at pinion reference circle [N]
F_{tH} - tangential load in a transverse plane for $K_{H\alpha}$ and $K_{F\alpha}$ [N]
h - tooth depth (between tip line and root line) [mm]
h_{a1} - addendum of pinion [mm]
h_{a2} - addendum of wheel [mm]
h_{fP} - dedendum of basic rack of cylindrical gears [mm]
i - number of rows in the bearing
K_A - application factor (gears)
$K_{D_{max}}$ - maximum ball diameter limiter
$K_{D_{min}}$ - minimum ball diameter limiter
K_{F_α} - transverse load factor (root stress of gears)
K_{F_β} - face load factor (tooth-root stress of gears)
K_{H_α} - transverse load factor (contact stress of gears)
K_{H_β} - face load factor of gears
K_V - dynamic factor (gears)
L_{10} - basic rating life, in million revolutions
m - module of gears [mm]; parameter for mobility conditions of bearings
m_n - normal module of gears [mm]
m_t - transverse module of gears [mm]
n_w - number of planet gears
odn - represents the boundary condition based on ratio $\frac{bs}{b}$
p_b - pitch on the base circle of gears [mm]
P_r - dynamic equivalent radial load in [N]
Φ_o - assembly angle in [rad]

Q - accuracy rate
R - effective radius of gears curvature
r_{ct} - hob tip radius of gears
r_i - inner raceway groove curvature radius [mm]
r_o - outer raceway groove curvature radius [mm]
S_F - factor of safety from breakage (gears)
S_H - factor of safety from pitting (gears)
S_r - gear rim thickness [mm]
S_{rn} - normal gear rim thickness [mm]
σ_F - effective tooth root stress of gears $[\frac{N}{mm^2}]$
σ_{FP} - permissible bending stress $[\frac{N}{mm^2}]$
$\sigma_{F_{lim}}$ - nominal stress number (bending) of gears $[\frac{N}{mm^2}]$
σ_{F0} - nominal tooth-root stress of gears $[\frac{N}{mm^2}]$
σ_H - effective contact stress of gears $[\frac{N}{mm^2}]$
σ_{HP} - allowable contact stress $[\frac{N}{mm^2}]$
$\sigma_{H_{lim}}$ - allowable stress number (contact) of gears $[\frac{N}{mm^2}]$
σ_{H0} - nominal contact of gears $[\frac{N}{mm^2}]$
u - gear ratio
v_7 - represent a ratio $\frac{F_t K_a / b}{100}$
W - bearing width [mm]
x_1 - profile shift coefficient of pinion
x_2 - profile shift coefficient of wheel
x_a, x_b, x_g - addendum modification of sun gear, ring gear and planet gear, - respectively X - dynamic radial load factor of bearings
y_a - running-in allowance for a gear pair [μ m]
Y - dynamic axial load factor of bearings
Y_β - helix angle factor of gears
Y_{drelT} - relative notch sensitivity factor of gears
Y_ε - contact ratio factor of gears
Y_F - tooth form factor of gears tooth form factor, for the influence on nominal tooth root stress with load applied at the outer point of single pair tooth contact
Y_{RrelT} - relative surface factor of gears
Y_S - stress correction factor of gears
Y_{ST} - stress correction factor of gears
Y_X - size factor of gears
Z - number of rolling elements in a single-row bearing; or number of rolling elements per row of a multi-row bearing with the same number of rolling elements per row
z_1 - number of teeth of the pinion
z_2 - number of teeth of the wheel
z_a, z_b, z_g - numbers of teeth of sun gear, ring gear and planet gear, respectively
Z_β - helix angle factor of gears
Z_E - elasticity factor of gears $\sqrt{\frac{N}{mm^2}}$

Z_ε - contact ratio factor of gears
Z_h - zone factor of gears
Z_L - lubricating factor of gears
Z_ν - speed factor of gears
Z_R - roughness factor of gears
Z_W - work hardening factor of gears
Z_X - size factor of gears

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This section gives mathematical background for optimisation problems, briefly describes the methods required to deal with hard instances of these problems, and shortly underlines how this mathematical tool helps in solving some practical problems in mechanical engineering. In addition, the main results obtained during the research connected with this work are presented at the end of this section.

1.1 Mathematical background

Intuitively speaking, optimisation problems require finding the best solution among all feasible solutions. One of the most famous optimisation problems is the traveling salesman problem (TSP) [Rothlauf 2011]. For a given weighted graph G, the problem is to find the shortest cycle in a graph by visiting each node exactly once. In Chapter 2 the definition of optimisation problems taken from [Cvetković 1996] is provided.

According to [Rothlauf 2011] (Chapter 2), the optimisation problem is pair (X, f), where X is a set of feasible solutions and $f : X \to R$ is an evaluation function that assigns a real value to every element x of the solution space. The condition $x \in X$ represents the constraints that a solution should satisfy in order to be feasible. The problem is to find an $x^* \in X$ for which

$$f(x^*) \geq f(x) \text{ for all } x \in X \text{ (maximisation problem)}$$

$$f(x^*) \leq f(x) \text{ for all } x \in X \text{ (minimisation problem).}$$

The solution x^* is called a *globally optimal solution* (or optimal solution if no confusion can occur) to the given problem [Rothlauf 2011].

The example of an objective function is to find $\min f(x)$, subject to a constraint $x \in X$. Here, X is a *set of feasible solutions* and f is the *objective function* to be optimised with respect to the *set of constraints* $x \in X$.

When solving an optimisation problem, it is necessary to find all feasible solutions x^* (or at least some of them) such that $f(x^*) = \min_{x \in X} f(x)$. Such solutions are called *(globally) optimal solutions*. By taking into account

$$\max_{x \in X} f(x) = -\min_{x \in X}(-f(x))$$

the maximisation problem is easily reduced to a minimisation problem, and it is sufficient to consider only one of these two problems. Formulated like this a solution can be also considered as a globally optimal solution. Therefore, in the rest of this work, only minimisation is analysed. If X is the finite or countable set, the optimisation problem is the problem of *combinatorial* or *discrete optimisation*. In the case when $X \subseteq R^n$ ($X \nsubseteq Z^n$), the optimisation problem is called *continuous*, otherwise it is referred to as *mixed optimisation problem*. The optimisation problems can be classified according to various criteria. For example, one could distinguish *linear* or *nonlinear* optimisation problems, based on the type of objective function f and constraints [Boyd 2004]. In addition, the optimisation problems may be *deterministic* (if the values of all relevant parameters are known in advance) or *stochastic* [Schneider 2006], [Boyd 2004].

In *local optimisation*, the compromise between solution quality and search efficiency is made. The idea is to give up seeking for the optimum x^* which minimises the objective over all feasible solutions. Instead, a solution that is only locally optimal, i.e., that minimises the objective function on a subset of feasible solutions, is determined [Boyd 2004]. In *global optimisation*, the true globally optimal solution of the given optimisation problem is tried to be found.

Some optimisation problems have multiple (conflicting) objectives. These type of problems are referred to as *multi objective* optimisation problems and they require special mathematical tools, since the best solution for one objective may not be the best for another.

If the mathematical model is given by the system of equations, the usual goal of optimisation is to find the best solution of that given system of equations. Mathematical model is described by variables and parameters. In the optimisation process, normally, the values of variables are to be determined, while the values of parameters are fixed. However, in some situations, the roles of variables and parameters may be changed. This is the case described in this research.

Methods used to solve optimisation problems can be divided into three categories: exact, heuristic (approximate, simulation) and meta-heuristic. Exact optimisation methods are providing the true optimal solution. On the other hand, they are usually computationally and memory consuming. Therefore, they could be applied only to a small size instances of the problem. Heuristic methods are usually problem oriented and designed to quickly provide solutions of high quality. For example, the greedy heuristic [Johnson 1995] for TSP problem is based on the shortest distance. The most deployed are the so called *iterative heuristics*, like local search, pattern search, etc. Iterative heuristics start from an initial solution and try to find better solutions in the neighborhood of the initial solution.

Meta-heuristics are general set of rules that can be applied to solve a variety of optimisation problems. It is expected that application of meta-heuristics (Genetic Algorithms (GA), Neural Networks, Tabu Search, Variable Neighborhood Search, Simulated Annealing, Bee Colony Optimization, Ant Colony Optimization, etc.) when resolving optimisation problems generate better solutions than the one that is given by classical, constructive heuristics with reasonable (acceptable) increase of the execution time. However, some special heuristics are reaching the solution in a shorter computing time. Applying meta-heuristics methods often significantly

improves an initial solution obtained by classical heuristical methods. Development and implementation of meta-heuristic methods is an individual research field.
Apart from Tabu Search and Simulated Annealing, GAs belong to the group of old meta-heuristics methods. They are based on biological evolution described by Darwin's theory of natural selection. The specific mechanisms of GA use the principles of microbiology and their implementation mimics the genetic process. The basic idea of GA approach to problem solving is how to move from current set of feasible solutions towards the better solutions within feasible boundaries. The initial set of solutions is usually randomly generated taking into account the given constraints in order to preserve the feasibility. The number of feasible solutions at the beginning of each generation is called *population size*. For any feasible solution the value of the corresponding objective function can be determined, whose quality is determined by the appropriate *fitness value*. The better fitness value ensures higher probability for the corresponding solution to be involved in GA operator. The basic operations of GA are: selection, crossover and mutation. These operations are executed in each generation until the fulfillment of a pre-determined stopping condition. At that point the best obtained solution is reported as the final one.
Among all optimisation methods, approximate and simulation methods are also recognised as important in solving optimisation problems. They are applicable in cases when mathematical model is not available. Instead, they use experimental or historical data to learn the behaviour of the considered system. Therefore, they are usually used for predicting the future states of the system. Typical representatives of this class of methods are Artificial Neural Networks (ANN) and Bayesian Networks (BN).

1.2 Optimisation in mechanical engineering

Many studies on modelling and design of various mechanical components involving optimisation processes can be found in the literature. For example, transmission errors, prediction of gear utilization with dynamic load, gear noise and optimal design are some of the major concerns for their designers. The design of roller bearings is a challenging task in the field of mechanical engineering. As the producers usually do not reveal the real aspects of the production of bearings, the optimisation of design parameters in bearings can improve the performance as compared to those that currently exist in the standards and catalogues.
Bearing related optimisation problems are solvable by using the optimisation methods. In order to have long working life and good performances of rolling bearings, as important components in mechanical engineering, a plenty of constraints must be satisfied during their design. In order to do that, optimal design methodology is employed by using GA. For example, to design rolling-element bearings, a constraint nonlinear optimisation procedure is used, based on GA.
Exact optimisation techniques are efficient in solving design problems in mechanical engineering. For example, they help in modeling and design of assemblies in

aerospace engineering. In addition, Bayesian Networks are used to predict the behaviour of various mechanical systems. An example of such a system is the transformer oil filtering machine.

Appearing contact and root bending stress on teeth of a gear pair is commonly investigated problem in mechanical engineering on which FEM model is usually applied. In modern gear design, static stress analysis in order to reduce stress concentration is investigated with an aim to minimise development of the initial cracks (by reducing the fatigue). Within the optimisation of the gearbox components, the biggest challenge is to optimise the gearbox with the smallest volume which can carry the system load. For solving these and similar kind of optimisation problems, meta-heuristics can be of a great help because they provide usable results in a shorter computing time with respect to other optimisation methods. In this research, therefore, meta-heuristics, mostly GA, are used as a tool for solving several optimisation problems.

Optimisation problem of working load is related to assessment of working capacity of machine parts, especially gear pairs. In that sense, GA-based procedure is used to optimise design of helical and spur gear pairs. Approximate and simulation methods are widely applicable for solving optimisation problems in the design of mechanical elements and assemblies.

1.3 Contributions

One of the problems investigated within this research is the optimisation problem of the transverse load distribution factor at helical and spur gears. Load transmission by gear pairs is followed by the non-uniform load distribution in the meshing process. The opposite assumption, where the load factor does not change over time along the line of contact, was made. The aim was to identify the parameters with the largest influence on violating this assumption. It was also necessary to determine the extent of their changes. For the purposes of developing this model, all parameters which determine transverse load factor, according to [ISO 6336-1], [ISO 6336-2], [ISO 6336-3], [ISO 1328], [ISO 53] and [ISO 21771] were considered as relevant. The developed model represents the main mathematical contribution in this example. The model is used within the optimisation algorithm based on GA that involves an additional local search optimisation procedure used at the end in order to improve the solution obtained by GA. Such a hybrid algorithm has 12 direct input variables affecting the objective function. The main procedure is divided into several sub-procedures: calculation of geometry, calculation of the stiffness and calculation of the value of total contact ratio. Since the mathematical model of this problem is nonlinear and continuous, the corresponding computational methods, such as Newton-Raphson method and the nonlinear interpolation of three-dimensional function, are implemented. The obtained results [Milojević 2013] showed that the proposed hybrid GA is useful and applicable for optimisation of helical and spur gears design. These results are improvements of research results obtained in [Sánchez 2013], [Pedrero 1996], [Pedrero 2011], [Sánchez 2013], [Simon 1988], [Zhang 2010], [Zhang 1999]. There are the following improvements of previous results:

- In the model presented in [Milojević 2013] most of parameters like helical angle or number of teeth are treated as non-constant which means there are less simplification than in models [Pedrero 2011], [Pedrero 2010], [Sánchez 2013].
- Load distribution is obtained by using GA instead of other methods usually used (i.e. finite element method). Settings of GA used in [Milojević 2013] are conducted respecting the nature of the problem which improves results provided in [Zhang 2010]. Results being obtained by GA in design parameters of helical and spur gears for different gear ratios improve performances of these mechanical elements more than the methods used in [Simon 1988], [Zhang 1999].

Other formulations of mathematical models, known from the state-of-the-art, are varying a maximum of 6 input parameters, while the work presented in [Milojević 2013] is optimising 12 different parameters, categorised in several modules. More precise, in [Zhang 2010], only six input parameters such as, face width, number of teeth on pinion, module, shaft diameters and distance between the bearings on reducer have been taken into consideration. In [Milojević 2013], 6 additional parameters influencing load distribution factor have been analysed with respect to the results obtained in [Zhang 2010], [Pedrero 2010] and [Pedrero 2011]. It is also concluded that 4 parameters that directly affect accurate determination of stiffness have been taken as fixed inputs in other investigations, which is pointing out to the advantage of the results in [Milojević 2013]. Results also show that the optimised value for one of the considered parameters is 7.5 times improved in [Milojević 2013] with respect to [Zhang 2010], while the same parameter in [Pedrero 2010] and [Pedrero 2011] is taken as fixed input and therefore, it is not optimised at all. Other significant result is twice better optimised value of the considered parameter in [Milojević 2013] comparing to the same result in [Zhang 2010].

Planetary gear trains take a very significant place among the gear transmissions which are used in many branches of industry such as automobile transmissions, aircrafts, marine vessels, machine tool gear boxes, gas turbine gear box, robot manipulators, etc. Planetary gear trains have a number of advantages over the transmission with fixed shafts. The relationship between nine influential parameters of planetary gear trains was formulated as the multi-objective nonlinear problem. The weighting method was used to approximate the Pareto set. This method transforms the multi-objective optimisation problem into single-objective optimisation problem by associating each objective function with a weighting coefficient and then minimising the weighted sum of the objectives. The gear contact minimum film thickness is calculated by the Dowson and Higginson's method. In [Rosić 2011a], [Rosić 2011b] the relationship between nine influential parameters is considered, totally (all nine) and pairwise. The proposed GA-based approach produced quite satisfactory results promptly supplying the designer with the preliminary design parameters of planetary gear train for different gear ratios.

Among the most significant contributions of [Rosić 2011b] is deployment of GA method for multi-objective optimisation of the planetary gear train efficiency which has not been done before in the literature. Very complex analysis, involving 9 conflicted objective functions and 6 constraints has been performed in [Rosić 2011b] and [Rosić 2011a]. For the same problem considered in [Qing-Chun 2008] and [Tripathi 2010], only 2 conflicted objective functions influencing the efficiency have been optimised. Optimisation in [Rosić 2011b] and [Rosić 2011a] covered 10 influential parameters, while in [Cho 2006], only 1 relationship between the inputs and outputs is analysed, covering 3 influential parameters. In [Tripathi 2010], only 6 constraint functions have been analysed, while in [Rosić 2011b] and [Rosić 2011a] 8 constraints have taken into account due to the complexity of minimisation of the elastohydrodynamic lubrication film. New models are developed to comprehend influential parameters and to provide simultaneous optimisation of larger number of objective functions.

The multi-objective optimisation of bearings dynamical load ratings and working life, having in mind that these objectives are not conflicting, has also been considered. GA has proven to be a suitable technique in situations when it is necessary to deal with continuous optimisation problems. However, two other meta-heuristic methods based on multi-start local search heuristic algorithms are applied to optimise the problem of dynamic load capacity and working life of ball bearings. The applied local search methods are Pattern Search and Active Set, and the resulting meta-heuristics are referred to as MPS and MAS. The comparative results with GA are given in [Milojević 2014]. Obtained results represent the improvements with respect to these presented in [Chakraborty 2003], [Costin 2010], [Gupta 2007], [Mendi 2010a], [Rao 2007], [Waghole 2014], [Shigley 1989]. Based on the obtained results, it appears that increase in the number of balls and adequate improvement in the terms of internal geometry can increase the dynamic capacity and working life of bearings compared to standard catalog values.

Comparing with the above mentioned papers, the optimisation methods proposed in [Milojević 2014] provided significant increase in a dynamic capacity with respect to the values from the available standards [Bowman] in all eight examples. The average percentage of the improvement is 9.4%, 12.2% and 12.6% for GA, MPS and MAS, respectively comparing to the values in [Bowman]. It is also concluded that optimisation in [Milojević 2014] is improving dynamic capacity for 13.41%, 20.91%, 18.43% than optimisation in [Gupta 2007] for 4 considered types of bearings. Also, the average percentage of improvement of the dynamic capacity with respect to [Rao 2007] for 4 cases of bearings is 22.22%, 30.3% and 27.64%. The work presented in [Gupta 2007] and [Rao 2007] is referring to the values from the [Shigley 1989].

If mathematical model is known and given in the form of algebraic equations then optimisation methods are applied. If it is not case, other techniques should be applied such as Bayesian or Neural Network. As a main contribution in mechanical engineering, Bayesian Network is used first time for reliability assessment of the mechanical system for painting of metal semi-products [Milojević 2012]. The functionality of the mechanical system is presented and the method for predicting the impact of the sub-systems' failure on the final product quality is developed. Based on historical data about system's behavior, probabilities of defined events are determined by applying BN and used for prediction and decisions making.

Another Bayesian model has been developed to predict the behaviour of the transformer oil filtering machine. The model is implemented in C♯ and tested on real problems. Some different scenarios were tested and the obtained results are presented in [Glišović 2013]. The results show that the improvement of product quality can be obtained by applying the proposed methods and therefore significantly support the application of BN in mechanical engineering.

The similar problems of fault prediction in Mechanical systems applying Bayesian networks have been considered in [Alamaniotis 2014], [Hernandez-Leal 2011] and [Boksteen 2014]. The quality of the prediction is estimated based on the difference between historical data and the data obtained after Bayesian network application. In [Milojević 2012], the prediction errors have been calculated and the best obtained result guaranteed the certainty of 97%. In estimation of the prediction results for the remaining useful life of turbine [Alamaniotis 2014], authors claimed 3.3% of minimum error and 42.5% of maximum error occurred. The maximal prediction error in [Milojević 2012] is 3% which improves for 0.3% with respect to [Alamaniotis 2014]. Maximal error obtained in [Glišović 2013] is 1% while in [Hernandez-Leal 2011] authors obtained the minimal error of 15.29% for BN diagnosis of the failures in the combined cycle power plant. This is pointing out that the results obtained in [Milojević 2012] and [Glišović 2013] are approximately 15 and 5 time precise than the results in [Hernandez-Leal 2011], respectively. In [Boksteen 2014], the prediction of plant power and efficiency as a function of ambient temperature is obtained with approximately 95% of certainty. This is leading to the approximate error of 5% which makes the results in [Milojević 2012] and [Glišović 2013] better for 2% and 4%, respectively. The conclusion is that the modeling of the problem is improved in [Milojević 2012] and [Glišović 2013] with respect to the literature and providing more certain prediction as such.

The manuscript consists of the following chapters. Chapter 2 introduces the optimisation problems in a general sense. Chapter 3 contains the description of various optimisation algorithms. An exact optimisation algorithm is the algorithm that provides the true optimal solution to a given optimisation problem. Heuristic, approximate and simulation methods are also presented in Chapter 3. In addition, Chapter 3 contains the description of meta-heuristic methods, where nature-inspired, hybrid and other types of meta-heuristics methods are detailed. Some relevant optimisation problems in mechanical engineering are presented in Chapter 4. The methods that have been applied in purpose of solving optimisation problems of machine elements and assemblies are also discussed. In Chapter 5 an example of optimisation of the transverse load distribution factor of helical and spur gears is described in detail. In order to maximise uniformity of load distribution the GA based method is proposed. In addition, in Chapter 5 the mathematical model for optimisation of pre-design parameters of planetary gears formulated in the form that is the most convenient for the application of meta-heuristic methods is presented. Chapter 6 contains the description of the mathematical model for optimisation of radial ball bearings in order to maximise basic working life and dynamic load capacity. Application of Bayesian Network to reliability assessment of two mechanical systems is described in Chapter 7. The final chapter (8) contains concluding remarks and directions for future research.

Chapter 2

Optimisation

Optimisation problems are common in many disciplines and various domains. In optimisation problems we have to find solutions which are optimal or near-optimal with respect to some goals. Usually, we are not able to solve problems directly (in one step), but we follow some iterative process which guides us through problem solving [Rothlauf 2011]. Often, the modelling process is separated into different steps which are executed one after the other, where the optimisation is related to one of the steps. Commonly used steps are: recognising and defining problems, constructing and solving models (optimisation), and evaluating and implementing solutions.

The optimisation problems can be found almost everywhere in real life. Some representative examples are routing, scheduling, resource allocation, etc. This chapter is devoted to the definition, classification and complexity of various optimisation problems.

2.1 Optimisation problems

In general, optimisation problems are defined as follows [Cvetković 1996]:
Definition 1. Let $f : S \to \mathbb{R}$ or a real function defined on the set S and let $X \subseteq S$ be some given set. The problem is to find

$$\min f(x),$$

subject to a constraint

$$x \in X.$$

Here, the domain S represents a *solution (search) space*, X is a *set of feasible solutions* and f is the *objective function* to be optimised with respect to the *set of constraints* $x \in X$. Each $x \in S$ is called *solution* and $x \in X$ is called *feasible solution*. In order to solve a given optimisation problem, it is necessary to find all feasible solutions x^* (or at least some of them) such that $f(x^*) = \min_{x \in X} f(x)$. The resulting solution x^* is called *optimal solution* or *global optimum*. It is obvious that $(\forall x \in X) f(x^*) \leq f(x)$. The solution $x^\sharp$ is called *local optimum* if $f(x^\sharp) \leq f(x)$ for $x \in X^\sharp \subset X$. Here, $X^\sharp = B(x^\sharp, \varepsilon) \subset X$, with $B(x^\sharp, \varepsilon)$ denoting a ball of small enough radius.

A special case of this formulation is given by [Rothlauf 2011, Boyd 2004]:

$$
\begin{aligned}
\text{minimise} \quad & z = f(x), \qquad (2.1)\\
\text{subject to} \quad & \\
& g_i(x) \geq 0,\ \ i = 1, ..., m\\
& h_i(x) = 0,\ \ i = 1, ..., p,\\
& x \in W_1 \times W_2 \times ... \times W_n,\ \ W_i \in \{\mathbb{R}, \mathbb{Z}, \mathbb{B}\},\ \ i = 1, ..., n,
\end{aligned}
$$

where x is a vector of n decision variables $x_1, ..., x_n$, $f(x)$ is the objective function that is used to evaluate different solutions, and $g(x)$ and $h(x)$ are inequality and equality constraints on the variables x_i. $\mathbb{B}$ indicates the set of binary values $\{0, 1\}$.

2.2 Classification of optimisation problems

Optimisation problems can be classified by several criteria and some of them are analysed in the following sections.

2.2.1 Optimisation problems according to domain type

Decision variables can be either continuous ($x \in \mathbb{R}^n$) or discrete ($x \in \mathbb{Z}^n$). Consequently, optimisation models are either *continuous* where all decision variables are real numbers, *combinatorial* where the decision variables are from a finite or at most countable, discrete set, or *mixed* where some decision variables are real and some are discrete.

2.2.2 Optimisation problems according to function and constraint type

Based on the function and constraint type, problems can be either *linear* or *nonlinear* optimisation problems.

Optimisation problems are *linear* if the objective and all constraint functions are *linear* [Boyd 2004]:

$$
\begin{aligned}
\text{minimise} \quad & c^T \cdot x\\
\text{subject to} \quad & a_i^T \cdot x \leq b_i, \quad i = 1, \ldots, m
\end{aligned}
$$

Here the vectors $c, a_1, ..., a_m \in R^n$ and scalars $b_1, ..., b_m \in R$ are problem parameters that specify the objective and constraint functions.

In order to solve this *linear* optimisation problem, one has to find a vector $(x_1 \ldots x_n)$, such that $c_1x_1 + c_2x_2 + \cdots + c_nx_n$ has the minimal value and the set of the following inequalities is fulfilled:

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &\leq b_1 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &\leq b_m
\end{aligned}
$$

Representative examples of *linear* problems are resource allocation problems, production problems, or network flow problems.

Nonlinear optimisation (or *nonlinear* programming) is the term used to describe an optimisation problem where the objective or constraint functions are not *linear* [Boyd 2004]. There are no effective methods for solving the general *nonlinear* programming problem. Even simple looking problems, with as few as ten variables, can be extremely challenging, while problems with a few hundreds of variables can be intractable. Methods for the general *nonlinear* programming problem therefore take several different approaches, each of which involves some compromise. One of the special cases is *convex* objective function.

2.2.3 Optimisation problems according to parameter values

According to the parameter values problems can be *deterministic* or *stochastic*. The deterministic model is the one in which values of parameters are "a priori" known. Therefore, deterministic models perform the same way for a given set of initial parameter values. Conversely, in a stochastic model, randomness is present, and parameter values are not described by unique values, but rather by probability distributions.

Stochastic optimisation problems include uncertain or dynamic information in their parameters. The objective function value and the violation of constraints of such problems are therefore random variables. Evaluating the objective function value and/or its feasibility can be done either exactly (if a closed-form expression is available), by approximation, by Monte Carlo simulation or by using Fuzzy sets. Meta-heuristics enriched by any of these possibilities have been proposed to solve different stochastic problems.

2.2.4 Classification according to solution type

A *globally* optimal solution is the one where there are no other feasible solutions with better objective function values. A *locally* optimal solution is one where there are no other feasible solutions in the small enough neighbourhood with better objective function values. You can visualize this as a point at the top of a "peak" or at the bottom of a "valley" which may be formed by the objective function and/or the constraints - but there may be a higher peak or a deeper valley far away from the current point.

In *local optimisation*, the compromise between solution quality and search efficiency is made. The idea is to abandon seeking for the optimum x^* which minimises the objective over all feasible points. A large fraction of the research on general nonlinear programming has focused on methods for local optimisation, which, as a consequence have been well developed. In an engineering design application, for example, local optimisation can be used to improve the performance of a design originally obtained by manual, or other design methods.

The goal of *global optimisation* is to find the true global optimal solution of the given optimisation problem. The possibility of finding the true global solution is higher for optimisation problems with a small number of variables where computing time is not critical. An example from engineering design is the worst-case analysis or verification of a high value or safety-critical system. Here, the variables represent uncertain parameters that can vary during manufacturing, or with the environment or operating condition. The objective function is a utility function, i.e., one for which smaller values are worse than larger values, and the constraints represent prior knowledge about the possible parameter values. The optimisation problem is the problem of finding the worst-case values of the parameters. If the worst-case value is acceptable, we can certify the system as safe or reliable (with respect to the parameter variations).

2.3 Combinatorial optimisation problem

The optimisation problem (2.1) is *combinatorial* if S is a finite or at most countable set. Combinatorial optimisation problems are concerned with the efficient allocation of limited resources to meet the desired objectives. The decision variables can take values from discrete sets with respect to additional constraints on basic resources, such as labour, supplies, or capital. These constraints are used to restrict the possibilities for solutions that are considered feasible. Usually, there are many possible alternatives to consider and a corresponding value of the objective function determines which of these alternatives is the best. Typical sets of solutions used for combinatorial optimisation models are integers, permutations, or graphs.

Some representative examples of combinatorial optimisation problems are *Integer Linear Problem (ILP)*, Quadratic Assignment Problem (QAP) [Loiola 2007], Discrete Network Design Problem [Wang 2013], Facility Location Problems p-median [Arya 2012], [Mladenović 2007] and p-center [Mladenović 2003], [Davidović 2011]), 3-SAT Problem [Kutzkov 2007], Knapsack Problem [Bretthauera 2002], etc.

2.3.1 Integer linear problems (linear combinatorial problems)

One of the most known combinatorial optimisation problem is Integer Linear Programming (ILP) problem. ILPs are used to model combinatorial optimisation problems where the decision variables are integers and the objective function and constraints are linear. ILPs in canonical form can be formulated as [Rothlauf 2011]:

$$
\begin{aligned}
\min \quad & c^T x \\
\text{subject to} \quad & Ax \leq b, \\
& x_i \in \mathbb{N}_o,
\end{aligned}
$$

where $\mathbb{N}_o$ is the set of non-negative integers $x \geq 0$. Problems are called Mixed Integer Linear Problems (MILPs) if, besides the discrete decision variables x_i, there are some continuous decision variables y_i. Their canonical form is:

$$
\begin{aligned}
\min \quad & c^T x + d^T y \\
\text{subject to} \quad & Ax + By \leq b, \\
& x_i \in \mathbb{N}_o, \\
& y_j \geq 0.
\end{aligned}
$$

If we drop the integer constraints of an ILP, we get an LP in canonical form which is called *relaxation* of the ILP or the relaxed problem. The solution of this problem represents lower bound for the original ILP in the case of minimisation, i.e., upper bound when the maximum is searched.

2.3.2 Complexity of discrete optimisation problems

Combinatorial optimisation problems that can be solved by polynomial algorithms are considered "easy" and belong to the so-called P-class. An example of *polynomially solvable combinatorial optimisation problem* is 2-SAT [Misra 2013]. However, most of the real life problems are hard. Some of them are known as "NP" hard problems.

In the theory of complexity, NP (non-deterministic polynomial time) is a set of the optimisation problems which one can solve in a polynomial time on the non deterministic Turing machine [Ognjanovic 2004]. Equivalently, NP is the set of problems whose solutions can be deterministically checked on Turing machine in polynomial time. NP-complete problems are the hardest problems in NP class in the sense that every problem in NP class can be reduced to them. Reducing in this context means that for every problem L from NP and the problem C in NP-complete there is a deterministic polynomial time algorithm that converts the cases $l \in L$ into cases $c \in C$, such that l is the optimum of L if and only if c is the optimum of C. To prove that the NP problem A is NP-complete, it is sufficient to show that some already known NP-complete problem is reducible to A.

2.4 Continuous optimisation problem

The optimisation problem (2.1) is *continuous* if $W = W_1 \times W_2 \times \cdots \times W_n \subseteq \mathbb{R}^n$. Actually, continuous optimisation problems are concerned with the optimal settings of continuous decision variables. Here, the main problem is to determine optimal values among uncountably many possibilities.

2.4.1 Convex optimisation problem

Convex optimisation problem, asks to minimise the function $f(x)$, subject to the constraints $g_i(x) \leq 0$, where $i = 1, \dots, m$, under the condition that the following functions $f, g_1, \dots, g_m : \mathbf{R^n} \to \mathbf{R}$ are convex. More precisely those functions should for all $x, y \in \mathbf{R^n}$ satisfy the following inequalities:

$$f(\alpha x + \beta y) \leq \alpha f(x) + \beta f(y),$$

$$g_i(\alpha x + \beta y) \leq \alpha g_i(x) + \beta g_i(y),$$

where $\alpha, \beta \in \mathbf{R^n}$ and $\alpha + \beta = 1, \alpha \neq 0, \beta \neq 0$. Each $x \in \mathbf{R^n}$ is optimisation variable, while f is the objective or optimisation function. Opposite to this, $g_i(i = 1, \dots, m)$ are inequality constraint functions [Boyd 2004].

2.4.1.1 Global optimality in convex problems

Before stating the description of global optimum, optimal point is defined. *Optimal value* of an convex optimisation problem is denoted f^*. This value is equal to minimum value of objective function when satisfying inequality constraints. We say that a x^* is *optimal point* if $f(x^*) = f^*$. It is said that a point x is local optimum if it satisfies all constraints of defined convex problem and if $f(x) \leq f(z)$, for all points z in the vicinity of x. More precisely, x is a *local optimum* if there exists $\varepsilon > 0$ and $\varepsilon \in \mathbf{R}$, such that all points $z \in \mathbf{R^n}$ satisfy the following two conditions:

$$||x - z|| \leq \varepsilon, \text{ and } f(x) \leq f(z).$$

It is said that a point x is *global optimum* if it satisfies all the constraints and it holds that $f(x) \leq f(z)$ for all $z \in \mathbf{R^n}$.

If the optimisation function is *convex* and *differentiable*, optimisation problem can be solved by *exact methods*. Moreover, if the optimisation function f is *twice differentiable*, then f is convex if the Hessian matrix $\mathbb{H}_f$ is a positive definite.

In this section, very widely known problems from a subclass of the convex optimisation called *linear programs* and *least-squares problems*, are presented.

2.4.1.2 Linear problem

(Continuous) Linear Programs (LP) are among the simplest continuous optimisation problems. Their characterization is that the objective function and the set of constraints are linear combination of decision variables. LPs in canonical form can be formulated as [Rothlauf 2011]:

$$\begin{aligned} \min \quad & c^T x \\ \text{subject to} \quad & Ax \leq b, \\ & x_i \in \mathbb{R}, \\ & x_i \geq 0. \end{aligned}$$

2.4.1.3 Least-squares problems for linear model

A least-squares problem, for linear model, is an optimisation problem (2.1) with no constraints [Boyd 2004] (i.e. $m = 0$) and with an objective function in the following form:

$$\text{minimise } f_0(x) = ||A^T x - b||_2^2 = \sum_{i=1}^{k} (a_i^T x - b_i)^2.$$

Here, A is a given matrix, and $x \in \mathbb{R}^n$ is the vector of optimisation variables. There are nonlinear models for which we define least-squares problems.

2.5 Multi-objective optimisation

Many optimisation problems have multiple (conflicting) objectives, essentially changing the concept of optimality, since the best solution for one objective may not be the best for another. In multi-objective optimisation the concept of dominance is therefore introduced. A solution is said to *dominate* another solution if its quality is at least as good on every objective and better on at least one. The set of all non-dominated solutions of an optimisation problem is called the *Pareto set* and the projection of this set onto the objective function space is called the *Pareto front*. The aim of multi-objective optimisation is to find (approximate) the Pareto front and therefore generate a set of *mutually non-dominated solutions* called the *Pareto optimal set* (*Pareto set approximation*).

In general, a multi-objective optimisation problem can be defined as determining a vector of decision variables within a feasible region to minimise a vector of objective functions $f = (f_1, \ldots, f_k)^T$ that usually conflict with each other [Eschenauer 1990]. Formally, the problem is stated as follows:

Find the values of n decision variables $(x_1, ..., x_n)$ which satisfy n upper and lower boundaries $x_{il}, x_{iu}, i = 1, ..., n$ and optimise (minimise or maximise) k objective functions. Since (as it has already been mentioned) the problems of minimising and maximising are equivalent $max f(x) = -min(-f(x))$, the general problem can be written as:

$$\begin{gathered} \text{minimise } \{f_1(x), ..., f_k(x)\} \\ \text{subject to } g(x) \leq 0, \\ x_{il} < x < x_{iu}, \end{gathered}$$

where x is a vector of decision variables, $f_i(x)$ is the i-th objective function, and $g(x)$ is a constraint vector function. The value x_{il} represents the lower boundary and x_{iu} the upper boundary of the decision variable x_i.

