

Bachelor Rapport Benchmarking

Toepasbaarheid van een Benchmark Methodologie in de Railsector

6 juni 2014



T.W. Duffhues

s1007181

Bachelor Technische Bedrijfskunde

1^e Begeleider Universiteit Twente:

2^e Begeleider Universiteit Twente:

Begeleider Lloyd's Register Rail Europe BV: Ir. G.J. van Silfhout

Bachelor Opdracht Benchmarking voor Universiteit Twente

Versie: 3

Faculteit Management en Bestuur

Dr. Ir. L.L.M. van der Wegen

Dr. J.G.M. Heerkens

UNIVERSITEIT TWENTE.

**Lloyd's
Register**
LIFE MATTERS

Voorwoord

Voor u ligt mijn bachelor rapport ter afsluiting van mijn bachelor Technische Bedrijfskunde aan de Universiteit Twente. Ik heb mijn onderzoek gedaan bij Lloyd's Register Rail Europe, waar ik in de periode van Augustus tot Oktober veel kennis heb opgedaan over een sector waar veel mensen dagelijks mee te maken hebben, zowel in positieve als negatieve zin, de railsector. Een sector waar ik van jongs af aan een interesse in heb gehad en waar ik tijdens mijn stage periode met veel plezier meer van de binnenkant heb mogen zien en meemaken.

Mijn onderzoek richtte zich op het ontwikkelen van een benchmark methodologie, omdat de benchmarking voor Lloyd's een interessante techniek is, maar eentje waar zij niet meer in thuis waren. Omdat de opdracht vrij breed geformuleerd was, duurde het even voordat er een duidelijke en heldere lijn in mijn onderzoek kwam. Hiervoor wil ik ten eerste mijn begeleider bij Lloyd's, Ir. G.J. van Silfhout, hartelijk bedanken. Je pragmatische en doelgericht aanpak was altijd fijn om mee te werken. Daarnaast wil ik mijn begeleiders van de Universiteit Twente, Dr. Ir. van der Wegen en Dr. J.G.M. Heerkens bedanken voor hun sturing en beschaving van mijn onderzoek, vooral in de laatste periode waar de tijdsplanning nogal krap was. Als laatste wil ik Ir. Jack Doomernik van Lloyd's bedanken voor de regelmatige besprekingen over DEA in de spoorsector. Het was altijd enerverend om je te spreken en te zien hoe jij de hogesnelheidssector in kaart wilde brengen voor je PhD aan de Universiteit van Antwerpen.

Als laatste wil ik mijn ouders bedanken voor alle kansen en mogelijkheden die zij mij hebben geboden. Ondanks dat het tempo niet altijd zo hoog lag als zou moeten, is het allemaal toch aardig goed gekomen, gezien het feit dat er nu echt een punt gezet gaat worden achter de bachelor fase.

Enschede, juni 2014

Tom Duffhues

Samenvatting

Vanuit het verleden is Lloyd's Register Rail Europe (LRRE) in aanraking geweest met benchmarking, echter is door de jaren heen de methode in onbruik geraakt. Om de kennis en kunde op het gebied van benchmarking weer op niveau te brengen is dit rapport geschreven. Het doel van het onderzoek is om een benchmark methodologie te ontwikkelen waarmee op een eenvoudige en gestructureerde manier een benchmark onderzoek kan worden opgezet. Belangrijk aan deze methodologie is dat het rekening houdt met de eigenschappen die inherent verbonden zijn aan het spoorwezen. De hoofdvraag die uit dit doel komt is:

“Hoe ziet een benchmark methodologie eruit die gericht is op de railsector en is deze toepasbaar?”

Het antwoord op de gestelde hoofdvraag is de methodologie zoals deze wordt gepresenteerd in dit rapport. De methodologie omvat de volgende acht stappen:

1. Keuze benchmark proces
2. Keuze benchmark partners en identificatie stakeholders
3. Ontwikkeling variabelen voor DEA
4. Verzamelen data
5. Gebruik DEA
6. Analyse resultaten
7. Opstellen verbeterplan
8. Implementatie

Middels een uitgebreid literatuur onderzoek zijn alle facetten van benchmarking aan bod gekomen en zijn er op basis van een aantal modellen deze 8 stappen naar voren gekomen. Omdat er tijdens een benchmark exercitie een vergelijking gemaakt wordt, is er een vergelijkingstechniek nodig. Uit de literatuur kwam naar voren dat Data Envelopment Analysis (DEA) het meest geschikt is voor de railsector vanwege de eigenschappen die het bezit, alsmede het feit dat het vaker is gebruikt in de railsector. Daarnaast is er software pakket gevonden, DEA Frontier Free™ van Zhu (2009), dat het gebruik van DEA eenvoudig maakt. Dit softwarepakket is gebruikt als basis om een Excel bestand te ontwikkelen waarmee op een eenvoudige manier DEA gebruikt kan worden, waarmee de kans op een foutieve invoer geminimaliseerd wordt. Omdat een benchmark valt of staat met de betrokkenheid van de benchmark partners, is er gekozen om binnen deze 8 stappen expliciet twee acceptatie stappen te hebben. De benchmark partners, die uiteraard stakeholders zijn, geven hier toestemming om door te gaan met huidige lijn van het onderzoek. Hiermee is het belang van de stakeholder gewaarborgd binnen de benchmark exercitie.

Om de ontwikkelde methodologie te testen in de praktijk op toepasbaarheid is er een test casus gebruikt. Deze casus is in samenspraak met LRRE opgezet. De casus behelst het klein onderhoud dat door de servicebedrijven van NedTrain wordt uitgevoerd aan de treinstellen van het type Sprinter Light Train (SLT).

Uit de casus blijkt dat de toepasbaarheid goed is. Echter komt ook naar voren dat het inschatten of de resultaten waarheidsgetrouw niet mogelijk is binnen de scope van dit onderzoek. De gevonden resultaten waren onverwacht voor LRRE. De data die zijn gebruikt als input voor de methodologie zijn dezelfde data die NedTrain, de Nederlandse Spoorwegen en LRRE al tijden gebruiken, maar de kwaliteit van deze data staat al jaren ter discussie volgens LRRE.

Om met zekerheid te kunnen zeggen waar het verschil tussen verwachte resultaten en daadwerkelijke resultaten vandaan komt, dient de methodologie als geheel gevalideerd te worden. Binnen de scope

van dit onderzoek is dit niet mogelijk, maar de belangrijkste aanbeveling is om middels een oude casus de methodologie te valideren. Door een oude casus opnieuw uit te voeren en te kijken of de resultaten overeenkomen of juist niet, kan er uitspraak gedaan worden over de validiteit van de benchmark methodologie.

Omdat de data kwaliteit een belangrijk knelpunt is, wordt er ook aanbevolen om te kijken hoe men van te voren kan inschatten wat de kwaliteit is van de data. Dit kan vervolgens worden gebruikt om te bepalen of men DEA gebruikt, of een techniek die minder gevoelig is voor lage kwaliteit data zoals Key Performance Indicatoren. Dit houdt in dat er in stap 3 in plaats van variabelen voor DEA, variabelen voor KPI's worden ontwikkeld en deze in stap 5 gebruik worden. In het nieuwe stappenplan bestaat dus een keuze tussen DEA en KPI's in het geval van hoge dan wel lage kwaliteit data.

Inhoudsopgave

1	Inleiding	8
1.1	Aanleiding	8
1.2	Lloyd's Register Rail Europe	8
1.2.1	Korte leeswijzer	8
1.2.2	Lloyd's Register Group	8
1.2.3	Lloyd's Register Rail Europe BV	8
1.3	Doelstelling	9
1.4	Probleemstelling	9
1.5	Algemene leeswijzer rapport	10
2	Theoretisch kader	11
2.1.1	Korte leeswijzer	11
2.2	Benchmarking	11
2.2.1	Definitie benchmarking	11
2.2.2	Verschillende typen benchmarks	12
2.2.3	Waarom moeten bedrijven benchmarken en wat zijn de voordelen hiervan	14
2.2.4	Wat zijn de voorwaarden en valkuilen van benchmarking	14
2.3	Benchmark modellen	15
2.3.1	Modellen	15
2.3.2	Overzicht	19
2.4	Methoden van prestatie evaluatie	22
2.4.1	DMU, efficiëntie en effectiviteit	22
2.4.2	Methoden van prestatie evaluatie	23
2.4.3	Keuze methode	24
2.5	Data Envelopment Analysis	25
2.5.1	Input/ output	26
2.5.2	Constant-Return-to-Scale/Variable-Return-to-Scale	26
2.5.3	DEA mathematische weergave	26
2.6	Stakeholders	30
2.6.1	Stakeholder Theorie	30
2.6.2	Belang stakeholder	31
2.7	Conclusie	32
3	Benchmark methodologie railsector	33
3.1	Leeswijzer	33
3.1.1	Gebruik van het theoretisch kader	33
3.2	Het model	33
3.2.1	Benchmark model	34
3.3	Benchmark methodologie	36
3.3.1	Stap 1 – keuze benchmark proces	36
3.3.2	Stap 2 – keuze benchmark partners en stakeholder analyse	36
3.3.3	Stap 3 – ontwikkeling van variabelen voor DEA	37

3.3.4	Stap 4 – verzamelen data	37
3.3.5	Stap 5 – gebruik van DEA	38
3.3.6	Stap 6 – analyse van resultaten	39
3.3.7	Stap 7 – opstellen verbeterplan	39
3.3.8	Stap 8 – implementatie	39
3.4	Excel Handleiding	40
4	Benchmark casus	41
4.1.1	Korte leeswijzer	41
4.2	Keuze casus	41
4.3	Casus	41
4.3.1	Stap 1 – keuze benchmark proces	41
4.3.2	Stap 2 – keuze benchmark partners en identificatie stakeholders	41
4.3.3	Stap 3 – ontwikkeling variabelen voor DEA	42
4.3.4	Stap 4 – data verzamelen	43
4.3.5	Stap 5 – gebruik van DEA	44
4.3.6	Stap 6 – analyse resultaten	45
4.3.7	Stap 7 – opstellen verbeterplan	47
4.3.8	Stap 8 – implementatie	47
4.4	Bevindingen casus	48
4.4.1	Scenario's	48
4.4.2	Toepasbaarheid van de methodologie	49
4.4.3	Discussie LRRE	50
4.4.4	Aanbevelingen	50
4.4.5	Conclusie	51
5	Conclusie en aanbevelingen	52
5.1	Conclusie	52
5.2	Aanbevelingen	53
	Literatuurlijst	54
Appendix A	Begrippenlijst	56
Appendix B	Excel Handleiding	57
B.1	Belangrijk	57
B.1.1	Excel Solver	57
B.2	Tabbladen	57
B.2.1	Voorblad	57
B.2.2	Uitleg	57
B.2.3	Formules	58
B.2.4	Data	58
B.2.5	Ruwe Data	58
B.2.6	Overige tabbladen	58
B.3	Overzicht	58
B.4	Modellen beschikbaar in DEA Frontier Free™	58

	B.4.1	Envelopment model	59
	B.4.2	Measure specific model	59
Appendix C	DEA in Excel		61
C.1	Voorblad		61
C.2	Uitleg		62
C.3	Formules		63
Appendix D	Casus scenario's		64
D.1	Scenario 2 – basis, genormeerd naar SLT/Overige treinen		64
	D.1.1	Beschrijving	64
	D.1.2	Variabelen	64
	D.1.3	Uitwerking	64
D.2	Scenario 3 – invloed reset		65
	D.2.1	Beschrijving	65
	D.2.2	Variabelen	65
	D.2.3	Uitwerking	65
D.3	Scenario 4 – invloed vervanging		66
	D.3.1	Beschrijving	66
	D.3.2	Variabelen	66
	D.3.3	Uitwerking	67
D.4	Scenario 5 – invloed drukte I, rustige periode		67
	D.4.1	Beschrijving	67
	D.4.2	Variabelen	68
	D.4.3	Uitwerking	68
D.5	Scenario 6 – invloed drukte II, drukke periode		68
	D.5.1	Beschrijving	68
	D.5.2	Variabelen	69
	D.5.3	Uitwerking	69
Appendix E	Aangepast model		70

1 Inleiding

1.1 Aanleiding

In tijden dat er steeds meer op de kosten en prestatie gelet moet worden, is het vinden van best-practices op een eenduidige en eenvoudige manier belangrijk. Benchmarking kan hiervoor gebruikt worden. Benchmarking houdt in dat verschillende partijen vergeleken worden op een objectieve en systematische manier en dat er lering getrokken wordt van de Best-In-Class (BIC). In het verleden heeft Lloyd's Register (LR) benchmark projecten uitgevoerd, maar de methodiek is de laatste jaren niet gebruikt. Het is voor Lloyd's Register Rail Europe (LRRE) een onbekende methodiek. LRRE heeft de vraag gesteld of benchmarking werkt en of het gebruikt kan worden als dienst voor haar klanten, omdat met relatief weinig inspanning ontdekt kan worden hoe de klant zich verhoudt tot haar branchegenoten of hoe afdelingen intern zich verhouden, of hoe productielocaties ten opzichte van elkaar presteren. Belangrijk is dat benchmark methodologie gebruikt kan worden in de railsector en er een handleiding beschikbaar is voor gebruikers.

1.2 Lloyd's Register Rail Europe

1.2.1 Korte leeswijzer

In § 1.2.2 wordt de missie van Lloyd's Register als geheel uitgelegd. In § 1.2.3 wordt specifiek Lloyd's Register Rail Europe BV kort geïntroduceerd. Lloyd's Register Rail Europe is onderdeel van Lloyd's Register Group.

1.2.2 Lloyd's Register Group

“Veiligheid is waar ons werk al sinds 1760 om draait en we besteden onze tijd, ons geld en onze middelen aan het vervullen van de missie van Lloyd's Register: het beschermen van levens en eigendommen, het bevorderen van onderwijs en onderzoek in transport en engineering. Door ons onafhankelijk advies, brede kennis, diepgaande ervaring en nauwe relaties dragen wij bij aan een veiligere wereld.”

“Wij bieden onafhankelijke garantie aan bedrijven die met hoge risico, kapitaalintensieve assets in de energie- en transportsectoren werken, om de veiligheid van levens, eigendommen en het milieu te verhogen. Zo kunnen onze klanten veilige, verantwoordelijke en duurzame supply chains creëren.”

“De Lloyd's Register Group is één van de wereldleiders op het gebied van het beoordelen van bedrijfsprocessen en producten aan de hand van internationaal erkende normen. De normen zijn afkomstig van grote onafhankelijke instanties of zijn normen die wij zelf hebben ontwikkeld. We streven ernaar om van ontwerp en nieuwbouw tot dagelijks gebruik en uit dienst name complete risicomanagementoplossingen te bieden voor de gehele levensduur en zodoende de veiligheid, integriteit en operationele werking van assets en systemen te garanderen.”

“Onze klanten beheren doorgaans grootschalige, hoogwaardige assets waar de prijs van fouten zeer hoog kunnen zijn, zowel financieel als wat betreft de impact op lokale gemeenschappen en het milieu. In dergelijke omstandigheden hebben organisaties behoefte aan advies en ondersteuning waar ze te allen tijden op kunnen vertrouwen.” (Lloyd's Register)

1.2.3 Lloyd's Register Rail Europe BV

Lloyd's Register Rail Europe BV is onderdeel van de Lloyd's Register Group en valt onder de divisie Transportation. LRRE is een rail consultant en de diensten die geleverd worden beslaan het hele spectrum van de rail sector. LRRE is verdeeld onder in 6 teams: Infra, Interface and Measurement, Design and Build, Maintain; Operate, Safety and Testing, Certification and ISA. Infra, Interface and Measurement houdt zich bezig met de infrastructuur van het spoor. Design and Build houdt zich bezig

met het ontwerp en het bouwen van nieuw aan te schaffen materieel voor de klant. Maintain richt zich op het onderhouden van rollend materieel. Operate houdt zich bezig met de operationele kant van het vervoerswezen. Safety and Testing houdt zich bezig met de veiligheid van materieel en het testen van verschillende systemen. Certification and ISA is namens de overheid bevoegd, om als Notified Body¹ de certificering voor toelating van nieuw materieel op het Nederlandse spoornet uit te voeren. Met de jarenlange ervaring in en expertise van de railsector is LRRE een goede en betrouwbare partner.

1.3 Doelstelling

Omdat LRRE op dit moment niet goed thuis is in benchmarking door het niet gebruiken van de methode in de laatste jaren, is het doel van het onderzoek om inzicht te geven in mogelijke beschikbare methoden en toepassingen. Bij het ontwikkelen van een benchmark methodologie is het belangrijk dat het duidelijk is hoe een benchmark wordt uitgevoerd, waar deze aan moet voldoen en wat de resultaten betekenen. Het onderzoek wordt gedaan aan de hand van academische literatuur, de ervaring van LRRE als rail consultant en de documentatie van voormalig onderdelen van LR. De methodologie beschrijft het hele proces van benchmarking, van welke methode gebruik wordt gemaakt en is toegespitst op de railsector.

De methodologie wordt getest door het uitwerken van een praktijk case. Dit om te kijken wat voor resultaten dit oplevert en een uitspraak te kunnen doen over de toepasbaarheid van de benchmark methodologie. De casus die wordt uitgewerkt, is het benchmarken van de service bedrijven (SB) van NedTrain BV (NTR) en specifiek naar het uitvoeren van klein onderhoud aan de Sprinter Light Train (SLT) van de Nederlandse Spoorwegen (NS). De casus is gekozen, omdat LRRE veel samenwerkt met NTR op dit gebied en er dankzij deze samenwerking data beschikbaar zijn over deze SB's. Door de benchmark uit te voeren, wordt er gekeken naar de toepasbaarheid van de benchmarkmethodologie en naar de eventuele uitkomsten. Deze kunnen aanleiding vormen voor vervolgonderzoeken naar de SB's en het verbeterpotentieel.

1.4 Probleemstelling

LRRE heeft zich in het verleden bezig gehouden met het thema benchmarking. Vanwege het feit dat er historische data beschikbaar zijn en dat er in de huidige tijd steeds meer behoefte is aan het zo efficiënt en effectief mogelijk inzetten van resources, is er de intentie om benchmarking weer aan te bieden aan klanten van LRRE. De hoofdvraag die volgt uit de gestelde intentie is:

“Hoe ziet een benchmark methodologie eruit die gericht is op de railsector en is deze toepasbaar?”

De hoofdvraag is op te splitsen in twee stukken. Het eerste gedeelte van de hoofdvraag is gericht op de benchmarkmethodologie zelf, toegespitst op de railsector. Het tweede gedeelte is gericht op het in de praktijk testen van de methodologie om te kijken wat de resultaten zijn en te kijken naar de toepasbaarheid.

Om de hoofdvraag te beantwoorden en meer structuur in de beantwoording aan te brengen, wordt gebruik gemaakt van deelvragen.

De eerste deelvraag is:

1. *“Wat is benchmarking?”*

Deze kennisvraag is gericht op het omschrijven van benchmarking vanuit de literatuur.

¹ Een notified body (NoBo) is een door een overheid aangewezen keurings- of testinstituut dat van producten moet testen of zij aan de daarvoor geldende (Europese) richtlijnen voldoen.

De tweede deelvraag luidt:

2. “Welke lering kan er getrokken worden uit de documentatie van voormalige onderdelen van LR?”

Er is informatie aanwezig over benchmarking vanuit documentatie van LR. Deze informatie wordt gebruikt om de kennis over benchmarking verder uit te breiden en lering te trekken over hoe de methodologie toegespitst wordt op de railsector. Het antwoord op de vraag bestaat uit delen van hoofdstuk 2 waarbij vermeld wordt dat de desbetreffende informatie vanuit de LR documentatie komt.

De derde deelvraag omvat de methodologie zelf:

3. “Hoe ziet de benchmark methodologie eruit bij gebruik van Data Envelopment Analysis in de railsector?”

De derde deelvraag is erop gericht om op basis van de theorie een model te presenteren. Dit model maakt gebruik van Data Envelopment Analysis (DEA), dat in de spoorsector vaker wordt gebruikt en relatief eenvoudig is. Er wordt, om antwoord te geven op deze deelvraag, uitgelegd welke methoden er allemaal zijn en vervolgens waarom DEA de voorkeur heeft.

Door het volgen van het stappenplan dat gepresenteerd wordt en het gebruik van DEA, is de gebruiker in staat om op een gestructureerde wijze een volledige benchmark op te zetten. Hierbij worden alle aspecten meegenomen die van belang zijn bij een benchmark, zoals besproken bij deelvraag 1 en 2, alsmede de unieke omgeving van de railsector.

De vierde deelvraag gaat over de praktijkcasus en heeft als doel om de methodologie te testen in de praktijk en te kijken naar de toepasbaarheid:

4. “Wat zijn de resultaten bij het uitvoeren van de benchmark bij NedTrain BV die uitgevoerd wordt op de Service Bedrijven van de Nederlandse Spoorwegen?”

De vraag is erop gericht om het model te testen op toepasbaarheid in de praktijk aan de hand van een casus. De casus zal bestaan uit het doen van een benchmark bij de SB's van NTR. Gekeken wordt op welke punten welk SB beter scoort. Er wordt niet gekeken naar de oorzaak van het verschil in prestatie, dit vormt een onderzoek op zich en kan een vervolgstap zijn op dit rapport.

De deelvragen vormen samen het algehele antwoord op de gestelde hoofdvraag die voortvloeit uit de behoefte van LRRE om benchmarking te onderzoeken.