The optimal solution in this case is not unique because the objectives can contradict each other. Therefore, a set of solutions that is called the Pareto optimal set is considered according to the following definition:

Definition 2. Pareto optimal: Consider a point x^* in the feasible solution space, X, $f(x^*)$ is multi-objective function, where $x^* \in X$. The point (the assigned values of decision variables) is Pareto optimal if and only if there does not exist another point, $x \in X$, that satisfies $f(x) \leq f(x^*)$ and $f_i(x) < f_i(x^*)$ [Arora 1989] for at least one function. In other words, this definition states that, for a minimisation problem, there is no other point which can cause a decrease in one objective function value without causing a simultaneous increase in at least one of the other objective function values. □

Definition 3. Dominated and non-dominated points: A vector of objective functions values, $f(x^*)$, is non-dominated if and only if there does not exist another vector, $f(x)$, that satisfies $f(x) \leq f(x^*)$ with at least one $f_i(x) < f_i(x^*)$. Otherwise, $f(x^*)$ is dominated. □

Definition 4. Pareto front: The set $X^* = \{x^* \in \mathbb{R}^n | x^* \text{ is non-dominated solution}\}$ which is composed of all the non-dominated Pareto optimal solutions that compromise the Pareto front of non-dominated solutions [Eschenauer 1990], is called Pareto front. □

Definition 5. The "feasible" domain is defined by the following: $D = \{x \in \mathbb{R}^n | g(x) \leq 0\}$.

Fig. 2.1 shows a mapping of the "feasible" domain D, given by Definition 5, into the criterion space X where the Pareto-optimal solutions lie on the curved section AB.

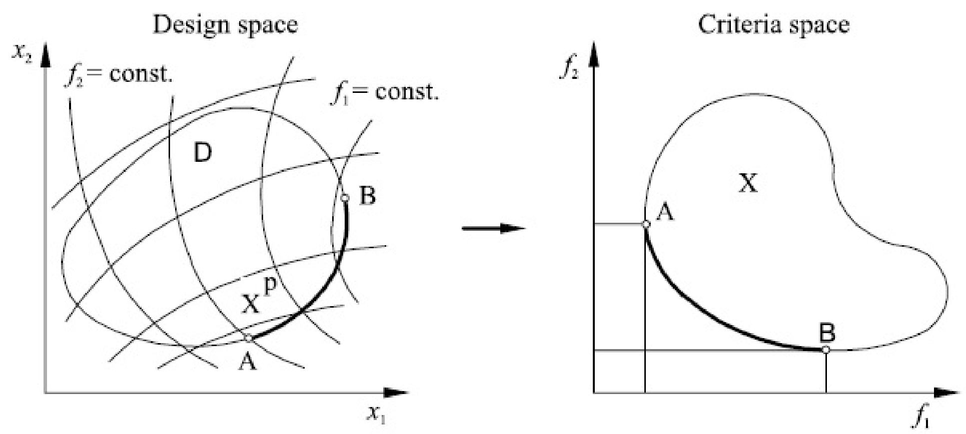

Figure 2.1: Mapping of a feasible set into the criterion space taken from [Rosić 2011a]

The *weighting sum method* is used to approximate the Pareto set. This method transforms the multi-objective optimisation problem into a single-objective optimisation problem by associating each objective function with a weighting coefficient and then minimising the weighted sum of the objectives, as follows:

$$\min f(x) = \Sigma_{i=1}^{k} w_i f_i(x), \tag{2.2}$$

where weights w_i are non-negative, such that:

$$\Sigma_{i=1}^{k} w_i = 1, \tag{2.3}$$

for k objective functions.

A subset of the Pareto optimal set can thus be generated through a systematic varying of the weights $w = \{w_1, w_2, \cdots, w_k\}$ and repeatedly solving the scalar form of the problem (2.2). The weights are modified after every certain number of iterations during the optimisation.

The weights are defined by the following equation:

$$w_i(l) = \frac{random_i(l)}{\Sigma_{j=1}^{k} random_j(l)}, \tag{2.4}$$

where l is the index of iteration, $random_j(l)$ is the function used to create a uniformly distributed random value in the range $[0, 1]$.

Chapter 3

Optimisation methods

This section describes the conventional (classical) and meta-heuristics methods used for solving optimisation common problems of machine elements and assemblies. Main contribution of this section is survey of some optimisation methods that are applicable to solving practical optimisation problems such as load distribution problem in a meshed gear pair. Additionally, there is a contribution with survey of methods that are efficient in solving optimisation problems of ball bearings geometry, prediction of potential machine faults and providing help with decision making. The methods are divided into the following groups: exact optimisation methods, heuristic, approximation and simulation methods, meta-heuristics methods. Main focus was in explanation of meta-heuristic methods, such as genetic algorithms, because they are used to provide one of the main contributions in this thesis.

3.1 Exact Optimisation methods

This section summarizes some exact optimisation methods based on the following criteria (adapted from [Schneider 2006], [Brunet 2010]):

1. Exact optimisation methods that include algorithms for solving discrete optimisation problems.
2. The set of methods that solve optimisation problems by using Newton's algorithm and its adaptations.
3. Methods that use global optimisation techniques.
4. Methods that use linear programming techniques in order to solve continuous problems.

3.1.1 Algorithms for Discrete Optimisation Problems

According to [Schneider 2006], [Brunet 2010], the group of methods that solve discrete optimisation problems includes the following: exhaustive enumeration, branch & bound, dynamic programming, and cutting plane.

Exhaustive enumeration method (EE) is a combinatorial optimisation technique. The method evaluates all the combinations of discrete variables [Schneider 2006]. The optimal solution obtained is the minimum over all objective function values calculated for the complete list of feasible solutions. This method assures the global

optimum, but it may be computationally or memory consuming. In order to reduce the computational time of EE, some other exact methods were developed.
The Branch & Bound (BB) uses various intelligent ways to avoid visiting solutions that do not have good quality [Schneider 2006]. Branch & Cut (BC) [Schneider 2006] is another exact optimisation method used instead of EE due to its reduced computational complexity. Dynamic programming is another exhaustive search method that intelligently enumerates all solutions of a combinatorial optimisation problem [Rothlauf 2011]. To be able to apply dynamic programming, the problem has to be formulated as a *multistage process*. In that sense, dynamic programming was proposed by Bellman [Bellman 1964] as an approach to solve multistage decision process.
Cutting Plane methods perform in an iterative manner and therefore (in general case) they require exponential effort for NP-complete problems [Rothlauf 2011] .

3.1.2 Differentiable functions optimisation

3.1.2.1 Newton's Method

This section introduces Newton's optimisation method. The Google Scholar search engine offers wide range of resources that explain Newton's method, as well as explain how to use the method in solving many optimisation problems. According to [Brunet 2010], the method has an iterative algorithm that minimises a function of the form $f : \mathbb{R}^n \to \mathbb{R}$. Each step of the Newton's algorithm consists in finding the minimum of the quadratic approximation of the function f around the current point x. As stated in [Brunet 2010], this principle can be expressed by using the second order Taylor expansion of the form:

$$\min_{\delta} f(x) + \nabla f(x)\delta + \frac{1}{2}\delta^T H_f(x)\delta. \tag{3.1}$$

In equation (3.1), H denotes Hessian regular matrix, δ is vector, $\nabla f(x)$ is gradient. A necessary condition for the optimum of the problem (3.1) is obtained by setting to zero the derivative of the cost function. This amounts to solve the following linear system of equations:

$$H_f(x)\delta = -\nabla f(x). \tag{3.2}$$

Based on [Brunet 2010], the search direction δ for the Newton method is defined as follows:

$$\delta = -(H_f(x))^{-1}\nabla f(x). \tag{3.3}$$

Note that, in practice, the inverse Hessian matrix in equation (3.3) does not need to be explicitly calculated. It is possible to compute δ from the equation (3.2) using an efficient solver of linear systems.

Fig. 3.1 gives an complete Newton's method. According to [Brunet 2010], the update $x = x + \delta$ can be replaced by the update $x = x + \gamma\delta$ where γ is a positive value smaller than 1.

Algorithm 3: Newton's Method

input : $f : \mathbb{R}^n \rightarrow \mathbb{R}$ a twice-differentiable function
$\mathbf{x}^{(0)}$ an initial solution

output: $\mathbf{x}^{\star}$, a local minimum of the cost function f.

begin
- $k \leftarrow 0$;
- **while** STOP-CRIT **and** $(k < k_{max})$ **do**
  - $\mathbf{x}^{(k+1)} \leftarrow \mathbf{x}^{(k)} + \boldsymbol{\delta}^{(k)}$;
  - with $\boldsymbol{\delta}^{(k)} = -\left(\mathbf{H}_f(\mathbf{x}^{(k)})\right)^{-1} \boldsymbol{\nabla} f(\mathbf{x}^{(k)})$;
  - $k \leftarrow k + 1$;
- **return** $\mathbf{x}^{(k)}$

end

Figure 3.1: Newton's Optimisation Algorithm (taken from [Brunet 2010])

Function f must satisfy certain conditions specified in [Culioli 1994]. In other words, if the Newton's method is initialized with $x^{(0)} \in N$, the convergence is quadratic. Outside of the neighbourhoods of the local minima of the cost function, there are no guarantee that Newton's method will converge.

3.1.2.2 Quasi-Newton methods

Quick convergence is an advantage of Newton's method [Brunet 2010]. This means that it does not take a lot of iterations to reach the minimum. Each iteration requires one to compute the Hessian matrix [Brunet 2010]. To reduce the cost of computation, the goal of quasi-Newton approach is to replace the Hessian matrix with appropriate approximation [Brunet 2010].

Quasi-Newton [Culioli 1994] relies on the second order Taylor expansion of the function to optimise, except that the Hessian matrix is replaced with an approximation A [Culioli 1994]:

$$f(x + \delta) \approx f(x) + \delta + \frac{1}{2}\delta^T A\delta. \tag{3.4}$$

The gradient of this approximation with respect to δ is given as follows :

$$\nabla f(x + \delta) \approx \nabla f(x) + A\delta. \tag{3.5}$$

The general principle of a quasi-Newton approach is to choose the matrix A such that [Culioli 1994]:

$$\nabla f(x + \delta) = \nabla f(x) + A\delta. \tag{3.6}$$

The difference between the different formulation are the properties that the matrix A satisfies at each iteration of the algorithm.

3.1.2.3 Newton-Raphson method

According to [Schneider 2006], Newton-Raphson method gives answer on the following question: *whether it is possible to find optimal operating points of systems with many variables and a definite cost for any operating point so that there exist a cost function whose value can be minimised?* [Schneider 2006]. To answer on this question, let's start with continuous function, denoted by $\mathbb{H}(x)$, with single variable x. Additionally, suppose that it is not expensive to compute derivatives of $\mathbb{H}(x)$ [Schneider 2006]. In this case, if the starting value of x is not too far from a minimum, the Taylor expansion of the form [Schneider 2006]:

$$\mathbb{H}(x) = \mathbb{H}(x_0) - b(x - x_0) + A(x - x_0)^2/2, \tag{3.7}$$

will allow the "stimulation" of the value of the minimum (the point at which the first derivative of $\mathbb{H}$ disappears), at the point where the following condition should hold [Schneider 2006]:

$$b = A(x - x_0). \tag{3.8}$$

By guessing the next approximation, x_1, for the point of the minimum will be [Schneider 2006]

$$x_1 = x_0 + b/A \tag{3.9}$$

Based on [Schneider 2006], the last formula will give a quadratically convergent iterations of approximations to the minimum [Schneider 2006]. In this case the error in minimising $\mathbb{H}(x)$ decreases by the square of the deviation remaining in x with each iteration [Schneider 2006].

3.1.2.4 Gauss-Newton method

The Gauss-Newton method [Culioli 1994] is an optimisation algorithm used to minimise a cost function that can be written as a sum of squares. The following formula expresses such function:

$$f(x) = \sum_{i=1}^{m} \left(f_i(x)\right)^2, \tag{3.10}$$

where $f : \mathbb{R}^n \to \mathbb{R}$, x is vector, and each f_i ($i \in [1, m]$) is a function of the form $f_i : \mathbb{R} \to \mathbb{R}$. The only hypothesis that must be satisfied for the Gauss-Newton method is that the functions f_i are all differentiable [Culioli 1994] . Equation (3.10) is the definition of a least-squares problem. Least-squares optimisation problems are important since the corresponding cost functions often arise when estimating the parameters of a parametric model [Culioli 1994]. The derivation of the Gauss-Newton method is more conveniently done by considering the minimisation of a vector-valued function $\mathcal{F} : \mathbb{R}^n \to \mathbb{R}^m$, in the following way [Culioli 1994]:

$$\min_x \|\mathcal{F}(x)\|^2, \tag{3.11}$$

where $\mathcal{F}$ is defined as:

$$\mathcal{F}(x) = \begin{pmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{pmatrix}. \tag{3.12}$$

In formula (3.12), $\mathbf{x}$ are vectors. Note that, solving the problem (3.11) is equivalent to minimising the function f as defined in equation (3.10) [Culioli 1994]. The Gauss-Newton method is an iterative algorithm where each iteration consists in minimising the first-order approximation of the function $\mathcal{F}$ around the current solution. The first-order approximation of $\mathcal{F}$ is given by [Culioli 1994]:

$$\mathcal{F}(x + \delta) = \mathcal{F}(x) + \mathsf{J}_{\mathcal{F}}(x)\delta. \tag{3.13}$$

In Gauss-Newton method, each iteration tries to determine the step δ, by solving the following minimisation problem [Brunet 2010]:

$$\min_{\delta} \|\mathcal{F}(x) + \mathsf{J}_{\mathcal{F}}(x)\delta\|^2. \tag{3.14}$$

Problem (3.14) is well known as a linear least-squares minimisation problem, that does not require many computations [Brunet 2010]. According to [Brunet 2010], Fig. 3.2 illustrates the complete principle of the Gauss-Newton method.

Algorithm 4: Method of Gauss-Newton

input : $f : \mathbb{R}^n \to \mathbb{R}$ a function such that $f(\mathbf{x}) = \sum_{i=1}^{m} (f_i(\mathbf{x}))^2$
where all the f_i are differentiable functions from $\mathbb{R}^n$ to $\mathbb{R}$
$\mathbf{x}^{(0)}$ an initial solution

output: $\mathbf{x}^{\star}$, a local minimum of the cost function f.

begin
  $k \leftarrow 0$;
  while STOP-CRIT **and** $(k < k_{max})$ **do**
    $\mathbf{x}^{(k+1)} \leftarrow \mathbf{x}^{(k)} + \boldsymbol{\delta}^{(k)}$;
    with $\boldsymbol{\delta}^{(k)} = \arg\min_{\boldsymbol{\delta}} \|\mathcal{F}(\mathbf{x}^{(k)}) + \mathbf{J}_{\mathcal{F}}(\mathbf{x}^{(k)})\boldsymbol{\delta}\|^2$;
    $k \leftarrow k + 1$;
  return $\mathbf{x}^{(k)}$
end

Figure 3.2: Gauss-Newton method (taken from [Brunet 2010])

The step δ is a descent direction [Björck 1996] (see page 38 in [Brunet 2010]). As stated in [Brunet 2010], if the algorithm converges then the limit is a stationary point of the function f (but not necessarily a minimum). However, with no assumptions on the initial solution, there is no guarantee that the algorithm will converge, even locally.

It is stated in [Brunet 2010] (see page 38), that the convergence speed of the Gauss-Newton method is almost quadratic, but only in case of acceptable starting point and a "nice" function f (i.e. a mildly nonlinear function). Even more, it can be worse than quadratic if the starting point is far from the minimum or if the matrix $J_{\mathcal{F}}^T J_{\mathcal{F}}$ is ill-conditioned [Brunet 2010].

Adapted Newton's methods, also, play an important role in solving least-squares minimisation problem [Levenberg 1944]. The following subsection examines methods [Marquardt 1963] for global optimisations, and discusses their relationships with Newton's method.

3.1.2.5 Levenberg-Marquardt method

The content and the presentation of this section is inspired by [Madsen 2004] and [Brunet 2010]. The Levenberg [Levenberg 1944] and the Levenberg-Marquardt method [Marquardt 1963] are an versions of the Gauss-Newton method [Brunet 2010]. The Levenberg-Marquardt algorithm is given in Fig. 3.3 [Brunet 2010].

Algorithm 5: Levenberg-Marquardt algorithm

input : $f : \mathbb{R}^n \rightarrow \mathbb{R}$ a function such that $f(\mathbf{x}) = \sum_{i=1}^{m}(f_i(\mathbf{x}))^2$
where all the f_i are differentiable functions from $\mathbb{R}^n$ to $\mathbb{R}$
$\mathbf{x}^{(0)}$ an initial solution

output: $\mathbf{x}^\star$, a local minimum of the cost function f.

begin
- $k \leftarrow 0$;
- $\lambda \leftarrow \max \mathrm{diag}(\mathbf{J}^\mathsf{T}\mathbf{J})$;
- $\mathbf{x} \leftarrow \mathbf{x}^{(0)}$;
- **while** STOP-CRIT **and** $(k < k_{max})$ **do**
  - Find $\boldsymbol{\delta}$ such that $(\mathbf{J}^\mathsf{T}\mathbf{J} + \lambda \mathrm{diag}(\mathbf{J}^\mathsf{T}\mathbf{J}))\boldsymbol{\delta} = \mathbf{J}^\mathsf{T}\mathbf{f}$;
  - $\mathbf{x}' \leftarrow \mathbf{x} + \boldsymbol{\delta}$;
  - **if** $f(\mathbf{x}') < f(\mathbf{x})$ **then**
    - $\mathbf{x} \leftarrow \mathbf{x}'$;
    - $\lambda \leftarrow \frac{\lambda}{\nu}$;
  - **else**
    - $\lambda \leftarrow \nu\lambda$;
  - $k \leftarrow k + 1$;
- **return** $\mathbf{x}$

end

Figure 3.3: Levenberg-Marquardt (taken from [Brunet 2010])

The Levenberg algorithm [Brunet 2010] solves a least-squares minimisation problem proposed by [Levenberg 1944]. Slight variation of the initial method of Levenberg algorithm, also known as the Levenberg-Marquardt algorithm, is given in [Marquardt 1963]. The step for the Levenberg-Marquardt algorithm, denoted as δ_{lm} where lm in index denotes short description of first words of name for Levenberg-Marquardt algotirhm, is defined as [Marquardt 1963]:

$$(J^T J + \lambda diag(J^T J))\delta_{lm} = J^T f. \tag{3.15}$$

The linear system of equations (3.15) is called *augmented normal equations*. The value λ is a positive value named the *damping parameter*, and J is the Jacobian matrix of the function f evaluated at x, and $f^T = (f_1(x), ..., f_m(x))$ [Marquardt 1963], where $m \geq 1$, $m \in N$. The matrix $(J^T J + \lambda diag(J^T J))$ has a property of being a positive definite. Therefore, δ_{lm} is necessarily a descent direction. For large values of λ, it holds $\delta_{lm} \approx -\frac{1}{m}\nabla f$ [Brunet 2010]. In this case, the Levenberg-Marquardt algorithm is almost a gradient descent method (with a short step) [Brunet 2010]. This strategy is appropriate when the current solution is far from the minimum [Brunet 2010]. On the contrary, if λ is a small value, then the

Levenberg-Marquardt step δ_{lm} is almost identical to the Gauss-Newton step δ_{gn} [Marquardt 1963], [Madsen 2004]. This is a desired behaviour for the final iterations of the algorithm since, near the minimum, the convergence of the Gauss-Newton method can be almost quadratic. The length and the direction of this step are affected by the damping parameter [Brunet 2010]. In this case one does not need for a line-search procedure in the iterations of this algorithm [Brunet 2010]. The value of λ is changed along with the iterations based on the following strategy [Brunet 2010]:

1: If the current λ results in an improvement of the cost function, then the step is applied and λ is divided by a constant ν (with, typically, $\nu = 2$).

2: On the contrary, for the iteration where the current λ increases the function, the step is discarded and λ is multiplied by ν.

Moves 1 and 2 are strategies for updating the damping parameter [Brunet 2010]. More details about the strategy are given in [Madsen 2004].
Next section introduces another algorithm that falls into exact optimisation methods, and it is known as Newton-Raphson [Schneider 2006]. The algorithm is solving linear-continuous problem [Schneider 2006], where state of a system can be described with an N-dimensional vector x with real-valued components [Schneider 2006].

According to [Brunet 2010], *Iteratively Reweighed Least Squares* and *Golden Section Search (GSS)* algorithms apply to minimisation of function $f : \mathbb{R} \to \mathbb{R}$ [Brunet 2010], over given interval. In case of second one, the function f should be continuous and unimodal on the given interval [Brunet 2010]. For the first case of minimisation, the function is $f(x) = \sum_{i=1}^{n} w(x) \|f_i(x) - y_i\|^2$, where w is a function from $\mathbb{R}^n$ to $\mathbb{R}$ [Brunet 2010]. The GSS algorithm works on the principle of refining a set of 3 locations $x_1 < x_2 < x_3$ with the assumptions that $f(x_2) \leq f(x_1)$ and $f(x_2) \leq f(x_3)$ [Brunet 2010].

3.1.2.6 Active Set Method

Active set method for nonlinear optimisation is based on sequential linear programming (SLP) methods.

If the problem is to find

$\min_x f(x)$ subject to a contraint $g(x) \geq 0$, where $f(x)$ is quadratic objective function and $f(x) : R^n \to R$.

SLP methods are solving a trust-region LP around the current iterate x^k, given by:

$\min_d f_k^T d$ subject to $g_k + D_k^T d \geq 0$ and $||d||_\infty \leq \Delta_k$,

where $h_k = \nabla f(x^k)$, $g_k = g(x^k)$, and $D_k = \nabla g(x^k)^T$ [Leyffer 2005].

The solution of this LP provides an estimate of the active inequality constraints, which is used to define an equality constrained quadratic programs(QP) to compute a second-order step.

Active Set (AS) methods are two-phase iterative methods that provide an estimate of the active set at the solution. In the first phase (the feasibility phase or phase 1), the objective is ignored while a feasible point is found for the constraints $h(x) = b$ and $D(x) \geq v$, where D is inequality constraint matrix, and v is constant vector. In the second phase (the optimality phase or phase 2), the objective is minimised while feasibility is maintained. For efficiency, it is beneficial if the computations of both phases are performed by the same underlying method. The two-phase nature of the algorithm is reflected by changing the function being minimised from a function that reflects the degree of infeasibility to the quadratic objective function.

Suppose that x_0 is feasible point. Active-set methods are computing a sequence of feasible iterates $\{x_k\}$ such that $x_{k+1} = x_k + \Delta_k p_k$ and $f(x_{k+1}) \leq f(x_k)$, where p_k is nonzero search direction, and Δ_k is nonnegative step length. Active-set methods are based on Farkas' Lemma [Wong 2011]. The lemma states that a feasible point x myst satisfy *first order optimally conditions* or to be starting point of a direction such that $B_a p \geq 0$ and $D(x)^T p < 0$, where B_a is active constraint matrix.

One technique to find a solution to the first order necessary optimality conditions here is to guess the subset of inequality constraints which will be active at the optimum, called the **active set** [Leyffer 2005], [Wong 2011], [Murty 1988].

Methods that can be considered as active set methods are

- Successive linear programming (SLP)
- Sequential quadratic programming (SQP)
- Sequential linear-quadratic programming (SLQP)
- Reduced gradient method (RG)
- Generalized reduced gradient method (GRG)

3.1.3 Algorithms for Linear-Continuous problems

3.1.3.1 Simplex method

The simplex algorithm solves the linear problems where the state of a system can be described with an n-dimensional vector x with real-valued components [Schneider 2006]. The cost function is given in the following form [Schneider 2006]:

$$\mathbf{f}(x) = c_1 x_1 + c_2 x_2 + ... + c_n x_n = cx \tag{3.16}$$

with $c \in \mathbb{R}^n$ and $x_i > 0$ $((\forall i)(i = 1, \ldots, n))$. The simplex algorithm performs over a set of linear inequalities. It means that, in case of two dimensions, optimisation problem can be solved geometrically [Schneider 2006] by determining the region of feasible solutions, also known as the polytope. Authors in [Schneider 2006] stated that graphical solution is not possible for dimensions higher than two. In all cases, set of inequalities can be transformed into set of equalities and solved by using matrix calculus.

3.1.3.2 Cholesky factorisation method

Cholesky factorisation method is well known tool that efficiently solves linear equations [Davis 1999]. The method supposes that every symetric matrix A can be written as

$$A = RR^T, \tag{3.17}$$

where R is called the Cholesky factor of A [Davis 1999] [Brunet 2010]. Additionally, we suppose that the matrix A is positive definite because all R should be real numbers, and not complex numbers. Suppose that the following equation has to be solved:

$$Ax = b, \tag{3.18}$$

with positive definite A of order n, and $b \in \mathbb{R}^n$. According to [Brunet 2010], the algorithm works as follows:

1. Factorise matrix A as RR^T.
2. Solve $RR^Tx = b$.

The costs of factorisation is $(\frac{1}{3})n^3$ flops. To solve equation $RR^Tx = b$ one should use *forward and back* substitutions [Davis 1999]. Although efficient, this approach works only if the problem is well-conditioned [Davis 1999]. Practically speaking, the Cholesky factorisation solves efficiently sparse matrices.

QR Factorisation method [Brunet 2010] is used to solve linear least-squares minimisation problem of the form [Brunet 2010]:

$$\min_x \|Fx - y\|^2 \tag{3.19}$$

with $F \in \mathbb{R}^{m \times n}$, $m \leq n$, and $y \in \mathbb{R}^m$. The method is based on Cholesky factorisation method [Brunet 2010] and it can be solved by using back-substitution algorithm.

3.2 Heuristic, Approximate and Simulation methods

This section covers four methods from the group of heuristics, approximate and simulation methods. According to the author's best knowledge and practical experience, these four methods are very applicable in solving problems that require a lot of memory to be solved in comparison to the exact optimisation methods analysed in previous section. Heuristics methods may produce results by themselves, improve the efficiency of other optimisation algorithms by generating good starting values.

Approximation methods obviously find solution that is close to exact one, while simulation methods are applicable in manufacturing engineering for simulation models for static and dynamic analysis of different types of objects.

3.2.1 Finite elements model method

The finite element method (FEM) is from the group of the simulation methods, very often used in solving the problems from mechanical engineering. FEM is used in building, modeling and simulation of advanced engineering systems [Liu 2003]. FEM helps in checking workability of the product to be finished, and more importance the cost of effectiveness [Liu 2003]. According to [Liu 2003], the FEM procedure of computational modeling includes the following steps:

- Modeling of the geometry.
- Meshing (discretization).
- Specification of material property.
- Specification of boundary, initial and loading conditions.

Engineers during modeling of engineering systems try to reduce its complexity by using techniques borrowed from geometry. Geometry of such complex engineering systems can be represented by geometry of its elements. For example, curved parts of the geometry and its boundary can be modeled by using curves and curved surfaces [Liu 2003]. Software for modeling geometry allows engineers to rotate and translate created lines and curves. According to the practical experience in manufacturing engineering, modeling geometry helps to understand model of creating complex system, use patterns in geometry to optimise some elements of the engineering system, and finally having documentation of the engineering system that is very important for later maintenance of the system and its improvements.
To solve complex modeling problem in manufacturing engineering, engineers try to divide problem domain in smaller elements (also known as cells) [Liu 2003]. The techniques of dividing domain, in FEM, by using set of *grids* and *nodes* is known as *meshing* [Liu 2003]. Solution within cells is approximated by using, for example, polynomials [Liu 2003].
Almost every engineering system has more than one material. For simulations in FEM, it is important to define its properties [Liu 2003]. During modeling of engineering system, obtaining these properties is not easy task. Mostly, engineers use available databases of material properties. By experience in particular domain, engineers already know many properties of materials to be used during FEM.
The most important role in solving simulation, during the application of FEM methods, is boundary, initial and loading condition [Liu 2003]. User of a system specify these conditions to the geometrical identities (such as points, lines or curves), or elements of grid [Liu 2003].

3.2.2 Pattern search

An overview of Pattern Search (PS) method based on Hooke-Jeeves (HJ) Pattern Search optimisation algorithm [Stanimirovic 1999], [Wetter 2003] is given in the following lines. The HJ algorithm is known as generalized pattern search algorithm. As stated in [Wetter 2003], [Lai 2007] the HJ constructs sequence of iteration, also known as moves, among which *exploratory moves* and *pattern moves* can be distinguished. The first move achieves as a result value of function $f(x)$ in the neighbourhood of the current basic point x^k. Each variable x_j^k from $x^k = (x_1^k, ..., x_n^k)$ of function $f(x^k)(k = 1, \dots, n)$ is changed for a given incremental Δ and the corresponding function value is calculated. If function value is reduced, also noted as if *a move is successful*, then new value of that variable will be retained. Upon analysing all available variables, new basic point x^{k+1} is reached. If the function reduction fails, then $x^{k+1} = x^k$ holds. The pattern move improves the speed of search by using information about function $f(x)$ in order to determine the most appropriate search direction [Lai 2007]. To explain in details how these moves work, steps of HJ algorithm described in [Stanimirovic 1999] are recalled. Instead of *moves*, authors [Stanimirovic 1999] use term *search*.

At the beginning, the initialisation of all coordinate values at some point x_0 is performed, and the starting value for incremental step Δ is determined. The initial point x_0 becomes the first basic point, i.e. $x_B^0 = x_0$. Next, exploratory move I is run where variables are changing by given incremental step, one at a time. In the k^{th} iteration, the basic point $x_B^k = (x_1^k, \dots, x_n^k)$ and $x_1^{k+1} = x_1^k + \Delta$. During the exploratory move I, if f is improved, then $x_1^{k+1} = x_1^k + \Delta$ is adopted as a new element of x_B^{k+1}. If f is not improved, then $x_1^{k+1} = x_1^k - \Delta$ and the objective value of f is checked for the improvement. The process of changing x_j^k $(j = 2, \dots, n)$ by some Δ is the same as for x_1^k, until all the independent variables have been changed and the exploratory move I is completed. Each time the current value of f is compared with the value of f at previous point [Stanimirovic 1999]. If the improvement of the function f by changing the variable x_j for $+\Delta$ or $-\Delta$ is not reached, then the previous value of coordinate is kept, i.e. $x_j^{k+1} = x_j^k$. The exploratory move I produces new basic point x_B^{k+1}.
After applying exploratory move I, pattern move is applied by defining the new starting point as follows: $x_B^{k+2} = x_B^{k+1} + (x_B^{k+1} - x_B^k)$. The alternative for determining x_B^{k+2} was also considered in [Stanimirovic 1999].

As a continuation of pattern move, exploratory move II is applied. Success of pattern move is provided, after success of exploratory move II. The same holds for failure of pattern move. It can be claimed that pattern move fails if f is not improved after exploratory move II. In this case Δ is reduced gradually. On the contrary, if f is improved after applying exploratory move II, the last point reached at this step is coined as new *basic point* x_B^{k+2}. The whole process is finished when Δ is less than some pre-specified value. An implementation of the HJ algorithm is given in [Stanimirovic 1999].

3.2.3 Bayesian networks

Bayesian networks (BN) belong to the class of artificial networks that deal with probabilistic models [Ben-Gal 2007]. In another words, they represent knowledge in uncertain domain by using graphs, more precisely, Directed Acyclic Graphs (DAGs). Each node in a DAG represents a random variable, and the edges between the nodes denotes probabilistic dependencies among the linked random variables [Ben-Gal 2007]. According to [Ben-Gal 2007], they became extremely applicable in last decade. The following, (non formal) definition of BN, is borrowed from [Ben-Gal 2007] (see page 1 in [Ben-Gal 2007]).
The network is defined by a pair $B = \langle G, \Theta \rangle$, where G is the DAG whose nodes $X_1, X_2, ..., X_n$ represent random variables, and whose edges represent the direct dependencies between these variables [Ben-Gal 2007]. The graph G encodes independence assumptions, by which each variable X_i is independent of its non-descendents given by its parents in G [Ben-Gal 2007]. The second component Θ denotes the set of parameters of the network. This set contains the parameter $\Theta_{x_i|\pi_i} = P_B(x_i|\pi_i)$ for each realization x_i of X_i conditioned on π_i, the set of parents of X_i in G. Finally, the following formula holds [Ben-Gal 2007]:

$$P_B(X_1, X_2, ..., X_n) = \Pi_{i=1}^{n} P_B(X_i|\pi_i) = \Pi_{i=1}^{n} \Theta_{X_i|\pi_i} \tag{3.20}$$

It is important to note that B defines a unique Joint Probability Distribution (JPD) over V [Ben-Gal 2007]. More generally, edge from node X_i to the node X_j states that a value taken by the variable X_j depends on the value taken by the variable X_i, or more generally it means that X_i "affects" X_j [Ben-Gal 2007]. It can be also claimed that there are parent-child relations among nodes in DAG, as follows: the node X_i is *parent of* node X_j and, inverse, X_j is *child of* X_i [Ben-Gal 2007]. It is important to note that parent-child relations among nodes in this type of networks are not reflexive. It means that no node is parent of itself.
Bayesian networks learning problem is defined as follows (see page 1 in [Ben-Gal 2007]): *Given training data and prior information (e.g., expert knowledge, casual relationships), estimate the graph topology (network structure) and the parameters of the JPD in the BN.*
There are four methods of learning BN and they are listed as stated in [Ben-Gal 2007]:

- *Maximum-likelihood estimation*
- *Expectation-maximisation (EM) (gradient ascent)*
- *Search through model space*
- *EM +search through model space.*

3.2.3.1 An example of Bayesian network

The example shown in Fig. 3.4 illustrates a BN. It states a person who might suffer from a back injury, an event represented by the variable *Back* (denoted by B). Such

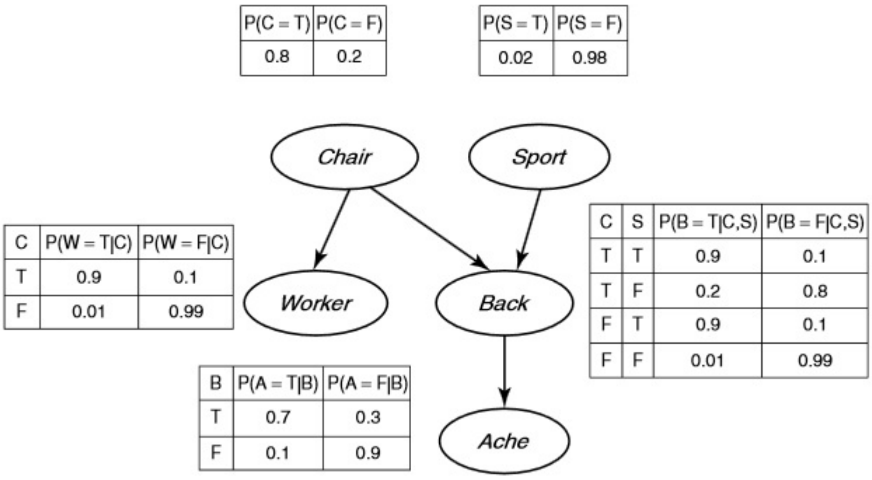

Figure 3.4: The backache BN example (taken from [Ben-Gal 2007])

an injury can cause a backache, an event represented by the variable *Ache* (denoted by A) [Ben-Gal 2007]. S denotes *Sport* . The back injury might result from a wrong sport activity. *Chair* is denoted by C. The back injury might also result from new uncomfortable chairs installed at the person's office [Ben-Gal 2007]. It is assumed that a $Worker$, denoted by W, may report a similar backache syndrome. All variables are binary and they are true (T), or false (F). Close to the each node, the conditional probability table (CPT) is shown [Ben-Gal 2007].
The example in Figure 3.4 shows parent-child relations between nodes, illustrated by arrows. The parents of the variable *Back* are the nodes *Chair* and *Sport*. The child of *Back* is *Ache*, but the parent of $Worker$ is *Chair*. Several independence statements are observed from this BN, by following the BN independence assumption. For example, the variables *Chair* and *Sport* are marginally independent, but when *Back* is given they are conditionally dependent [Ben-Gal 2007]. This relation is known as *explaining away*. When *Chair* is defined, $Worker$ and *Back* are conditionally independent. When *Back* is defined, *Ache* is conditionally independent of its parents *Chair* and *Sport* [Ben-Gal 2007].
Joint distribution of all the variables is defined by using the chain rule as follows:

$$P(C,S,W,B,A) = P(C)P(S|C)P(W|S,C)P(B|W,S,C)P(A|B,W,S,C) \quad (3.21)$$

Previous formula is recognized as factorisation [Ben-Gal 2007]. Instead of the formula above, JPD can be defined in a so called factored form as follows:

$$P(C,S,W,B,A) = P(C)P(S)P(W|C)P(B|S,C)P(A|B) \quad (3.22)$$

The inference over BN in practical usage of BN is very important. Main inference task in BN is to derive unobserved variable, by using probabilistic inference [Ben-Gal 2007].