1.5 Algemene leeswijzer rapport

Het rapport is opgebouwd uit 5 delen. Hoofdstuk 1 gaat over de aanleiding, doelstelling, probleemstelling en opzet van het rapport. Hoofdstuk 2 gaat om begripsvorming en uiteenzetting van informatie over benchmarking en lering uit de documentatie van voormalig onderdelen van LR. Het derde gedeelte gaat over de benchmark methodologie in hoofdstuk 3. Dit hoofdstuk is een handleiding en kan gevolgd worden bij het uitvoeren van een benchmark. Door de methodologie te volgen is men in staat een benchmark op een gestructureerde en theoretisch onderbouwde wijze te doen. Het vierde gedeelte is de uitwerking van een casus om de methodologie te testen in de praktijk in hoofdstuk 4. Tenslotte worden in hoofdstuk 6 de conclusies en aanbevelingen van het rapport gegeven, in deze aanbeveling zitten ook de verbeterpunten voor de benchmark methodologie en wordt een aanbeveling gedaan hoe de methodologie gevalideerd kan worden. Bijgevoegd zijn de literatuurlijst en de appendix met een verklarende begrippenlijst, een handleiding voor het gebruik van Excel en eventueel ander materiaal waar naar verwezen wordt.

2 Theoretisch kader

2.1.1 Korte leeswijzer

Voor het beantwoorden van de eerste twee deelvragen (1. Wat is benchmarking en 2. Welke lering kan er getrokken worden uit de documentatie van voormalig onderdelen van LR), is een theoretisch kader geschreven. In dit theoretisch kader wordt informatie gegeven vanuit de academische literatuur, alsmede informatie uit de documentatie van de voormalige onderdelen van LR. Dit dient als basis voor de benchmark methodologie die in hoofdstuk 3 wordt gepresenteerd.

Het kader wordt gestructureerd door de deelvraag op te splitsen in verschillende onderwerpen om zo een gehele beantwoording mogelijk te maken van de deelvraag. De deelvragen worden per paragraaf behandeld. In § 2.2 wordt uitgelegd wat benchmarking inhoudt, welke typen er zijn en wat de voor en nadelen zijn. In § 2.3 worden verschillende modellen behandeld die te vinden zijn in de literatuur. Dit om inzicht te krijgen wat er al is en hoe deze eventueel van elkaar verschillen. In § 2.4 wordt uitgelegd welke verschillende manieren er zijn om prestaties met elkaar te vergelijken op een kwantitatieve manier. In § 2.5 wordt de vergelijkingsmethode Data Envelopment Analysis verder uitgewerkt. In § 2.6 wordt het belang van de stakeholder uitgelegd.

2.2 Benchmarking

2.2.1 Definitie benchmarking

Benchmarking is een methode die gebruikt wordt om vergelijkbare bedrijven, onderdelen van bedrijven of zelfs hele industrieën, regio's of landen met elkaar te vergelijken. Het woord 'benchmark' werd voor het eerst gebruikt in de landmeetkunde waar het vaststellen van een referentiepunt, een benchmark, gebruikt werd voor landbepalingen rond 1842 (Moriarty & Smallman, 2009).

De eerste organisatie die een benchmark daadwerkelijk toepaste op een bedrijfskundige manier was de Xerox Corporatie in 1979 (Camp, 1989). Het doel was om de toenemende kosten, de verminderende kwaliteit en de verhoogde concurrentie van de concurrenten in Japan het hoofd te bieden. Dat deed Xerox door zichzelf te vergelijken met de concurrenten en hier de best-practices uit te destilleren.

Benchmarking leunt op twee principes, namelijk de voorbeeldfunctie van de Best-In-Class (BIC) en de wens als organisatie om de stand van zaken van de BIC te bereiken en te overtreffen (Moriarty & Smallman, 2009). Daarnaast is het belangrijk dat benchmarking niet een eenmalige bezigheid is, maar onderdeel is van een continu verbeterproces (Moffett, Anderson-Gillespie, & McAdam, 2008).

Vanuit de literatuur zijn er verschillende definities beschikbaar over benchmarking.

De eerste definitie is van Camp (1989). Hij borduurde voort op het werk dat Xerox had gedaan en was de eerste die de thematiek benchmarking op de academische kaart zette.

"Benchmarking is de zoektocht naar de best-practices die leiden naar superieure prestaties."

De component continu is een belangrijk onderdeel van benchmarking (Vaziri, 1992):

"Benchmarking is het proces van continu de prestaties van een bedrijf op kritieke klanteisen te vergelijken tegen de beste van haar concurrenten of industrie om te bepalen wat verbeterd moet worden."

De benchmark definitie van Keeyley et al (1997) richt zich op het vinden van de best-practices:

"Benchmarking is een proces voor het identificeren en importeren van best-practices voor het verbeteren van prestaties."

Benchmarking kan ook gedefinieerd worden als een management techniek voor het verbeteren van performance (Yasin, 2002):

“Benchmarking is een management techniek om performance gaps te identificeren en operationele performance te verbeteren.”

De definitie van Moriarty and Smallman (2009) legt de nadruk op het verbeterproces van benchmarking:

“Benchmarking is een modelvoorbeeld gedreven teleologisch² proces dat plaatsvindt binnen de organisatie met als doel om bewust de huidige stand van zaken te veranderen in een superieure stand van zaken.”

Alle definities hebben de overeenkomst dat benchmarking zich richt op het analyseren van de huidige situatie en leren van een ander, als blijkt dat deze beter presteert. Dit gebeurt op een gestructureerde manier. Op basis van deze definities is er een definitie geformuleerd die aansluit bij het doel van het onderzoek:

“Benchmarking is een methode waar op een continue en eenduidige manier een objectieve vergelijking gemaakt wordt van een proces, met als doel om de spelers te identificeren die de best-practices beheersen en te stimuleren om van hen te leren.”

2.2.2 Verschillende typen benchmarks

Om een keuze te maken wat voor benchmark er gedaan wordt, is het van belang om te weten wat voor soort benchmark relevant is voor de situatie. Er zijn verschillende typen benchmarks op verschillende niveaus in een organisatie. Ten eerste zijn er 4 spanningsvelden te onderscheiden waar de benchmark wordt uitgevoerd, deze geven aan waar de benchmark plaats vindt (Moriarty & Smallman, 2009):

1. *Interne Benchmark.* Bij dit type benchmark wordt de prestatie van verschillende bedrijfsonderdelen vergeleken. Hier kan vervolgens intern lering uit worden getrokken. Verder kan een bedrijf ook zichzelf jaar op jaar vergelijken.
2. *Concurrerende Benchmark.* Het bedrijf wordt vergeleken met zijn directe concurrenten binnen de sector waar het bedrijf actief is. Door te leren van de BIC kan er een verbetering bereikt worden.
3. *Functionele Benchmark.* Het bedrijf wordt vergeleken met niet-concurrenten. Dit houdt in dat de benchmark algemene componenten vergelijkt die voorkomen bij de bedrijven die vergeleken worden. Dit alles om hier lering uit te trekken.
4. *Generieke Benchmark.* Bij dit type benchmark worden bedrijven vergeleken uit verschillende industrie sectoren om verbeterpotentieel op te sporen.

Vervolgens zijn er 3 niveaus waarop een benchmark wordt uitgevoerd (Moriarty & Smallman, 2009):

1. *Strategische Benchmark.* De vergelijking wordt gemaakt op strategisch niveau.
2. *Proces Benchmark.* De vergelijking wordt gemaakt op tactisch niveau door processen met elkaar te vergelijken om de meest effectieve operationele processen te identificeren.
3. *Operationele Benchmark.* De vergelijking wordt gemaakt op operationeel niveau door middel van het vergelijken van directe prestaties.

In Tabel 1 staat de relevantie van elke soort spanningsveld ten opzichte van het niveau in de organisatie. Op basis hiervan kan een keuze worden gemaakt op welk niveau een type benchmark de

² Teleologie: ‘leer van de doelen’. In context: er is een reden waarom de een beter presteert dan een ander.

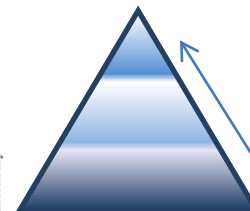
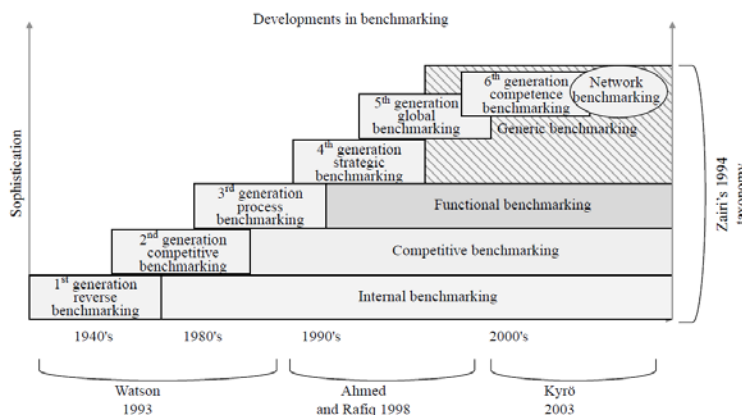
voorkeur heeft. Dit komt de bruikbaarheid van een benchmark ten goede omdat de juiste benchmark wordt gebruikt. Daarnaast moet er gekeken worden of er een momentopname gemaakt wordt van de huidige situatie, of dat er gekeken wordt over een tijdsspanne.

Tabel 1 - Benchmark Relevantie (Moriarty & Smallman, 2009)

	Intern	Extern		
		Concurrerend	Functioneel	Generiek
Strategisch	Laag	Hoog	Laag	Hoog
Proces	Gemiddeld	Laag	Hoog	Hoog
Operationeel	Gemiddeld	Hoog	Gemiddeld	Laag

Naarmate een bedrijf meer ervaren wordt met benchmarking bestaat de mogelijkheid om de scope op te voeren. Dit houdt in dat een bedrijf dat relatief nieuw is met de methode begint op het operationele

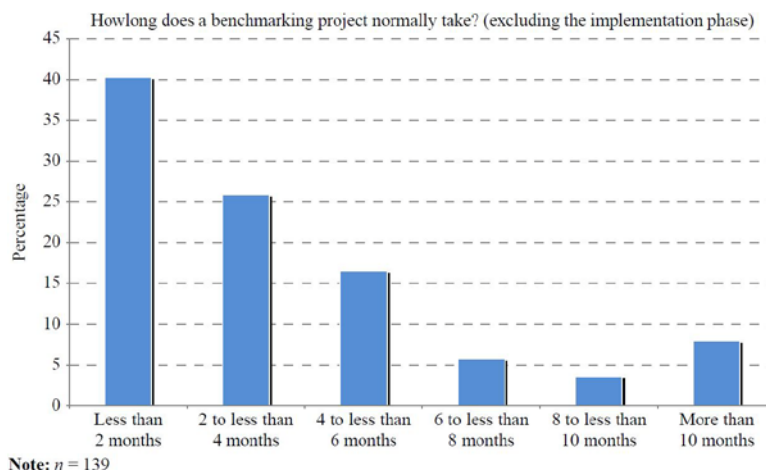
niveau intern gericht. Naarmate de ervaring toeneemt in benchmarking, kan een bedrijf de scope van de benchmark vergroten en meer op strategisch niveau de vergelijking maken (McAdam et al., 2008). Dit wordt ook onderschreven door Moriarty and Smallman (2009) op basis van hun onderzoek naar een meer uniforme benchmark theorie zoals te zien in Figuur 1 en Figuur 2.



Figuur 1 - Organisatie Driehoek

Figuur 2 - Ontwikkeling Benchmark Niveau (Moriarty & Smallman, 2009)

De duur van een benchmark project varieert meestal tussen minder dan 2 tot 6 maanden, zoals te zien in Figuur 3. Dit is exclusief de implementatie fase van de gevonden oplossingen. De grootte van een benchmark team wordt als ideaal gezien bij 4 of minder personen. Het team bestaat meestal uit personen van het senior/middle management en geselecteerde medewerkers met kennis van zaken (Adebanjo, Abbas, & Mann, 2010).



Figuur 3 - Lengte van Benchmark (Adebanjo et al., 2010).

2.2.3 Waarom moeten bedrijven benchmarken en wat zijn de voordelen hiervan

Er zijn verschillende voordelen die benchmarking een organisatie kan opleveren. Om overzicht te houden is er een lijst gegeven met de voordelen, met korte uitleg, die genoemd worden door meerdere auteurs (K. Anderson & McAdam, 2005; Isoraite, 2004; Joo, Nixon, & Stoeberl, 2011; Moffett et al., 2008; Training, 2003; van Hoorn, 2005; Yasin, 2002):

1. *Benchmarking zoekt naar de best presterende werkwijzen.*

Benchmarking draait om het zoeken naar de best werkwijzen. Dit geeft een organisatie de mogelijkheid om zich bewust te worden van deze best-practices, die anders niet bekend waren. Door intern te kijken waar beter gepresteerd wordt, kan verbetering door de hele organisatie worden doorgevoerd.

2. *Benchmarking geeft een objectief beeld waar een organisatie staat ten opzichte van anderen.*

Benchmarking is een methode waarbij organisaties op een zelfde manier worden vergeleken. Het voordeel hiervan is dat er een objectief beeld ontstaat, waarin duidelijk wordt waar de organisatie op dit moment staat ten opzichte van de groep waarmee deze vergeleken wordt.

3. *Benchmarking stelt een organisatie in staat om haar zwakke plekken te identificeren en deze te verbeteren.*

Door op de juiste manier organisaties te vergelijken, krijgt een organisatie inzicht in waar haar zwakke plekken liggen. Deze informatie vormt de basis voor het uitvoeren van een verbeterslag. Dit geeft een streven voor de organisatie om zichzelf te verbeteren. Het blijkt dat organisaties die benchmarking geïmplementeerd hebben en dit als een continu proces beschouwen, beter scoren op operationeel- en werkniveau.

4. *Benchmarking maakt het mogelijk om externe best-practices te vinden.*

Omdat meten op een eenduidige manier wordt gedaan, is het mogelijk om buiten het werkgebied van de organisatie best-practices te vinden. Dit kan verbreding opleveren van inzichten en handelswijzen.

5. *Besluitvorming op basis van feiten.*

Benchmarking beoogt objectief en systematisch te meten. Dit levert feiten op, waardoor de besluitvorming objectiever kan plaatsvinden. Hierdoor wordt getracht de subjectieve mening van partijen te verkleinen en daardoor de vergelijking beter te maken.

6. *Benchmarking maakt het mogelijk om beter in te spelen op de wensen van de klant.*

Door eenduidige en objectieve meting kan een organisatie op een duidelijk manier haar dienstverlening laten zien. Doordat benchmarking gestoeld is op continue verbetering heeft de klant zo ook de tools om inzicht te krijgen in dit proces.

7. *Benchmarking dwingt een organisatie om haar eigen processen in kaart te brengen.*

Om te kunnen vergelijken moet een organisatie een goed beeld hebben van haar eigen processen. Wanneer men zich bewust is van de eigen processen geeft dit zowel inzicht als een betere vergelijking met anderen. Tevens houdt dit ook in dat er binnen een bedrijf meer transparantie ontstaat over werkwijzen.

2.2.4 Wat zijn de voorwaarden en valkuilen van benchmarking

Er is een aantal fundamentele voorwaarden waar rekening gehouden mee moet worden om een benchmark succesvol te laten zijn. Deze voorwaarden zijn vaak ook de grootste valkuilen waardoor een benchmark kan mislukken.

De belangrijkste voorwaarde voor het slagen van een benchmark is het betrokkenheid van de partners. Daarnaast is het belangrijk dat de vergelijking eenduidig gebeurt. Dat wil zeggen dat alle partijen op dezelfde manier vergeleken worden. Dit om het “appel-peer” verhaal te voorkomen (Braadbaart &

Yusnandarshah, 2008). Verder is het van belang dat de kwaliteit van de data zoveel mogelijk gewaarborgd zijn, slechte data leveren slechte vergelijkingen op (Luu, Kim, & Huynh, 2008; Merkert, Smith, & Nash, 2010).

Het doel van benchmarking is om continue verbetering mogelijk te maken. Om op de lange termijn de volle potentie van benchmarking te benutten moet er na de eerste exercitie niet direct gestopt worden met benchmarking (Adebanjo et al., 2010; Luu et al., 2008). Benchmarking is een continue proces wat zichzelf herhaald.

Benchmarking kan voor de betrokken partijen naming, blaming en shaming in de hand werken. Dit kan weer zorgen voor een verminderende betrokkenheid bij het proces, wat niet bevorderlijk is voor de resultaten van de benchmark (van Hoorn, 2005). Dit terwijl participatie essentieel is voor het slagen van de benchmark, omdat openheid nodig is voor het identificeren van verbeterpotentieel (Moriarty, 2011).

Deze openheid kan de indruk wekken dat er bedrijfsgevoelige informatie bekend wordt bij andere partijen, die eventueel concurrent zijn. Benchmarking kan dit ondervangen door het omzetten van absolute waarden naar index waarden (tussen 0 en 1). Hierdoor blijft potentieel bedrijfsgevoelige informatie privé (Caldas, 2013).

2.3 Benchmark modellen

2.3.1 Modellen

Om inzicht te verkrijgen in welke stappen belangrijk zijn voor het uitvoeren van een benchmark, wordt in deze paragraaf een overzicht gegeven van zeven modellen die vanuit de literatuur beschikbaar zijn. Daarnaast zijn er twee modellen die vanuit de documentatie van LR beschikbaar zijn. Dit zijn de laatste twee modellen. Geen van de modellen gaat specifiek over de railsector gaat en is een concreet stappenplan, daarom heeft het overzicht als doel om inzicht te krijgen in welke stappen vaker voorkomen en daardoor belangrijk zijn. Deze belangrijke stappen worden vervolgens aangevuld met de behoefte vanuit de railsector en de stakeholder theorie (2.6).

De modellen kunnen grofweg in twee categorieën verdeeld worden: concrete stappenplannen en conceptuele modellen. De concrete stappenplannen zullen de basis vormen en de conceptuele modellen dienen als bron voor informatie over wat belangrijk is vanuit de literatuur. Een overzicht van alle modellen met de bijbehorende stappen is gegeven in Figuur 9 in 2.3.2 (R. J. H. Anderson, Robin C; Trompet, Mark; Adeney, William E., 2003; Camp, 1989; Finnigan, 1996; Laird, 2011; Spendolini, 1992; Training, 2003; Underground, 2012; van Hoorn, 2005) en Benchmark Approach.

2.3.1.1 BoB model

Tabel 2 - BoB Model (Training, 2003).

Stap	Omschrijving
1.	Identification of areas of benchmarking
2.	Identification of relevant dimension
3.	Identification of indicators and of data needs
4.	Collection and collation of data
5.	Identification of benchmarks and choice of indicators
6.	Analysis of the reasons for performance differences
7.	Analysis of possible remedial measures
8.	Proposals for action and continuous improvement programs
9.	Monitoring results

Het model dat in BoB wordt besproken heeft drie fasen, met elk drie stappen. Het model is ontwikkeld door NEA transport research en training instituut. De eerste drie stappen draaien om het vaststellen van het benchmarkgebied. De volgende drie stappen draaien om het analyseren van de data. De laatste stappen gaan over de keuze van oplossingen en de implementatie. Het rapport van BoB is geschreven voor de spoorsector en is gericht op het vergelijken van passagiersvervoer. Het model bevat concrete stappen.

2.3.1.2 Camp

Tabel 3 - Model van Camp (Camp, 1989).

Stap	Omschrijving
1.	Keuze Proces
2.	Keuze partners
3.	Bepaal methode voor gegevens verzameling en verzamelen data
4.	Bepaal de kloof tussen eigen prestaties en prestaties van partners
5.	Bepaal de gewenste toekomstige prestaties
6.	Publicatie rangen en standen
7.	Functionele doelen opstellen
8.	Opstellen verbeterplannen
9.	Implementatie
10.	Toets benchmark opnieuw aan toekomstige ontwikkeling

Het model van Camp is gebaseerd op Xerox. Het model bevat 10 concrete stappen. Het model is gericht op transparantie dat zich uit in stap 6, publicatie van rangen en standen.

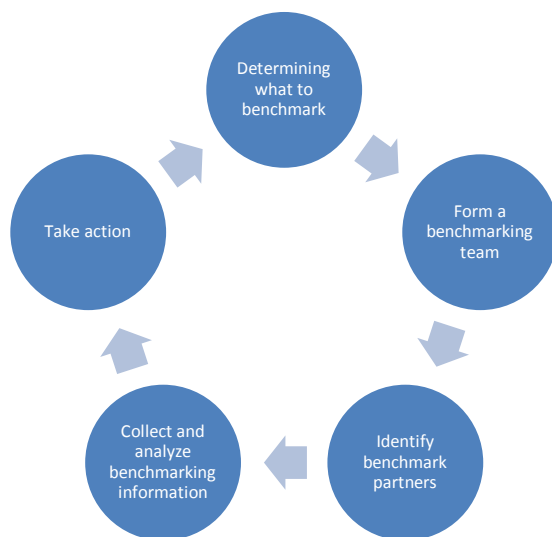
2.3.1.3 Van Hoorn

Tabel 4 - Model van Van Hoorn (van Hoorn, 2005).