3.2.4 Artificial neural networks

To solve optimisation problems, engineers often use artificial neural networks (ANN). According to [Prridy 2005], ANN can be defined as a parallel distributed structure for processing data (information) which has the following properties:

- It is a mathematical model inspired by the biological nervous system.
- It consists of numerous linked processing elements called *nodes*.
- The node dynamically responds to the input stimulus, and its answer completely depends on the local information contained in its environment.
- ANN has the property to learn, remember and generalise based on the training set.

ANN consists of a large number of neurons (connected nodes) that work in parallel (simultaneously) and are organized by some regular architectures [Liu 2003]. Updated from [Liu 2003], Fig. 3.5 shows a simple mathematical model of biological neuron, so called M-P neuron. In the ANN model i-th node calculates the weighted

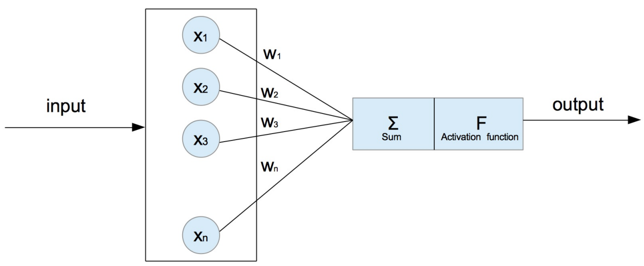

Figure 3.5: Neuron sketch (updated from [Liu 2003])

sum of input signals and as an output generates signal $y_i = 1$ or $y_i = 0$ depending on whether the weighted sum is greater or smaller than a predefined threshold θ_i [Liu 2003]:

$$y_i(t+1) = a(\sum_{j=1}^{m} w_{ij} x_j(t) - \theta_i) \tag{3.23}$$

where the activation function $a(f)$ is a unit step function [Liu 2003]:

$$a(f) = \begin{cases} 1 & \text{za } f \geq 0; \\ 0 & \text{za } f < 0. \end{cases}$$

Training an neural network is important for problems to be solved. For a specific task to solve a class of functions F, training denotes to find $f^* \in F$ which solves the task in some optimal sense, by using a set of observations [Liu 2003].

3.3 Meta-heuristics methods

Meta-heuristics have become extremely efficient in solving real problems based on optimisation. The basic requirement is to obtain solutions close to optimal in a reasonable time. In addition they can be used to accelerate other methods e.g., by providing a good initial solution [Davidović 2006].

3.3.1 Meta-heuristics in general

Generally speaking, meta-heuristics proved to be very successful as integral parts of the knowledge discover systems within the artificial intelligence [Davidović 2006]. According to [Davidović 2006], [Talbi 2009], [Glover 2003], [Crainic 2010], the most important characteristics of meta-heuristic methods include :

- *Simplicity*: They should be based on simple and easily understandable rules;
- *Precision*: The steps that describe meta-heuristic methods should be formulated in a precise mathematical terms, if possible;
- *Consistency*: All steps of methods should be in accordance with the rules which define meta-heuristics;
- *Efficiency*: The application of meta-heuristics to a specific problem needs to ensure getting solutions close to optimal for most of real examples, specially for official test examples(benchmarks) available in that class;
- *Effectiveness*: The method must provide an optimal or a solution close to the optimum in a reasonable executive time, for each specific problem;
- *Robustness*: The method should give equally good results for wide range of examples from the same class, and not only for some selected test examples;
- *Clarity*: It should be clearly described to be easily understood and, more importantly, easily implemented and used;
- *Universality*: The principles which define the methods should be general to assure easy application to new problems.

Meta-heuristics methods orchestrate an interaction between local improvement procedures and higher level strategies to create a process capable of escaping from local optima and performing a robust search of a solution space [Talbi 2009]. The procedures utilise one or more *neighbourhood structures* as the *means of defining admissible moves to transition from one solution to another, or to build or destroy solutions in constructive and destructive processes* [Talbi 2009].
Based on type procedure, the degree to which neighbourhoods are exploited varies. In the case of certain population-based procedures, such as genetic algorithms, neighbourhoods are implicitly (and somewhat restrictively) defined by reference to replacing components of one solution with those of another, by variously chosen rules of

exchange popularly given the name of "crossover." In other population-based methods, based on the notion of path relinking, neighbourhood structures are used in their full generality, including constructive and destructive neighbourhoods as well as those for transitioning between (complete) solutions.

3.3.2 Background to meta-heuristics

Meta-heuristics are designed to help in solving complex optimisation problems where other optimisation methods have failed to be effective or efficient [Crainic 2010]. They represent general sets of rules (the recipes) to build efficient optimisation methods [Crainic 2010], [Davidović 2006]. Most important challenge is adapting the meta-heuristics to a particular problem or a problem class [Talbi 2009]. Some of the widely used meta-heuristics methods are tabu search, genetic algorithms, simulated annealing. The structure of the search has many common elements over different methods. In each step of the search algorithm, there is always a solution (or a set of solutions) x_k, which represents the current state of the algorithm [Crainic 2010], [Davidović 2006].

3.3.3 Types of meta-heuristics methods

There are three fundamental classification criteria for meta-heuristics [Crainic 2010], [Talbi 2009], [Voss 2001]: solution treatment, number of solutions and development inspiration. According to the solution treatment, meta-heuristics are classified as constructive and improvement. According to number of solutions: single-solution or population based meta-heuristics are distinguished. Based on the development inspiration, nature-inspired methods are considered opposed to mathematically founded meta-heuristics [Talbi 2009], [Voss 2001]. These classes are not mutually exclusive [Voss 2001].

3.3.3.1 Constructive meta-heuristics

Constructive meta-heuristics construct solutions from their constituting elements [Voss 2001]. This is done by adding one element at a time to a partial solution, an operation that is also called *a move*. To improve the quality of the final solutions, some constructive meta-heuristics include a local search phase after the construction phase [Talbi 2009], [Voss 2001].

Ant Colony Optimization (ACO) is term used to denote constructive meta-heuristics that build solutions by mimicking the foraging behaviour of ants [Talbi 2009]. The ACO utilizes multiple artificial agents (known as ants) that construct solutions in parallel [Talbi 2009]. The ants construct solutions utilising learning, in such a way that elements that were present in high quality solutions will receive a larger probability of being selected as a result of their higher pheromone levels [Talbi 2009]. BCO and GRASP are also examples of constructive meta-heuristic methods.

3.3.3.2 Improvement

Improvement based meta-heuristics operate on complete solutions and apply various transformations to enhance them. In most of the cases, local search procedure is used as the enhancement transformation with an addition of appropriate mechanisms to escape from local optima. Some of the well known improvement based meta-heuristics are Simulated Annealing (SA), Tabu Search (TS), Genetic Algorithm (GA), Variable Neighborhood Search (VNS), recently developed versions of Bee Collony Optimization (BCO).
Tabu Search (TS) is among the most popular local search based methods. It was created by Fred W. Glover in 1986 [Glover 1986] and formalized in 1989 [Glover 1989]. TS is a meta-heuristic employing local search methods used for mathematical optimisation [Glover 1986]. In tabu search it is important to distinguish how solutions are selected from the neighbourhood [Glover 1986]. In each step of the algorithm, there is a list L_k of solutions that have recently been visited and are therefore considered as *tabu* for the next moves in order to avoid cycling.

3.3.3.3 Single solution meta-heuristics

Single solution based meta-heuristics are efficient in solving various optimisation problems in different domains. At each step, they operate with a single solution trying either to construct it using basic components and attained knowledge or to enhance the current best solution by the appropriate transformations. According to [Talbi 2009], some of them are: Variable Neighborhood Search, Fitness Landscape Analysis, Simulated Annealing, Tabu Search, Iterated Local Search, Guided Local Search.
Variable Neighborhood Search (VNS) [Hansen 2010], [Mladenović 1997], are designed for approximating solutions of discrete and continuous optimisation problems. It computes distant neighbourhoods of the current solution, and moves to a new one iff an improvement was made [Hansen 2010]. VNS is successfully applied to linear programming problems, integer programming problems, mixed integer programming problems, nonlinear optimisation problems [Hansen 2010].

3.3.3.4 Population-based meta-heuristics

Population-based meta-heuristics can be understandable as an operator on a population of solutions [Talbi 2009]. At the beginning, the population is initialized. Second step is that a new population of solutions is generated [Talbi 2009]. Finally, this new population is integrated into the current one using some selection procedures. The search process is stopped when a given condition is satisfied (stopping criterion) [Talbi 2009]. Algorithms such as Evolutionary Algorithms (EAs), Scatter Search (SS), Estimation of Distribution Algorithms (EDAs), Particle Swarm Optimization (PSO), Ant Colony Optimization(ACO), Bee Colony Optimization (BCO), and Artificial Immune Systems (AISs) belong to this class of meta-heuristics [Talbi 2009].

3.3.3.5 Nature inspired meta-heuristics

Many meta-heuristics are inspired by natural processes. According to [Talbi 2009], Evolutionary Algorithms and Artificial Immune Systems from biology; Swarm Intelligence Methods (Ant Colony Optimization, Bee Colony Optimization, and Particle Swarm Optimization) are proposed by analyzing processes in social sciences; and Simulated Annealing is taken from physics [Talbi 2009].

3.3.3.6 Mathematically based meta-heuristics

Contrary to the nature inspired, there are mathematically-founded meta-heuristic methods. They use the concepts of metric functions to measure the distance between various solutions [Talbi 2009]. More precisely, they utilize various mathematical transformations to modify existing solutions. Some of them are Tabu Search, Iterated Local Search, Variable Neighborhood Search, Fitness Landscape Analysis, Guided Local Search [Talbi 2009].

3.3.4 Further enhancement of meta-heuristics

3.3.4.1 "Hybrid" meta-heuristics

Hybrid meta-heuristics [Talbi 2002] combine algorithms such as population-based meta-heuristics, single-solution meta-heuristics, mathematical programming, constraint programming (CP), and machine learning techniques. As stated in [Talbi 2009], there are four different types of combinations:

- Combining meta-heuristics with (complementary) meta-heuristics.
- Combining meta-heuristics with exact methods from mathematical programming approaches that are mostly used in operations research.
- Combining meta-heuristics with constraint programming approaches developed in the artificial intelligence community.
- Combining meta-heuristics with machine learning and data mining techniques.

Hybridisation of meta-heuristics have two major issues such as design and implementation [Yang 2008]. Design includes functionality and architecture of the algorithm. The implementation includes the hardware platform, programming, model, and environment on which the algorithm is to be run [Yang 2008].

A typical example of hybrid meta-heuristic is *Greedy Randomized Adaptive Search Procedure* (*GRASP*) [Yang 2008]. It has two parts:

- Constructive.
- Improvement (local search).

Constructive is related to search diversification process, while improvement (local search) realizing intensification phase [Yang 2008].

3.3.4.2 Parallelization

Many difficulties arise during solving combinatorial optimisation problems, where the number of feasible solutions usually grows exponentially with the number of objects in the initial set. Parallelization of search procedures [Crainic 2010], is one approach that can remove this shortcoming. Based on [Talbi 2009], parallel and distributed computing optimise the design and implementation of meta-heuristic methods in order to reach the following goals :

- Speeding up the search (i.e., reducing the search time).
- Improving the quality of the obtained solutions.
- Improving the robustness;
- Solving large-scale problems.

To improve the final solution quality within a smaller amount of execution time, the parallel execution is enabled overefficient search through different regions of the solution space [Crainic 2010, Crainic 2005].

3.3.5 Multi-start meta-heuristics

One of the simplest way to create meta-heuristic method is to restart some heuristic optimisation procedure from randomly generated initial points [Glover 2003]. For example, Multi-start Local Search (MLS) is obtained when local search is used as the heuristic procedure [Glover 2003].

The best solution in a neighbourhood is called a local optimum (as opposed to a global optimum, which is the best among all solutions in the feasible domain) [Glover 2003]. When the current solution is a local optimum, a meta-heuristic uses a strategy to "skip" from this local optimum [Glover 2003]. Depending on the optimisation problem as a local search, one of the following methods: Pattern search, Newton's method, Levenberg-Marquardt method, Golden section search can be used [Glover 2003].

3.3.6 Genetic algorithm

Genetic algorithms (GA) are inspired by biological evolution based on Darwin's theory of natural selection. The GA tries to move from one set of feasible randomly generated solutions to another by applying genetic operation [Haupt 2004]. From the initial set of the randomly generated solutions, a subset of feasible solutions with the advantage of containing the best members is selected. Random processes are constantly used to generate new feasible solutions using solutions obtained in the previous iteration. The size of permissible set of solutions is the same in each iteration. For the next iteration, the best solutions are selected with larger probability in order to provide more accurate solutions. The process is repeated until the predetermined conditions of stopping are satisfied. Some of the advantages of GA are [Haupt 2004]:

- Can optimise both discrete and continuous problems;
- Does not require the original information or the nature of the problems which it resolves;
- Simultaneously searches from a wide sampling of the cost surface;
- Can be applied to large scale problems (with a large number of variables);
- Is suitable for parallel processing;
- Optimise variables with extremely complex admissible field (avoiding local minimum);
- May encode the variables so that the optimisation is done with the encoded variables;
- Can work with numerically generated data, experimental data, or analytical functions.

These benefits allow GAs to obtain excellent results when traditional optimisation methods are not able to do the same.

3.3.6.1 Definitions of terms

According to [Kumara 2007] the following fundamental (basic) terms in GA are defined:

- *Population*: Represents a set of solutions in the corresponding iteration, or a set of feasible solutions. N_p is the number of feasible solutions in the population, also called the population size.
- *Generation*: A single iteration in the process of GA is called a generation. Each generation is dealing with population of a size N_p (which is not the case in biology).
- *Chromosome*: A solution from the population is called a chromosome. For a given problem it represents the set of variables in a given solution. The variables in a chromosome do not have actual values; Instead they are coded [Whitley 1994]. The most commonly used coding is binary, however, integer or real values codes can also appear.
- *Gen*: The encoded value of each variable in the chromosomes is called a gen. Depending on the values that variables can take, GAs are divided into binary, integer and continuous. Gene of binary GA can take only values 0 and 1, while generally it can be any real number.

3.3.6.2 The structure of GA

There are three basic operators of GA [Davidović 2006]:

- selection
- crossover and
- mutation

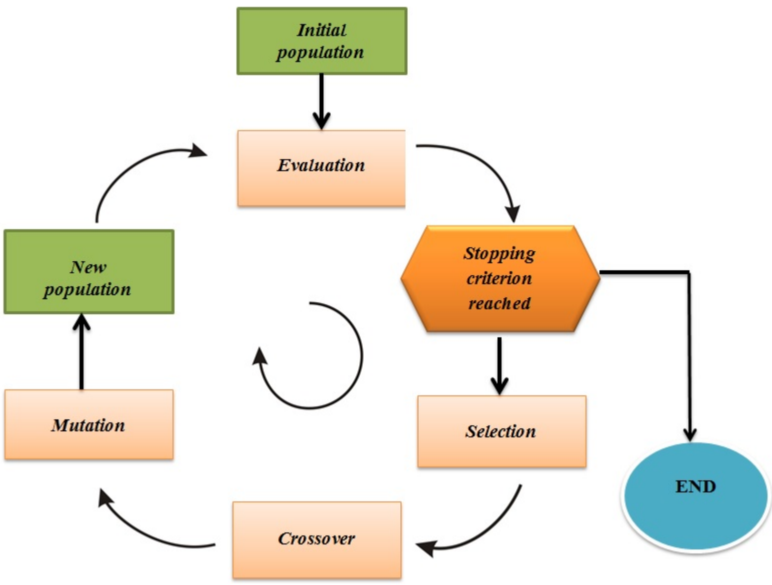

Figure 3.6: The basic operators of GA (updated from [Davidović 2006])

These operators are executed in each generation until the fulfillment of a predetermined stop conditions Fig. 3.6.

Selection is a process in which individual chromosomes are copied onto the next generation according to their fitness values. Fitness function $f(i)$ is assigned to each of the individuals in the population in order to define the quality of the corresponding solution. The increased value of this function indicates higher quality of the individual. This function can be either linear, nonlinear, differentiable or non-differentiable, with or without discontinuities. However, it has to be a positive function (because the algorithm looks only its value and no other property). A selection operation is usually implemented using the roulette wheel (wheel of fortune) Fig. 3.7.

Example: Considering a sequence of six chromosomes (binary strings) whose values of the fitness function are given in the Table 3.1:

As a total performance of the set of chromosomes is equal to 50, in order to select the individual which will be transferred to the next generation, a number from the

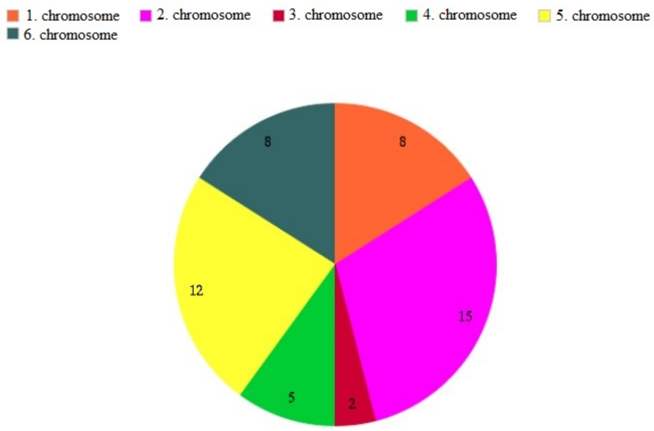

Figure 3.7: The wheel of fortune (updated from [Davidović 2006])

Table 3.1: Population of six individuals and their fitness values

ID	Chromosome	Value of the fitness function	Cumulative value
1	01110	8	8
2	11000	15	23
3	00100	2	25
4	10010	5	30
5	01100	12	42
6	00011	8	50

Table 3.2: Illustration of the selection process

Random number	26	2	49	15	40	36	9
Chosen chromosome	4	1	6	2	5	5	2

interval [0,50] should be randomly generated. The possible selections are illustrated in Table 3.2. The described process is entirely equivalent to turning the wheel of fortune. The value of fitness function directly affects the selection of chromosomes by using the roulette wheel.

Crossover: Selection takes into account good individuals of the population. Crossover is a GA operator with the role to create better individuals from the existing ones. In nature, the progeny has two parents and inherits genes from both of them. Similarly, the crossover operator takes two chromosomes (parents) and combines them with the probability p_c. The crossover chooses two chromosomes randomly and then, again randomly, chooses the intersection (*crossover*) *points*.

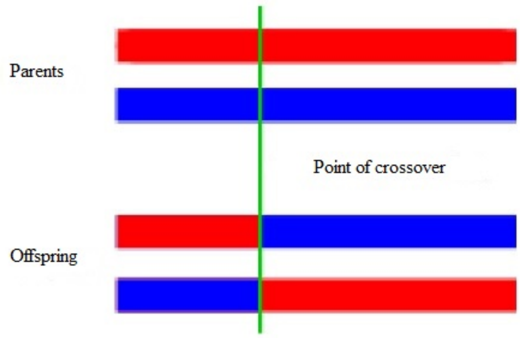

Figure 3.8: Crossover (updated from [Davidović 2006])

Finally, a random number from interval 0-1 is selected and, if it is smaller than crossover probability p_c, the offsprings are generated: part of genes are exchanged starting from the intersection point (Fig. 3.8).

Mutation: Although selection and crossover generally lead to better solutions, they do not bring new quality or information at the level of gene. As a source of different values of genes (some bits) mutation operation is used. With a low probability p_m, each gene is inverted in a chromosome (Fig. 3.9).

1100110111 → 1100100111

Figure 3.9: Mutation (updated from [Davidović 2006])

The parameters that define GA are [Davidović 2006]:

- N_p-population size;
- p_c-probability of crossover (crossover rate);
- p_m-probability of mutation.

As in the nature, mutation may lead to degenerative individuals (which will be quickly eliminated by a selection process), or may generate a completely new quality. Degree of mutations should be carefully selected because it is a random search operator.

Chapter 4

Optimisation Problems in Mechanical Engineering

The results from the recent literature related to the optimisation problems with gears, bearings, gear boxes, shafts and some other mechanical elements and assemblies will be described in this section. A detailed description of some optimisation problems related to the mechanical elements and assemblies is given below. In addition, in some cases, formal description i.e., mathematical models represented by a system of equations are also presented. Those models are important since the solutions lead to the optimal parameter values for a given element and/or an assembly. Each section discusses one class of problems, and gives an overview of works that are in some way involved in these problems.

4.1 Gears related problems

In connection with the gear optimisation, there is a large number of results in the recent literature. Gears are among the most important mechanical transmission elements, thanks to their high efficiency (which can reach over 99% in a gear pair). Limited resources on our planet force the efficiency related to the usage of fuel and force increasing of the working life and the power of mechanical systems.

One can find many studies on modelling and design of gears which are followed by the implementation of the corresponding optimisation process. Transmission errors, prediction of gear utilization with dynamic load, gear noise and optimal design are some of the major concerns for the designers.

Errichello [Errichello 1979], Ozguven and Houser [Ozguven 1998] investigated development of simulation models for static and dynamic analysis of different types of gears. Harris [Harris 1958] was the first one to study transmission errors. He showed that the behaviour or the gears at low speeds can be summarized in a set of statistical errors in the transfer curve.

In later years, Mark [Mark 1978], [Mark 1979] investigated the theoretical aspects, related to the vibration of gears. Expressions for the coefficients of Fourier-series for all components of the static transmission error are derived. They are obtained by two-dimensional Fourier transformations of local tooth-pair stiffness and stiffness-weighted deviations of tooth faces from perfect involute surfaces. Results are valid for arbitrary, specified tooth-face contact regions and include spur gears as the special case of helical gears with zero helix angle.

Kohler and Regan [Kohler 1985] discussed the error transmission of the gear from the part of errors transfer in the frequency domain using the analytical approach. Kubo and co-authors [Kubo 1991] evaluated the transmission error of cylinder gears using the contact model.

Friction gear substantially varies depending on the position of the load. Improvement of the analysis of compliance is made based on the works of Weber [O'Donnell 1974] and O'Donell [Heywood 1952]. Sensitivity analysis of friction is a modified version of Haywood's method [Kasuba 1961]. These results show the improvement of compliance sensitivity analysis of friction whit the estimates, using the tests, FE analysis and results of analytical transformations. Cockerham [Cockerham 1967] made a computer program for designing the 20-degree pressure angle of the gear, which ignores the cut off in gear flanks. In order to minimise size and weight, the optimisation models are presented in [Kamenatskaya 1975], [Andersson 1973].

The paper [Zhang 1999] provides analysis of teeth in contact and their load distribution for gears with crossed axes. This approach is based on a model of teeth in contact with the influences of modification on the teeth, with production errors and surface deformations. The load on teeth in contact is distributed along the lines of the tooth surface which coincides with the relative primary direction on surfaces of teeth in contact. Compared with the existing analyses of teeth in contact which assume that the contact teeth are solid, the proposed model gives a more realistic analysis of the errors at gear train, contact patterns and the distribution of the load. As a numerical example, a contact pair of helical gear with small angle of rotation is analysed using a computer program that implements the above described approach. The analysis makes the conclusion that helical gears with a small angle of rotation have similar coupling characteristics and load distribution as gears with parallel axes.

In [Flodin 2001] a simplified model for simulate wear in helical gear has appeared. The helical gear model is designed considering several independent thin sun gears. Sun gears have the common central axes along which they gradually turn relative to one another for the size of the angle of inclination of tooth. Load distribution along the tooth depends on the strength of the tooth, and the tooth strength is determined by the empirical model which was developed in 1988 in [Simon 1988]. The deformation on the tooth and the deviation from the ideal form caused by the tooth friction and wear, also affects for the load distribution. The modified Arcard's wear model and basic principle called "single point observation principle" are used in [Flodin 2001]. This principle means that the conditions in the particular points on the sides of gears are observed. It is assumed that the contact pressure in this simplified model is constant and equal to the middle Hertzian pressure. Sliding distance, the point at the sides of teeth compared with interactive surface during the working cycle is determined by analytic expressions previously used for worm gear pairs. Results obtained by simulating the simplified model are compared with the corresponding results obtained from the previously made simulation of an extended model.

In [Atanasovska 2006] the description of the procedure developed for investigating the influence of the addendum modification coefficient value on the load capacity of cylindrical involute gear is given. The model of meshed gears teeth contact in FEM is made to enable simultaneously monitoring of tooth flanks stress state and tooth roots stress state. In order to compare the stress states of meshed teeth's flanks and roots during the meshing period for gear pairs with different values of addendum modification coefficients, the comparative diagrams are made and shown.

A 3D FEM model for simultaneously monitoring strain and stress state of teeth flanks, teeth roots and parts of gears is simulated in accordance with the analysis of mathematical solution in [Atanasovska 2007]. Special types of contact finite elements that define contact of two deformable bodies are used for teeth flanks contact simulation.

Atanasovska [Atanasovska 2010] investigated the effects of the nominal load value on load distribution of simultaneously meshing gear teeth pairs, and on the involute gear load capacity. The presented results confirmed that the nominal load value has a significant influence on the gear load capacity calculations. In addition, a detailed description of the iterative numerical method, developed to support the modelling and analysis of load distribution in meshed gears using FEM, is provided.

In 2010, Zhang and co-autors [Zhang 2010] established a mathematical model for optimisation in designing gear pair used for the purposes of reducing the load (gear reducer). The model was developed with an aim to maximise the working life of bearings and minimise the size. Using a genetic algorithm and genetic tools in MATLAB for faster and more precise obtaining of optimal solutions, design efficiency and quality are improved.

In [Jabbour 2009] the analysis of plastic helical gears has been performed. This analysis takes into account the non-uniform load distribution along the contact line. The model proposed for the calculation of the load is based on the actual gear ratio in the meshed gear pairs. In addition, this model provides tooth bending and contact stress calculations along the contact line.

In [Pedrero 1996] a simple, analytical method for the estimation of the addendum modification factors for gears designed to have a specific balanced sliding is presented. The analysed model is valid for every value of the pressure angle and the addendum. The proposed method requires neither iterations nor tabular values, and that makes it efficient for computer applications.

J. Pedrero [Pedrero 2010] presented a model of non-uniform load distribution along the contact line, obtained from the minimum elastic potential energy criterion. This model, combined with the equations of Navier and Hertz, yields more realistic values of the bending and contact stresses. An approximate, accurate equation for the inverse unitary potential, allowing analytic calculations of the load per unit length at any point of the contact line and any position of the cycle of meshing, is also presented. The same equation, with a slight modification of the coefficients, is also valid for undercut teeth. Results have been validated by the comparison with some studies carried out by FEM.

The results related to a balanced model of load distribution along the line of contact obtained in [Pedrero 2010] are resumed in [Pedrero 2011]. The extension of model is established according to the criterion of the minimum elastic potential. This model combined with the Hertz equation yields more accurate values of the contact stress. The load per unit length at any point of the line of contact and any position of the meshing cycle has been described by a very simple analytic equation. Therefore, it was possible to carry out a complete study of the location and value of the critical contact stress. A recommendation for the calculation of the pitting load capacity of spur and helical gears is given.

In [Sánchez 2013], a non-uniform model of load distribution along the line of contacts in spur and helical gears, obtained from the minimum elastic potential criterion, combined with the equations of the linear elasticity has been used to evaluate the fatigue tooth-root stress. The critical value of the stress and the critical load conditions have been obtained, and a complete analysis of the tooth bending strength has been carried out. The load per unit length at any point of the line of contact and any position of the meshing cycle has been described by a very simple analytic equation. Therefore, it was possible to carry out a complete study of the location and value of the tooth-root bending stress. A recommendation for the calculation of the bending load capacity of spur and helical gears is given.

3D FEM for the conduction of the surface contact stress and the root bending stress calculations of a pair of spur gears with manufacturing errors, assembly errors and tooth modifications is developed in [Li 2007]. Positions of a pair of parallel-shaft spur gears are defined in a 3D coordinate system. Then, a tooth contact of the pair of gears is assumed on a reference face around the geometrical contact line. This tooth contact on the reference face is called a face-contact model of the gears. With this face-contact model, contact reference points on the reference face of one gear are firstly assumed, then a geometrical method that can consider the effects of manufacturing errors, assembly errors and tooth modifications are presented. Manufacturing errors, assembly errors, and tooth modifications on tooth contact are presented to find the responsive contact reference points on the tooth surface of the second gear (the mating gear).

A set of modern tools for the design of the gear trains is presented in [Ciavarella 1999]. Kinematic optimisation (minimisation and balancing specific slip), static stress analysis (in order to reduce stress concentration) and development of the initial cracks (fatigue assessment of previously existing effects) are taken into account. All three mentioned aspects are integrated into the software developed by the authors. In particular, boundary element method and FEM network are automatically generated to match the gears produced with standardised tools. Boundary element method is used only for automatic subcritical field propagation of initial cracks. On the other hand, FE networks are used only for cases with no cracks.

Gologlu and Zeyveli [Gologlu 2009] minimised the value of two-phase spiral gear wheel train taking normal module. The number of gear teeth, gear teeth width and the flexibility have been considered as variables while the contacts have been treated as constraints. A stochastic approach based on GA has been applied. The results have been compared with a developed deterministic design procedure.

Faggioni and coauthors [Faggioni 2011] presented the basic method of optimisation focused on reducing gear vibrations by teeth profile modifications. A multi-objective optimisation problem concerned with minimising the noise at a cylindrical gear pair was investigated in [Szabó 2005]. The results showed that in the case of the optimal geometry, noise level was reduced for 10% compared to the simpler geometry case.

In [Li 2008] the adaptive GA approach to multi-objective optimisation problem of gear train which has been used as a reducer has been proposed. GA in combination with suitable objective functions has been used in [Barbieri 2008] to find the corresponding modification on the gear teeth profile.

The face load factor is a common coefficient used in gear design standards that takes into account the uneven distribution of load across the face width of the gears caused by the mesh misalignment. In [Roda-Casanova 2013], a FEM that considers the gears and the corresponding shafts is proposed. The results obtained from the application of FE analysis to this model are compared with those obtained from application of the ISO Standard 6336 coefficient-based method (Method C). Finally, the influence of the gear shafts length, the face width of the gears, the relative position of the gears over their shafts, the ratio between the pitch radii of the gears and the their shafts radii, and the relation between the mesh misalignment and the face load factor, are investigated.

An approach for optimising geometry for the flank of a tooth by minimising the equivalent contact stress is given in [Guyonnea 2013]. The stress calculation method is based on Hertz theory. The geometric variation of the flank of the tooth is achieved relative to the involute profile. The optimum profile is obtained by Monte Carlo simulation. During this optimisation, a polynomial expression of the tooth geometry is used. The four characteristic contact points are the parameters influencing the simulation. The Monte Carlo simulation is compared with analytical propagation.

Rafiee and co-authors [Rafiee 2010] introduced an automatic feature extraction system for gear and bearing fault diagnosis using wavelet-based signal processing. Vibration signals recorded from the two experimental set-ups were processed for gears and bearing conditions. Four statistical features were selected: standard deviation, variance, kurtosis, and the fourth central moment of continuous wavelet coefficients of synchronised vibration signals (CWC-SVS). The mother wavelet selection was broadly discussed.

In [Mendi 2010a], the dimensional optimisation of motion and force transmitting components of a gearbox has been performed by GA. The aim was to obtain the optimal dimensions for gearbox shaft, gear and the optimal rolling bearing. With the optimisation of the gearbox components, the design with the smallest volume which can carry the system load was obtained. The results obtained by GA optimisation were compared to those generated by analytical methods. The comparison indicates that GA can be a reliable tool in machine element design problems.

4.2 Bearings related problems

A long fatigue life is one of the most important criteria in the optimum design of needle roller bearings (NRBs) [Waghole 2014]. Therefore, the dynamic capacity of the bearing is optimised. The nonlinear optimisation model has been formulated and threaded with Artificial Bee Colony Algorithm (ABCA), Differential Search Algorithm (DSA), Grid Search Method (GSM) and Hybrid Method (HM, a novel approach of combination of the ABCA/DSA and GSM). A total of four design variables corresponding to bearing geometry, which include the roller diameter, roller length, pitch diameter and number of rollers, were considered. In addition, three constraint parameters have been optimised. The constraint violation study is carried out to prioritise the constraints. The effect of the tolerance of design variables on the dynamic capacity were investigated by sensitivity analysis. The dynamic capacity of optimised bearings is found better than those specified in bearing catalogues.

Rolling bearings are widely used as an important component in most of the mechanical and aerospace engineering applications. The design of rolling bearings has to satisfy various constraints, e.g., the geometrical, kinematics and the strength, while delivering excellent performance, long life and high reliability. This invokes the need of an optimal design methodology to achieve these objectives collectively, i.e., the multi-objective optimisation. In [Gupta 2007], three primary objectives for a rolling bearing, namely, the dynamic capacity (C_d), the static capacity (C_s) and the elastohydrodynamic minimum film thickness (H_{min}) have been optimised separately, pair-wise and simultaneously using an advanced multi-objective optimisation algorithm: non-dominated sorting based genetic algorithm (NSGA II). These multiple objectives are conflicting and therefore Pareto optimality was investigated. A sensitivity analysis of various design parameters has been performed, to see changes in bearing performance parameters, and results show that, except for the inner groove curvature radius, no other design parameters have adverse affect on performance parameters.

A constraint nonlinear optimisation procedure based on GAs for designing rolling-element bearings has been developed in [Rao 2007]. Based on maximum fatigue life, the objective function and associated kinematic constrains have been formulated. The design parameters include the bearing pitch diameter, the rolling element diameter, number of rolling elements and inner and outer-race groove curvature radii. The constraints contain unknown constants, which have been given ranges based of parametric studies through initial optimisation runs. In the final run of the optimisation, these constraint constants are also included as design parameters. The optimised design parameters yield better fatigue life as compared to those listed in standard catalogues. A convergence study has been performed to ensure that the optimised design variables do not suffer from local extremes.

A new methodology for predicting crack initiation life was presented and validated experimentally in [Liu 2008]. The methodology considers the total fatigue life as the sum of crack initiation life and crack propagation life. It has been established that the crack propagation life can be estimated based on a modified Paris' law when the size of the crack is larger than a certain value. However, a verified method for estimating the crack initiation life with good accuracy doesn't exist. The proposed methodology for predicting the crack initiation life is based on a dislocation model, and the constants for the model are determined by the crack initiation lives obtained by a new approach. This new approach determines the crack initiation life by subtracting the predicted crack propagation life from the experimentally obtained total fatigue life. The developed crack initiation life model is combined with the crack propagation life model for the prediction of fatigue life. Experimental valuation shows that the developed model results in the predictions that are improved 14% with respect to the International Standard.

4.3 Optimisation of the roller bearings design

Different objective functions of rolling-element bearings may be proposed based on the operating requirements. The most important requirement is the length of a bearing life which is a consequence of a fatigue life. In a normal operating conditions of rolling-element bearings, the main reason of failure is contact fatigue.

To solve this problem for a given size of the bearing outline dimensions (i.e., bearing bore, d, and outside diameter, D), the dynamic load rating C_r should be maximised. The dynamic load rating C_r is defined as the constant radial load which a group of apparently identical bearings can endure for a rating life of one million revolutions of the inner ring (stationary load and stationary outer ring). The fatigue life, L, of the bearing (in millions of revolutions) subject to any other applied load P_r is given by:

$$L = \left(\frac{C_r}{P_r}\right)^a, \tag{4.1}$$

where $a = 3$ for radial ball bearings, while in the case of barrel bearing $a = \frac{10}{3}$.

Five design parameters for the given problem are represented as [Chakraborty 2003]:

$$X = [D_b, Z, D_m, f_o, f_i]^T,$$

with:
D_b- ball diameter
$Z-$ number of balls
D_m- pitch diameter
$f_o = r_o D_b-$ outer raceway curvature coefficient
$f_i = r_i D_b-$ inner raceway curvature coefficient
r_o and r_i outer and inner raceway groove curvature radius, respectively (see Fig. 4.1 and Fig. 4.2).