Stap	Omschrijving
1.	Keuze en afbakening van het proces
2.	Keuze benchmarking partners
3.	Beschrijving en analyse van proces en indicatoren
4.	Ontwikkeling KPI
5.	Stakeholder keuze KPI
6.	Metten van KPI
7.	Analyse resultaten
8.	Opstellen verbetertrajecten
9.	Implementatie

Van Hoorns model is gebaseerd op dat van Camp. Het heeft echter een iets andere opbouw en is er op gericht om de belangen van de stakeholders te verenigen met het doel van de benchmark. Om deze reden wordt een duidelijk onderscheid gemaakt tussen het ontwikkelen van de KPI's en de keuze van de KPI's. Door deze twee stappen duidelijk gescheiden te houden, vinden er feitelijk twee analyses van de KPI's plaats. Doordat de stakeholders kiezen welke KPI's er daadwerkelijk gebruikt gaan worden, zit er een stuk kwaliteitswaarborging in de KPI's. Tevens zijn de KPI's belangrijk bevonden door de stakeholders en dit creëert draagvlak onder de stakeholder voor de benchmark. Dit is, zoals aangeduid in 2.2.4 essentieel voor het succes van een benchmark.

2.3.1.4 Spendolini

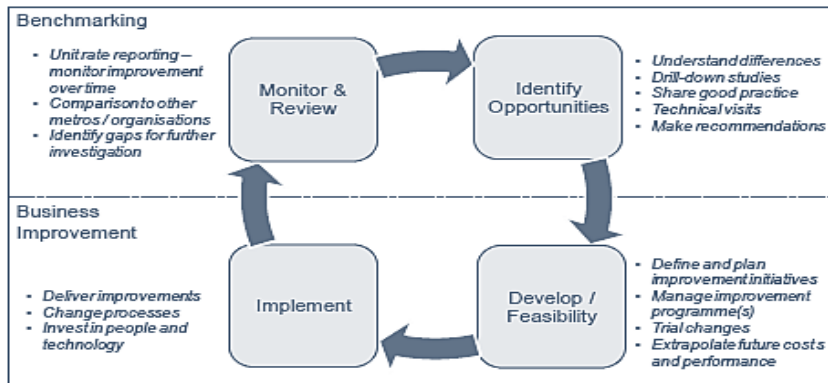


Het model van Spendolini richt zich vooral op de organisatiekant van benchmarking en is een conceptueel model. Dit houdt in dat het meer een richtlijn is die aangehouden wordt en dat het de benchmark op strategisch niveau beschrijft zonder concrete invulling. Het continue aspect van het model komt sterk naar voren, omdat het model in de vorm van een cirkel is.

Figuur 4 - Model van Spendolini (Spendolini, 1992).

2.3.1.5 TFL Rail and Underground

Het model dat in het rapport van tfl rail and underground wordt gebruikt bestaat uit 4 fasen met in het totaal 16 deeltappen. De vier fasen zijn strategisch en de 16 deeltappen zijn de meer geoperationaliseerde stappen die doorlopen worden. Het is een conceptueel model.



Figuur 5 - Model van Tfl and Underground (Underground, 2012).

2.3.1.6 Anderson et al

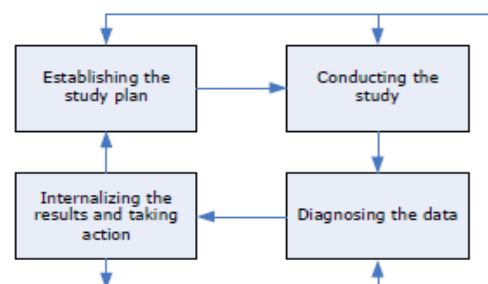


Figuur 6 - Model van Anderson et al (2008).

Het conceptueel model van Anderson et al komt voort uit het rapport dat geschreven is voor de Europese Commissie ten behoeve van Sustainable Transport gericht op de infrastructuurkant van het spoorwezen. De success dimensions zijn er voor om de factoren te identificeren die essentieel zijn voor succes. Op basis hiervan worden de KPI's bepaald. Het model voorziet nog een harmonisatiestap, hiermee kunnen waarden beter vergeleken worden.

2.3.1.7 Finnigan

Het model van Finnigan is een conceptueel model met vier stappen. Volgens Finnigan kan het model meer gezien worden als 4 fasen waar elke benchmark op gebaseerd is. Finnigan was zelf betrokken bij de implementatie van benchmarking bij Xerox in de beginjaren.



Figuur 7 - Model van Finnigan (1996).

2.3.1.8 ERA Model (LR Documentatie)

Het model van de European Railway Agency (2011) komt uit de documentatie van LR. Het is een concreet stappenplan met vijf expliciete stap en nog een suggestie in de vorm van een zesde stap.

1. Wat wordt er gebenchmarkt?
2. Welke niveau van aggregatie wordt er gebruikt?
3. Welke input en output indicators worden er gebruikt?
4. Welke partijen worden vergeleken?
5. Van welke techniek wordt er gebruik gemaakt?
6. (Aandragen van oplossingen)

De vragen omvatten de keuze wat er gebenchmarkt wordt, op welke niveau (strategisch, tactisch, operationeel), welke KPI's en welke partijen. Het model gaat daarna verder met de keuze van de vergelijkingstechniek. Helemaal op het einde in het rapport worden nog suggesties gegeven op basis van de bevindingen, dit kan nog als een 6^e stap, aandragen van oplossingen, worden gezien, hoewel dit niet als zodanig wordt gepresenteerd.

2.3.1.9 Benchmark Approach (LR Documentatie)

De Benchmark Approach is een model wat vanuit de documentatie van LR naar voren komt als een model om een benchmark te structureren. Het bestaat uit vier stappen.



Figuur 8 - Benchmark Approach

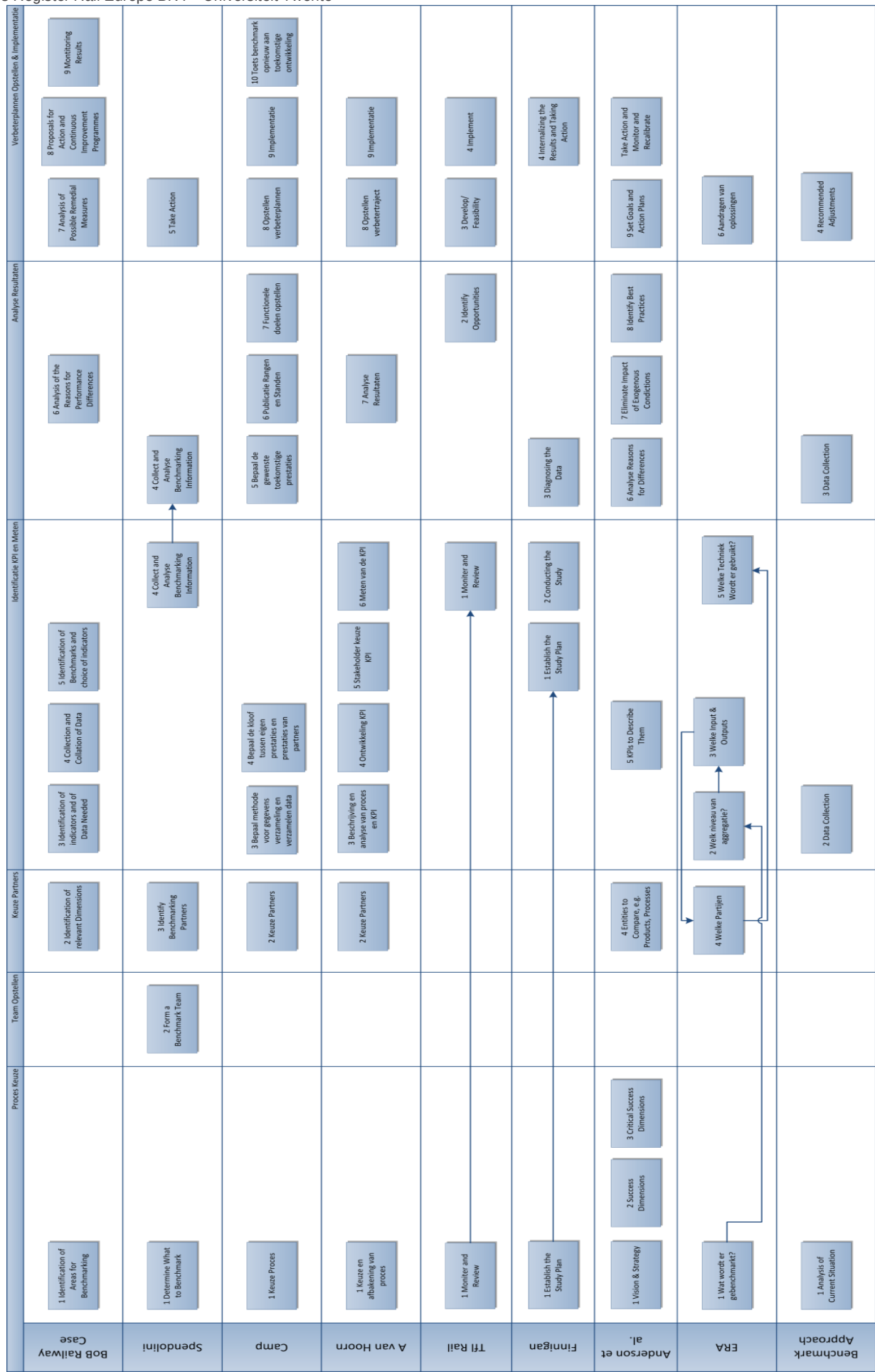
In stap 1 wordt de scope van het project bepaald tezamen met de KPI's, hier wordt gebruik gemaakt van een lijst van KPI's, zoals die destijds werden gebruikt. In de tweede stap worden alle data verzameld die nodig zijn voor het uitvoeren van de benchmark. Vaak wordt hier gebruik gemaakt van uitkomsten en data van vorige onderzoeken die gedaan zijn. Deze worden echter wel anoniem gemaakt om de privacy van die partijen te waarborgen. In stap 3 wordt de daadwerkelijke benchmark uitgevoerd. De vergelijking en normalisatie technieken zijn ontwikkeld door voormalige onderdelen van LR. De normalisatie van gegevens is bij de meeste KPI's al gebeurd in de lijst. Veel gebruikte omzettingen zijn het naar aantal FTE, naar landenniveau met PPP of naar aantal km/jaar. Dit om partijen vergelijkbaar te maken.

2.3.2 Overzicht

Om ordening te krijgen in de overeenkomsten en de verschillen zijn de stappen in de modellen is er in Figuur 9 een overzicht van de verschillende modellen. Om een keuze te maken welke stappen relevant zijn, zijn de stappen verdeeld in 6 groepen. Deze dienen puur ter overzichtelijkheid om onderscheid te maken tussen de verschillende stappen uit de modellen. De groepen zijn de volgende: proceskeuze, team opstellen, keuze partners, identificatie en meten, analyse resultaten en verbeterplannen opstellen en implementatie. Deze groepen vertegenwoordigen stappen die redelijk hetzelfde zijn en maken hierdoor het meer inzichtelijk waar de modellen overeenkomen of verschillen. Vervolgens wordt er met de kennis van de literatuur bepaald welke stappen belangrijk zijn omdat deze stappen vaker voorkomen in de verschillende modellen, of omdat de literatuur aangeeft dat ze belangrijk zijn.

De literatuur geeft meerdere keren aan dat commitment zeer belangrijk is voor het slagen van de benchmark (Braadbaart & Yusnandarshah, 2008; Mitchell et al., 1997; Moriarty, 2011; van Hoorn, 2005). Om de belangen en betrokkenheid van de stakeholder extra te waarborgen kan men werken met

acceptatie moment van de stakeholder (van Hoorn, 2005). Twee logische momenten om dit te doen zijn bij het opstellen van de lijst prestatie indicatoren en bij het opstellen van het verbeterplan. Juist in de spoorsector waar meerdere partijen afhankelijk van elkaar zijn, maar niet perse dezelfde belangen nastreven, zijn deze acceptatie momenten belangrijk. Echter biedt geen een model twee acceptatie momenten, hierdoor is er gekozen om een nieuw stappenplan te construeren wat in 3.2 wordt gepresenteerd. Om het model praktisch op te bouwen is worden de concrete modellen als basis gebruikt. De conceptuele modellen worden gebruikt om te kijken welke stappen vaak voorkomen en daardoor belangrijk zijn. Er zijn een aantal stappen die bij meerdere modellen terugkomen: keuze proces, keuze partners, prestatie indicator ontwikkeling, data verzameling, resultaten, analyse en implementatie. Deze stappen worden als belangrijk beschouwd en dus meegenomen in het model in. Doordat er negen modellen zijn gebruikt als basis en de hieruit meest voorkomende stappen als belangrijk zijn bestempeld is de kans klein dat er belangrijke stappen ontbreken. De negen modellen spannen een behoorlijke tijdsperiode van academisch onderzoek en daardoor wordt er aangenomen dat zij alle relevante ontwikkelingen op het gebied van benchmarking behandeld hebben.



Figuur 9 - Overzicht van Benchmark Modellen

2.4 Methoden van prestatie evaluatie

Bij het uitvoeren van een benchmark wordt gebruik gemaakt van een methode om te vergelijken op prestaties. Deze stap komt terug in modellen in de vorm van bepalen van de kloof in prestaties, meten van prestatie of analyse van data. Deze evaluatie van de prestaties kan middel verschillende technieken uitgevoerd worden. Uitgangspunt in dit onderzoek is dat de methode Data Envelopment Analysis (DEA) wordt gebruikt. Er wordt onderbouwd waarom DEA de meest voor de hand liggende techniek is om te gebruiken in dit onderzoek. DEA vormt ook het uitgangspunt waar het model in hoofdstuk 3 op gebaseerd is.

Er worden eerste verschillende technieken benoemd met voor- en nadelen, op basis van deze eigenschappen en de karakteristieken van de railsector wordt de keuze voor DEA onderbouwd. DEA wordt vervolgens verder uitgewerkt om de werking inzichtelijk te maken.

Een aantal termen wordt vaak gebruikt bij benchmarking. Deze termen worden eerst uitgelegd om duidelijkheid te creëren.

2.4.1 DMU, efficiëntie en effectiviteit

Begrippen die vaak voorkomen in de benchmark literatuur zijn Decision Making Unit (DMU), Efficiëntie en Effectiviteit. Ter volledigheid worden deze begrippen kort uitgelegd zodat er geen verwarring ontstaat over de betekenis in de context benchmarking.

2.4.1.1 DMU

Een Decision Making Unit (DMU) is een term die gebruikt wordt bij benchmarking. Het stelt een organisatie voor die geëvalueerd wordt en dus de mogelijkheid heeft om beslissingen te nemen die de prestaties beïnvloeden (Charnes, 1978; Ozcan, 2008).

2.4.1.2 Efficiëntie

Efficiëntie, of efficiency, is de ratio van output over input. Deze ratio geeft aan in hoeverre een organisatie in staat is zijn input om te zetten in output door middel van een bedrijfsproces. Om efficiëntie te verhogen moet men of de input verlagen ten opzichte van de output, of de output verhogen ten opzichte van de input of allebei verhogen dan wel verlagen, zolang de output maar toeneemt ten opzichte van de input (Ozcan, 2008). Efficiency bestaat vaak uit meerdere soorten efficiency. Deze hebben elk een oorzaak. De meest gebruikte termen in benchmarking gedefinieerd (McCoullough, 2007; Ozcan, 2008):

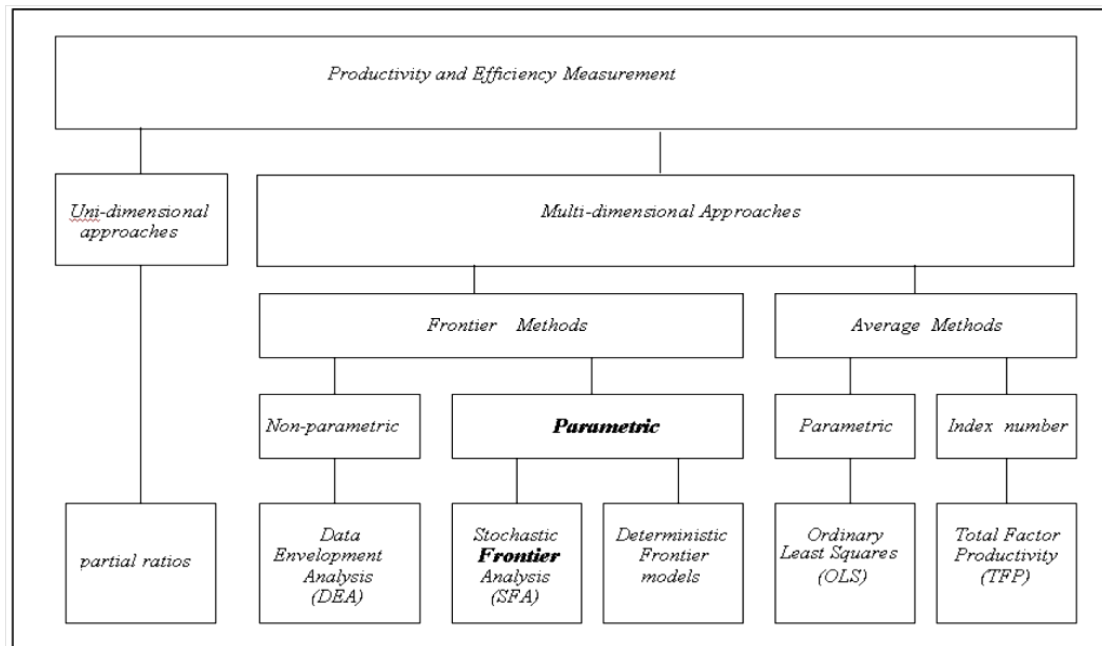
- *Productive Efficiency*. Een bedrijf kan niet meer goederen of diensten produceren zonder dat dit ten koste gaat van een ander product of dienst. Dit gebeurt vaak wanneer een bedrijf op minimale gemiddelde kosten produceert.
- *Allocative Efficiency*. Een bedrijf kan de klanttevredenheid niet verhogen zonder dat dit ten koste gaat van een andere klanttevredenheid. Dit gebeurt vaak wanneer een bedrijf gericht is op marginale kosten. Het doel is om met een optimale mix van inputs de gewenste output te genereren.
- *Scale Efficiency*. De efficiency die optreedt bij verschil in grootte door gebruik te maken van schaalvergroting. Dus dezelfde output met verminderde input.
- *Technical Efficiency*. De verhouding tot wat maximaal te produceren is en wat er geproduceerd wordt binnen de beschikbare middelen. Maximale output ten opzichte van een gegeven input.

2.4.1.3 Effectiviteit

Effectiviteit geeft aan in hoeverre de gewenste uitkomst wordt bereikt. Effectiviteit gaat dus niet over het bedrijfsproces zelf, maar over de uitkomst hiervan ten opzichte van de gewenste uitkomst.

2.4.2 Methoden van prestatie evaluatie

Uit de literatuur blijkt dat er 5 prestatie evaluatie methoden zijn die vaak gebruikt worden in de railsector bij benchmarking (Laird, 2011; Merkert et al., 2010; Ozcan, 2008). Elke methode heeft zijn sterke en minder sterke punten. Van elke methode worden deze sterke en minder sterke punten kort toegelicht. De methode gekozen in dit onderzoek is Data Envelopment Analysis (DEA) aan de hand van een overzichtstabel. Deze methode wordt dan ook verder uitgewerkt in 2.5. De verschillende methoden hebben verschillende statistische eigenschappen. In Figuur 10 is te zien welke eigenschappen de verschillende modellen hebben.



Figuur 10 - Overview of Productivity and Efficiency Measurement (Laird, 2011)

2.4.2.1 Ratio Analyse

De ratio analyse is de basis voor het vergelijken van prestaties. Door het delen van een output over een input krijgt men een efficiency ratio, vaak beter bekend als Key Performance Indicators (KPI). Het voordeel van de ratio analyse is dat er op een eenvoudige en snelle manier gekeken kan worden hoe input wordt omgezet in output en dat de ratio analyse niet heel gevoelig is voor uitschieters. Nadeel is dat er echter maar met één input en één output tegelijk gerekend kan worden. Wanneer er meer ratio's berekend worden, kunnen deze ook op verschillende manier worden geïnterpreteerd, hetgeen tot subjectieve uitspraken kan leiden (Ozcan, 2008).

2.4.2.2 Least Square Regression

De Least Square Regression (LSR) is een populaire techniek die veel wordt gebruikt bij het berekenen van efficiency. LSR kan omgaan met meer inputs en outputs en deze methode is tevens bestand tegen ruis in de data. Daarnaast is het nuttig bij tijdsreeksen en kan gekeken de invloed is van scale efficiency. Tevens is het mogelijk om statistische test toe te passen voor bijvoorbeeld de betrouwbaarheid. Het model heeft ook een aantal voorwaarden en nadelen. Er wordt aangenomen dat alle betrokken partijen efficiënt zijn. Verder wordt er verondersteld dat alle partijen onafhankelijk zijn van elkaar. Gezien de verwevenheid die er is in de railsector is echte onafhankelijkheid moeilijk aan te nemen. Daarnaast is LSR een techniek die stoelt op de neiging naar het gemiddelde, wat niet altijd efficiënte relaties zijn. Zo kan het dat een partij die efficiënt is, volgens LSR minder efficiënt moet worden om beter te scoren. Verder is het niet mogelijk om individuele inefficiënties op te sporen (Ozcan, 2008).