Based on the dynamic load rating, the objective function can be expressed as follows:

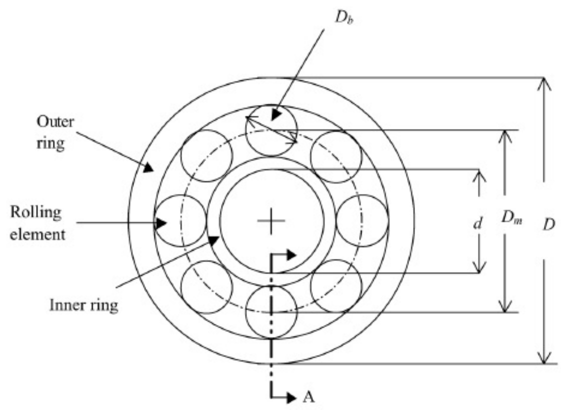

Figure 4.1: Ball bearing geometry taken from [Chakraborty 2003]

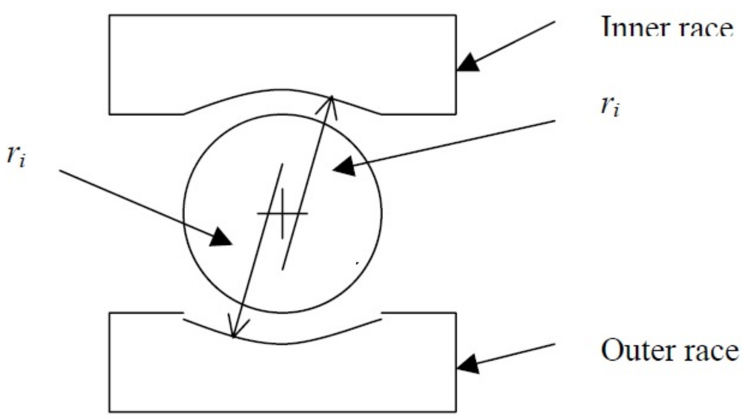

Figure 4.2: Cut sections of Bearing Races (Sectional plane A, see Fig. 4.1) taken from [Chakraborty 2003]

$$max[f(x)] = \begin{cases} \max[f_c Z^{23} D_b^{1.8}], & D_b \leq 25.4\ mm, \\ \max[3.647 f_c Z^{23} D_b^{1.4}], & Db > 25.4\ mm, \end{cases}$$

The problem is then maximised with respect to the constraints which are formulated based on the ISO standards. All constraints will be explained in more details in Chapter 6.

To satisfy various constraints in optimum design of rolling bearings, application of GAs is proposed in [Gupta 2007]. The challenge includes optimal design methodology in order to achieve multi-objective optimisation of dynamic capacity (C_d), static capacity (C_s), and the elastohydrodynamic minimum film thickness (H_{min}). Pair-wise and advanced multi-objective optimisation algorithm is used, also known as NSGA II (non-dominated sorting based genetic algorithm). The algorithm marks the objectives that need to be optimised, and generates initial population of size N. In [Gupta 2007], crossover probability was set to 0.85, mutation probability to 0.2, population size $N = 4500$ and a maximum number of generations $n_{max} = 50$.

In [Rao 2007], a constraint non-linear optimisation procedure based on genetic algorithm is developed for designing rolling element bearings. The algorithm starts by choosing population size, maximum number of generations, crossover probability, and mutation probability. After generation initial population, an increment $n = 1$ is defined. Reproduction, crossover, and mutation over population is applied while $n \leq n_{max}$. Termination of algorithm is forced by satisfying condition $n > n_{max}$.

4.4 Optimisation of the rotor system

Design of a rotor-bearing system is a challenging task due to various conflicting design requirements, which should be fulfilled. The paper [Costin 2010] considers an automatic optimisation approach for the design of a rotor supported on tilting-pad bearings.

A number of geometrical characteristics of the rotor, including the parameters defining the configuration of tilting pad bearings, are considered as design variables for the automatic optimisation process. The power loss in bearings, stability criteria, and unbalance responses are defined as a set of objective functions and constraints.

The assembled motion equation for the rotor-bearing system is formulated as:

$$[M]\{\ddot{q}\} + (\Omega[G] + [C])\{\dot{q}\} + [K]\{q\} = \{f(t)\} \tag{4.2}$$

where $\{q\}$ is the displacement vector, Ω is the angular speed of the rotor, $[M]$ is the inertia matrix, $[G]$ is the gyroscopic matrix, $[C]$ is the damping matrix, $[K]$ is the stiffness matrix, and f is external force vector.

The FEM is used to discretise the above given motion equation. In FEM code, a node with four degrees of freedom (two displacements and rotations in the lateral plane) is defined at each disk location, and the inertial properties of disks are introduced in the equations of motion of lateral rotor vibrations. The Euler-Bernoulli beam theory is employed to determine the mass and stiffness matrices of beam elements.

The code uses a vector approach to transform the second order equation (4.2) into the the first order equation:

$$\{\dot{x}\} = [A]\{x\} + \{b\} \tag{4.3}$$

where

$$[A] = \begin{bmatrix} 0 & I \\ -M^{-1}K & -M^{-1}(\Omega G + C) \end{bmatrix}, \tag{4.4}$$

$$\{x\} = \begin{pmatrix} q \\ \dot{q} \end{pmatrix} \tag{4.5}$$

and

$$\{b\} = \begin{pmatrix} 0 \\ M^{-1}f \end{pmatrix} \tag{4.6}$$

which can be easily solved.

Therefore, the considered single-objective optimisation problem can be presented in the following standard form:

$$\min f(x) \text{ where } x = \left[x^1, x^2, \cdots, x^n\right]^T \tag{4.7}$$

such that: $a_j \leq g_j(x) \leq b_j$, $j = \overline{1,m}$ and $h_k(x) = 0$, $k = \overline{1,l}$. The functions $f(x)$, $g_i(x)$ and $h_k(x)$ are the functions of the independent variables x^i. Function $f(x)$, known as the objective function, identifies the quantity to be minimised. The functions $g_i(x)$ and $h_k(x)$ are used to define constraints. The feasible space where one can find the best solution is usually n-dimensional rectangle defined by using their upper and lower bounds,

$$x_l^i \leq x^i \leq x_u^i, \text{ for } i = \overline{1,n}.$$

Formulated design optimisation problem is solved using the heuristic optimisation algorithms, namely GA and PSO. Numerical example for demonstration of the rotor modelling is given as: the centrifugal compressor, which is approximately 2.8 m in length and 954kg in weight, has been discretised into 34 beam elements (sections) with 35 nodes (stations) using an "in-house" code for the rotor dynamics analysis. The presented computational results show that both heuristic algorithms have found design solutions with better performance than the nominal design and GA have exhibited the fastest convergence.

4.5 Optimisation problem of working load

For the assessment of working capacity of machine parts, components and assemblies it is essentially important to analytically and experimentally determine the workload. For gear pairs it is very complex to determine the characteristics and the intensity of the load due to the strong influence of kinematic, geometric conditions and accuracy grade. Workload, based on the power and movement speed is referred to as *nominal load* [Ristivojević 2002]:

$$F_{nom} = f_1(P;\omega) \text{ or } F_{nom} = f_2(P;v)$$

Exact determination of the workload characteristic (**the actual workload**) is very complex in most cases. In addition, it is sometimes economically unjustified, because, apart from the theoretical research, it also requires the experimental evaluation. Therefore, the analysis is usually based on the applicable load of machine parts or assemblies. Nominal load is translated into relevant, based on the **factors of working conditions**, and **the applicable load factor** K.

$$F_{mer} = F_{nom}K;\ K \geq 1$$

In [Ristivojević 2002], this factor does not include the entire range of operating conditions. Instead, the theoretical and experimental investigations are conducted for a specific class of working conditions. The results are presented in the form of tables and/or diagrams. Typical examples of this procedure are the gears and roller bearings factors of working conditions:

In the teeth of gear pairs, there are two load distribution [Ristivojević 2002]:
(1) The load distribution at the meshed pair of teeth;
(2) The load distribution along the current line of contact.

Optimising the impact of these two distributions on the relevant tooth load is very important. Load transfer using the gear pair, implies a non-uniform load distribution during the meshing process. As a result of load transfer, contact and tooth-root stresses appear. These stresses are the main component of the budget in the design process calculation and the working life of a gear pair depends on them. According to the variation of gear parameters from nominal values and based on the number of teeth in contact, a non-uniform distribution is created and it takes into account various values along the line of contact as illustrated in the Fig. 4.3.

The equation for the maximum contact stress (4.8) and tooth-root stress (4.9) in the distribution loads, which are then taken in the calculations, and according to the standards [ISO 6336-2, ISO 6336-3], are as follows:

$$\sigma_H = Z\sigma_{H0}\sqrt{K_A K_V K_{H\beta} K_{H\alpha}} \leq \sigma_{HP} \tag{4.8}$$

$$\sigma_F = \sigma_{F0} K_A K_V K_{F\beta} K_{F\alpha} \leq \sigma_{FP} \tag{4.9}$$

In the above mentioned equations for contact stress calculation:

- σ_{H0} represents the nominal contact stress,
- σ_{F0} represents the nominal tooth-root stress caused by errors in the gearing teeth that are strained with static nominal torque;
- Z is a contact factor that turns the contact stress at the point of teeth contact during meshing to the internal contact points (different for pinion and wheel);

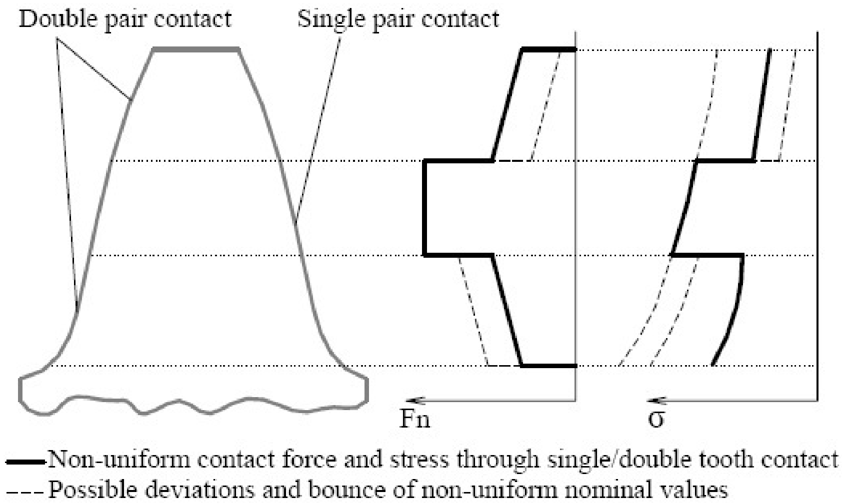

Figure 4.3: Contact model of a gear pair taken from [Milojević 2013]

- K_A is the application factor, which takes into account the load increment due to externally influenced variations of input or output torque;
- K_V is the dynamic factor, which takes into account the load increment due to internal dynamic effects;
- $K_{H\beta}$ is the face load factor for contact stress;
- $K_{H\alpha}$ is the transverse load factor for contact stress;
- $K_{F\beta}$ is the face load factor for tooth-root stress;
- $K_{F\alpha}$ is the transverse load factor for tooth-root stress;
- σ_{HP}/σ_{FP} is the permissible contact/bending stress.

The load distribution factor in the gears of spur and helical gear, according to the ISO standard [ISO 6336-1], takes into account many parameters. In addition, it is assumed to be variable along the contact line. Models provided in standards are not consistent with the experimental results due to some changes in gear meshing along the contact line that produce a non-uniform load distribution. Therefore, within calculations it is necessary to take into account some additional influencing factors for the contact and bending stress.

The factors responsible for the characteristic ratio of non-uniform load distribution at a meshed gear pair are $K_{H\alpha}$ and $K_{F\alpha}$ [Zhang 1999]. According to the standard [ISO 6336-1], these factors are calculated using the following equations:

$$K_{H\alpha} = K_{F\alpha} = \frac{\varepsilon_\gamma}{2}\left(0,9 + 0,4\frac{c_{\gamma\alpha}(f_{pb} - y_a)}{F_{tH}/b}\right), \tag{4.10}$$

for gears with total contact ratio $\varepsilon_\gamma \leq 2$,

$$K_{H\alpha} = 0,9 + 0,4\sqrt{\frac{2(\varepsilon_\gamma - 1)}{\varepsilon_\gamma}\frac{c_{\gamma\alpha}(f_{pb} - y_a)}{F_{tH}/b}}, \tag{4.11}$$

for gears with $\varepsilon_\gamma > 2$.

Howard [Howard 2001] details a simplified gear dynamic model aimed at exploring the effect of friction on the resultant gear case vibration. The model incorporates the effect of variations in gear tooth torsional mesh stiffness, developed using FE analysis, as the gears mesh together. The method of introducing the frictional force between teeth into the dynamic equations is given. The comparison between the results with friction and without friction is investigated using MATLAB and Simulink models developed from the differential equations. The effects of a single tooth crack on the frequency spectrum and on the common diagnostic functions of the resulting gearbox component vibrations are also shown.

Kahraman [Kahraman 1992] developed a FEM model of a geared rotor system on flexible bearings. The developed model includes the rotary inertia of shaft elements, the axial loading on shafts, flexibility and damping of bearings, material damping of shafts and the stiffness and damping of gear mesh. Coupling between the torsional and transverse vibrations of gears was considered in the model. A constant mesh stiffness was assumed. The analysis procedure can be used for forced vibration analysis of geared rotors by calculating the critical speeds and determining the response of any point on the shafts to mass unbalances, geometric eccentricities of gears, and displacement transmission error excitation at the mesh point. The dynamic mesh forces due to these excitations can also be calculated. The model has been applied to several systems for the demonstration of its accuracy and for studying the effect of bearing compliances on system dynamics.

Sabot and Perret-Laudet [Perret-Liaudet 1994] analysed the noise in the gearbox. Problems with the noise in the cabin of the car or truck can be attributed to transmission errors. That errors impact the increase of dynamic load on the gears, shafts, bearings and housing. Such problems were solved using FEM.

In [Sánchez 2008] a GA-based optimisation procedure for the design of gear transmissions is presented. For gear design, simultaneous discrete and continuous nonlinear related variables were used. The approach presented uses GAs as a tool to achieve the optimal (or at least near-optimal) designs.

4.6 Reliability assessment in mechanical systems

A application of Bayesian networks (BN) to the problem of reliability assessment of power systems is presented in [Yu 1999]. Efficient probabilistic inference algorithms in BNs not only permit computation of the loss of load probability, but also answer various probabilistic queries about the system. The advantages of BN models for power system reliability evaluation are demonstrated through examples. Results of a reliability case study of a multi-area test system are also reported.

The application of BN to reliability reassessment of structural system, with the incorporation of two important features combined with the branch-and-bound method to improve its efficiency is proposed in [Mahadevan 2011]. BNs for agricultural use are presented in [Kristensen 2002] and a prototype decision support system for growing malting barley with one main and two sub modules is given. The result of this system (network) is the set of probabilities for plant growing. In [Lauría 2006] Bayesian Belief Networks for real-world data representing an information technology environment, based on combining the techniques like artificial intelligence, statistics, and computer-based decision making were used. In [Chen 2012] a BN is developed to model an open press electric shock accident as a result of charged press enclosure. It offers a comparative tool for various safety designs of a machine system to guarantee the machine inherent safety.

In [Richard 2008] the use of multilevel neural networks is presented for Bayesian probabilities estimation with analysing the influence of network complexity, the amount of training data, and the degree to which training data reflect true likelihood distributions to results. The paper [Marquez 2010] presents the nodes of BNs extended with fault trees by defining the time-to-failure of the fault tree constructs as deterministic functions of the corresponding input components' time-to-failure. This approach is used for solving any configuration of static and dynamic gates with general time-to-failure distributions and can be used with any parametric or empirical distribution for the time-to-failure of the system components.

In [Krstić 2009], the probabilistic reasoning is used to analyse the preventive maintenance of the motor vehicles clutch. The main advantage of probabilistic reasoning is the possibility of rational decisions making, even when there is not enough information to completely describe a system.

The paper [Langseth 2007] gives a review of BNs used for reliability evaluation, especially, the discussion on the properties of the modelling framework that make BNs particularly well suited for reliability applications. In addition, it points to an on-going research that is relevant for practitioners in reliability.

Chapter 5

Optimisation Problems in Helical, Spur and Planetary gears

For the assessment of machine parts, components and assemblies, it is very important to determine their experimental and analytical characteristics depending on the set of relevant parameters. For the proper operation of assemblies and machines, it is extremely important to investigate all parameters in the gears pre-design. It is also of great importance to provide balanced distribution of workloads.

Today, optimisation is a common method to improve the properties of mechanical devices and it has received a significant attention. Many engineering problems have multiple objectives, including engineering system design and nonlinear optimisation. Solving engineering problems, especially design optimisation, in most of the cases involves multiple and conflicting objectives. Within this research, meta-heuristics are applied on solving optimisation problems at helical, spur and planetary gears.

One of the investigated problems within this research is the optimisation problem of the transverse load factor at helical and spur gears. Load transmission by gear pairs is followed by non-uniform load distribution in the meshing process. The opposite assumption, where the load factor does not change over time along the line of contact, was made. The goal was to identify parameters with the largest influence on violating this assumption. It was also necessary to determine the extent of their changes. For the purposes of developing this model, all parameters which determine transverse load factor, according to [ISO 6336-1], [ISO 6336-2], [ISO 6336-3], [ISO 1328], [ISO 53] and [ISO 21771] were considered as relevant. The proposed optimisation algorithm was based on GA and involved an additional local search optimisation procedure called at the end in order to improve the solution obtained by GA. Such a hybrid algorithm has 12 direct input variables affecting the objective function. The main procedure was divided into several modules: Calculation of geometry, Calculation of the stiffness and Calculation of the value of total contact ratio. Since the mathematical model of this problem is nonlinear and continuous, the corresponding computational methods, such as Newton-Raphson method and interpolation of three-dimensional function, were implemented. The obtained results are presented in [Milojević 2013].

Planetary gear trains take a very significant place among the gear transmissions which are used in many branches of industry such as automobile transmissions, aircrafts, marine vessels, machine tool gear boxes, gas turbine gear boxes, robot manipulators and other. Planetary gear trains have a number of advantages over the transmission with fixed shafts. The multi-objective nonlinear optimisation of planetary gear trains is considered here. The weighting method is used to approximate the Pareto set. This method transforms the multi-objective optimisation problem into a single-objective optimisation problem by associating each objective function with a weighting coefficient and then minimising the weighted sum of the objectives. The proposed GA-based approach has produced quite satisfactory results promptly supplying the designer with the preliminary design parameters of a planetary gear train for different gear ratios. The obtained results [Rosić 2011a, Rosić 2011b] showed that the genetic algorithm is useful for application in optimisation of planetary gears design.

5.1 Optimisation of transverse load distribution factor of helical and spur gears

The model of meshing, based on the assumption that the transverse load factor does not change over the time and along the line of contact, is adopted in this research. Moreover, the assumption includes that these load factors have the same value $K_{H\alpha} = K_{F\alpha} = 1$, for both double and single pair tooth-contact, in order to determine if there are still some deviations from the assumptions and the extent of their changes.

In addition, the rigidity of the pair of teeth was taken into consideration as a very meaningful parameter. Rigidity is influenced by a lot of input data that are used for optimisation (it is adopted that the gears are made from steel). A new approach to calculating the best values of all relevant parameters for meshing gears, in such a way that the load is uniform at any point of the line of contact, has been presented. All influential parameters were varied with the time, but the parameters of the basic rack were pre-approved from [ISO 53] and as such were considered to be constant input parameters.

This new method allows finding the optimal geometry with respect to many other relevant parameters based on meshing of helical and spur gears Fig. 5.1. The model is evaluated as the simulation of gear meshing along the line of contact.

5.1.1 Load distribution model of helical and spur gears

Load transmission by gear pairs is followed by non-uniform load distribution in the meshing process. As a result of load transmission, the root stress on the teeth contact surface occurs. These stresses are the main parameters in gear calculations, design procedures and period of exploitation.

Due to gear parameters deviations from nominal values and depending on the number of teeth pairs in contact, the appearing stress was followed by non-uniform load distributions and different values along the line of contact, as shown in Fig. 4.3.

The maximal contact stress and the tooth-root stress in load distribution, that

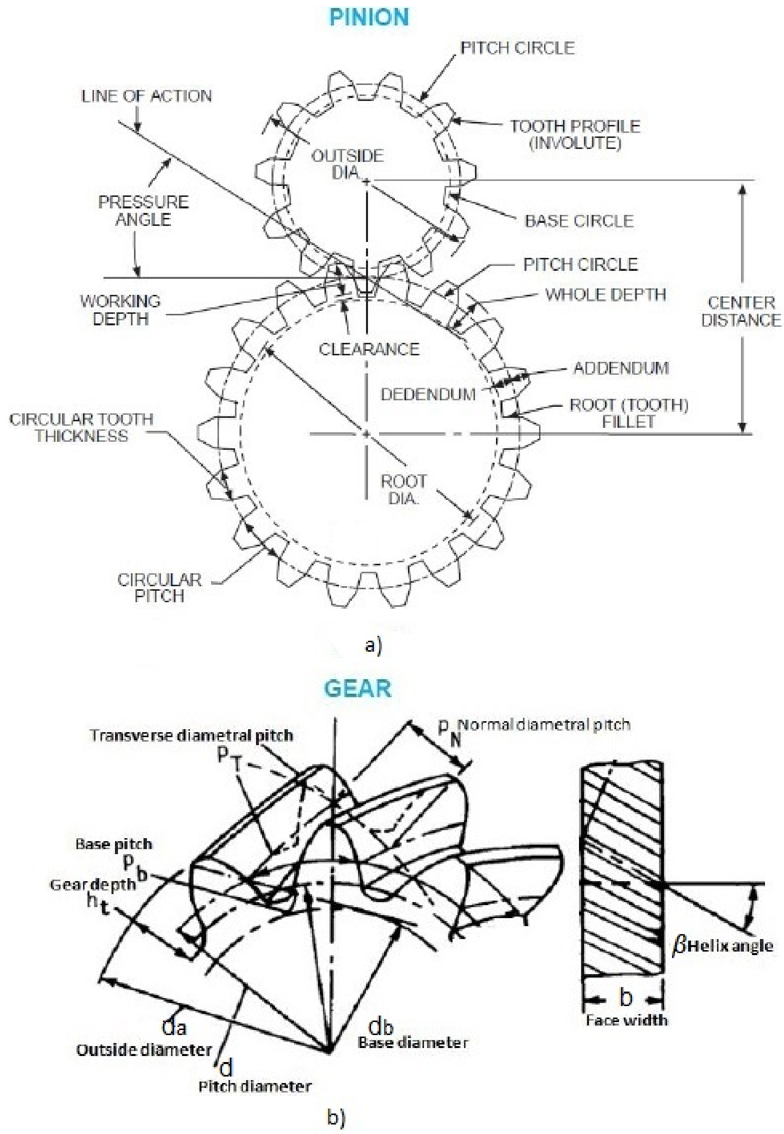

Figure 5.1: (a) Spur Gears Geometry (b) Helical Gear Geometry taken from [Akinnuli 2012]

were taken into consideration for further calculations, are calculated by Eqs. (4.8) and (4.9) according to [ISO 6336-2] and [ISO 6336-3].

Based on ISO standard [ISO 6336-1], transverse load factor of helical and spur gears depends on many parameters and it is assumed to be variable along the contact line. Models given by standardisation are not consistent with the experimental results since the changing meshing stiffness of the pair of teeth along the contact line produces a non-uniform load distribution. As a consequence, some additional load distribution factors need to be computed in order to determine bending and contact stresses [Zhang 1999]. According to [ISO 6336-1], these factors are calculated by equations (4.10) and (4.11).

The set of equations required for the computations of transverse load factors in helical and spur gears is given below.

$$z_2 = z_1 u \tag{5.1}$$

$$S_{rn} = S_r m_n \tag{5.2}$$

$$\tan(\alpha_w) = 2(x_1 + x_2)\frac{\tan(\alpha)}{(z_1 + z_2)} + \tan(\alpha) - \alpha \tag{5.3}$$

$$y_F = \arccos\left(\frac{z_1 + z_2}{2}\left(\frac{\cos(\alpha)}{\cos(\alpha_w)} - 1\right)\right) \tag{5.4}$$

$$\begin{aligned} C_{th} = (0.04723 + \frac{0.1551}{z_1} + \frac{0.25791}{z_2} - 0.00635x_1 - 0.11654\frac{x_1}{z_1} - \\ 0.00193x_2 - 0.24188\frac{x_1}{z_2} + 0.00529x_1^2 + 0.00182x_2^2)^{-1} \end{aligned} \tag{5.5}$$

$$C_r = 1 + \left(\frac{\log(odn)}{5^{\frac{S_r}{5m_n}}}\right); \tag{5.6}$$

$$d_1 = m z_1 \tag{5.7}$$

$$f_{pb} = \text{ adopted from the Table } K_v \text{ [Ristivojević 2006] based on the parameters } (x_1, m_n, d_1) \tag{5.8}$$

$$y_a = 0.075 f_{pb} \tag{5.9}$$

$$h_{fp} = 1.25m \tag{5.10}$$

$$C_b = \left(1 + 0.5\left(1.5 - \frac{h_{fp}}{m_n}\right)\right)(1 - 0.02(0.348888 - \alpha_{p_n})) \tag{5.11}$$

$$C_m = 0.8 \tag{5.12}$$

$$C = \begin{cases} C_{th}C_mC_rC_b, & x_7 \geq 100, \\ C_{th}C_mC_rC_bx_7^{0.25}, & x_7 < 100 \end{cases} \tag{5.13}$$

$$h_{a1} = (1 + y_F - x_2)m \tag{5.14}$$

$$h_{a2} = (1 + y_F - x_1)m \tag{5.15}$$

$$a = \left(\frac{(z_1 + z_2)}{2} + y_F\right) m \tag{5.16}$$

$$d_2 = mz_2 \tag{5.17}$$

$$d_{b1} = d_1 \cos(\alpha) \tag{5.18}$$

$$d_{b2} = d_2 \cos(\alpha) \tag{5.19}$$

$$c_1 = 0.2 \tag{5.20}$$

$$c_2 = 0.2 \tag{5.21}$$

$$h = (2.25 + y_F - (x_1 + x_2))\, m_n \tag{5.22}$$

$$d_{a1} = d_1 + 2h_{a1} \tag{5.23}$$

$$d_{a2} = d_2 + 2h_{a2} \tag{5.24}$$

$$d_{f1} = d_{a1} - 2h \tag{5.25}$$

$$d_{f2} = d_{a2} - 2h \tag{5.26}$$

$$\varepsilon_\beta = 0.9 \tag{5.27}$$

$$\varepsilon_\alpha = \frac{\sqrt{(\frac{d_{a1}}{2})^2 - (\frac{d_{b1}}{2})^2} + \sqrt{(\frac{d_{a2}}{2})^2 - (\frac{d_{b2}}{2})^2}}{m\pi\cos(\alpha)} \tag{5.28}$$

$$F_{th} = \frac{F_T}{b} = \frac{\frac{F_T K_A}{b}}{K_A} \tag{5.29}$$

$$c\gamma_\alpha = \begin{cases} C(0.75\varepsilon_\alpha + 0.25), & \varepsilon_\alpha \geq 1.2 \text{ and } \beta < 0.5233 \\ 0.9C(0.75\varepsilon_\alpha + 0.25), & \text{otherwise} \end{cases} \tag{5.30}$$

The presented equations cover helical and spur gears regardless of the value of helix angle β. The set of equations that differ with respect to the value of β is given below. For $\beta = 0$:

$$\alpha = \alpha_n \tag{5.31}$$

$$\alpha_p = \alpha_{p_n} \tag{5.32}$$

$$m = m_n \tag{5.33}$$

$$\varepsilon_\gamma = \varepsilon_\alpha. \tag{5.34}$$

For $\beta > 0$:

$$\alpha = a\tan\left(\frac{\tan(\alpha_n)}{\cos(\beta)}\right) \tag{5.35}$$

$$\alpha_p = a\tan\left(\frac{\tan(\alpha_{p_n})}{\cos(\beta)}\right) \tag{5.36}$$

$$m = \frac{m_n}{\cos\beta} \tag{5.37}$$

$$\varepsilon_\gamma = \varepsilon_\alpha + \varepsilon_\beta. \tag{5.38}$$

The above described equations are used within the optimisation procedure proposed in this research.

More precisely, the optimisation function is

$$\min f(V) = (1 - K_{H\alpha}(V))^2, \tag{5.39}$$

where

$$V = \{z_1, u, \beta, b, m_n, \frac{b_s}{b}, S_r, \frac{F_T K_A}{b}, Q, x_1, x_2, K_A\} \tag{5.40}$$

with respect to the following constraints:

$$x_1 \geq x_2, \tag{5.41}$$

$$-0.5 \leq x_1 + x_2 \leq 2.0 \tag{5.42}$$

The solution representation is given as a real valued vector of 12 components, with respect to the lower and upper bound. To complete the model, namely, to properly deal with the nonlinearity, some additional numerical procedures are required. They are described in the next sub-section.

5.1.2 Additional numerical calculations

5.1.2.1 Interpolation of three-dimensional function

In order to find a proper value of the transverse base pitch deviation f_{pb} (see (5.8)), it was necessary to perform interpolation based on the following variables: the accuracy grade, standard modulus and pitch diameter. Values for the appropriate base pitch deviation, corresponding to the above-mentioned variables are given in [ISO 1328], and the three-dimensional functionality is illustrated in Fig. (5.2). The accuracy grade and standard modulus are direct input values, while the pitch diameter is obtained by calculation. The interpolation is performed through a separate procedure (module), and the values obtained for the transverse base pitch are dynamically transferred to the main program.

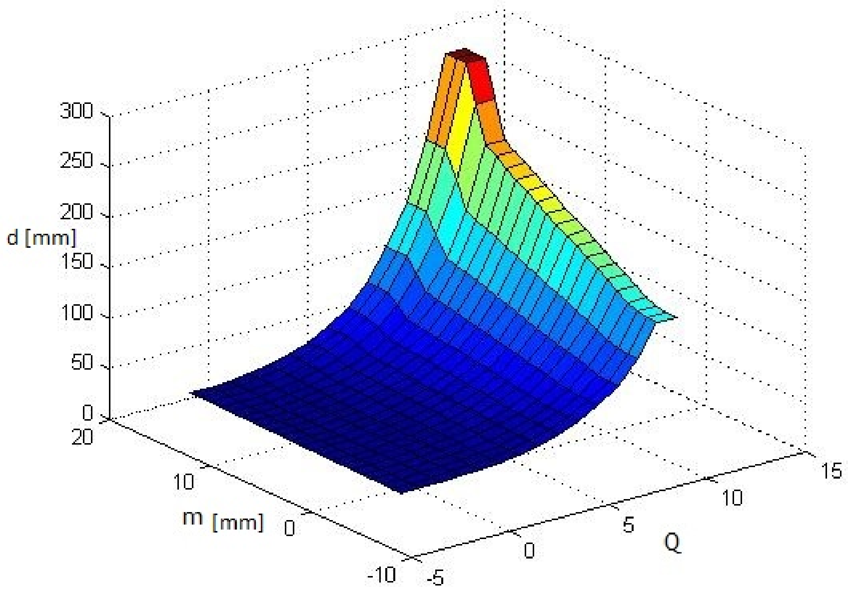

Figure 5.2: Functionality of accuracy grade, standard modulus and pitch diameter taken from [Milojević 2013]

Three-dimensional or trilinear interpolation is the name given to the process of linearly interpolating points within a box (3D) given values at the vertices of the box (see Fig. 5.3).

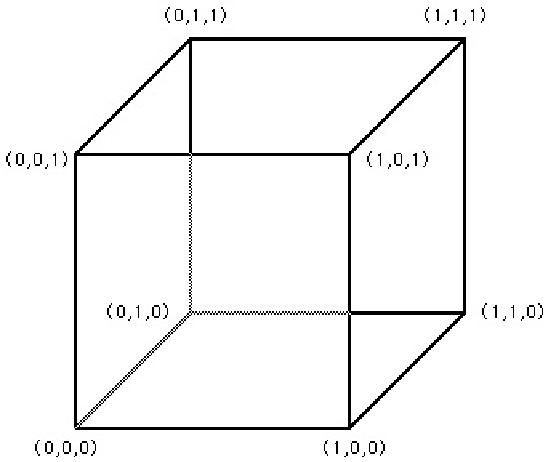

Figure 5.3: Trilinear interpolation box

The value at position (x, y, z) within the cube will be denoted V_{xyz} and is given by

$$\begin{gathered} V_{xyz} = \\ V_{000}(1-x)(1-y)(1-z)+ \\ V_{100}x(1-y)(1-z)+ \\ V_{010}(1-x)y(1-z)+ \\ V_{001}(1-x)(1-y)z+ \\ V_{101}x(1-y)z+ \\ V_{011}(1-x)yz+ \\ V_{110}xy(1-z)+ \\ V_{111}xyz \end{gathered}$$

5.1.2.2 Newton-Raphson numerical method for solving nonlinear equation

It is very difficult to find a root of a nonlinear equation algebraically. Using some basic concepts of calculus, there are several ways of numerically evaluating roots of complicated equations. For this purpose, commonly used Newton-Raphson method was selected.

In this research, Newton-Raphson numerical method was used to solve the nonlinear equation of the working pressure angle (5.3), which depends on the number of teeth on the pinion gear z_1, the number of teeth on the wheel gear z_2, profile shift coefficient of the pinion x_1 and profile shift coefficient of the wheel gear x_2.

All of these values are direct inputs and they dynamically change their values (in order to find optimum values). Solving of this nonlinear equation is performed in a separate procedure (module), and the values obtained for the angle are dynamically transferred in the main program.

5.1.3 Main procedure

Hybrid algorithm of this procedure, has 12 direct input variables affecting the output function, as shown in Fig. 5.4, where the main procedure is divided into several modules and each module will be explained in detail (see Fig. 5.5 and Fig. 5.6). The different parameters which impact transverse load factor of spur and helical gears are considered. These parameters are related to geometry, specific load distribution $\frac{Ft}{b}$, stiffness C, the application factor K_A and the accuracy grade Q.

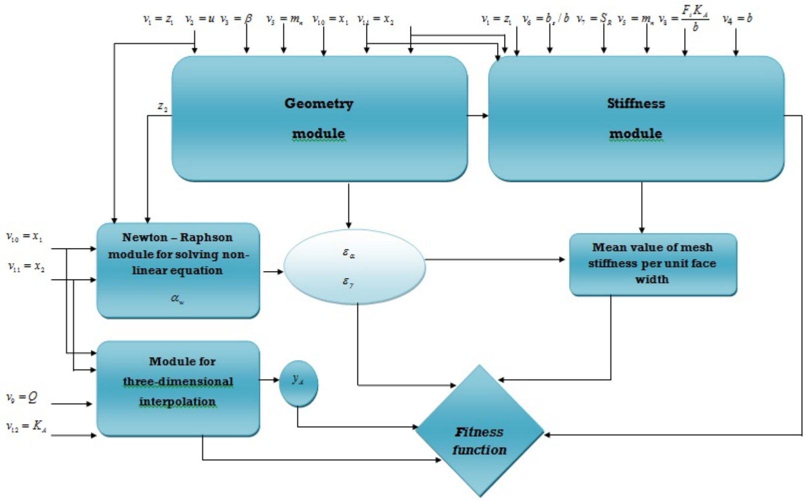

Figure 5.4: Main procedure algorithm (updated from [Milojević 2013])

The term *geometry* means optimisation against the number of teeth on the pinion gear z_1, gear ratio u (that gives the number of teeth on the wheel gear z_2, when multiplied by the number of teeth on the pinion gear), standard modulus m_n, profile shift coefficients of the pinion (x_1) and the wheel (x_2) and the helix angle β. All of these factors, together with the pressure angle and the normal pressure angle, directly impact the calculations of pitch diameters, addendum diameters, base diameters and root diameters.

To be more specific, the calculation of the geometry parameters performed by a geometry module (illustrated in Fig. 5.5), proceeds as follows: Starting from six (out of 12) direct input values v_1, v_2, v_3, v_5, v_{10} and v_{11}, the required output values for z_2, d_{f1} and d_{f2} are calculated according to the presented model. The intermediate dependencies are illustrated by solid arrows in Fig. 5.5. For gears with the helix angle $\beta > 0$, an additional input α is required and it is marked with the dashes line in Fig. 5.5

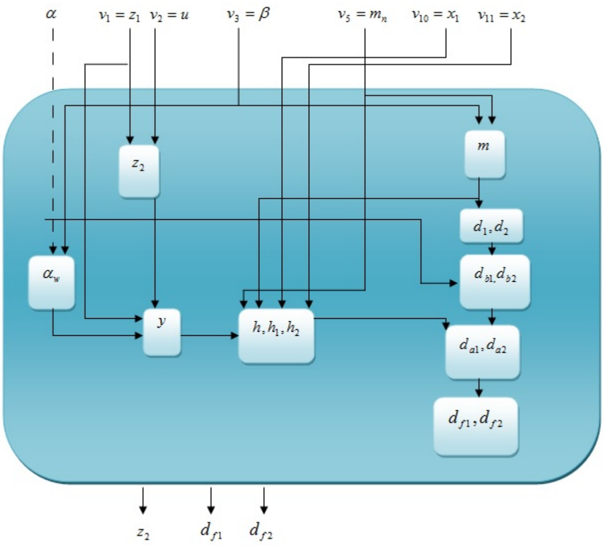

Figure 5.5: Geometry algorithm (updated from [Milojević 2013])

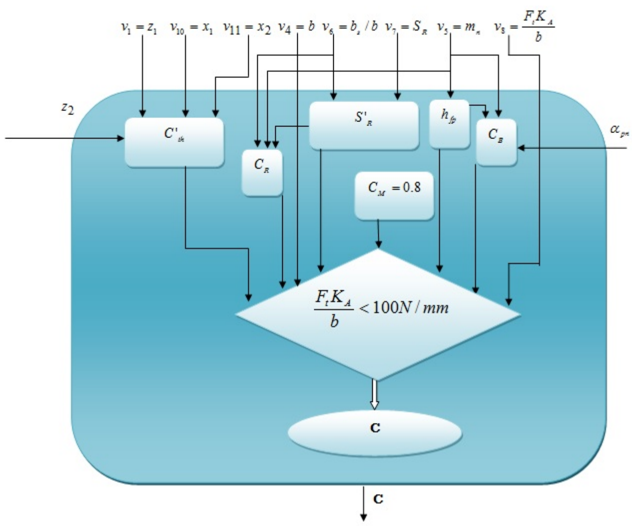

Figure 5.6: Stiffness algorithm (updated from [Milojević 2013])

The term *stiffness* means the optimisation against the basic rack factor C_B, correction factor C_M, gear blank factor C_R, theoretical single stiffness C_{th} and the helix angle β. The stiffness value also depends on whether the specific load is less than $100\frac{N}{mm}$ or not (see Fig. 5.6). Therefore, it is evident that specific load distribution is one of the most significant input values for the optimisation. In optimizing the basic rack factor C_B, the following parameters are taken into account: a standard modulus m_n, the normal pressure angle of the basic rack α_{pn} and addendum of the basic rack h_{fp}. In this optimisation, the gear blank factor C_R is presented as a function of gear rim thickness S_R, and ratio of central web thickness and gear width (b_s/b).

C_{th} is appropriate for solid disc gears and for the specified standard basic rack tooth profile. C_{th} for a helical gear is the theoretical single stiffness, relevant to the appropriate virtual spur gear [ISO 6336-1]. According to [ISO 6336-1], in this research, C_{th} is taken into consideration as a function of the number of teeth on the pinion gear z_1, the number of teeth on the wheel gear z_2, the profile shift coefficient of the pinion x_1 and the profile shift coefficient of the wheel gear x_2.

The obtained value for the stiffness is then used for calculation of the mean value of mesh stiffness per unit face width $C_{\gamma\alpha}$ within *Mean value of mesh stiffness per unit width* module (see Fig. 5.4). For the calculation of $C_{\gamma\alpha}$, additional intermediate value, the total contact ratio ε_γ is needed. Apart from the geometry, the working pressure angle α_w has a great influence on the total contact ratio ε_γ. The value of α_w is obtained from the *Newton-Raphson* module. Geometry module outputs enabled the calculation of another intermediate value, the transverse contact ratio of gears ε_α. Transverse base pitch deviation f_{pb} (the values may be used for calculations in accordance with ISO 6336, using tolerances complying with [ISO 1328]) is obtained in the *Module for interpolation of three-dimensional function*. The value of f_{pb} is used for the calculation of running-in allowance for a gear pair y_a and tangential load in a transverse plane, F_{tH}. Tangential load in a transverse plane, F_{tH} is a function of tangential load, application factor K_A, dynamic factor K_V and face load factor $K_{H\beta}$. The differences between the helical and spur gears are taken into account through a branch in the simulation depending on the value of the helix angle β.

At this point all is prepared to calculate the $K_{H\alpha}$, $K_{F\alpha}$ and K_V, i.e., the value of the objective function (5.39) can be determined (see Fig. 5.4).

5.1.4 GA implementation

The proposed optimisation algorithm is based on GA and involves an additional local search optimisation procedure called at the end in order to improve the solution obtained by GA. Therefore, it is referred to as *hybrid optimisation algorithm*, HGA (Fig 5.7).

Parameter settings of HGA during the optimisation process are shown in Table 5.1. The simulation was iterated three times, with different HGA setups in order to find the best convergence of the process. As the stopping criterion for HGA the maximum number of generations was set. In addition, the values for the maximum

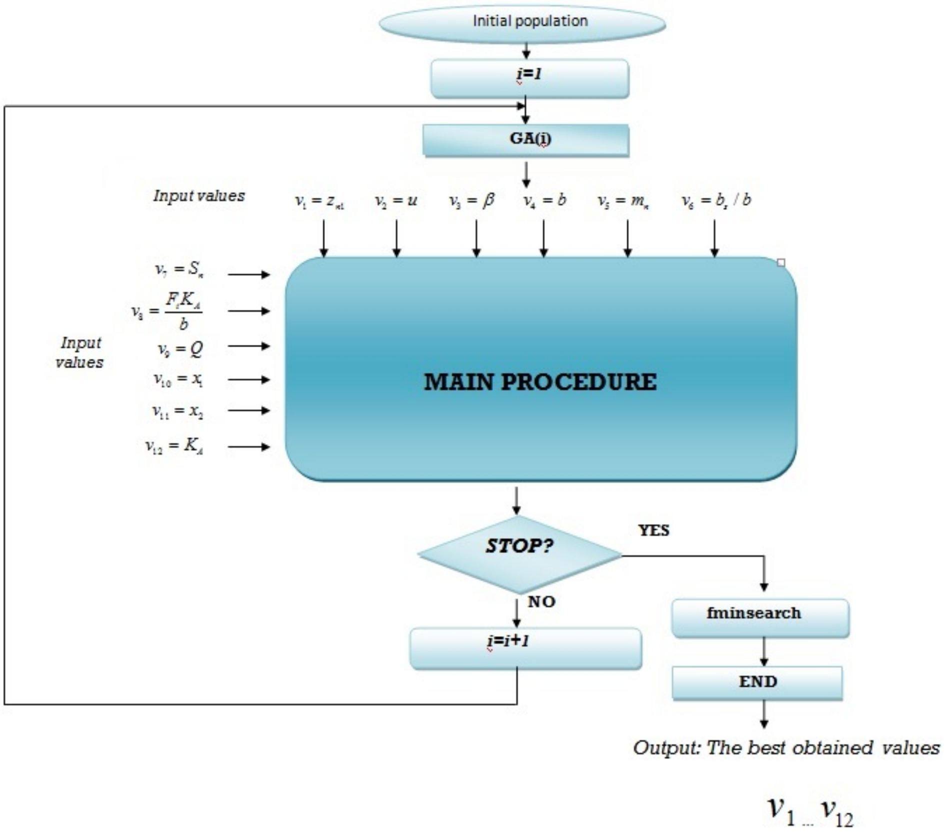

Figure 5.7: Hybrid optimisation algorithm (updated from [Milojević 2013])

number of stall generations and the function tolerance were specified. Chromosome is given in formula (5.40) and and it consists of 12 variables $v_1 \cdots v_{12}$.

Crossover combines two individuals, or parents, to form the offspring for the next generation [Matlab 2014]. In three execution of the genetic algorithm for solving these problem, different types of crossover were applied: Scattered, Single point, Two point.

Scattered crossover initially creates a random binary vector. Then, it combines genes with the value 1 from one parent and the genes where the vector is a 0 from the second parent, to form the child.

Single point crossover initially chooses a random integer n between 1 and the number of variables (which is 12). To form the child, it selects the vector entries with the indices less than or equal to n from the first parent and completes the chromosome with the remaining genes taken from the second parent.

Two point crossover establishes two random integers m and n between 1 and the number of variables l ($l = 12$). Furthermore, it selects first m genes and the last $l - n$ from the first parent, while the remaining genes are taken from the second. Finally, it merges these genes to form a single chromosome.

Mutation is deployed to make small random changes in the individuals of the population, which provide genetic diversity and enable the GA to search a broader space. In three execution of the GA, three different types of mutation were applied: Uniform, Adaptive feasible and Gaussian.

Uniform Gaussian mutation deploys a two-step process. First, it selects a fraction of the vector entries of an individual for mutation, where each entry has the same probability as the mutation rate of being mutated. Then, it replaces each selected entry by a random number selected uniformly from the range for that entry.

Adaptive feasible mutation randomly generates directions that are adaptive with respect to the last successful or unsuccessful generation. A step length is chosen along each direction so that linear constraints and bounds are satisfied.

Gaussian mutation adds a random number to each vector entry of an individual. This random number is taken from a Gaussian distribution centered on zero. The standard deviation of this distribution can be controlled with two parameters. The Scale parameter determines the standard deviation at the first generation. The Shrink parameter controls how standard deviation shrinks as generations go by. If the Shrink parameter is 0, the standard deviation is constant. If the Shrink parameter is 1, the standard deviation shrinks to 0 linearly as the last generation is reached.

5.1.5 Results

As shown in Table 5.1, three executions of the same simulation were performed with different settings of HGA to determine the best possible convergence of the obtained solution to the required minimum. In all executions HGA has reached the same final results, which is $f = 0$. The execution time and number of the stall generations were different (see Table 5.2). The obtained values for each variable are shown in

Table 5.1: Selected parameter values for HGA taken from [Milojević 2013]

Name of parameter	*First execution*	*Second execution*	*Third execution*
Population type	Double vector	Double vector	Double vector
Encoding	Binary	Binary	Binary
Scaling func.	Proportional	Rank	Top w. q. 0.4
Selection	Roulette	Stochastic uniform	Uniform
Elite count for reproduction	4	15	30
Crossover func.	Scattered	Single point	Two point
Population size	40	100	300
Mutation	Uniform	Adaptive feasible	Gaussian
Probability of mutation	Rate 0.1	-	Scale 1.0, Shrink 1.0
Max number of generations	1000	1000	1000
Hybrid func.	fminsearch	fminsearch	fminsearch
Func. tolerance	$10^{(-15)}$	$10^{(-15)}$	$10^{(-15)}$
	The remaining	parameters are set	to their default values
Stopping criteria	Maximal number of generations or number of stall generations (1000)		

the Table 5.3.

Table 5.2: HGA solver simulation properties taken from [Milojević 2013]

	First execution	*Second execution*	*Third execution*
Stopped in	1001	1001	1001
Final time of process	127 s	130 s	104 s
Convergence	Yes	Yes	Yes
Stopp. criteria	Stall generations	Stall generations	Stall generations
Stall gen.	900	86	924
Stall time	21 sec	23 sec	20 sec

Optimisation terminated: average change in the fitness value less than options

Convergence obtained in the first, second and the third execution is given in Figs. 5.8-a, 5.8-b, 5.8-c, respectively. The graph on the top of each figure illustrates changing in the value of the fitness function. The corresponding values of the variables generated as the output of HGA, are presented in the diagram below.

Table 5.3: Final results taken from [Milojević 2013]

Variable	Name	*First execution*	*Second execution*	*Third execution*
v_1	z_1	44	36	27
v_2	u	3.5	3	4
v_3	β	21.5°	28°	30°
v_4	b	69 mm	131 mm	54 mm
v_5	m_n	10	3	15
v_6	b_s/b	$0,34$	1.048	1.109
v_7	S_r	1	2.9	3
v_8	$\frac{F_t K_A}{b}$	1260 N/mm^2	1410 N/mm^2	1472 N/mm^2
v_9	Q	1	3	1
v_{10}	x_1	0.905	0.946	0.951
v_{11}	x_2	0.806	0.811	0.795
v_{12}	K_A	1	1.6	1
$f(x)$	$K_{H\alpha}$	1	1	1

Optimisation terminated: average change in the fitness value less than options

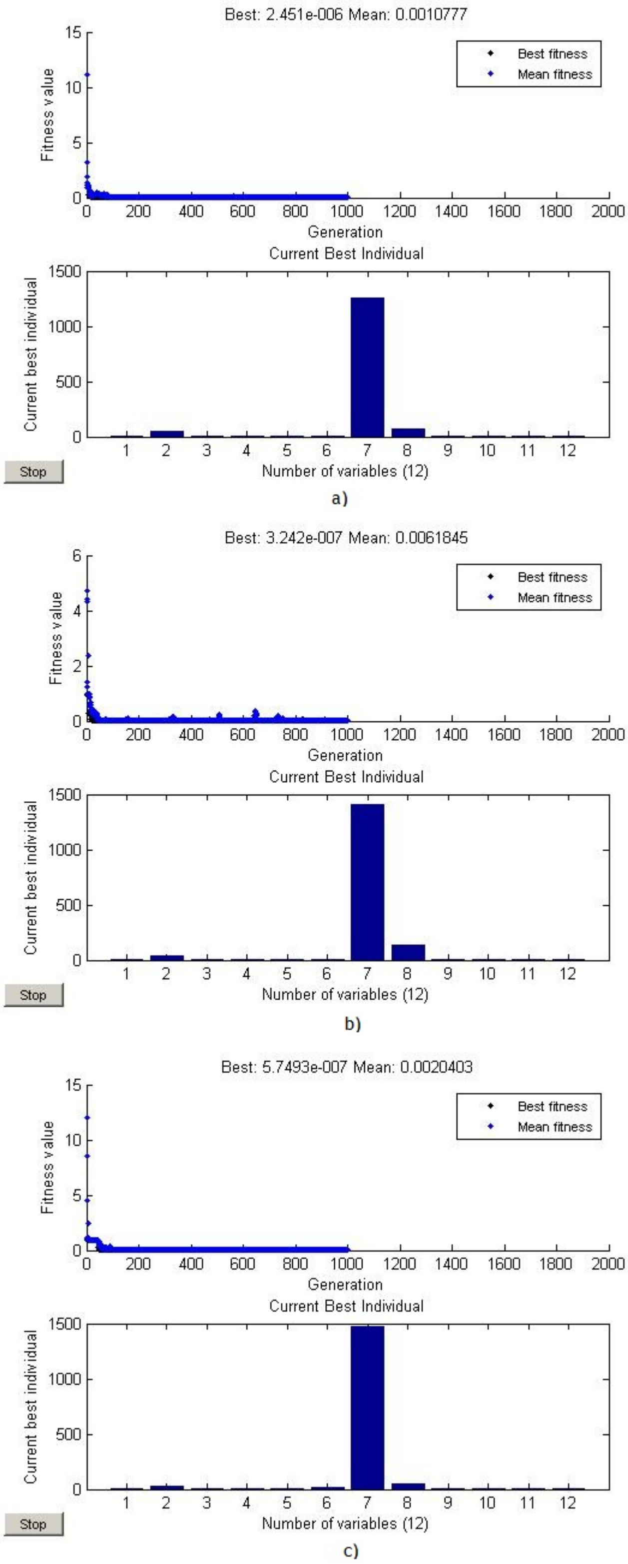

Figure 5.8: Convergence in generations taken from [Milojević 2013]

Although the load transverse factor could take any possible value, it always converged to the value 1, which affected the distribution of load and made it uniform, while the influential variables on the load transverse factor took the corresponding values.

The presented results show that the most influential variable on the value of load transverse factor is helix angle β, but, in addition, the profile shift coefficients x_1 and x_2 also affected changing the value of load transverse factor. HGA could have chosen the values for the helix angle β from the range $0° - 30°$ and it was usually assigned the values between $20° - 30°$ to make the load transverse factor value converging to 1. It that sense, it can be concluded that the value of helix angle is influencing the value of the load transverse factor. It is noted that for any number of teeth (from the range $18 - 54$) and any gear ratio (from the range $1 - 5$), this method achieves a value 1 of the load transverse factor, which therefore corresponds to the uniform load distribution.

The similar situation is with the specific load: at low values of the specific load distribution, load transverse factor converges to value 1. If the values of the specific load are higher, the load transverse factor converges to 0.5.

The influence of the profile shift coefficients of the pinion and wheel to the value of the load transverse factor is as follows: higher differences between x_1 and x_2 (with respect to the constraints (5.41) and (5.42)) provide convergence of the load transverse factor to 1. The described results are presented in [Milojević 2013]. Comparing to the state-of-the art results, it is shown that there are less simplification than in models [Pedrero 2011], [Pedrero 2010], [Sánchez 2013], [Zhang 2010]. This is the first time in the literature that all the geometrical parameters have been optimised in order to establish uniform load distribution. In most of the investigations, geometry parameters have been given as fixed inputs. In other papers, a maximum of 6 input parameters have been varied, while the work presented in [Milojević 2013] is optimising 12 different parameters, categorised in several modules. In [Pedrero 2010], non-uniform load distribution along the line of contact of spur and helical gear teeth, obtained from the minimum elastic potential energy criterion, has been investigated. Calculation of load distribution along the line of contact line has been validated by deployment of FEM. The minimum of elastic potential energy has been taken into account during calculations. For the purposes of this study, the geometry parameters such as: number of teeth on pinion and wheel, pressure angle, helix angle, module, profile shift coefficient, face width and center distance has been given as fixed inputs, while in [Milojević 2013] this coefficient has been varied. Also, in [Pedrero 2010] the combinations giving undercut teeth were avoided, which gives the specific contribution of the work in [Milojević 2013]. The study in [Pedrero 2011] has been restricted only to gears with a transverse contact ratio between 1 and 2, with non-undercut teeth. In [Milojević 2013], the algorithm has taken into consideration the differences in calculation based on the changing values of transverse contact ratio. Thanks to deployment of GA, results obtained for the load distribution are making it more uniform than the load distribution obtained in [Simon 1988], [Zhang 1999], which improves these results. In

[Zhang 2010], only six input parameters as face width, number of teeth on pinion, module, shaft diameters and distance between the bearings on reducer has been taken into consideration. In so far known investigations about load distribution at helical and spur gears, most of the authors have been considering only few isolated cases, while in [Milojević 2013] it is supposed that all geometrical characteristics can be varied. The combination of artificial intelligence methods such as GA, enabled discovery of the most influential parameters for creating uniform load distribution. Table 5.4 quantitatively compares results of 12 parameters taken as inputs/outputs for the optimisation of transverse load distribution factor of helical and spur gears. Criteria taken in this comparison are based on the common parameters used in [Milojević 2013], [Zhang 2010], [Pedrero 2010] and [Pedrero 2011]. The conclusions are that [Milojević 2013] analysed 4 additional parameters with respect to the publications [Zhang 2010], [Pedrero 2010] and [Pedrero 2011]. Ratio of central web thickness and gear width, gear rim thickness, accuracy rate, and application factor are not taken to be optimised in given research papers. These 4 parameters are the key parameters for accurate determination of stiffness. Due to the large number of influential parameters the calculation of teeth stiffness is very complex and mostly is not taken into consideration. This is pointing out to the advantage of the work presented in [Milojević 2013]. Columns [Milojević 2013], [Zhang 2010], [Pedrero 2010] and [Pedrero 2011] represent the ratio of input parameters in each execution. Results are showing that the optimised value of module m_n in [Milojević 2013] is 7.5 times improved with respect to [Zhang 2010]. In [Pedrero 2010] and [Pedrero 2011] the parameter m_n is taken as fixed input and therefore, it is not optimised at all. There is two times better optimisation of parameter z_1 with respect to the same one optimised in [Zhang 2010]. The best optimisation output in [Milojević 2013] is 44, but in [Zhang 2010] the best optimisation result is 34.

Table 5.4: Comparison of [Milojević 2013] with [Zhang 2010], [Pedrero 2010] and [Pedrero 2011]

Parameter	[Milojević 2013]	Improvement ([Milojević 2013])	[Zhang 2010]	Improvement ([Zhang 2010])	[Pedrero 2010]	Improvement ([Pedrero 2010])	[Pedrero 2011]	Improvement ([Pedrero 2011])
z_1	[3,100]	Ex. 1: 44. Ex. 2: 36 Ex. 3: 27	[17,50]	Ex. 1: 24. Ex. 2: 34	4:12	NO	Ex. 1: 15 Ex. 2: 30	NO
u	[1,5];h:0.5	Ex. 1: 3.5 Ex. 2 :3. Ex. 3:4	Fixed input	NI	Fixed in all Ex.	NO	Fixed in all Ex.	NO
β	[0,45]	Ex. 1:21.5 Ex. 2: 28 Ex. 3: 30	Fixed input	NO	Ex. 1: 0. Ex. 2:15 Ex. 3:15 . Ex. 4 :15.	NO	Fixed	NO
b	[10,200]	Ex. 1: 69mm. Ex. 2: 131mm Ex. 3: 54mm	NI	NO	Ex. 1: 20 . Ex. 2: 8 Ex. 3: 32. Ex. 4: 20.	NO	Fixed	NO
m_n	[1,50)	Ex. 1: 10 Ex. 2: 3. Ex. 3: 15	NI	Ex. 1: 2 Ex. 2: 2	All Ex.=4	NO	All Ex.: 5	NO
b_s/b	[0,1.3]	Ex. 1 :0.34 Ex. 2: 1.048 Ex. 3: 1,109	NA	NO	NA	NO	NA	NO
S_R	[1,3]	Ex.1 : 1 Ex. 2: 2.9 Ex. 3: 3	NA	NO	NA	NO	NA	NO
$\frac{F_t K_A}{b}$	[1,+00)	Ex. 1 : 1260 Ex. 2:1410 Ex. 3:1472	NA	NA	All Ex. :1000	NO	All Ex. : 3840	NO
Q	[1,8]	Ex. 1 : 1 Ex. 2 : 3 Ex. 3: 1	NA	NO	NA	NO	NA	NO
x_1	[-1,+1]; h=0.001	Ex. 1: 0.905 Ex. 2: 0.946 Ex. 3: 0.951	NO	NO	All Ex.: 0	NO	All Ex.: 0.1	NO
x_2	[-1,+1]; h=0.001	Ex. 1: 0.806. Ex. 2: 0.811. Ex. 3: 0.795	NO	NO	Fixed on : 0;0	NO	All Ex.: 0.1	NO
K_a	[1,2]	Ex. 1 :1 Ex. 2: 1.6 Ex. 3:1	NA	NO	NA	NO	NA	NO
α_w	Fixed on: 20	NO	NO	NO	Fixed on : 20	NO	Fixed on: 20	NO
Legend: Ex.=Execution NO= No output NA = Not analysed NO = Not optimised NI = No information								

5.2 Optimisation of planetary gear train using multi-objective genetic algorithm

The purpose of this research is to present a method for solving multi-objective nonlinear optimisation of planetary gear trains. Optimisation of gear train is based on the genetic algorithm (GA). The weighting method is used to approximate the Pareto set. This method transforms the multi-objective optimisation problem into a single-objective optimisation problem by associating each objective function a weighting coefficient and then minimising the weighted sum of the objectives.

5.2.1 Planetary gear trains

Planetary gear trains are among the most significant gear transmissions. They are used in many branches of industry such as automobile, aircrafts, marine vessels, machine tool gear boxes, gas turbine gear box, robot manipulators, etc. Planetary gear trains have a number of advantages over the transmission with fixed shafts [Xu 2007].

The advantages of such a transmission are following:

- Under similar operating conditions the planetary transmissions serve longer and produce less noise compared to the fixed shaft transmissions;
- This power transmission unit can handle larger torque loads relative to its compact size than any other gear combination in standard transmission;
- Improved efficiency of a gearing system can reduce the requirements on the capacity of the lubrication system and the gearbox lubricant, thereby reducing the operation costs of the system;
- Efficiency prediction can assist in estimating the power requirements during the design stage of a machine and thus ensuring that the system operates reliably. It can also assist in estimating the power output for a given power input;
- The input and output shafts are concentric, so no bending moments or torques are created from radial forces;
- High overall transmission ratio speed.

A planetary gear train consists of three different types of gears: *planet, sun* and *internal* gear (see Fig 5.9).

The effect of instantaneous efficiency of an involute gear drive was studied in [Radzimovsky 1973] and [Anderson 1986]. The problem of efficiency of the planetary gear train was studied experimentally in [Kasuba 1962]. Gearbox efficiency optimisation was not normally included in the gearbox design methodology in the past due to the lack of a technique that could assess a large number of design variables.

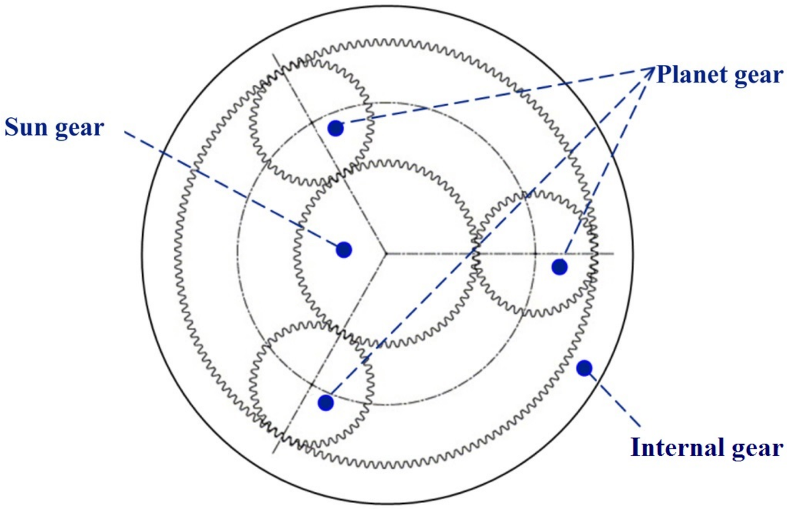

Figure 5.9: Presentation of an elementary planetary gear set having three planet gears taken from [Sun 2013].

The introduction of a larger number of criteria considering the desirable performances, even the conflicting ones (axial distance - efficiency), represents a significant step towards the reality of a planetary gear train model solved by the multi-objective optimisation methods.

There are many papers that present various approaches to finding the Pareto optimal front, mostly based on the evolutionary algorithms. Articles [Haupt 2004, Martinez 2009, Holland 1975] are related to a numerical analysis based upon GA, where the Pareto GA method is defined as a method based on the characteristic of GA to search for non-dominated solutions. GA starts by generating a random initial population. Using an iterative procedure, the current population is updated and the next population is created by using GA operators, namely: selection, mutation and crossover.

In this GA application, *single point crossover* is applied in order to combine two parents to form a child, for the next generation. *Single point* type of crossover is working in the way that, initially choose a random integer n between 1 and the number of variables (which is 10 for this case). To form the child, it selects the vector entries with indices less than or equal to n from the first parent and selects remaining genes from the second parent. To perform a mutation process, the *adaptive feasible mutation* was deployed. This type of the mutation randomly generates directions that are adaptive with respect to the last successful or unsuccessful generation. A step length is chosen along each direction so that linear constraints and bounds are satisfied [Matlab 2014].

The design of planetary gear trains requires a range of geometrical and kinematic conditions in order to perform the mounting and appropriate meshing of the gears during their work. It is necessary to express the above requirements in terms of the corresponding functional constraints, whereby all the relevant values of the gears, and planetary gear trains as a system, are defined in advance. In developing the optimisation model, one must start with very strict engineering requirements, which a planetary gear train should fulfil with respect to the efficiency, volume, factor of safety, etc. Based upon the defined requirements, it follows that it is practically impossible to describe a planetary gear train regarding the desirable performances with a single criterion.

5.2.2 Planetary gear train efficiency

According to their kinematic structure, planetary gear trains are complex toothed mechanisms which can be decomposed into external and internal toothed gears with the corresponding interaction. Each *planet* gear has a supporting link, called the carrier or arm, which keeps the centre distance between the two meshing gears constant. Planets are free to rotate with respect to the carrier. The gear trains in operation are characterised by losses in the mechanical energy arising as a consequence of the friction between the contact surfaces of the meshing teeth and the friction in the bearings. The analysis considers sliding losses, which are the result of the friction forces developed as the teeth slide across each other, the rolling losses resulting from the formation of an elastohydrodynamic film. The contact starts at the intersection of the tip diameter of the internal gear with the path of contact at A_2 (Fig. 5.10). The path of contact is tangent to the base circles of two gears. The contact ends at the intersection of the tip diameter of the external gear with the path of contact at E_2. In this boundary case, the addendum circles of radii r_{a1} and r_{a2} cross the motionless points A_2 and E_2, respectively. In order to evaluate the efficiency of an internal gear pair, one must consider the equilibrium of the gears. Figure 5.10 shows the normal forces F_n, the rolling friction forces F_R, and the sliding friction forces F_m, which suffices 1 for teeth in the path of approach and 2 for teeth in the path of recess.

On the basis of the models developed for a gear pair with external and internal gearing, the efficiency of a planet gear train can be determined. The power losses within the gears are expressed by means of the efficiency. The instantaneous efficiency for internal gear at any particular instant, from the relevant input torque T_1, is determined according to the expression:

$$\eta_i = \frac{T_2}{T_1}\frac{1}{u_{gb}^H}. \tag{5.43}$$

where T_2 is the output torque acting on the wheel; T_1 - input torque acting on the pinion; u_{gb}^H - relative gear ratio. The overall efficiency for gearing under consideration may be written as follows:

$$\eta_{gb}^H = \frac{1}{l}\int_{A_2}^{E_2} \eta_i d_\xi. \tag{5.44}$$

where l is the length of the path of contact A_2E_2; ξ - path of contact distances. The instantaneous frictional force due to the sliding of two gear teeth against each other

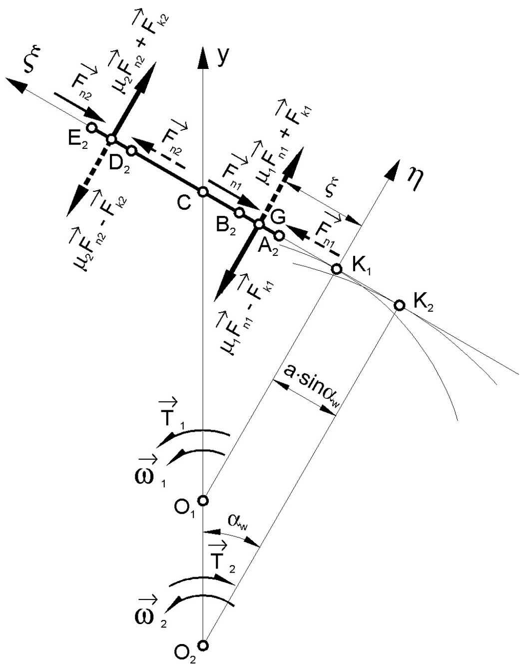

Figure 5.10: Forces between gear teeth taken from [Rosić 2011a]

is:

$$F_\mu(\xi) = F_n K_\alpha(\xi). \tag{5.45}$$

The friction coefficient is calculated by the method of Benedict and Kelley for mineral oil [Benedict 1961]:

$$\mu(\xi) = 0,0127 \log \left(\frac{\frac{29,66}{b} F_n}{\eta \nu_{kl} \nu_{ko}^2} \right). \tag{5.46}$$

where ν_R is the rolling velocity, ν_S is the sliding velocity, h is the fluid dynamic viscosity and b is the face width of gear. Since elastohydrodynamic lubrication film is minimal, the corresponding instantaneous force is given by the following equation [Anderson 1986]:

$$F_r = C_1 h(\xi) b. \tag{5.47}$$

where $C_1 = 9 \cdot 10^7$ is a constant of proportionality.

The gear contact minimum film thickness is calculated by the method of Dowson and Higginson [Dowson 1977]

$$h(\xi) = 1,6\alpha^{0,6}(\eta \nu_R)^{0,7} E^{0,003} \frac{R^{0,43}}{F_n^{0,13}} \tag{5.48}$$

where α is the viscosity-pressure coefficient of lubricant; R - the effective radius of curvature; E - the Young modulus of gear material. For convenience, the output torque of the train is assumed to be constant. From the equilibrium of gears, we have:

$$F_{n1} = \frac{T_1 - p_1 F_{R2} - \xi F_{R1}}{d_{b1} + \mu_2 p_1 - \mu_2 \xi}, \tag{5.49}$$

where

$$p_1 = \xi + p_b;$$

$$p_2 = a \sin \alpha_w + \xi + p_b;$$

$$p_b = m\phi \cos \alpha - \text{ base pitch };$$

$$p_3 = a \sin \alpha + \xi.$$

The output torque on the driven gear at any any point of time can be expressed in the form:

$$T_2 = F_{n1} d_{b2} + p_2(\mu_2 F_{n1} - F_{R2}) - (\mu_1 F_{n1} + F_{R1}) p_3. \tag{5.50}$$

On the basis of the models developed for a gear pair with external and internal gearing, the efficiency of a planetary gear train may be expressed as follows:

$$\eta^b_{aH} = \frac{1 - \eta^H_{ab} u^H_{ab}}{1 - u^H_{ab}}, \tag{5.51}$$

where

$$\eta^H_{ab} = \eta^H_{ab} \eta^H_{gb}. \tag{5.52}$$

η^H_{ab} is the relative efficiency for gear pair $a - b$; η^H_{gb} - relative efficiency for gear pair $g - b$; u^H_{ab}-relative gear ratio.

Based upon the developed models, computer programs for instantaneous efficiency determination have been devised. The numerical results used for the determination of the instantaneous efficiency of a gear pair with internal gearing are shown in Fig. 5.11.

In order to achieve the maximal possible value of efficiency, the following objective functions have to be considered:

- the centre distance

$$f_1(X) = \frac{m_n z_a}{\cos \beta}(1 + u^H_{a-g}) \frac{\cos \alpha_t}{\cos \alpha_{wt}} \tag{5.53}$$

- the efficiency

$$f_2(X) = \frac{1 - \eta^H_{ab} u^H_{ab}}{1 - u^H_{ab}} \tag{5.54}$$

- the contact ratio

$$f_3(X) = \varepsilon_{\alpha_{(a-g)}}(x) \tag{5.55}$$

- the pressure angle

$$f_4(X) = \alpha_{w_{(a-g)}}(x) \tag{5.56}$$

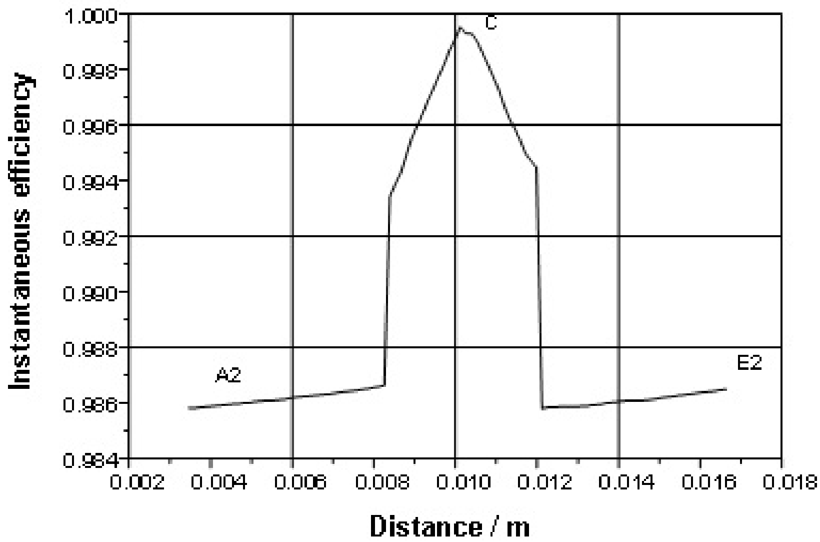

Figure 5.11: Instantaneous efficiencies during the contact period taken from [Rosić 2011a]

- the safety factor for bending stress

$$f_5(X) = S_{F(a)}(x) = \frac{[\sigma_f]_{M(a)}}{\sigma_{F(a)}} \tag{5.57}$$

- the safety factor for contact stress

$$f_6(X) = S_{H(a)}(x) = \frac{[\sigma_H]_{M(a)}}{\sigma_{F(a)}} \tag{5.58}$$

- the volume of material used for gears

$$f_7(X) = V(x) \tag{5.59}$$

- the safety factor for bending stress

$$f_8(X) = S_{F(b)}(x) = \frac{[\sigma_F]_{M(b)}}{\sigma_{F(b)}} \tag{5.60}$$

- the outer diameter

$$f_9(X) = D_{out}(x) \tag{5.61}$$

The intermediate parameter values are calculated as follows:
- the transverse contact ratio

$$\varepsilon_\alpha = \frac{0.5(\sqrt{d_{a(a)}^2 - d_{a(b)}^2} + \sqrt{d_{a(g)}^2 - d_{b(g)}^2}) - a \sin \alpha_{wt}}{\pi m_t \cos \alpha_t} \tag{5.62}$$

- the function of the pressure angle

$$\text{inv}\alpha_{wt} = 2\frac{x_a + x_g}{z_a + z_g}\tan\alpha_n + \text{inv}\alpha_t \tag{5.63}$$

- the tooth root stress for the sun gear

$$\sigma_{F(a)} = \frac{F_t}{bm_n}Y_{Fa}Y_{Sa}Y_{\varepsilon}Y_{\beta}K_A K_V K_{\alpha}K_{\beta} \tag{5.64}$$

- the critical root stress

$$[\sigma_F]_M = \sigma_{Flim}Y_{ST}Y_{NT}Y_{\delta relT}Y_{RrelT}Y_X \tag{5.65}$$

- the factor of safety from breakage

$$S_{F(a)} = \frac{[\sigma_F]_{(a)}}{\sigma_{F(a)}} \geq S_{Fmin} \tag{5.66}$$

- the effective contact stress

$$\sigma_H = Z_H Z_E Z_{\varepsilon} Z_{\beta}\sqrt{\frac{F_t}{bd_{(a)}}\frac{u+1}{u}K_A K_V K_{H\alpha}K_{H\beta}} \tag{5.67}$$

- for the sun gear - planet gear

$$\sigma_{Ha} = \sigma_{Hg} < [\sigma_H]_M = min\{[\sigma_H]_{(a)}, [\sigma_H]_{(g)}\} \tag{5.68}$$

- the critical contact stress

$$[\sigma_H]_M = \sigma_{Hlim}Z_{NT}(Z_L Z_V Z_R)Z_V Z_X \tag{5.69}$$

- the factor of safety from pitting (sun gear - planet gear)

$$S_{H_{a,g}} = \frac{[\sigma_H]_{(a,g)}}{\sigma_H} \geq S_{H_{min}} \tag{5.70}$$

In addition, it is also necessary to include the functional constraints:
- the factor of safety from bending

$$g_{1,2,3} = \frac{[\sigma_F]_{M(a,g,b)}}{\sigma_{F(a,g,b)}} - S_F > 0 \tag{5.71}$$

- the factor of safety from pitting

$$g_4 = \frac{[\sigma_H]_{M(a,g)}}{\sigma_H} - S_H > 0 \tag{5.72}$$

$$g_5 = \frac{[\sigma_H]_{M(g,b)}}{\sigma_H} - S_H > 0 \tag{5.73}$$

- the radial interference

$$g_6 = \delta x > 0 \tag{5.74}$$

- space requirements

$$g_7 = 2a \sin\left(\frac{\phi}{n_W}\right) - f - d_{a-g} \geq 0 \tag{5.75}$$

- the condition for an assembly

$$h_1 = \frac{z_a z_b}{n_W D(z_g z_b)} - INT = 0 \tag{5.76}$$

Based on the given objective functions and on the functional constraints, all the relevant values of the planetary gear train have also been identified.