2.4.2.3 Total Factor Productivity

Total Factor Productivity (TFP) is een non-parametrische³ vergelijkingstechniek die daardoor niet afhankelijk is van een soort kansverdeling. De techniek kan met meerdere inputs en outputs werken. De techniek is gebaseerd op index cijfers en geeft daardoor de mogelijkheid om bedrijfsgevoelige informatie te maskeren. De techniek kan of tijdreeksen of verschillen tussen verschillende partijen duiden. Beide tegelijkertijd kan niet. Dit is het grote nadeel van TFP (Ozcan, 2008)

2.4.2.4 Stochastic Frontier Analysis

Stochastic Frontier Analysis (SFA) is een parametrische vergelijkingsmethode. SFA heeft niet nodig dat alle partijen efficiënt zijn met de beschikbare middelen. SFA kan tevens omgaan met ruis in de data. SFA kan gebruikt worden voor het meten van verschillende soorten efficiency, zoals beschreven in 2.4.1.2. Nadeel is wel dat SFA geen onderscheid kan maken tussen technical en allocative efficiency als oorzaak van inefficiëntie. Dit is echter wel nodig als er acties aan de gevonden resultaten verbonden moeten worden. Bovendien kan er door de parametrische eigenschappen een betrouwbaarheidstoets worden gedaan. SFA is daarnaast een relatief lastige methode (Laird, 2011; Ozcan, 2008).

2.4.2.5 Data Envelopment Analysis

DEA is een non-parametrische vergelijkingstechniek. Omdat het een relatieve vergelijkingstechniek is wordt er uitspraak gedaan over de relatieve prestatie en niet over de absolute prestatie. Voordeel hiervan is dat bij het gebruik van DEA het niet nodig is dat alle partijen, in absolute zin, efficiënt zijn. Verder laat DEA meerdere inputs en outputs toe. Omdat lineair programmeren (LP) wordt gebruikt, geeft DEA de mogelijkheid een score te geven op verschillende soorten efficiency zoals genoemd in 2.4.1.2 (Bottasso, 2011; Ozcan, 2008; Roberts, 2012; Winston, 2004). DEA heeft als voordeel dat het zeer bruikbaar is bij weinig vergelijkingspartners. Nadeel van DEA is dat het gevoelig is voor ruis in de data. Dit komt doordat de techniek stoelt op extreme waarde en deze zijn gevoelig voor uitschieters. Daarnaast kan DEA in eerste instantie niet overweg met eventuele betrouwbaarheidstoetsen. Dit kan wel worden ondervangen (Growitsch & Wetzel, 2009).

2.4.3 Keuze methode

Om de verschillende eigenschappen van de methoden inzichtelijk en overzichtelijk te maken is in Tabel 5 een overzicht gegeven. Er zijn verschillende eigenschappen die de railsector typeren en de methode moet geschikt zijn om hiermee om te gaan. Verder is het van belang dat de methode kan omgaan met niet beïnvloedbare variabelen en dat weinig vergelijkingspartners geen probleem hoeft te zijn.

Zoals te zien in Tabel 5 bezitten zowel DEA als SFA veel eigenschappen en zijn hiermee de twee preferente manieren. Echter is het zo dat SFA moeilijker is dan DEA. DEA heeft daarnaast nog een aantal andere voorkeuren. Hansen (2013) beredeneert dat DEA geschikt is voor de railsector vanwege een aantal eigenschappen. Zo is er een quasi monopolistische structuur in de railsector en daardoor zijn er weinig vergelijkingspartners. DEA kan hier goed mee omgaan (Caldas, 2013; Growitsch & Wetzel, 2009; Merkert et al., 2010). Daarnaast is de formele link tussen input en output vaak niet logisch of duidelijk en zijn input en output vaak niet gebaseerd op economische of financiële factoren. Dit komt door de service industrie die de spoorsector is waardoor de beoordeling van de prestatie niet gebeurd op de gangbare manier als winst per passagier. DEA is geschikt voor het gebruik van niet financiële variabelen, alsmede het ontbreken van een functionele relatie tussen inputs en outputs (Zhu, 2009). DEA kan ook omgaan met de invloed van omgevingsvariabelen. Deze zijn veel voorkomend in de vorm van invloed van de overheid (Zhu, 2009). Omdat betrokkenheid essentieel is, is het voordeel ook van DEA dat er eenvoudig met index getallen gewerkt kan worden, waardoor bedrijfsgevoelige informatie niet openbaar wordt (Caldas, 2013).

³ Onafhankelijk van een kansverdeling. Kansverdeling hoeft niet bekend te zijn.

Dit alles bij elkaar wordt geconcludeerd dat DEA het meest geschikt is voor de benchmarkmethodologie in de railsector. Ondanks dat in Tabel 5 SFA meer eigenschappen heeft, bezit DEA eigenschappen die het meer geschikt maken voor de railsector. Daarnaast is DEA vaker gebruikt in de railsector (Hansen, 2013).

Naast voordelen zijn er ook nadelen die bijkomend zijn bij DEA. Zo is de methode weliswaar geschikt, maar hiervoor moeten wel extra handelingen gedaan worden die lastig zijn. Dit maakt dat uitschieters in de data de resultaten kunnen beïnvloeden (Bottasso, 2011; Roberts, 2012). Daarnaast is het nadeel van het gebruik van een non-parametrische test dat uitspraak doen over statistisch significantie moeilijker is. Dit is echter wel te ondervangen door het gebruik van een bootstrap methode (Growitsch & Wetzel, 2009).

Tabel 5 - Overzicht van eigenschappen Methoden van Prestatie Evaluatie (gebaseerd op Ozcan & Zhu)

Eigenschap:	DEA	SFA	TFP	LSR	Ratio
Kan meerder inputs en outputs gebruiken	✓	✓	✓	✓	✗
Heeft geen prijs of gedragsinformatie nodig	✓	✓	✗	✓	✓
Kan onderscheid maken tussen technische, allocatieve en schaal efficiencies	✓	✓	✗	✗	✗
Robuust voor gebruik bij jaar op jaar fluctuaties	✓	✓	✗	✓	✗
Robuust tegen de heterogeniteit van de DMU's en de stand van de technologie	✓	✓	✓	✓	✗
Bezit van veel gebruikte statistische eigenschappen	✓	✓	✗	✓	✗
Geschikt voor gebruik ondanks dat sommige DMU's niet op de efficiency frontier zitten	✓	✓	✓	✗	✓
Robuust tegen uitschieters en ruis	✓	✓	✗	✓	✗
Omgang met zowel productie efficiency en effecten van de externe omgeving	✓	✗	✗	✗	✗
Voorkomt mogelijk mis match van gekozen functie en data	✗	✓	✗	✗	✗
Voorkomt mogelijke arbitraire aannames gemaakt door de methode (zwarte doos probleem)	✗	✓	✗	✓	✗
Voorkomt mogelijke correlatie problemen tussen variabelen	✓	✗	✓	✗	✓
Produceert voorspellingen van elasticiteiten en kosten	✗	✓	✗	✓	✗
Eenvoudige Techniek	✗	✗	✗	✓	✓

2.5 Data Envelopment Analysis

Roberts (2012) en Bottasso (2011) geven beiden een overzicht van voor en nadelen, deze zijn in Tabel 6 weergegeven en geven extra inzicht in de voor- en nadelen van het gebruik van DEA. Vervolgens wordt verder uitgewerkt hoe Data Envelopment Analysis in zijn werk gaat op het technische aspect.

Tabel 6 - Voor/nadelen DEA, afgeleid van Roberts (2012) en Bottasso (2011).

DEA voordelen	DEA nadelen
Laat meerdere inputs en outputs toe	DEA is een extreme waarden techniek en kan daardoor fouten geven bij ruis en uitschieters in de data.
Vereist niet een directe functionele relatie tussen inputs en outputs.	Aangezien het een non-parametrische methode is, is het testen van hypothese met betrouwbaarheidssinterval lastig.
Kan externe invloeden van de omgevingen die prestaties beïnvloeden meenemen in de evaluatie.	Efficiency scores kunnen gevoelig zijn voor de keuze van input en output variabelen.
De DMU wordt direct geëvalueerd tegen partners, niet via een statistische methode.	Doordat voor elke DMU de score apart wordt berekend kan het voor grote benchmarks met veel partijen een rekenintensieve klus zijn.
Het is een redelijk transparante methode die gemakkelijk gereproduceerd kan worden.	Bij het toevoegen van een niet beïnvloedbare variabele moet bekend zijn wat het effect van deze variabele is. Dit om te bepalen of het een input of een output variabele is.

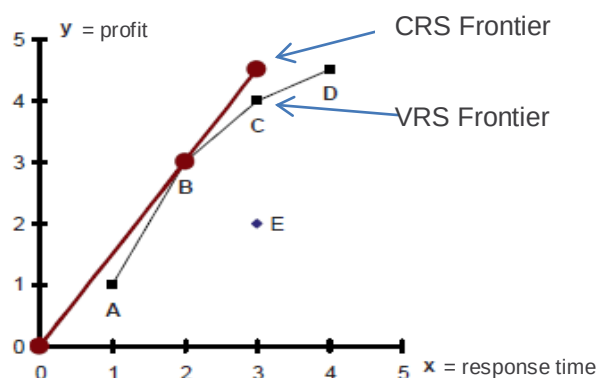
2.5.1 Input/ output

Binnen DEA zijn er verschillende mogelijkheden om prestatie evaluaties te maken. De eerste keuze die gemaakt kan worden is of er naar de input gekeken wordt, of naar de output. Als er naar de input gekeken wordt, wordt er getracht de efficiency te verhogen door de input niveaus te verlagen. Bij focus op de output wordt getracht om de efficiency te verhogen door de output niveaus te verhogen (Caldas, 2013; Joo et al., 2011; Ozcan, 2008; Winston, 2004; Wu, 2012; Zhu, 2009). Door gebruik te maken van de mathematische eigenschappen van DEA zoals Zhu (2009) ze representeert, kan er gebruik gemaakt worden van de bijbehorende DEAFrontier™ software (Zhu).

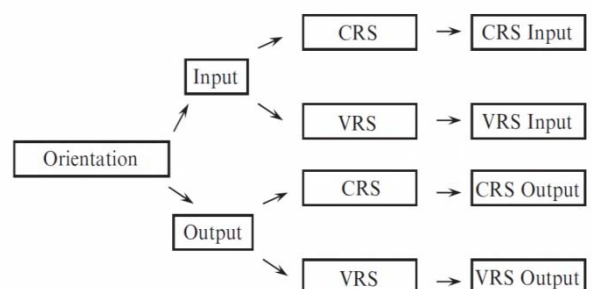
2.5.2 Constant-Return-to-Scale/Variable-Return-to-Scale

Er zijn twee modellen die gehanteerd kunnen worden. Beide modellen zijn geschikt voor zowel input als output focus. Het eerste model is het zogenaamde Constant>Returns-to-Scale (CRS) model (Caldas, 2013; Ozcan, 2008; Winston, 2004; Zhu, 2009), ontwikkeld door Charnes (1978). Het tweede is het Variable>Returns-to-Scale (VRS) model (Caldas, 2013; Ozcan, 2008; Zhu, 2009), ontwikkeld door Banker (1984).