5.2.3 Optimisation procedure

The optimisation function is

$$\min y = \sum_{i=1}^{9} w_i f_i(X),$$

where

$$X = \{m_n, z_a, z_g, x_a, x_g, x_b, a_1, n_w, h, r_{ct}\},$$

with respect to the constraints presented with equations (5.71) - (5.76).

The purpose of assigning weighting coefficients to each objective function is to transform the multi-objective optimisation problem into single-objective optimisation problem. Nine functions are optimised simultaneously and the corresponding weighting coefficients w_i, for $i = 1, ..., 9$ have randomly assigned values.

The considered optimisation problem is addressed by GA based approach. In GA variables are represented as coded strings in Table (5.5). The coding discretises the search space of the optimisation problem. Thus, GAs are able to work with both discrete and continuous functions. GA is useful optimisation method when other techniques, such as gradient descent or direct analytical discovery, are not applicable, providing great flexibility to a wide range of applications.

The solution is coded in binary code, and decoded after the optimisation process. Each integer or real value is represented by a given number of binary digits. This actually means that the corresponding binary strings may have different length. Therefore, in order to ensure that crossover operations result with feasible solutions, the cross-sites should be properly determined. Although some of the variables x_1-x_{10} are continual, a discrete values in specific intervals have been assigned to them with a determined step (see the Table 5.6). A chromosome is presented with 42 digits values. A more sophisticated changes in the observed parameters can be achieved as a part of the future research which would require longer binary codes.

Table 5.5: GA coding of design variables taken from [Rosić 2011a]

Design variables vector	Variables	Values	Random binary digits	String length l
Module m_n	x_1	23	10101	5
Number of teeth z_a	x_2	18	1101	4
Number of teeth z_g	x_3	31	11010	5
x_a	x_4	0.4	11100	5
x_g	x_5	0.4	11100	5
x_b	x_6	0.2	111	3
a_1	x_7	144	11101	5
n_w	x_8	6	101	3
h	x_9	5.3	11111	5
r_{ct}	x_{10}	6.0	1111	4
A single 42-bit				1010111011101
individual				011100111001
(chromosome)				11111011111111011111

Mutation is a random process where one genotype is replaced by another to generate a new chromosome. Each genotype has the probability of mutation, to change from 0 to 1 and vice versa. Mutation is used to change the elements (genes) in strings which are generated by a crossover operator. It may improve an existing gene string and lead the search towards better solutions regions. Mutation is an important part of GA as it helps to prevent the population from stagnating at any local optima. The settings of GA used for the considered optimisation problem are given in Table 5.7.

5.2.4 Results

Pareto optimality can be illustrated graphically by considering the set of all feasible objective values, i.e., the set of all points in the objective space corresponding to at least one setting of the design variable. The experimental results are shown in the form of diagrams, representing pair-vise relationship between the objective functions. Diagrams are derived by using implemented MatLab library. Fig. 5.12 presents the relationship between the axial distance and efficiency, while Fig. 5.13 illustrates the axial distance - the outer diameter of the planetary gear train correlation.

Based upon a geometrical interpretation of the results in the criterion space, the following conclusions can be drawn:

- the criteria f_1 and f_2 (axial distance - efficiency) are mutually conflicting;
- there exists a very strong correlation between the criteria f_1 and f_9 (axial distance - outer diameter of the planetary gear train).

Table 5.6: Coding patterns

Variable	Lower border	Upper border	Step
x_1	2	33	1
x_2	5	20	1
x_3	5	36	1
x_4	-1	0.55	0.05
x_5	-1	0.55	0.05
x_6	-0.15	0.2	0.05
x_7	86	148	2
x_8	1	8	1
x_9	2	5.3	0.1
x_{10}	4.5	6.0	0.1

Table 5.7: Parameters settings for GA taken from [Rosić 2011a]

Name of parameter	*Value*
Population type	Double vector
Population size	50
Encoding	Binary
Scaling func.	Rank
Selection	Uniform
Elite count for reproduction	10
Crossover func.	Single point
Mutation	Adaptive feasible
Max number of generations	1000
Func. tolerance	$10^{(-15)}$
The remaining parameters are set	to their default values
Stopping criteria	Maximal number of generations or number of stall generations (1000)

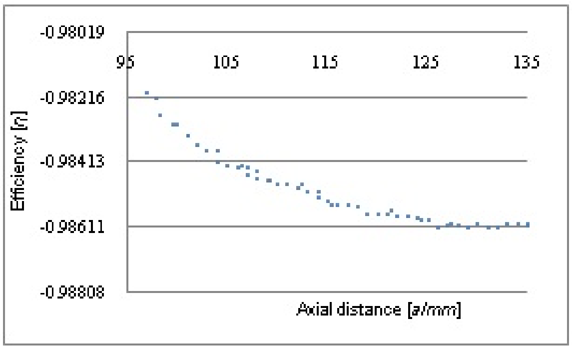

Figure 5.12: The criterion space for the axial distance - efficiency taken from [Rosić 2011a]

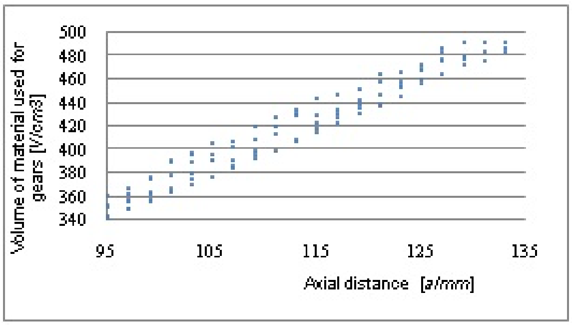

Figure 5.13: The criterion space for the axial distance - volume of material used for gears taken from [Rosić 2011a]

Multi-objective modelling reflects very well the design process in which usually several conflicting objectives have to be satisfied such as the efficiency of planetary gear trains and the distance between centres of a sun gear and a planetary gear. The effect of changes of the design parameters gives useful information regarding the sensitivity of various features in the model. A Pareto set, presented as a plot of the efficiency and axial distance of the planetary gear train, gives a quantitative description of the compromise between the efficiency and size. The results illustrate the importance of formulating the problem as a multi-objective optimisation.

First time in the literature, Genetic Algorithm method is applied for optimisation of the planetary gear train efficiency in [Rosić 2011b] and [Rosić 2011a]. Complexity of analysis can be seen in optimisation of 9 functions, while in [Castillo 2002] only 3 relationships have been analysed analytically. In [Qing-Chun 2008], MATLAB optimisation toolbox for multi-objective optimisation design of planetary gear train was used to optimise 2 conflicted functions: minisation of volume (weight) and maximisation of the efficiency. In [Cho 2006], only one relationship between the inputs and outputs is analysed, covering 3 influential parameters. In [Tripathi 2010], a multi-objective optimization of multi-stage planetary gear train is done. Optimisation in [Tripathi 2010] has covered only 2 conflicting functions of multi-stage planetary gear train: minimisation of the surface fatigue life factor and minimisation of volume of gear box. In [Tripathi 2010], only 6 constraint functions have been analysed, while in [Rosić 2011b] and [Rosić 2011a] apart from the 9 objective function, 8 constraints have taken into account due to the complexity of minimisation of the minimal elastohydrodynamic lubrication film.

The results obtained within this research are useful for the planetary gear predesign. A designer supplied with this preliminary design values has a chance to improve geometrical characteristics of the planetary gear train in order to achieve high efficiency. The described results are presented in [Rosić 2011b] and [Rosić 2011a].

Chapter 6

Optimisation of Ball Bearing Dynamical Load Ratings and Rating Life

A new method for optimisation of dynamical load ratings and rating life, as a function of 10 different parameters, of radial ball bearings was developed using meta-heuristics. The aim of this research was to determine which parameters have the largest influence on achieving the maximum working life. In order to find the proper values for dynamic radial and axial load factors, it was necessary to perform interpolation of three-dimensional nonlinear function.

Dynamical load ratings and rating life of a radial ball bearing, based on ISO standard [ISO 281, ISO 75], depend on many factors, and it is assumed that the rating life can be extended by optimising the influencing parameters. The goal of the optimisation was to find the optimal inner geometry of bearings based on the outer geometry, using three meta-heuristic methods.

The optimisation is performed in the form of a numerical simulation. Apart from the formulas and procedures from [ISO 281, ISO 75], the values of some specific parameters were varied in order to find the appropriate combination of geometry, rating factor (the value of which varies with a bearing type and design), static and dynamic radial load rating, the value of the parameter for mobility conditions, the dynamic radial and axial factor and the factor which depends on the geometry of the bearing components and the material [Milojević 2014].

In order to simplify the model, the radial component of the actual bearing load and the axial component of the actual bearing load are set as constant. Optimised parameters are mostly related to the bearings geometry, and the term *geometry* refers to an optimisation against the number of rolling elements in a single row-bearing, the nominal ball diameter, the pitch diameter of the bearing and the radial contact angle of bearings. All of these factors, together with the factor of geometry and material, directly impact the calculation of basic dynamic load rating.

6.1 Problem formulation

The use of the term *basic rating life* L_{10} refers to the optimisation against the load rating and the equivalent load, which is formulated in the equation:

$$L_{10} = \left(\frac{C_r}{P_r}\right)^a, \tag{6.1}$$

which is already defined in Chapter 4 with equation (4.1). In the case of radial rolling bearings, $a = 3$, while in the case of the barrel bearings, $a = \frac{10}{3}$. The equation (6.1) only holds if the number of revolutions is equal or greater than 10 min^{-1} and it is only certain with probability of 90%.

Load rating P_r is given by the equation:

$$P_r = XF_r + YF_a \tag{6.2}$$

where X denotes the dynamic radial load factor and Y stands for the dynamic axial load factor.

To find the optimum value of load rating P_r, it is necessary to find the optimal values of factors X and Y, which will be explained in the next sections. Basic dynamic radial load rating for radial ball bearings is given by the following equations [ISO 281]:

$$C_r = \begin{cases} b_m f_c (i \cos\alpha)^{0.7} Z^{2/3} D_b^{1.8}, & D_b \leq 25.4mm \\ 3,647 b_m f_c (i \cos\alpha)^{0.7} Z^{2/3} D_b^{1.4}, & D_b > 25.4mm \end{cases} \tag{6.3}$$

Therefore, the optimisation function to be maximised is

$$L_{10}(S),$$

for $S = \{K_{D_{min}}, K_{D_{max}}, \varepsilon, m, \beta, Z, D_b, D_m, f_i f_o\}$.

6.1.1 Constraints

$$c_1(S) = \frac{\Phi_o}{2\arcsin(D_b/D_m)} - Z + 1 \geq 0 \tag{6.4}$$

$$c_2(S) = 2D_b - K_{D_{min}}(D-d) \geq 0 \tag{6.5}$$

$$c_3(S) = K_{D_{max}}(D-d) - 2D_b \geq 0 \tag{6.6}$$

$$c_4(S) = D_m - (0,5-m)(D+d) \geq 0 \tag{6.7}$$

$$c_5(S) = (0,5 + m)(D + d) - D_m \geq 0 \tag{6.8}$$

$$c_6(S) = \frac{d_i - d}{2} - \frac{D - d_o}{2} \geq 0 \tag{6.9}$$

$$c_7(S) = 0,5(D - D_m - D_b) - \varepsilon D_b \geq 0 \tag{6.10}$$

$$c_8(S) = \beta W - D_b \geq 0 \tag{6.11}$$

$$c_9(S) = f_i \geq 0,515 \tag{6.12}$$

$$c_{10}(S) = f_o \geq 0,515 \tag{6.13}$$

$$c_{11}(S) = \left[\frac{[U^2 + (D/2 - T - D_b)^2 - (d/2 + T)^2]}{2U(D/2 - T - D_b)}\right] + 1 \geq 0 \tag{6.14}$$

$$c_{12}(S) = 1 - \left[\frac{[U^2 + (D/2 - T - D_b)^2 - (d/2 + T)^2]}{2U(D/2 - T - D_b)}\right] \geq 0 \tag{6.15}$$

$$c_{13}(S) = (D_b/D_m) + 1 \geq 0 \tag{6.16}$$

$$c_{14}(S) = 1 - (D_b/D_m) \geq 0 \tag{6.17}$$

6.1.2 Constraint explanations

For the convenience of the bearing assembly, the number Z and the diameter D_b of balls should satisfy the requirement given by the relation (6.4), where Φ_o is the maximum tolerable assembly angle calculated by the equation [Gupta 2007, Rao 2007]:

$$\Phi_o = 2\pi - 2\arccos\left[\frac{[U^2 + (D/2 - T - D_b)^2 - (d/2 + T)^2]}{2U(D/2 - T - D_b)}\right] \tag{6.18}$$

where

$$T = (D - d - 2D_b)/4; \tag{6.19}$$

$$U = (D - d)/2 - 3T. \tag{6.20}$$

The diameter of the rolling element should be chosen from certain bounds, i.e., the following inequality must hold (as shown in Fig. 6.1):

$$K_{D_{min}} \frac{D-d}{2} \leq D_b \leq K_{D_{max}} \frac{D-d}{2} \tag{6.21}$$

Where $K_{D_{min}}$ and $K_{D_{max}}$ are unknown constants (which decide the possible minimum and maximum diameters of the rolling element) and D and d are the outside and bore diameters of bearings, respectively. For the relation (6.21), the corresponding constraint conditions are given as (6.5) and (6.6).

In order to guarantee the running mobility of bearings, the difference between the pitch diameter and the average diameter in a bearing should be less than a certain given value. Therefore, the constraints given by the relations (6.7) and (6.8) are to be satisfied, where m is an unknown input variable. In practice, the inner ring is always exposed to more stresses than the outer ring, this necessitated the need to put a constraint on the ring thickness that should be larger than or equal to the outer ring thickness. This requirement is given by the condition (6.9), where d_i and d_o are the inner and outer raceway diameters at the grooves. The thickness of the bearing ring at the outer raceway bottom should not be less than εD_b, where ε is an unknown constant. Therefore the constraint condition is according to [Gupta 2007, Rao 2007] given by the relation (6.10).

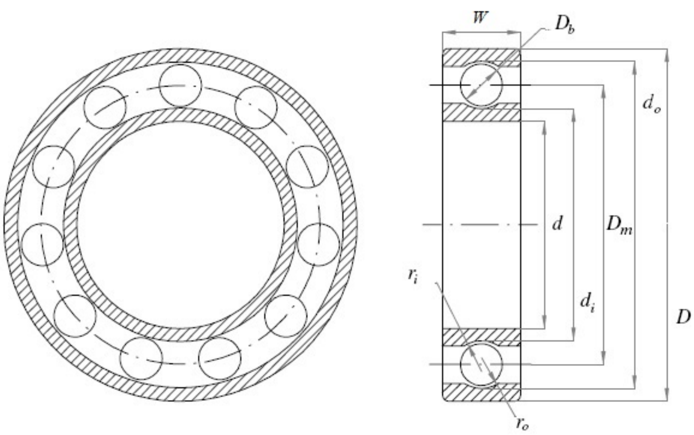

Figure 6.1: Radial ball bearing macro-geometries taken from [Gupta 2007]

The width of a bearing, (W) generates the constraint on the diameter of the ball that can be written as a relation (6.11), where β is an unknown constant. $K_{D_{max}}$ and β are the factors which decide the upper bound of the rolling element diameter. Groove curvature radii of inner f_i and outer f_o raceways in a bearing should be larger than $0,515 D_b$. If this was not true the dynamic load rating of the bearing would decrease. Therefore, two more constraint conditions are given by the relations (6.12) and (6.13).

In addition to the presented constraints, some geometrical characteristics are expressed by the given lower and upper bounds:

$$3 \le Z \le 500$$

$$1 \le D_m \le 500$$

$$1 \le D_b \le 500$$

$$0,6 \le K_{D_{max}} \le 0,7$$

$$0,4 \le K_{D_{min}} \le 0,5$$

$$0,3 \le \varepsilon \le 0,35$$

$$0,03 \le m \le 0,08$$

$$0,7 \le \beta \le 0,85.$$

Arguments of trigonometric functions arccos and arcsin have to take the values between -1 and 1, so the following equations have to be satisfied:

$$-1 \le \left[\frac{[U^2 + (D/2 - T - D_b)^2 - (d/2 + T)^2]}{2U(D/2 - T - D_b)}\right] \le 1 \tag{6.22}$$

$$-1 \le (D_b/D_m) \le 1 \tag{6.23}$$

which produces the constraints: (6.14), (6.15), (6.16), (6.17).

6.1.3 Additional parameter settings

Values of b_m for radial ball bearings, as given by [ISO 281], are summarised in the table below.

Bearing type	b_m
Radial and angular contact ball bearings (except filling slot bearings), insert bearings and self-aligning ball bearings	1,3
Filling slot bearings	1,1

In this study $b_m = 1,3$ has been adopted, since only radial contact ball bearings are investigated.

The f_c can be calculated as follows:

$$f_c = 37,91 \cdot f_d,$$

$$f_d = \left\{1 + \left[1,04 f_g^{1,72} f_{io}^{0,41}\right]^{10/3}\right\}^{-0,3} \left[\frac{\gamma^{0,3}(1-\gamma)^{1,39}}{(1+\gamma)^{1/3}}\right] \left[\frac{2f_i}{2f_i - 1}\right]^{0,41}$$

where

$$\gamma = D_b cos\alpha / D_m,$$

$$f_g = \frac{1-\gamma}{1+\gamma},$$

$$f_{io} = \frac{f_i(2f_o - 1)}{f_o(2f_i - 1)},$$

and α is the free contact angle that depends upon the type of a bearing.

In order to simplify the model, the following working conditions are given: $F_a = 100N$, $F_r = 1500N$.

For the purpose of this study, several bearing types are selected. They are described by the outer dimensions D, d, W, as it is presented in Table 6.1, while the corresponding inner dimensions are obtained by optimisation methods.

Table 6.1: Bearing types and the corresponding data

Values from catalogue			
Bearing type	Num. Spec.		
	D	d	W
6200	30	10	9
6201	32	12	10
6202	35	15	11
6203	40	17	12
6204	47	20	14
6205	52	25	15
6206	62	30	16
6207	72	35	17

The selected bearing types belong to the class of the single-row-deep-groove ball bearings, and therefore it holds that $i = 0$ and $\alpha = 0°$.

6.1.4 Interpolation of three-dimensional nonlinear function

To find a proper value of the dynamic radial load factor X and the dynamic axial load factor Y, it was necessary to perform an interpolation based on the relative axial load (given by the relations $\frac{f_o F_a}{C_{or}}$ and $\frac{F_a}{iZD_b^2}$) and e (the limiting value of $\frac{F_a}{F_r}$). Values of the appropriate dynamic radial and axial load factors, for the ranges of above-mentioned three values, are given in [ISO 281]. The relations $\frac{f_o F_a}{C_{or}}$ and $\frac{F_a}{iZD_b^2}$ are input values of the separate interpolation module, and e is a result of calculations based on the working conditions F_a and F_r where:

$$C_{or} = f_o iZD_b^2 \cos\alpha. \tag{6.24}$$

The data relevant for this study are given in Table 6.2.

Table 6.2: ISO 76 data taken from [ISO 281]

Bearing type		Relative axial load		Single row bearings $i = 1$				Double row bearings $i = 2$				e
				$\frac{F_a}{F_r} \leq e$		$\frac{F_a}{F_r} > e$		$\frac{F_a}{F_r} \leq e$		$\frac{F_a}{F_r} > e$		
				X	Y	X	Y	X	Y	X	Y	
Radial contact ball bearing $\alpha = 0°$		$\frac{f_o F_a}{C_{or}}$	$\frac{F_a}{iZD_b^2}$	1	0	0,56		1	0	0,56		
		0,172	0,172				2,3				2,3	0,19
		0,345	0,345				1,99				1,99	0,22
		0,689	0,689				1,71				1,71	0,26
		1,03	1,03				1,55				1,55	0,28
		1,38	1,38				1,45				1,45	0,3
		2,07	2,07				1,31				1,31	0,34
		3,45	3,45				1,15				1,15	0,38
		5,17	5,17				1,04				1,04	0,42
		6,89	6,89				1				1	0,44
Angular contact ball bearing	$\alpha = 20°$	-	-	1	0	0,43	1	1	1,09	0,7	1,63	0,57
	$\alpha = 25°$	-	-			0,41	0,87		0,92	0,67	1,41	0,68
	$\alpha = 30°$	-	-			0,39	0,76		0,78	0,63	1,24	0,8
	$\alpha = 35°$	-	-			0,37	0,66		0,66	0,6	1,07	0,95
	$\alpha = 40°$	-	-			0,35	0,57		0,55	0,57	0,93	1,14
	$\alpha = 45°$	-	-			0,33	0,5		0,47	0,54	0,81	1,34

6.2 Meta-heuristics implementations

The above described optimisation problem contains the continuous nonlinear objective function, with linear continuous constraints. Therefore, various optimisation methods, designed for this case can be applied. In this study, the optimisation is performed using three different meta-heuristic methods:

- Genetic algorithms (GA)
- Multi-start Pattern Search (MPS)
- Multi-start Active Set (MAS)

Implementation of Active Set method, for the purposes of this research has been performed within Fmincon [Venkataraman 2009] optimisation function built in MATLAB optimisation tool. Active set multistart execution is one of the optimisation methods incorporated into proposed procedure shown in Fig 6.2 [Milojević 2014] within the Fmincon function. GA, MPS and MAS methods use the same main procedure, as well as the same input values. Extensive experimental evaluation is performed in order to determine the best parameter settings for each of the used methods. The resulting settings are described in Tables 6.3, 6.4 and 6.5.

Table 6.3: The parameter settings in GA taken from [Milojević 2014]

Population	Double vector
Population size	50
Initial population	Randomly generated
Scaling function	Proportional
Selection function	Tournament
Elite count	10
Crossover fraction	0.6
Mutation	Adaptive feasible
Crossover function	Two point
Number of generation	1000 generation

The stopping criterion is either of the two: the maximum number of generations or the maximum stall time.

Since PS is an iterative (local search) heuristic method, the parameter settings (as presented in Table 6.4) are given for a single execution (without restarts).

The stopping criterion is either of the six: Mesh tolerance (10^{-6}) or Max iterations (100*number of variables), or Max function evaluations (2000*number of variables), or X tolerance (10^{-6}), or Function tolerance (10^{-6}), or Nonlinear constraint tolerance (10^{-6}). The starting points are created in the form of the feasible solutions set. The number of PS restarts is determined by reaching a GA stopping time.

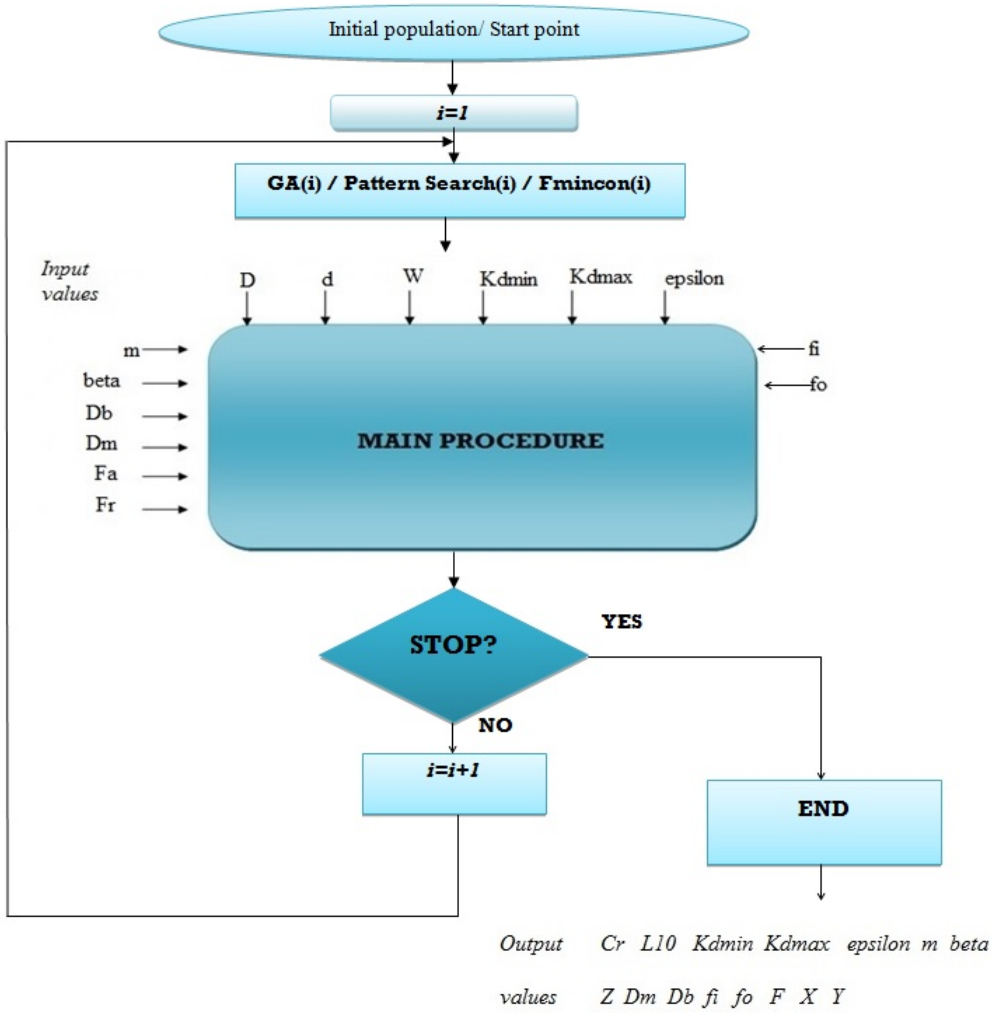

Figure 6.2: Optimisation algorithm taken from [Milojević 2014]

Table 6.4: The parameter settings in PS taken from [Milojević 2014]

Poll method	GPS Positive basis 2N
Complete poll	On
Polling order	Success
Complete search	On
Search method	GPS Positive basis 2N
Mesh Initial size	1.0
Mesh Max size	Inf
Accelerator	On
Rotate	On
Scale	On
Expansion factor	2.0
Contraction factor	0.5
Initial penalty	10
Penalty factor	100
Bind tolerance	10^{-3}
Cache	On
Tolerance	eps
Size	10^4

The parameter settings given in Table 6.5 correspond to a single execution of Fmincon function with incorporated Active Set method. The Active Set (AS) is an

Table 6.5: The parameter settings in Fmincon taken from [Milojević 2014]

Algorithm	Active set
Derivatives	Approximated by solver
Max iterations	400
Max function evaluations	1000
X tolerance	10^{-6}
Nonlinear constraint tolerance	10^{-6}
SQP constraint tolerance	10^{-6}
Function value check	None
User-supplied derivatives	None
Minimum pertubation	10^{-8}
Maximum pertubation	0.1
Type	Forward differences
Hessian	None
Typical X values	ones(10,1)

iterative method and the results obtained with AS depend on the starting point. Different local optimum solutions are obtained by varying the starting points. Therefore, in order to achieve the best possible solution, a Multi-start of the AS method is performed from different starting points. The initial solutions for each restart and the number of restarts are determined in the same way as for MPS. Evidently, MAS is implemented by repeatedly calling Fmincon function from different initial solutions until a stopping criterion is satisfied. In this case it is the CPU time required by GA to find its best solution.

6.3 Results

The simulation of dynamical load ratings of normal contact ball bearings is conducted by changing the various influential parameters. As a result, the corresponding values of the rating life and dynamic capacity are measured and optimised by the applied meta-heuristic methods. At the same time, the extent to which specific factors influence the best value of the rating life is determined.

Two functions, the dynamic load capacity and maximum working life (under a certain conditions), are optimised simultaneously which means that the problem should be treated as multi-objective. However, since the two objective functions are directly proportional, they are not conflicted and the optimisation is conducted like the problem is single-objective. Apart from this characteristic, this optimisation problem is also nonlinear problem with inequality constraints.

The comparative results for three methods are given in tables 6.6 and 6.8. It appears that, by increasing the number of balls in a ball bearing, it is possible to increase the dynamic load capacity with comparison to the available standards. Increasing of the dynamic load capacity leads to the increase in the value of the dynamic working life.

The MAS method requires less computing time with respect to GA and MPS. It generates the best values for the dynamic load capacity in six out of eight bearings type. GA gives the best results for working life in five out of eight cases. MPS provided the largest dynamic capacity for bearing types 6200 and 6202, while the longest working life reported for bearing types 6202, 6205, and 6206. The general conclusion is that for the considered examples of ball bearings MAS performed the best. It provided largest dynamic capacity, it was the fastest method and, at the same time, the reported working life was very close to the best obtained value.

During optimisation process, the variable Z is taken as a continuous variable. Once the optimisation is done, the further calculations are conducted with rounded values of Z to obtain the working life and dynamic capacities that correspond to integer values for number of balls. Achieved results show that the best value of the dynamic load capacity and the largest working life are obtained for $Z = 7$, $Z = 7$, $Z = 8$, $Z = 8$, $Z = 8$, $Z = 9$, $Z = 9$ and $Z = 9$, respectively, for the considered eight types of bearings.

The proposed optimisation methods provide the increase in a dynamic capacity with respect to the values from the available standards in all eight examples. The average percentage of the improvement is 9.4%, 12.2% and 12.6% for GA, MPS and MAS, respectively comparing to the values in [Bowman]. The average percentage of the improvement with respect to the [Gupta 2007] is 13.41%, 20.91%, 18.43% for four considered cases (See Table 6.7). The average percentage of improvement of the dynamic capacity with respect to [Rao 2007] for 4 cases of bearings is 22.22%, 30.3% and 27.64% (See Table 6.7). The work presented in [Gupta 2007] and [Rao 2007] is compared with the values from [Shigley 1989].

Table 6.6: Comparative results for three optimisation methods taken from [Milojević 2014]

Catalogue values [Bowman]		Heuristic optimisation								
Bearing type	Dynamic Capacity [N]	GA results			MPS results			MAS results		
		Dyn. Cap [N]	Work. life [h]	CPU time [s]	Dyn. Cap [N]	Work. life [h]	CPU time [s]	Dyn. Cap [N]	Work. life [h]	CPU time [s]
6200	5070	6842.4	**540.5016**	43.4931	**6848.8**	531.1687	41.746	6848.7	531.8595	**2.6832**
6201	6890	7223.7	**647.61**	21.2473	7223.7	627.2416	83.7723	**7238.6**	631.5472	**0.6084**
6202	7800	8079.6	940.6012	18.3301	**8319.3**	**989.2423**	8.5864	8263	971.6416	**0.39**
6203	9560	10636	**2351.4**	17.1445	10667	2181.8	87.7193	**10703**	2186.7	**0.2808**
6204	12700	14211	**6178**	100.761	14243	5589.6	83.3353	**14291**	5589.6	**0.234**
6205	14000	15307	7914.30	29.7962	15998	**8233.3**	41.5429	**16085**	8232.6	**0.4212**
6206	19500	19974	19215	45.2101	21756	**22942**	42.3698	**21878**	22941	**1.3416**
6207	25500	28241	**60961**	55.4272	28285	55027	83.8503	**28447**	55028	**0.4056**

Table 6.7: Comparative results for three optimisation methods with [Gupta 2007] and [Rao 2007]

Values from [Gupta 2007] and [Rao 2007]			Heuristic optimisation		
Bearing type	Dynamic Capacity [N] [Gupta 2007]	Dynamic Capacity [N] [Rao 2007]	GA results Dyn. Cap [N]	MPS results Dyn. Cap [N]	MAS results Dyn. Cap [N]
6200	6029.54	5942.36	6842.4	**6848.8**	6848.7
6202	7057.92	6955.35	8079.6	**8319.3**	8263
6204	12098.9	10890.9	14211	14243	**14291**
6206	18111.9	16387.4	19974	21756	**21878**

Table 6.8: Values for design parameters generated by three optimisation methods taken from [Milojević 2014]

Bearing type	Opt. method	Design parameters										Calculated values		
		$K_{D_{min}}$	$K_{D_{max}}$	ε	m	β [rad]	Z	D_m [mm]	D_b [mm]	f_i	f_o	F [rad]	X	Y
6200	GA	0.431	0.699	0.301	0.047	0.848	7	6.999	18.772	0.515	0.515	4.015	0.56	2.02
	MPS	0.4	0.7	0.3	0.066	0.85	7	6.999	18.8	0.515	0.515	4.015	0.56	2.02
	MAS	0.4	0.7	0.3	0.08	0.778	7	7.000	18.8	0.515	0.515	4.015	0.56	2.06
6201	GA	0.411	0.7	0.3	0.0483	0.7094	7	7.000	20.800	0.515	0.515	3.892	0.56	2.06
	MPS	0.4	0.7	0.3	0.066	0.849	7	7.000	20.799	0.515	0.515	3.89	0.56	2.06
	MAS	0.4	0.7	0.3	0.0380	0.85	7	7.000	20.8	0.515	0.515	3.892	0.56	2.06
6202	GA	0.401	0.700	0.302	0.067	0.819	8	6.853	24.000	0.515	0.515	3.732	0.56	2.06
	MPS	0.4	0.7	0.3	0.066	0.85	8	7.000	23.799	0.515	0.515	3.7926	0.56	2.10
	MAS	0.4	0.7	0.3	0.08	0.85	8	7.000	23.1	0.515	0.515	3.76	0.56	2.06
6203	GA	0.487	0.700	0.3104	0.0583	0.7640	8	8.050	26.8764	0.515	0.515	3.7705	0.56	2.22
	MPS	0.4	0.7	0.3	0.0665	0.8484	8	8.050	27.11	0.515	0.515	3.7705	0.56	2.22
	MAS	0.4001	0.7	0.3	0.08	0.7	8	8.050	27.12	0.515	0.515	3.7705	0.56	2.22
6204	GA	0.4008	0.7000	0.3001	0.0666	0.8096	8	9.4341	31.9035	0.515	0.515	3.7668	0.56	2.3
	MPS	0.4	0.7	0.3	0.0663	0.8484	8	9.450	31.8799	0.515	0.515	3.7693	0.56	2.3
	MAS	0.4	0.7	0.3	0.08	0.85	8	9.450	31.88	0.515	0.515	3.7693	0.56	2.3
6205	GA	0.4037	0.6974	0.3051	0.0536	0.7909	9	9.1936	36.9667	0.515	0.515	3.6268	0.56	2.3
	MPS	0.4	0.7	0.3000	0.0665	0.8484	9	9.450	36.8797	0.515	0.515	3.6585	0.56	2.3
	MAS	0.4	0.7	0.3	0.08	0.7	9	9.450	36.88	0.515	0.515	3.6585	0.56	2.3
6206	GA	0.4999	0.7	0.3490	0.0645	0.8461	8	11.200	42.6044	0.515	0.515	3.6529	0.56	2.3
	MPS	0.4	0.7	0.3000	0.0663	0.8484	9	11.200	44.0799	0.515	0.515	3.6529	0.56	2.3
	MAS	0.4	0.7	0.3	0.0318	0.85	9	11.200	44.08	0.515	0.515	3.6529	0.56	2.3
6207	GA	0.4197	0.6999	0.3101	0.0535	0.8053	9	12.946	50.9664	0.515	0.515	3.6486	0.56	2.3
	MPS	0.4000	0.7	0.3000	0.0663	0.85	9	12.950	51.2799	0.515	0.515	3.6489	0.56	2.3
	MAS	0.4	0.7	0.3	0.08	0.85	9	12.950	51.28	0.515	0.515	3.6489	0.56	2.3

Chapter 7

Reliability Assessment of Mechanical Systems by Bayesian Networks

The purpose of this study is to present a Bayesian network (BN) based prediction method for the quality of production during the operational process if failure occurs in the system. The considered two mechanical systems are a semi-automatic system plant for painting and varnishing metal products and a system for oil filtering. The input data for BNs are historical and have been collected through several years of systems operation. If a product failure is detected, based on Bayesian probabilities generated by the proposed BN, it is possible to determine which subsystem is probably responsible for significant slowing down of the operational process or for the defects in the final product.

On the other hand, when diagnosing a failure of the machine it is necessary to make the appropriate decision whether it is justified to continue the process of manufacturing or to stop it in order to eliminate failures. The aim of BN systems is to predict quality of a semi-product when some of the failures of the subsystems happen.

The proposed BN-models receive information of systems failures at the input and gives, with some probability, a prediction of the quality of products at the output. Based on the output probabilities, a user of the production system can decide whether to terminate the operational process or to leave the machine in the operational mode. A comparison of the historical probabilities for the considered mechanical system against the results obtained by BN method is given.

The maintenance of a system is divided into prevention, correction and investment. The investment refers only to the purchase of a new, better equipment that appears on the market. The corrective maintenance refers to an unpredictable event of the failure in the system. It means that the entire production should be stopped in order to remove the defect, which can take a lot of time, human energy and create high economic costs. More precisely, the required are, among other things, overtime involvement of a large number of people, buying new parts, the loss of money, stopping production flow. Therefore, it is necessary to pay special attention to the preventive maintenance.

The preventive maintenance refers to the manufacturer's recommendations in terms of what should be done based on the number of working hours after releasing machine at work. For example, the service life of the bearing is *a priori* known and therefore, bearings should be replaced before the working life period expires.

The preventive maintenance is performed on a daily, weekly, monthly and annually basis, and all of this is based on the manufacturer's recommendations, law regulations, guidelines and other technical data. In that way, the preventive maintenance serves to postpone the corrective maintenance, as much as possible, in order to minimise the number of failures.

If the mechanical system comes to the failure and the need for the corrective maintenance appears, then a proposed BN probability model can help in making the decision to postpone the corrective maintenance. Based on the BN obtained probability, the production manager can decide that failure will not cause a deterioration in the quality of the final product and will continue the production process. In that way, the repair can be postponed for some other, more convenient, opportunity. The purpose of the proposed BN procedures is the prediction of the quality of the final product before the end of the production process, when a failure appears. If it is necessary to make a quick decision whether to continue the production process or not, the Bayesian procedure could substitute the production manager's decision. Or, for example, if the production manager is on holiday, this procedure could also be helpful for someone with less experience and knowledge who has to make a decision instead of the manager. Bayesian procedure relieves the production manager of responsibility for decision making. Since the reparation of the failure and stopping of the production process is considered expensive, this procedure could help in making the appropriate decisions to postpone stopping of the production process and thus save money and reduce the number of working hours.

7.1 Reliability assessment of semi-automatic system plant for painting and varnishing metal products

7.1.1 Problem formulation

The semi-automatic system plant for painting and varnishing metal products (Fig. 7.1) in a wide sense, consists of seven subsystems with the following names:

- The chamber for cleaning and degreasing a product (x_1)
- The chamber for applying the base colour (x_2)
- The chamber for applying the final colour (x_3)
- Drying chamber (x_4)
- The Conveyor (x_5)
- The control panel that controls all subsystems, through the power and control. It contains the main switch, the main fuse, as in a house or a flat. So here one can stop a subsystem or the whole plant (x_6)
- g) The ventilation (x_7).

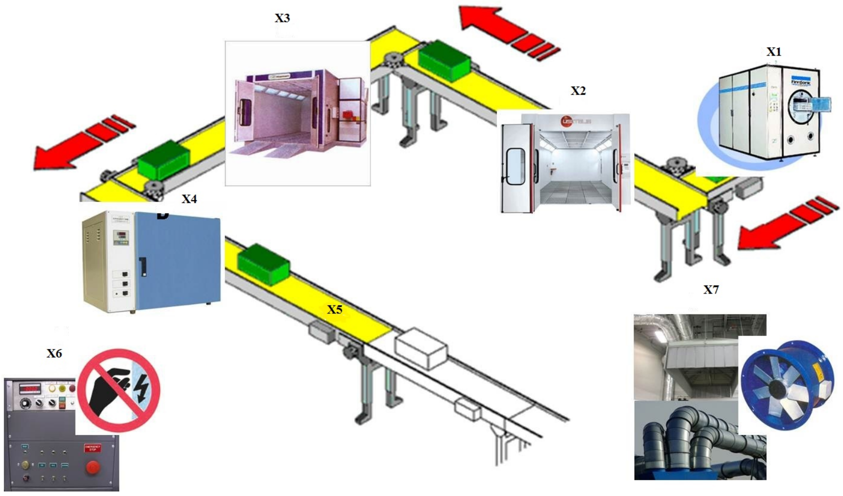

Figure 7.1: The considered mechanical system taken from [Milojević 2012]

For putting the system into operation it is necessary to turn on the main switch at the chamber x_6. First, the working pieces are placed on a conveyor belt in front of the chamber x_1. Second, the working piece is cleaned and degreased using the pressurised equipment with warm water and a degreasing agent is applied in the chamber x_1. Then, the working piece passes from chamber x_1 to the chamber x_2 using the conveyor x_5. Using the gearing pumps of the chamber x_2, the base colour from a buried tank is applied to the working piece by pouring. The excess paint is returned to the reservoir by gravity. At the same time, the washing process of other working pieces is performed in the chamber x_1. Again, moving of the conveyor x_5 leads the third working piece into the chamber x_1, while the second one goes from the chamber x_1 to the x_2, and the first working piece is somewhere between the chambers x_2 and x_3. Than, the conveyor leads the fourth working piece to enter the chamber x_1, the third working piece moves from the chamber x_1 to x_2, the second working piece leaves the chamber x_2, and the first piece comes in the chamber x_3 where the final painting is done. Powered by the conveyor, the fifth working piece goes into the chamber x_1 and the first piece into the drying chamber x_4. After the drying process, the first working piece goes out of the chamber x_4 and it should be removed from the conveyor while the sixth working piece goes into the chamber x_1. These 7 subsystems consist of a large number of parts which make the machine very complex and create the potential danger from a various failures on a daily basis. For example, if the failure of the light in the chamber x_1 occurs, it is logical that the production process does not have to be stopped because of that. On the other hand, the failure of a cleaning and degreasing device in the chamber x_1, or a pump for watering colour filure in the chamber x_2 will cause certain degradation of the product quality or the production speed and, therefore, the production process must be stopped. The probabilities that the product will be good, obtained by BN in that case would certainly be minimal (tending to zero).

7.1.2 The proposed Bayesian model

The Bayesian networks are graphical models that can be used to model stochastic systems [Pearl 1985]. They can be "built" on the basis of an expert knowledge or automatically from the data (or both). The problem analysed in this paper is suitable to be modelled by BN, since behaviour of the discussed system is probabilistic.

In this example, the Bayesian network is applied to a set of variables

$$X = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$$

where:

- x_1: the chamber for cleaning and degreasing a product
- x_2: the chamber for applying the base colour
- x_3: the chamber for applying the final colour
- x_4: drying chamber
- x_5: the conveyor
- x_6: the control panel
- x_7: the ventilation

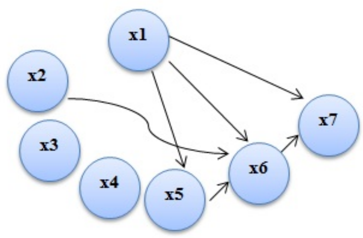

Figure 7.2: A part of the implemented BN taken from [Milojević 2012]

The Bayesian networks have a direct graph structure (Fig. 7.2) with independent conditional statements about X, and a set of local probabilities P. The nodes represent X variables and arches describe their relationship. Building of the Bayesian model for domain application involves three main steps:

- Identification of the variables that are of importance, along with their possible state values.
- Identification of of the relationships between the variables and expressing them in a graphical structure.

- The assessment of the probabilities required for its quantitative part. The above three steps are, in principle, performed one after the other.

However, building a BN usually requires a careful trade-off between the desire for a large and rich model on one hand, and the required effort for construction, maintenance, and probabilistic inference in the network on the other hand. In practice, therefore, the building of Bayesian models is a creative process that iterates over these steps until a desired network is achieved.

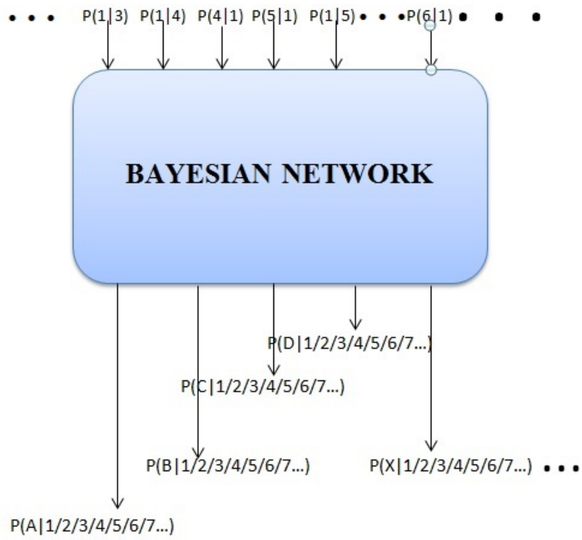

Figure 7.3: The BN inputs and outputs taken from [Milojević 2012]

In this research, the behavior and possible scenarios of the mechanical system are examined. A comparative analysis was performed between the results obtained by the computer program and historical probabilities. At the input of the computer system, the information about the resulting failure of the mechanical system is given, and at the output, the BN model gives a prediction of the product quality with some probabilities. The input was partly provided empirically, but most of the input data were collected from the documentation for maintenance of this machine in the last 20 years. By simply counting the number of failures form the data tables, it can be easily concluded what is the most common, or the rarest failure and what can never fail (like construction walls made of steel). In this model, the output of the system are events A, B, C, D and F, where A denotes a well done product with no errors, B means that the product was not done at all, the meaning of C is that the product was done slowly, the D states that the product was done incorrectly (with errors), and F means that there was a failure of the entire machine. An example of the input data is $P(B|x_2) = 0.1$ with the meaning that the probability of B (the product was not done at all), provided by the event x_2 (failure of the chamber x_2) is 0.1. An example of the input data connections in BN is shown in Fig. 7.2, while the schema of inputs and outputs is shown in Fig. 7.3. The actual Bayesian model for this study is shown in Fig.7.4.

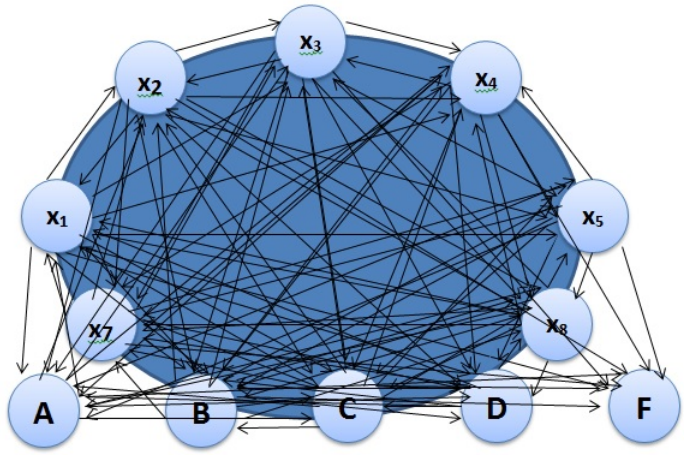

Figure 7.4: Bayesian procedure taken from [Milojević 2012]

7.1.3 Results

Table 7.1 gives the best BN probabilities for the cases A-F, compared with the empirical probabilities. It can be seen that the historical probability that a product will be correct is 90%, while the corresponding probability obtained by the proposed BN is 89%. The empirical probability that a product is not done at all is 5%, which was also obtained by BN. The empirical probability that the product is done slowly is 30% and by BN it was obtained that it will happen in 27% of cases. The empirical probability that the product is done with some errors is 5% and the corresponding BN-obtained result is 4%. The probability of failure of the whole machine is 5% and 4.5% obtained by historical data and BN, respectively. It can be concluded that these two probabilities differ in less than 3%, which confirms accurate modelling of realistic system conditions. A measure of quality of BNs prediction is reflected in the difference between the historical probabilities and the probabilities obtained by modeling. If the ratio is less than 3%, which is proofed in this research, it is considered that the prediction is very satisfactory [Darwiche 2009].

Table 7.2 presents the estimated failures probabilities of the subsystems. If any type $(B, ..., F)$ of bad products is detected, it is possible to find the cause of the particular case with BN model. If the product was not done at all (case B), the chambers x_1, x_2, x_3 and x_4 are equally responsible. The conveyor and the control panel have larger and mutually equal Bayesian probabilities. The relative ratio, which shows that each chamber has the same influence on the case B, corresponds to the actual event. A similar situation happens with the results of the cases C, D and F.

Table 7.1: Probabilities of the product quality in the general case [Milojević 2012]

Outcome	Bayesian probabilities	Empirical conclusion
Case A-the part is good	89%	90%
Case B-the part is not finished	5%	5%
Case C-the part is done slowly	27%	30%
Case D-the part is done with errors	4%	5%
Case F-the failure of the whole system	4,5%	5%

Table 7.2: The failure probabilities of a subsystem if the outcome is known from [Milojević 2012]

Outcome/Chamber	x_1	x_2	x_3	x_4	x_5	x_6	x_7
Case B-the part is not finished	2%	2%	2%	2%	5%	5%	3%
Case C-the part is done slowly	6%	6%	6%	6%	15%	15%	9%
Case D-the part is done with errors	2%	2%	2%	2%	5%	5%	3%
Case F-the failure of the whole system	2%	2%	2%	2%	5%	5%	3%

The most important role of the proposed BN, and the corresponding software, is to provide outputs in the cases of mutual failure of several subsystems. The corresponding BN obtained results are shown in Table 7.3. This table illustrates the possibility to calculate probabilities of A-F if some combinations of events x_1-x_7 happen. This provides an estimation of the quality of the products in the case of the machine failures. If cases x_1 and x_2 happen, there is a drastic difference between A and B, where the chance of securing good parts is only 4.5% while the bad product will appear in 45% of cases. This suggests that the production has to be stopped and the defects repaired.

Table 7.3: The outcome probabilities for cases with more than one failure from [Milojević 2012]

Failure/Outcome	A	B	C	D	F
x1, x2	4,5%	45%	13,3%	45%	45%
x3, x4	4.4%	45.6%	14%	45.6%	45.6%
x1, x2, x6	2.25%	65%	15.5%	65%	65%
x1	10%	39%	12.8%	39%	39%
x2, x4	4. 6%	45,3%	13,7%	45,3%	45,3%
x1, x3, x7	2.5%	64%	15%	64%	64%
x4, x5	3%	49%	13.9%	49%	49%

Based on the obtained results, it can be concluded that there is no significant difference between the results obtained by BN and historical probabilities. In the modern business environment, when enterprises are exposed to permanent changes, the theory of maintenance gains a new dimension in management, which is reflected through the development and implementation of a new concept in enterprise maintenance. The Bayesian model can equally effectively address the problems of traditional maintenance. This study have proved the feasibility of the model through the simulation experiments. The results are presented in [Milojević 2012].

7.2 Decision support system for oil filtering problem

Risk management in mechanical engineering is a continual process that requires the application of the appropriate tools, procedures and methodologies to avoid the risk or to keep it within certain limits. This study advances the currently available formal models of risk management in mechanical engineering through the BN implemented in $C\sharp$. The Bayesian approach is selected because it successfully solves the problems that are characterised by a risk analysis. There are three ways to incorporate Bayesian model in risk analysis. The first way is to take the full assessment and decision making. Another "forcing" Bayesian model [Jensen 2001], [Wenbin 2008] can only be used for the assessment of risk allocation. Finally, the third method uses Bayesian model as the means to select or to input distribution parameterisations for a model risk. The aim of this work is to highlight the advantages and disadvantages of Bayesian models.

7.2.1 Problem statement

The oil pump draws the oil with the help of an overhead tank from the storage of crude oil in an underground barrel, with 50 tons capacity. The overhead tank has a capacity of 2 tons. From the tank, the oil is transported to the machine where filtering occurs through a circular filtering process. The oil is circulated in a process consisting of circulator pumps, centrifugal pumps, vacuum pumps, a unit for degassing, heaters, filters and water cooling. Once the circular process is started, the oil takes the maximum flow from the reservoir of 2 tones, passes the filtering process through the machine and goes back into the tank. Once everything is filtered out, the oil is drained into a containers for use. The parts of this machine are as follows:

1. the warehouse crude oil in an underground barrel, of 50 tons (x_1);
2. the overhead tank of 2 tons (x_2);
3. the machine for a filtering process (x_3);
4. the unit for degassing (x_4);
5. the rough vacuum pump (x_5);

6. the fine vacuum pump (x_6);
7. the filters for the separation of mechanical purity (x_7);
8. the heaters (x_8);
9. the water tank for cooling (x_9).

7.2.2 Bayesian implementation

In this research, the behavior and possible scenarios of the oil filtering machine were examined. A comparative analysis is performed between the results obtained by the computer program and historical probabilities. At the input of the computer system, the information about the resulting failure of the oil filtering machine is given, and on the output, the BN model gives a prediction of the oil quality with some probabilities.

The output of the system are events A, B, C, D and F, denoting a result of the oil filtering. Case A denotes that the oil is filtered, case B denotes that the oil is not filtered, case C means that the oil is slowly filtered, case D means that the oil is filtered with errors and F stands for the cancellation of the entire installation.

The actual Bayesian model for this study is shown in the Fig.7.5.

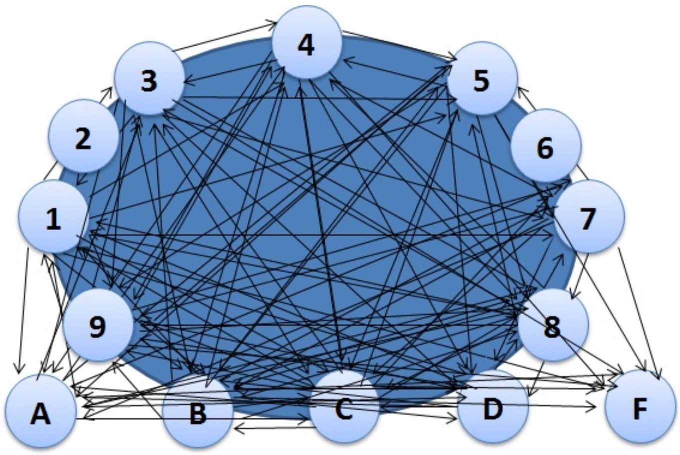

Figure 7.5: Inputs and Outputs of the Bayesian network taken from [Glišović 2013]

7.2.3 Results

Bayesian model was applied to predict the behaviour of an oil filtering machine. The model is implemented in C♯ and tested on real problems. Various scenarios were tested and the results are shown in Tables 7.4, 7.5 and 7.6. As in the previous case the BN obtained results are consistent with the empirical data. The provided predictions in the case of simultaneous failures of several subsystems enabled the identification of the most important parts of the considered machine that can then be given a special attention by supervisors.

The proposed BN model can be used for the development of the modern software for prediction in the area of mechanical engineering and decision making, even when there is not enough information. The extension of this work in the future is seen in developing of an expert system for automation of the production process in order to reduce the failures, all based on Bayesian probabilities. The results are presented in [Glišović 2013].

Table 7.4: Showing results of the historical conclusion of machines parts and the results obtained by Bayesian approach from [Glišović 2013]

	Bayesian network	Historical conclusion
The oil is filtered out A	91%	90%
The oil is not filtered out B	9.8%	10%
The oil is slowly filtered out C	29%	30%
The oil is poorly filtered out (with errors) D	10.2%	10%
A total installation cancelled X	9.9%	10%

Table 7.5: Representation of the failure probabilities of a subsystem if the outcome is known from [Glišović 2013]

The final outcome/ Actuator	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
The oil is filtered out A	100%	100%	90%	0%	0%	0%	0%	0%	10%
The oil is not filtered out B	0%	0%	10%	0%	0%	0%	90%	80%	10%
The oil is slowly filtered out C	0%	0%	0%	90%	90%	90%	10%	20%	10%
The oil is poor filtered out (with errors) D	100%	100%	100%	0%	0%	0%	0%	0%	20%

Table 7.6: The outcome probabilities for multiple failures simultaneously from [Glišović 2013]

	A	B	C	D	F
1,2	2%	10%	1%	0%	4,5%
2,3	87%	1%	3.2%	0%	4.8%
7,9	89.8%	8%	5.2%	10%	9.7%
4,5,6	9%	0.01%	0.02%	12.9%	0.01%
1, 9, 5	81%	5%	0.9%	4%	3.6%

Different type of BN are used for prediction of power plants, such as temporal BN [Hernandez-Leal 2011] or fuzzy BN [Alamaniotis 2014]. Bayesian calibration of power plant is also analysed in [Boksteen 2014]. In all of given cases [Hernandez-Leal 2011], [Alamaniotis 2014], [Boksteen 2014], prediction errors, with comparison to the input in the form of historical probabilities are analysed. In case of fuzzy BN application [Alamaniotis 2014] for calibration power plant, linear regression is used. In estimation of the prediction results for the remaining useful life of turbine (table 2 in [Alamaniotis 2014]), authors claimed 3.3% of minimum error and 42.5% of maximum error occurred. The maximal prediction error in [Milojević 2012] and [Glišović 2013] is 3%. These results [Milojević 2012, Glišović 2013] give better prediction for 39.5%, comparing to the result in [Alamaniotis 2014]. Temporal BN method is applied for the diagnosis of the failures in the combined cycle power plant [Hernandez-Leal 2011] and the minimal obtained error is 15.29%. This is pointing out that the results obtained in [Milojević 2012] and [Glišović 2013] are almost 5 time precise than the results in [Hernandez-Leal 2011]. In [Boksteen 2014], the prediction of plant power and efficiency as a function of ambient temperature is obtained with approximately 95% of certainty. This is leading to the approximate error of 5% which makes the results in [Milojević 2012] and [Glišović 2013] better for 2%.

Chapter 8

Concluding Remarks and Directions for Further Research

For the assessment of working capacity of machine parts, components and assemblies, analytical and experimental modelling of their characteristics based on the values of the adjustable parameters is essential. It is possible to determine desired parameter values by solving the optimisation problems over defined model. The most common optimisation problems related to mechanical elements and assemblies are the optimisation of the design of rolling bearings, optimisation of rotor system, gears design optimisation, workload optimisation of gears and/or gear pairs, etc.

If the problem involves the optimisation of objective function and constraints which solution is not achievable in a reasonable time, it s necessary to develop a more efficient and more reliable techniques for solving these kind of problems. Therefore, meta-heuristics can be of a great help because they are significantly different from traditional optimisation methods and represent the only possible tools to deal with optimisation problems that cannot be solved deterministically. For the optimisation problems in mechanical engineering, neural networks, genetic algorithms, particle swarm optimisation etc, are usually used.

Models from the available literature are taking into account a group of influential parameters on the observed target value, while all other parameters are allocated with fixed values in the purpose of simplification. Parameters which should be fixed can be selected in different ways, taking into account the importance to assure generality of such a model. For example, mathematical model of meshed cylindrical gear pair, determined in accordance with ISO standards, depends on several tenths of geometric design parameters. Selection of the most influential parameters is an important research topic in which meta-heuristics can be of great help. In situations when mathematical model is not known or it is too complicated for calculations, other techniques as Bayesian or Neural network could be applied.

Within this research, several topics are investigated, namely optimisation of helical and spur gears [Milojević 2013], optimisation of planetary gears [Rosić 2011b, Rosić 2011a], optimisation of dynamical load ratings and rating life at bearings [Milojević 2014] and reliability assessment of mechanical systems by Bayesian Networks [Milojević 2012, Glišović 2013].

Optimisation of transverse load distribution factor of helical and spur gears. One of the investigated problems related to the gears is the optimisation of the transverse load factor at helical and spur gears. Load transmission by gear pairs is not constant and it is followed by non-uniform load distribution in the meshing process. The opposite assumption, where the load factor does not change over time along the line of contact, was made. The goal was to identify parameters with the largest influence on violating this assumption. It was also necessary to determine the extent of their changes. For the purposes of developing this model, all parameters which determine transverse load factor, according to [ISO 6336-1], [ISO 6336-2], [ISO 6336-3], [ISO 1328], [ISO 53] and [ISO 21771] were considered as relevant. The proposed optimisation algorithm is based on GA and involves an additional local search optimisation procedure called at the end in order to improve the solution obtained by GA. Such a hybrid algorithm has 12 direct input variables affecting the objective function. The main procedure is divided into several sub-procedures: Calculation of geometry, Calculation of the stiffness and Calculation of the value of total contact ratio. Since the mathematical model of this problem is nonlinear and continuous, the corresponding computational methods, such as Newton-Raphson method and interpolation of three-dimensional nonlinear function, are implemented. This implementation is available as publicly available MatLab library for computation of transversal load distribution factor optimisation at helical and spur gears, based on proposed framework and using ISO standards. Library includes implementation of proposed method for optimisation of transversal load distribution factor optimisation at helical and spur gears. The presented results [Milojević 2013] showed that the most influential variable to the value of load transverse factor is helix angle, and in addition, the profile shift coefficients also affected changing the value of load transverse factor. It is noted that for any number of teeth (from the range $18 - 54$) and any gear ratio (from the range $1 - 5$), this method achieves a value 1 of the load transverse factor, which therefore corresponds to uniform load distribution. Helix angle can take any value in the range of standard values from $0° - 30°$, but generally was assigned the values between $20° - 30°$ to make the convergence of load transverse factor to 1. The similar method can be applied to other problems in mechanical engineering. The following includes list of contributions published in [Milojević 2013] and representing significant part of this thesis. Each item in the list has annotation written in parenthesis, that denotes in which domain is particular contribution.

- Mathematics: Optimisation model for load distribution at helical gears is given, based on ISO standards such that enable easy deployment of metaheuristic methods.

- Mechanical Engineering: For the first time in the literature, the Newton-Raphson method is applied for computation of non-linear operating (working) pressure angle equation at helical and spur gears.

- Mechanical Engineering: Definition of functional relation between the accuracy grade, standard modulus and pitch diameter, by applying ISO standards.

- Mechanical Engineering: A framework for computation transversal load distribution factor at helical and spur gears.
- Mathematics & Mechanical Engineering: Construction and definition of geometric module for computation geometry of helical and spur gears, according to ISO standards. Geometric module can be used as independent component, with six input values, for solving optimisation problems in domains different than one analysed in this thesis. For example, other types of gears or mechanical elements.
- Mechanical Engineering: In [Milojević 2013] 4 additional parameters influencing load distribution factor have been analysed with respect to the publications [Zhang 2010], [Pedrero 2010] and [Pedrero 2011].
- Mechanical Engineering: There is two times better optimisation of parameter z_1 with respect to the same one optimised in [Zhang 2010]. The best optimisation output in [Milojević 2013] is 44, but in [Zhang 2010] the best optimisation result is 34.
- Mechanical Engineering: Results are showing that the optimised value of module m_n in [Milojević 2013] is 7.5 times improved with respect to [Zhang 2010]. In [Pedrero 2010] and [Pedrero 2011] the parameter m_n is taken as fixed input and therefore, it is not optimised at all.

Future work in solving optimisation problem of transversal load distribution factor of helical and spur gears can be seen in application of other optimisation methods, beside GA. Comparative analysis of the achieved results should be provided. Development of transversal load distribution factor models based on ISO standards should be performed for other types of gears, such as swarm or crown. Application of GA or any other method for solving these kind of problems should be performed.

Optimisation of planetary gear trains. Planetary gear trains take a very significant place among the gear transmissions which are used in many branches of industry such as automobile transmissions, aircrafts, marine vessels, machine tool gear boxes, gas turbine gear box, robot manipulators, etc. Planetary gear trains have a number of advantages over the transmission with fixed shafts. The multi-objective nonlinear optimisation of planetary gear trains was considered. The weighting method is used to approximate the Pareto set. This method transforms the multi-objective optimisation problem into single-objective optimisation problem by associating each objective function a weighting coefficient and then minimising the weighted sum of the objectives. Nine functions are optimised simultaneously and the corresponding weighting coefficients w_i, for $i = 1, ..., 9$ have randomly assigned values. This kind of modelling reflects very well the design process in which usually several conflicting objectives have to be satisfied such as the efficiency of planetary gear trains and the distance between centers of sun gear and planetary gear. The effect of changes of the design parameters gives useful information regarding the sensitivity of various features in the model. A Pareto set, presented as a plot of the efficiency and axial distance of the planetary gear train, gives a quantitative description of the compromise between efficiency and size. The proposed GA-based approach produced quite satisfactory results promptly supplying the designer with the preliminary design parameters of planetary gear train for different gear ratios. The obtained results [Rosić 2011a, Rosić 2011b] showed that the genetic algorithm is useful and applicable for optimisation of planetary gears design. The concrete scientific contributions related to this topic are:

- Mathematics & Mechanical Engineering: It is proposed a formal model of planetary gear train in operation with considered losses in the mechanical energy arising as a consequence of the friction between the contact surfaces of the meshing teeth and the friction in the bearings.
- Mechanical Engineering: Application of GA and weighted coefficient method for approximation of Pareto set in order to transform the multi-objective optimisation problem into a single-objective optimisation on an example of solving planetary gear train problem.
- Mechanical Engineering: The proposed model considers sliding losses, which are the result of the friction forces developed as the teeth slide across each other, the rolling losses resulting from the formation of an elastohydrodynamic film.
- Mechanical Engineering: Based upon a geometrical interpretation of the results in the criterion space, it is concluded that axial distance and efficiency are mutually conflicting. Also, it is determined that there exists a very strong correlation between the axial distance and the outer diameter of the planetary gear train.
- Mechanical Engineering: First time in the literature, GA method is applied for optimisation of the planetary gear train efficiency.
- Mechanical Engineering: Complexity of analysis can be seen in optimisation of 9 functions, while in [Castillo 2002] only 3 relationships have been analysed analytically.
- Mechanical Engineering: 7 additional conflicted functions have been analysed comparing to the work in [Qing-Chun 2008] and [Tripathi 2010].
- Mechanical Engineering: Optimisation of 10 influential parameters have been covered, while in [Cho 2006], only one relationship between the inputs and outputs is analysed, covering 3 influential parameters.
- Mechanical Engineering: In [Tripathi 2010], only 6 constraint functions have been analysed, while this investigation is considering 8 constraints have taken into account due to the complexity of minimisation of the minimal elastohydrodynamic lubrication film.

For the future work in solving the multi-objective nonlinear optimisation of planetary gear trains, the author recommends the changing of the values of weighting coefficients w_i, for $i = 1, ..., 9$ with respect to the experiments that are already performed. It is expected that changing the values assigned to weighting coefficients will influence the values of solutions for nine conflicting objectives. Also, the possible application of other methods, beside GA, should lead to achieving a successful comparative results.

Optimisation of ball bearing dynamical load ratings and rating life. The multi-objective optimisation of bearings dynamical load ratings and working life, having in mind that these objectives are not conflicting, has also been considered. Changing the internal geometry, for example number of balls, leads to changes in dynamic capacity and in the length of bearing life in comparison to standard catalog values. Three meta-heuristic methods are applied for the considered problem, namely: Genetic Algorithm (GA), Multistart Pattern Search (MPS) and Multistart Active Set (MAS). MAS method obtained results in a shorter computing time than GA or MPS. MAS obtained the best results for dynamic load capacity in six of eight bearings type. GA gave the best results for working life in five of eight cases. The proposed optimisation methods provide the increase in dynamic capacity with respect to the values from catalog in all eight examples. The average percentage of the improvement is 9.4%, 12.2% and 12.6% for GA, MPS and MAS respectively. Even assuming the producers left some degree of safety, the percentage of the improvement obtained by meta-heuristic optimisation can be considered as significant [Milojević 2014]. The concrete scientific contributions related to this topic are:

- Mathematics: Standard ISO model of the selected radial ball bearings has been transformed to provide effective application of meta-heuristics methods.
- Mathematics & Mechanical Engineering: Comparative analysis of three optimisation methods within optimisation framework proposed in this thesis, with respect to execution time and solution quality. The analysis enabled increasing of dynamic load capacity and working life for eight types of radial ball bearings.
- Mechanical Engineering: Two functions, the dynamic load capacity and maximum working life (under a certain conditions), are optimised simultaneously. However, since the two functions are not conflicted, optimisation is conducted like the problem is single objective.
- Mechanical Engineering: MPS and MAS are for the first time applied for solving optimisation problems of dynamic load capacity and working life.
- Mechanical Engineering: Achievement of increased dynamical load capacity with respect to the standard catalogue values, by applying GA, MPS, and MAS.
- Mechanical Engineering: Proposed geometry structure (geometric parameter values) for improvement of dynamic capacity and rating life of a bearing.

Formal framework is developed according to ISO standards, and experimental evaluation is based on GA, MPS, MAS.

- Mechanical Engineering: The average percentage of the improvement in dynamic capacity with respect to the values from [Bowman] is 9.4%, 12.2% and 12.6% for GA, MPS and MAS, respectively, in 8 cases.

- Mechanical Engineering: The average percentage of the improvement in dynamic capacity with respect to the values from [Gupta 2007] is 13.41%, 20.91%, 18.43% for GA, MPS and MAS, respectively, in 4 cases.

- Mechanical Engineering: The average percentage of the improvement in dynamic capacity with respect to the values from [Rao 2007] is 22.22%, 30.3% and 27.64% for GA, MPS and MAS, respectively, in 4 cases.

Within the future work, a comparative analysis of meta-heuristic methods with respect to other optimisation applied to increasing dynamic load capacity and working life for eight types of radial ball bearings should be performed. The optimisation of the dynamic load capacity and working life should be performed for each bearing type presented in the market.

Bayesian Network (BN) prediction of the reliability assessment of the mechanical systems. As one of the examples, the functionality of the system for painting and varnishing metal product is presented and the method for predicting the impact of the sub-systems' failure on the final product quality is developed. Several different scenarios are tested: probability of system failure if product quality state is known, probability of the product quality conditions if failure of one subsystem occurs or if several subsystems fail simultaneously. The obtained results, also presented in [Milojević 2012], are consistent with the available historical data confirming the usability of the proposed approach. Another Bayesian model was developed to predict the behavior of machine for filtering transformer oil. The model is implemented in C♯ and tested on real problems. By using C♯, implementation and evaluation of model for prediction quality of products in case of machine for filtering transformer oil [Glišović 2013] is provided. Various scenarios were tested and the obtained results are presented in [Glišović 2013]. One of the main results obtained in this study is comparative analysis between the results obtained by BN and historic probabilities and it is concluded that there is no significant difference between them. The proposed approach enables the development of the modern software for prediction in the area of mechanical engineering and decision making even when there is no enough information. The concrete scientific contributions related to this topic are [Milojević 2012], [Glišović 2013]:

- Mathematics: A development of the mathematical model, based on Bayesian Network, for fault prediction in the cases of mechanical plants: for painting and varnishing metal products and for filtering transformer oil.

- Mechanical Engineering: A proposed Bayesian model of prediction quality of products if one or more subsystems or whole system failed, for both mechanical plants: for painting and varnishing metal products and for filtering transformer oil.
- Mechanical Engineering: The development of decision support systems with application of BN in order to predict quality of products if one or more subsystems or whole system failed.
- Mechanical Engineering: A comparative analysis of semifinal products quality prediction if one or more subsystems or whole system failed, based on historical data and BNs.
- Mechanical Engineering: A software system for automation of the production process in order to reduce the failures, all based on Bayesian probabilities.
- Mechanical Engineering: In estimation of the prediction results for the remaining useful life of turbine (table 2 in [Alamaniotis 2014]), authors claimed 3.3% of minimum and 42.5% of maximum error. The maximal prediction error in [Milojević 2012] is 3% which improve the result in [Alamaniotis 2014] for 39.5%.
- Mechanical Engineering: A minimal obtained error while diagnosing the failures, obtained in [Hernandez-Leal 2011] is 15.29%. Comparing to this result, [Milojević 2012] and [Glišović 2013] are almost 5 and 15 times precise, respectively.
- Mechanical Engineering: The prediction of the plant power efficiency, as a function of ambient temperature, is obtained in [Boksteen 2014] with approximate error of 5% which makes the results in [Milojević 2012] and [Glišović 2013] improved for 2% and 4%, respectively.

The future extension of this work includes the developing of an expert system for automation of the production process in order to reduce the failures, based on the Bayesian probabilities.

The successful applications of meta-heuristic and predictive methods on the considered problems encourage the future generation of new software systems useful for supporting the decision making in design and supervising of the complex mechanical systems. This thesis results with several technical contributions in mechanical engineering. It covers implementation of proposed models and algorithms by using C♯ and MatLab.

Bibliography

[Akinnuli 2012] B. O. Akinnuli, T. I. Ogedengbe and K. O. Oladosu *Computer Aided Design and Drafting of Helical Gears*. Journal of Emerging Trends in Engineering and Applied Sciences, vol. 3, no. 6, pages 959–968, 2012. (Cited on pages xxv and 59.)

[Alamaniotis 2014] M. Alamaniotis, A. Grelle and L. H. Tsoukalas *Regression to fuzziness method for estimation of remaining useful life in power plant components*. Mechanical Systems and Signal Processing, vol. 48, pages 188–198, 2014. (Cited on pages 7, 113 and 121.)

[Andersson 1973] S. A. E. Andersson. *On the design of internal Involute Spur Gears*. Transactions of Machine Elements Division: Lund Technical University, Lund, Sweden, 1973. (Cited on page 44.)

[Anderson 1986] N. E. Anderson and S. H. Loewenthal. *Contact Ratio Involute Spur Gears*. J. of Mechanisms, Transmissions and Automation in Design, vol. 108, no. 424, 1986. (Cited on pages 76 and 79.)

[Arora 1989] J. S. Arora. *Introduction to Optimal Design*. McGraw-Hill, New York, 1989. (Cited on page 16.)

[Arya 2012] V. Arya, N. Garg, R. Khandekar, A. Meyerson, K. Munagala, and V. Pandit. *Local Search Heuristics for k-Median and Facility Location Problems*. SIAM J. Comput., vol. 33, no. 3, pages 544–562, 2012. (Cited on page 12.)

[Atanasovska 2006] I. Atanasovska, V. Nikolić. *Influence of addendum modification coefficient on the gear load capacity*. Proceedings of "the 16th European Conference of Fracture (ECF16)", Alexandroupolis, Greece, pages 787–788, 2006. (Cited on page 45.)

[Atanasovska 2007] I. Atanasovska, V. Nikolić. *3D spur gear FEM model for the numerical calculation of face load factor*. The Scientific Journal: Facta Universitatis, Series: Mechanics, Automatic Control and Robotics, vol. 6, no. 1, pages 131–143, 2007. (Cited on page 45.)

[Atanasovska 2010] I. Atanasovska, R. Mitrović, D. Momčilović and A. Subić. *Analysis of the nominal load effects on gear load capacity using the finite element method*. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, vol. 224, no. 11, pages 2539–2548, 2010. (Cited on page 45.)

[Barbieri 2008] M. Barbieri and G. Scagliarini. *Optimization methods for spur gear dynamics*. ENOC, Saint petersburg, Russia, 2008. (Cited on page 47.)

[Bellman 1964] R. E. Bellman. *Dynamic programming: a bibliography of theory and application*. Rand Corporation, 1964. (Cited on page 20.)

[Benedict 1961] G. H. Benedict and B. W. Kelley. *Instantaneous Coefficient of Gear Tooth Friction*. ASLE Transactions, vol. 4, no. 1, pages 59–70, 1961. (Cited on page 79.)

[Ben-Gal 2007] I. Ben-Gal. *in Bayesian Networks by F. Ruggeri, F. Faltin and R. Kenett*. Encyclopedia of Statistics in Quality & Reliability, Wiley & Sons, 2007.

[Boyd 2004] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, New York, 2004. (Cited on pages xxv, 30 and 31.)

[Boksteen 2014] S. Z. Boksteen, J. P. V. Buijtenen, R. Pecnik, D. V. D. Vecht *Bayesian calibration of power plant models for accurate performance prediction*. Energy Conversion and Management, vol. 83, pages 314–324, 2014. (Cited on page 14.)

[Boyd 2004] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004. (Cited on pages 7, 113 and 121.)

[Bretthauera 2002] K. Bretthauer and B. Shetty. *The nonlinear knapsack problem - algorithms and applications*. European Journal of Operational Research, vol. 138, no. 3, pages 459–472, 2002. (Cited on pages 2, 9, 10, 11 and 15.)

[Brunet 2010] F. Brunet. *Contributions to Parametric Image Registration and 3D Surface Reconstruction*. PhD thesis, Université d'Auvergne, 2005. (Cited on page 12.)

[Björck 1996] A. Björck. *Numerical Methods for Least Squares Problems*. Society for Industrial and Applied Mathematics, 1996. (Cited on pages xxv, 19, 20, 21, 23, 24, 25 and 27.)

[Castillo 2002] J. M. del Castillo. *The analytical expression of the efficiency of planetary gear trains*. Mechanism and Machine Theory, vol. 37, no. 2, pages 197–214, 2002. (Cited on page 23.)

[Chakraborty 2003] I. Chakraborty, V. Kumar, B. S. Nair and R. Tiwari. *Rolling element bearing design through genetic algorithms*. Engineering Optimization, vol. 35, no. 6, pages 649–659, 2003. (Cited on pages 87 and 118.)

[Chen 2012] W. Chen, Y. Niu. *Open Press Electrical Accident Risk Assessment based on Bayesian Network*. Procedia Engineering, vol. 43, pages 542–546, 2012. (Cited on pages x, xvi, xxv, 6, 49 and 50.)

[Cho 2006] S. Cho, K. Ahn, J. M. Lee. *Efficiency of the planetary gear hybrid powertrain*. Journal of Automobile Engineering, vol. 220, vol. 220, no. 10, pages 1445–1454, 2006. (Cited on page 56.)

[Ciavarella 1999] M. Ciavarella and G. Demelio. *Numerical methods for the optimisation of specific sliding, stress concentration and fatigue life of gears.* International Journal of Fatigue, vol. 21, no. 5, pages 465–474, 1999. (Cited on pages 6, 87 and 118.)

[Cockerham 1967] G. Cockerham, G. *Computer-Aided Design of Spur or Helical Gear Train.* Computer-Aided Design, vol. 8, no. 2, pages 84–88, 1967. (Cited on page 46.)

[Costin 2010] D. Costin and A. Untaroiu. *Constrained Design Optimization of Rotor-Tilting Pad Bearing Systems.* Journal of Engineering for Gas Turbines and Power, vol. 132, 2010. (Cited on page 44.)

[Crainic 2010] T. Crainic, M. Toulouse. *Parallel Meta-heuristics in M. Gendreau, J. Potvin (Eds.), Handbook of metaheuristics.* Springer, pages 497–541, 2010. (Cited on pages x, xvi, 6 and 51.)

[Crainic 2005] T. G. Crainic, N. Hail. *Parallel meta-heuristics applications, in: E. Alba (Ed.), Parallel Metaheuristics.* John Wiley & Sons, Hoboken, NJ. pages 447–494, 2005. (Cited on pages 33, 34 and 37.)

[Culioli 1994] J. C. Culioli *Introduction á l'optimisation.* Ellipses Marketing, 1994. (Cited on page 37.)

[Cvetković 1996] D. Cvetković and et. al. *Kombinatorna optimizacija (Matematička teorija i algoritmi).* DOPIS, Beograd, 1996. (Cited on pages 21, 22 and 23.)

[Davidović 2006] T. Davidovic. *Scheduling Tasks to Multiprocessor Systems by Applying Metaheuristics.* Doctoral Thesis, Mathematical Faculty, Universisty of Belgrade, 2006. (Cited on pages 1 and 9.)

[Davidović 2011] T. Davidović, D. Ramnjak, M. Šelmić, D. Teodorović. *Bee colony optimization for the p-center problem.* Computers and Operations Research, vol. 38, no. 10, pages 1367–1376, 2011. (Cited on pages xxv, 33, 34, 39, 40 and 41.)

(Cited on page 12.)

[Davis 1999] T. A. Davis and W. W. Hager. *Modifying a sparse cholesky factorization.* SIAM Journal on Matrix Analysis and Applications, vol. 20, pages 606–627, 1999. (Cited on page 27.)

[Darwiche 2009] A. Darwiche. *Modeling and reasoning with Bayesian networks.* Cambridge University Press, 2009. (Cited on pages xi, xvi and 108.)

[Dowson 1977] D. Dowson and G. R. Higginson. *Elastohydrodynamic lubrication, 1st edition.* 2nd edition, Pergamon, Oxford, 1977. (Cited on page 79.)

[Errichello 1979] R. Errichello. *State-of-art review: gear dynamics.* Trans. ASME, J. Mech. Des., vol. 101, no. 3, pages 368–372, 1979. (Cited on page 43.)

[Eschenauer 1990] H. Eschenauer, J. Koski and A. Osycka. *Multicriteria Design Optimisation*. Springer-Verlag, 1990. (Cited on pages 15 and 16.)

[Faggioni 2011] M. Faggioni and F. S. Samani. *Dynamic optimization of spur gears*. Mechanism and machine theory, vol. 46, pages 544–557, 2011. (Cited on page 47.)

[Johnson 1995] D.S. Johnson, and L.A. McGeoch. *The traveling salesman problem: a case study in local optimization*. November, 1995. (Cited on page 2.)

[Flodin 2001] A. Flodin and S. Andersson. *A simplified model for wear prediction in helical gears*. Wear, vol. 249, pages 285–292, 2001. (Cited on page 44.)

[Glišović 2013] N. Glišović and M. Milojević. *Decision support system for mechanical engineering*. International conference on applied internet and information technologies, Zrenjanin, Serbia, 2013. (Cited on pages x, xi, xii, xvi, xvii, xxvi, xxvii, 7, 111, 112, 113, 115, 120 and 121.)

[Glover 1986] F. Glover. *Future Paths for Integer Programming and Links to Artificial Intelligence*. Computers and Operations Research, vol. 13, no. 5, pages 533–549, 1986. (Cited on page 35.)

[Glover 1989] F. Glover. *Tabu Search - Part 1*. ORSA Journal on Computing, vol. 1, no. 2, pages 190–206, 1989. (Cited on page 35.)

[Glover 2003] F. Glover and A. Kochenberger. *Handbook of Meta-heuristics*. Springer, 2003. (Cited on pages 33 and 37.)

[Gologlu 2009] C. Gologlu and M. Zeyveli *A genetic approach to automate preliminary design of gear drives*. Computers and Industrial Engineering, vol. 57, no. 3, pages 1043–1051, 2009. (Cited on page 46.)

[Gupta 2007] S. Gupta, R. Tiwari and S. B. Nair. *Multi-objective design optimisation of rolling bearings using genetic algorithms*. Mechanism and Machine Theory, vol. 42, pages 1418–1443, 2007. (Cited on pages x, xvi, xxv, xxvii, 6, 48, 51, 91, 92, 100, 101 and 120.)

[Guyonnea 2013] D. Guyonneau, E. Mermoz, J. Mailhé, J. M. Sprauel and J. M. Linares. *Stress optimization and study of the sensitivity to geometric variations of a spur gear tooth profile*. Mechanics and Industry, vol. 14, pages 31–41, 2013. (Cited on page 47.)

[Hansen 2010] P. Hansen, N. Mladenovic and J.A.M. Perez. *Variable neighbourhood search: methods and applications*. Annals of Operations Research, vol. 175, pages 367–407, 2010. (Cited on page 35.)

[Harris 1958] S. L. Harris. *Dynamic load on the teeth of spur gears*. Proc. Instn Meth. Engrs, vol. 172, pages 87–112, 1958. (Cited on page 43.)

[Haupt 2004] R. L. Haupt and S. E. Haupt. *Practical Genetic Algorithms*. John Wiley and Sons, Inc. Publication, 2004. (Cited on pages 37 and 77.)

[Hernandez-Leal 2011] P. Hernandez-Leal, L. E. Sucar, J. A. Gonzalez, E. F. Morales and P. H. Ibarguengoytia. *Learning temporal Bayesian networks for power plant diagnosis*. Modern Approaches in Applied Intelligence, Springer, pages 39–48, 2011. (Cited on pages 7, 113 and 121.)

[Heywood 1952] R. B. Heywood. *Designing by Photoelasticity*. Chapman and Hall, Ltd, 1952. (Cited on page 44.)

[Holland 1975] J. Holland. *Adaptation In Natural and Artificial Systems*. The University of Michigan Press, Ann Arbour, 1975. (Cited on page 77.)

[Howard 2001] I. Howard and S. Jia. *The Dynamic Modeling of Spur Gear in Mesh Including Friction and A Crack*. Mechanical System and Signal Processing, vol. 15, no. 5, pages 831–853, 2001. (Cited on page 55.)

[Jabbour 2009] T. Jabbour and G. Asmar. *Stress calculation for plastic helical gears under a real transverse contact ratio*. Mechanism and Machine Theory, vol. 44, no. 12, pages 2236–2247, 2009. (Cited on page 45.)

[Jensen 2001] F. V. Jensen. *Bayesian Networks and Decision Graphs*. Springer, New York, USA, 2001. (Cited on page 110.)

[Kahraman 1992] A. Kahraman. *Dynamic Analysis of Geared Rotors by Finite Elements*. Journal of Mechanical Design, vol. 114, pages 507–514, 1992. (Cited on page 55.)

[Kamenatskaya 1975] M. P. Kamenatskaya. *Computer Aided Design of Optimal Speed Gearbox Transmission Layouts*. Machines and Tooling, vol. 46, no. 9, pages 11–15, 1975. (Cited on page 44.)

[Kasuba 1961] R. Kasuba. *An Analytical and experimental Study of dynamic Loads on Spur Gear Teeth*. Docoral Thesis, University of Illinolis, 1961. (Cited on page 44.)

[Kasuba 1962] R. Kasuba and E. I. Radzimovsky. *Bending Stresses in the Bolts of a Bolted Assembly*. Experimental Mechanics, vol. 2, no. 264, 1962. (Cited on page 76.)

[Kohler 1985] H. Kohler and R. Regan. *The derivation of gear transmission error from pitch error records*. Proc. Instn Mech. Engrs, Part C, Journal of Mechanical Engineering Science, vol. 199(C3), pages 195–201, 1985. (Cited on page 44.)

[Kristensen 2002] K. Kristensen and I. A. Rasmussen. *The use of a Bayesian network in the design of a decision support system for growing malting barley* *without use of pesticides.* Computers and Electronics in Agriculture, vol. 33, pages 197–217, 2002. (Cited on page 56.)

[Krstić 2009] B. Krstić, V. Lazić, R. Nikolić, V. Raičević, V. Jovanović and I. Krstić. *Optimal Strategy for Preventive Maintenance of the Motor Vehicles Clutch.* Journal of the Balkan Tribological Association, vol. 4, 2009. (Cited on page 56.)

[Kubo 1991] A. Kubo and et. al. *Estimation of transmission error of cylindrical involute gears by tooth contact pattern.* JSME int. J., Ser. III, vol. 34, no. 2, pages 252–259, 1991. (Cited on page 44.)

[Kutzkov 2007] K. Kutzkov. *New upper bound for the ♯-SAT problem.* Information Processing Letters, vol. 185, no. 1, pages 1–5, 2007. (Cited on page 12.)

[Kumara 2007] N. S. Kumara. *Genetic Algorithms and genetic programming for multiscale modeling: applications in materials science and chemistry and advances in scalability.* Doctoral disertation, 2007. (Cited on page 38.)

[Lai 2007] L. L. Lai and T. F. Chan. *The Method of Hooke and Jeeves.* in Distributed Generation: Induction and Permanent Magnet Generators, John Wiley & Sons, Ltd. ISBN: 978-0470-06208-1, 2007. (Cited on page 29.)

[Langseth 2007] H. Langseth and L. Portinale. *Bayesian networks in reliability.* Reliability Engineering and System Safety, vol. 92, pages 92–108, 2007. (Cited on page 56.)

[Lauría 2006] E.J.M Lauría and P.J. Duchessi. *A Bayesian Belief Network for IT implementation decision support.* Decision Support Systems, vol. 42, pages 1573–1588, 2006. (Cited on page 56.)

[Levenberg 1944] K. Levenberg. *A method for the solution of certain non-linear problems in least squares.* The Quarterly of Applied Mathematics, vol. 2, pages 164–168, 1944. (Cited on page 24.)

[Leyffer 2005] S. Leyffer. *The return of the active set method.* Oberwolfach Reports vol. 2, no. 1, pages 107–109, 2005. (Cited on pages 25 and 26.)

[Li 2008] R. Li, T. Chang, J. Wang and X. Wei. *Multi-objective optimization design of gear reducer based on adaptive genetic algorithm.* 978-1-4244-1651-6, IEEE, 2008. (Cited on page 47.)

[Li 2007] S. Li. *Finite element analyses for contact strength and bending strength of a pair of spur gears with machining errors, assembly errors and tooth modifications.* Mechanism and Machine Theory, vol. 42, pages 88–114, 2007. (Cited on page 46.)

[Loiola 2007] E. M. Loiola, N. M. Maia de Abreu, P. O. Boaventura-Netto, P. Hahn and T. Querido. *A survey for the quadratic assignment problem.* European Journal of Operational Research, vol. 176, no. 2, pages 657–690, 2007. (Cited on page 12.)

[Liu 2008] C. R. Liua and Y. Choi. *A new methodology for predicting crack initiation life for rolling contact fatigue based on dislocation and crack propagation.* International Journal of Mechanical Sciences, vol. 50, no. 2, pages 117–123, 2008. (Cited on page 48.)

[Liu 2003] G. R. Liu and S. S. Quek. *The Finite Element Method: A Practical Course.* Elsevier Science Ltd, 2003. (Cited on pages xxv, 28 and 32.)

[Madsen 2004] K. Madsen, H. B. Nielsen and O. Tingleff. *Methods for non-linear least squares problems, 2nd edition.* Technical University of Denmark, April 2004. (Cited on pages 24 and 25.)

[Mahadevan 2011] S. Mahadevan, R. Zhang, N. Smith. *Bayesian networks for system reliability reassessment.* Structural Safety, vol. 23, pages 231–251, 2011. (Cited on page 56.)

[Mark 1978] W. D. Mark. *Analysis of the vibratory excitation of gear system, Basic theory.* J. Acoust. Soc. Am., vol. 63, pages 1409–1430, 1978. (Cited on page 43.)

[Mark 1979] W. D. Mark. *Analysis of the vibratory excitation of gear system, tooth error representations, approximations and application.* J. Scouts, Soc. Am., vol. 66, pages 1758–1787, 1979. (Cited on page 43.)

[Martinez 2009] M. Martinez, M. Garcia-Nieto, J. Sanchis and X. Blasco. *Genetic Algorithms Optimisation for Normalized Normal Constraint Method under Pareto Construction.* Advances in Engineering Software, vol. 40, no. 260, 2009. (Cited on page 77.)

[Marquardt 1963] D. Marquardt. *An algorithm for least-squares estimation of nonlinear parameters.* SIAM Journal on Applied Mathematics, vol. 11, pages 431–441, 1963. (Cited on pages 24 and 25.)

[Marquez 2010] D. Marquez, M. Neil and N. Fenton. *Improved reliability modeling using Bayesian networks and dynamic discretization.* Reliability Engineering and System Safety, vol. 95, pages 412–425, 2010. (Cited on page 56.)

[Mendi 2010a] F. Mendi, T. Başkal, K. Boran and F. E. Boran. *Optimization of module, shaft diameter and rolling bearing for spur gear through genetic algorithm.* Expert Systems with Applications, vol. 37, pages 8058–8064, 2010. (Cited on pages x, xvi, 6 and 47.)

[Milojević 2013] M. Milojević. *Optimization of transverse load factor of helical and spur gears using genetic algorithm.* Applied Mathematics and Information Sciences, vol. 4, pages 1323–1331, 2013. (Cited on pages ix, xi, xv, xvii, xxv, xxvii, 4, 5, 54, 57, 63, 65, 66, 68, 70, 71, 72, 73, 74, 75, 115, 116 and 117.)

[Milojević 2012] M. Milojević, Z. Ognjanović and N. Glišović. *Application of Bayesian Network to Reliability Assessment of Mechanical System.* Probabilistic logics and applications, Sept. 27–28, Mathematical Institute of the Serbian Academy of Sciences and Arts, Belgrade, Serbia, 2012. (Cited on pages x, xi, xii, xvi, xvii, xxv, xxvi, xxvii, 7, 105, 106, 107, 108, 109, 110, 113, 115, 120 and 121.)

[Milojević 2014] M. Milojević-Jevrić, T. Davidović. *Meta-heuristics application to optimise bearings dynamical load ratings and rating life.* SYM-OP-IS, XLI Symposium on Operational Research, Divcibare, Serbia, 2014. (Cited on pages x, xi, xvi, xvii, xxv, xxvii, 6, 89, 96, 97, 98, 101, 102, 115 and 119.)

[Misra 2013] N. Misra, N. S. Narayanaswamy, V. Raman, B. S. Shankar. *Solving min ones 2-SAT as fast as vertex cover.* Theoretical Computer Science, vol. 506, no. 30, pages 115–121, 2013. (Cited on page 13.)

[Mladenović 2007] N. Mladenović, J. Brimberg, P. Hansen and José A. Moreno-Pérez. *The p-median problem: A survey of metaheuristic approaches.* European Journal of Operational Research, vol. 179, no. 3, pages 927–939, 2007. (Cited on page 12.)

[Mladenović 1997] N. Mladenović and P. Hansen. *Variable neighborhood search.* Computers and Operations Research, vol. 24, no. 11, pages 1097–1100, 1997. (Cited on page 35.)

[Mladenović 2003] N. Mladenović, M. Labbé and P. Hansen. *Solving the p-Center problem with Tabu Search and Variable Neighborhood Search.* Networks, vol. 42, no. 1, pages 48–64, 2003. (Cited on page 12.)

[Murty 1988] K. G. Murty and Feng-Tien Yu. *Linear complementarity, linear and nonlinear programming.* Berlin, Heldermann, 1988. (Cited on page 26.)

[O'Donnell 1974] W. J. O'Donnell. *Stress and Deflection of Built-in Beams.* ASME Paper, No. 62-WA-16, 1974. (Cited on page 44.)

[Ognjanovic 2004] Z. Ognjanović, N. Krdžavac. *Uvod u teorijsko racunarstvo.* Beograd-Kragujevac, 2004. (Cited on page 13.)

[Matlab 2014] MatLab R2014b. *Optimization Toolbox^{TM} User's Guide.* The MathWorks, Inc., 2014. (Cited on pages 69 and 77.)

[Ozguven 1998] H. N. Ozguven and D. R. Houser. *Mathematical models used in gear dynamics.* Sound Vibr. vol. 121, pages 383–411, 1988. (Cited on page 43.)

[Pearl 1985] J. Pearl. *A model of activated memory for evidential reasoning.* Proceedings of the Cognitive Science Society, pages 329–334, 1985. (Cited on page 106.)

[Pedrero 1996] J. I. Pedrero and M. Artés. *Approximate equation for the addendum modification factors for tooth gears with balanced specific sliding.* Mechanism and Machine Theory, vol. 31, no. 7, pages 925–935, 1996. (Cited on pages ix, xv, 5 and 45.)

[Pedrero 2011] J. Pedrero, M. Pleguezuelos and M. Muñoz. *Critical stress and load conditions for pitting calculations of involute spur and helical gear teeth.* Mechanism and Machine Theory, vol. 46, pages 425–437, 2011. (Cited on pages ix, xv, xxvii, 5, 46, 73, 74, 75 and 117.)

[Pedrero 2010] J. Pedrero, M. Pleguezuelos, M. Artés, A. J. Antona. *Load distribution model along the line of contact for involute external gears.* Mechanism and Machine Theory, vol. 45, pages 780–794, 2010. (Cited on pages ix, xv, xxvii, 5, 45, 46, 73, 74, 75 and 117.)

[Perret-Liaudet 1994] J. Perret-Liaudet and J. Sabot. *Dynamics of Truck Gearbox.* 6th International Power Transmission and Gearing Conference, Scottdale, Arizona, USA, pages 249–258, 1994. (Cited on page 55.)

[Prridy 2005] K. L. Priddy and P.E. Keller. *Artifical Neural Networks. An introductions.* The Society of Photo-Optical Instrumentation Ingineers. (Cited on page 32.)

[Qing-Chun 2008] H. Qing-Chun, M. Rui and D. Fu-Hai. *Multi-objective optimal design study of the two-stage planetary geared systems based on the MATLAB.* Modern Manufacturing Engineering, vol. 3, 2008. (Cited on pages 6, 87 and 118.)

[Radzimovsky 1973] E. I. Radzimovsky, A. Mirarefi, W. E. Broom. *Instantaneous Efficiency and coefficient of friction of an involute gear drive.* Jornal of Engineering for Industry, vol. 1131 (1973). (Cited on page 76.)

[Rafiee 2010] J. Rafiee, M.A. Rafiee and P.W. Tse. *Application of mother wavelet functions for automatic gear and bearing fault diagnosis.* Expert Systems with Applications vol. 37, pages 4568–4579, 2010. (Cited on page 47.)

[Rao 2007] B. R. Rao and R. Tiwari. *Optimum design of rolling element bearings using genetic algorithms.* Mechanism and Machine Theory, vol. 42, no. 2, pages 233–250, 2007. (Cited on pages x, xvi, xxvii, 6, 48, 51, 91, 92, 100, 101 and 120.)

[Richard 2008] M. D. Richard and R. P Lippmann. *Neural Network Classifiers Estimate Bayesian a posteriori Probabilities.* Neural Computation, vol. 3, no. 4, pages 461–483, 2008. (Cited on page 56.)

[Ristivojević 2006] M. Ristivojević. *Zupčanici 2 - Opterecećenje i nosivost*. Zavod za udžbenike, Beograd, 2006. (Cited on page 60.)

[Ristivojević 2002] M. Ristivojević and R. Mitrović. *Raspodela opterećenja*. Mašinski fakultet, Beograd, 2002. (Cited on pages 52 and 53.)

[Roda-Casanova 2013] V. Roda-Casanova, F. T. Sanchez-Marin, I. Gonzalez-Perez, J. L. Iserte and A. Fuentes. *Determination of the ISO face load factor in spur gear drives by the finite element modeling of gears and shafts*. Mechanism and Machine Theory vol. 65, pages 1–13, 2013. (Cited on page 47.)

[Rosić 2011a] B. Rosić, S. Radenović, Lj. Janković and M. Milojević. *Optimization of planetary gear train using multiobjective genetic algorithm*. Journal of the Balkan Tribological Association, vol. 17, no. 3, pages 462–475, 2011. (Cited on pages ix, xi, xv, xvii, xxv, xxvii, 5, 6, 16, 58, 79, 81, 84, 85, 86, 87, 115 and 118.)

[Rosić 2011b] B. Rosić, Lj. Janković and M. Milojević. *Multicriteria optimization of planetary gear train using evolutionary strategies*. Proceedings of the 12th International Conference on Tribology, Serbian Tribology Society, Kragujevac, Serbia, 2011. (Cited on pages ix, xi, xv, xvii, 5, 6, 58, 87, 115 and 118.)

[Rothlauf 2011] F. Rothlauf. *Design of Modern Heuristics*. Principles and Application, Springer, 2011. (Cited on pages 1, 9, 12, 14 and 20.)

[Sánchez 2008] C. V. Sánchez, R. R. Colomer, F. A. V. Pla and S. S. Martínez. *Cylindrical parallel gear trains design by means of genetic optimization*. Selected Proceedings from the 12th International Congress on Project Engineering, Zaragoza, 2008. (Cited on pages viii, xiv and 55.)

[Sánchez 2013] M. B. Sánchez, M. Pleguezuelos, and J. Pedrero. *Enhanced model of load distribution along the line of contact for non-standard involute external gears*. Meccanica, Springer, vol. 48, no. 3, pages 527–543, 2013. (Cited on pages ix, xv and 5.)

[Sánchez 2013] M. B. Sánchez, J. I. Pedrero and M. Pleguezuelos. *Critical stress and load conditions for bending calculations of involute spur and helical gears*. International Journal of Fatigue, vol. 48, pages 28–38, 2013. (Cited on pages ix, xv, 5, 46 and 73.)

[Schneider 2006] J. J. Schneider and S. Kirkpatrick. *Stochastic Optimization*. Springer, 2006. (Cited on pages viii, xiv, 2, 19, 20, 22, 25, 26 and 27.)

[Shigley 1989] J. E. Shigley and C. R. Mischke. *Mechanical Engineering Design*. McGraw-Hill Book Company, New York, 1989. (Cited on pages x, xvi, 6 and 100.)

[Simon 1988] V. Simon. *Load and Stress Distributions in Spur and Helical Gears*. J. of Mech., Trans, and Automation, vol. 110, no. 2, pages 197–202, 1988. (Cited on pages ix, xv, 5, 44 and 73.)

[Sun 2013] H. Sun, Y. Zi, Z. He, J. Yuan, X. Wang and L. Chen. *Customized Multiwavelets for Planetary Gearbox Fault Detection Based on Vibration Sensor Signals*. Sensors, vol. 13, no. 1, pages 1183–1209, 2013. (Cited on pages xxv and 77.)

[ISO 281] Standard ISO 281. *Rolling bearings*. Dynamic load ratings and rating life, 2007. (Cited on pages xxvii, 89, 90, 93, 94 and 95.)

[ISO 6336-1] Standard ISO 6336-1. *Calculation of load capacity of spur and helical gears - Part 1*. Basic principles, introduction and general influence factors, 2009. (Cited on pages viii, xiv, 4, 54, 57, 60, 67 and 116.)

[ISO 6336-2] Standard ISO 6336-2. *Calculation of load capacity of spur and helical gears - Part 2*. Calculation of surface durability (pitting), 2006. (Cited on pages viii, xiv, 4, 53, 57, 60 and 116.)

[ISO 6336-3] Standard ISO 6336-3. *Calculation of load capacity of spur and helical gears - Part 3*. Calculation of tooth bending strenght, First edition, 1996. (Cited on pages viii, xiv, 4, 53, 57, 60 and 116.)

[ISO 75] Standard ISO 76. *Rolling bearings*. Static load ratings, 2006. (Cited on page 89.)

[ISO 1328] Standard ISO 1328-1:1995. *Cylindrical gears - ISO system of accuracy - Part 1*. Definitions and allowable values of deviations relevant to corresponding flanks of gear teeth, 1995. (Cited on pages viii, xiv, 4, 57, 63, 67 and 116.)

[ISO 53] Standard ISO 53:1998. *Cylindrical gears for general and heavy engineering*. Standard basic rack tooth profile, 1998. (Cited on pages viii, xiv, 4, 57, 58 and 116.)

[ISO 21771] Standard ISO 21771:2007. *Gears - Cylindrical involute gears and gear pairs*. Concepts and geometry, 2007. (Cited on pages 4, 57 and 116.)

[Bowman] Ball Bearings 6200 series. *Bearings - Standard* Olitite Bearings. (Cited on pages 6, 100, 101 and 120.)

[Stanimirovic 1999] P. S. Stanimirovic, M. B. Tasic and M. Ristic. *Symbolic implementation of the Hooke-Jeeves method*. Yugoslav Journal of Operations Research, vol. 9, no. 2, pages 285–300, 1999. (Cited on page 29.)

[Szabó 2005] F. J. Szabó. *Multidisciplinary optimization during gear design*. 6th world congresses of structural and multidisciplinary optimization, 2005. (Cited on pages viii, xiv and 47.)

[Talbi 2002] E. G. Talbi. *A taxonomy of hybrid metaheuristics.* Journal of Heuristics, vol. 8, pages 541–564, 2002. (Cited on page 36.)

[Talbi 2009] E. G. Talbi. *Metaheuristics-From Design to Implementation.* John Wiley and Sons, 2009. (Cited on pages 33, 34, 35, 36 and 37.)

[Tripathi 2010] V. K. Tripathi and H. M. Chauhan. *Multi Objective Optimization of Planetary Gear Train.* Lecture Notes in Computer Science, vol. 6457, pages 578–582, 2010. (Cited on pages 6, 87 and 118.)

[Venkataraman 2009] P. Venkataraman. *Applied Optimization with MATLAB Programming.* John Wiley & Sons, Hoboken, New Jersey. (Cited on page 96.)

[Voss 2001] S. Voss. *Meta-heuristics: the state of the art.* Local Search for Planning and Scheduling, LNAI, 2001. (Cited on page 34.)

[Xu 2007] H. Xu and A. Kahraman. *Prediction of Mechanical Efficiency of Parallel – Axis Gear Pairs.* Transactions of ASME, vol. 129, no. 58 (2007). (Cited on page 76.)

[Waghole 2014] V. Waghole and R. Tiwari. *Optimization of needle roller bearing design using novel hybrid methods.* Mechanism and Machine Theory, vol. 72, pages 71–85, 2014. (Cited on pages x, xvi, 6 and 48.)

[Wang 2013] S. Wanga, Q. Mengb and H. Yangc *Global optimization methods for the discrete network design problem.* Transportation Research Part B: Methodological, vol. 50, pages 42–60, 2013. (Cited on page 12.)

[Wenbin 2008] F. Wenbin and C. Yun. *Construction of Evaluation Index System of Operational Benefit Risk of Expressway Enterprises.* In Proceedings of International Conference on Information Management, Innovation Management and Industrial Engineering, pages 210–212, 2008. (Cited on page 110.)

[Wetter 2003] M. Wetter and J. Wright. *Comparison of a generalized pattern search and a genetic algorithm optimization method.* In Proceedings of Eighth International IBPSA Conference Eindhoven, Netherlands, August 11-14, 2003. (Cited on page 29.)

[Whitley 1994] D. Whitley. *A Genetic Algorithm Tutorial, Statistics and Computing.* vol. 4, pages 65–85, 1994. (Cited on page 38.)

[Wong 2011] E. L. S. Wong. *Active-set methods for quadratic programming.* Doctoral Thesis, University Of California, San Diego, 2011. (Cited on page 26.)

[Yang 2008] S. Yang. *Nature - Inspired Metaheuristics Algorithms.* Luniver Press, UK, 2008. (Cited on page 36.)

[Yu 1999] D. C. Yu, T. C. Nguyen and P. Haddawy. *Bayesian network model for reliability assessment of power systems.* IEEE Transactions on Power Systems, vol. 14, no. 2, 1999. (Cited on page 55.)

[Zhang 1999] J. J. Zhang, I. I. Esat and Y. H. Shi. *Load analysis with varying mesh stiffness*. Computers and Structures, vol. 70, no. 3, pages 273–280, 1999. (Cited on pages ix, xv, 5, 54 and 60.)

[Zhang 2010] X. Zhang, Z. Hu, C. Lun and J. Yu. *Optimization design of spur gear reducer based on genetic algorithm*. 978-1-4244-7161-4/10, IEEE, 2010. (Cited on pages ix, xv, xxvii, 5, 45, 73, 74, 75 and 117.)

[Zhang 1999] Y. Zhang and Z. Fang. *Analysis of tooth contact and load distribution of helical gears with crossed axes*. Mechanism and Machine Theory, vol. 34, pages 41–57, 1999. (Cited on pages ix, xv, 5, 44 and 73.)