Het verschil tussen CRS en VRS is de toevoeging van een voorwaarde bij de toegekende gewichten bij VRS. Dit komt omdat het CRS model niet goed werkt als er meer dan 1 optimale oplossing is, terwijl VRS dan wel werkt. Om het duidelijker te maken is in Figuur 12 het visueel weergegeven. Het verschil tussen CRS efficiënt en VRS efficiënt is dat bij CRS de frontier een rechte lijn is en deze niet kan veranderen qua getallen, terwijl deze bij VRS langs de optimale oplossingen en dit hoeft geen rechte lijn te zijn (Ozcan, 2008; Zhu, 2009). De gebogen lijn in Figuur 12 loopt langs meerdere punten is de output van een VRS model. Alle punten op de frontierlijn zijn efficiënt behalve E. De rechte lijn die van de oorsprong naar boven loopt is de CRS Frontier. Alleen punt B ligt op deze lijn en is efficiënt. A, C, D en E zijn dus inefficiënt. Figuur 11 geeft schematisch aan in het model in welke volgorde wordt gemaakt.



Figuur 12 - CRS/VRS Frontier (Afgeleid van Zhu (2009)).



Figuur 11 - Basic DEA Model Classification (Ozcan, 2008).

2.5.3 DEA mathematische weergave

Om het rekenwerk te vergemakkelijken en om de kans op fouten te minimaliseren wordt er gebruik gemaakt van door Zhu (2009) ontwikkelde DEA Frontier Free™. Om consistentie te waarborgen wordt gebruik gemaakt van de mathematische representatie zoals gegeven door Zhu.

In 2.5.3.1 wordt uitgelegd hoe de verschillende variabelen moeten worden opgevat. Verder wordt er vanwege het gebruik van DEA Frontier Free™ alleen naar de input gekeken, omdat de optie output niet beschikbaar is in de gebruikte versie.

2.5.3.1 Data Envelopment Analysis model

Het basis model van DEA bestaat uit twee rekenstappen. Dit model wordt in de DEA Frontier Free software het 'Envelopment Model' genoemd. De eerste stap is het berekenen van de efficiency van elke DMU. De tweede stap bestaat uit het uitrekenen van welke verbetering er nog mogelijk is (Ozcan, 2008;

Roberts, 2012; Zhu, 2009). In Tabel 7 staat de mathematische weergave vanuit Zhu. Hieronder worden eerst de variabelen uitgelegd met hoe het model mathematisch in elkaar zit.

DEA bestaat uit twee rekenstappen. De eerste stap is het bepalen van de Best-In-Class (BIC) voor elke DMU op basis van de referentie set. De referentie set is de verzameling van alle DMU's die mee worden genomen in de vergelijking. Elke DMU_i ($i = 1, \dots, n$) heeft m inputs x_{ij} ($i = 1, \dots, m$) en s output y_{rj} ($r = 1, \dots, s$). Elke DMU_j heeft een efficiency niveau θ dat van 0 tot 1 loopt. Hierbij is 1 efficiënt. Elke DMU wordt geëvalueerd ten opzichte van de referentie set. Deze referentie set bestaat uit alle DMU's die meegenomen worden in de vergelijking. De referentie set maakt het mogelijk om de coëfficiënt λ_j uit te rekenen waarmee duidelijk wordt welke van de DMU's in de referentieset de Best-In-Class (BIC) is van de desbetreffende DMU.

De tweede stap bestaat uit het evalueren van de verbetering die nog mogelijk is. Deze achterstand die de geëvalueerde DMU nog heeft, wordt slack genoemd, in de formule s_i^- . Het kan ook zijn dat een DMU efficiënt is, maar dat er des ondanks nog een verbetering mogelijk is. Wanneer een DMU een θ heeft van 1 en alleen maar slacks met de waarde 0, dan is de DMU 'DEA Efficiënt'. Wanneer een DMU een θ heeft van 1, maar nog slacks heeft die groter zijn dan 0, dan is deze DMU 'zwak efficiënt'. De slacks worden dus berekend nadat de efficiency bekend is.

De 'efficient target' geeft aan naar hoeveel de input verminderd moet worden van de DMU om het niveau te bereiken van de BIC.

Tabel 7 – Basis Envelopment Model (Zhu, 2009).

Frontier Type	Input-Oriented	Output-Oriented
	$\min \theta - \varepsilon (\sum_{i=1}^m s_i^- + \sum_{r=1}^s s_r^+)$	$\max \phi + \varepsilon (\sum_{i=1}^m s_i^- + \sum_{r=1}^s s_r^+)$
	subject to	subject to
CRS	$\sum_{j=1}^n \lambda_j x_{ij} + s_i^- = \theta x_{io} \quad i = 1, 2, \dots, m;$	$\sum_{j=1}^n \lambda_j x_{ij} + s_i^- = x_{io} \quad i = 1, 2, \dots, m;$
	$\sum_{j=1}^n \lambda_j y_{rj} - s_r^+ = y_{ro} \quad r = 1, 2, \dots, s;$	$\sum_{j=1}^n \lambda_j y_{rj} - s_r^+ = \phi y_{ro} \quad r = 1, 2, \dots, s;$
	$\lambda_j \geq 0 \quad j = 1, 2, \dots, n.$	$\lambda_j \geq 0 \quad j = 1, 2, \dots, n.$
VRS	Add $\sum_{j=1}^n \lambda_j = 1$	
NIRS	Add $\sum_{j=1}^n \lambda_j \leq 1$	
NDRS	Add $\sum_{j=1}^n \lambda_j \geq 1$	
Efficient Target	$\begin{cases} \hat{x}_{io} = \theta^* x_{io} - s_i^{-*} & i = 1, 2, \dots, m \\ \hat{y}_{ro} = y_{ro} + s_r^{+*} & r = 1, 2, \dots, s \end{cases}$	$\begin{cases} \hat{x}_{io} = x_{io} - s_i^{-*} & i = 1, 2, \dots, m \\ \hat{y}_{ro} = \phi^* y_{ro} + s_r^{+*} & r = 1, 2, \dots, s \end{cases}$

2.5.3.2 Measure specific model

Wanneer er inputs zijn waar een DMU geen invloed overheeft, bijvoorbeeld een externe input van een overheid, kan DEA hiermee omgaan. Wanneer dit het geval is kan met de toevoeging van een extra voorwaarde dit eenvoudig toevoegen aan het Envelopment Model. Het model heet dan een 'Measure Specific Model'. Het model wordt dan zoals in Tabel 8. De voorwaarde dat een input constant blijft is de tweede voorwaarde in Tabel 8 (Zhu, 2009).

Tabel 8 - Measure Specific Model (Zhu, 2009).

Frontier		Type	
		Input-Oriented	Output-Oriented
CRS	min	$\theta - \varepsilon(\sum_{i=1}^m s_i^- + \sum_{r=1}^s s_r^+)$	$\max \phi - \varepsilon(\sum_{i=1}^m s_i^- + \sum_{r=1}^s s_r^+)$
	subject to		subject to
		$\sum_{j=1}^n \lambda_j x_{ij} + s_i^- = \theta x_{io} \quad i \in I;$	$\sum_{j=1}^n \lambda_j x_{ij} + s_i^- = x_{io} \quad i = 1, 2, \dots, m;$
		$\sum_{j=1}^n \lambda_j x_{ij} + s_i^- = x_{io} \quad i \notin I;$	$\sum_{j=1}^n \lambda_j y_{rj} - s_r^+ = \phi y_{ro} \quad r \in O;$
		$\sum_{j=1}^n \lambda_j y_{rj} - s_r^+ = y_{ro} \quad r = 1, 2, \dots, s;$	$\sum_{j=1}^n \lambda_j y_{rj} - s_r^+ = y_{ro} \quad r \notin O;$
		$\lambda_j \geq 0 \quad j = 1, 2, \dots, n.$	$\lambda_j \geq 0 \quad j = 1, 2, \dots, n.$
VRS		Add $\sum_{j=1}^n \lambda_j = 1$	
NIRS		Add $\sum_{j=1}^n \lambda_j \leq 1$	
NDRS		Add $\sum_{j=1}^n \lambda_j \geq 1$	
Efficient			
Target		$\begin{cases} \hat{x}_{io} = \theta^* x_{io} - s_i^{-*} & i \in I \\ \hat{x}_{io} = x_{io} - s_i^{-*} & i \notin I \\ \hat{y}_{ro} = y_{ro} + s_r^{+*} & r = 1, 2, \dots, s \end{cases}$	$\begin{cases} \hat{x}_{io} = x_{io} - s_i^{-*} & i = 1, 2, \dots, m \\ \hat{y}_{ro} = \phi^* y_{ro} + s_r^{+*} & r \in O \\ \hat{y}_{ro} = y_{ro} + s_r^{+*} & r \notin O \end{cases}$

2.5.3.3 Voorbeeld

Om de mathematische weergave tastbaar te maken wordt er een voorbeeld uitgewerkt waarin duidelijk wordt hoe de resultaten gelezen moeten worden. Dit voorbeeld is rechtstreeks overgenomen uit Zhu (2009).

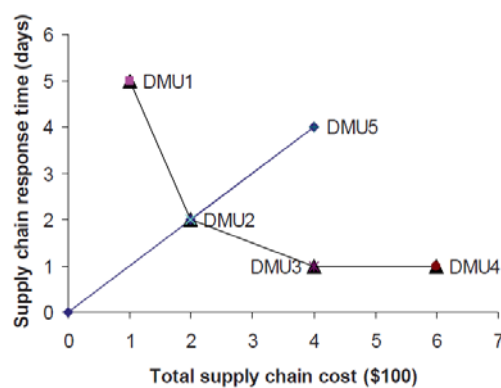
Tabel 9 - Voorbeeld DEA

DMU	Cost (\$100)	Response time (days)	Profit (\$1,000)
1	1	5	2
2	2	2	2
3	4	1	2
4	6	1	2
5	4	4	2

In Tabel 9 zijn 5 DMU's gegeven met elk een aantal inputs, cost en response time en een output, profit.

Het model, VRS, ziet er als volgt uit:

DMU 5
Min θ
Subject to
 $1 \lambda_1 + 2 \lambda_2 + 4 \lambda_3 + 6 \lambda_4 + 4 \lambda_5 \leq 4 \theta$
 $5 \lambda_1 + 2 \lambda_2 + 1 \lambda_3 + 1 \lambda_4 + 4 \lambda_5 \leq 4 \theta$
 $2 \lambda_1 + 2 \lambda_2 + 2 \lambda_3 + 2 \lambda_4 + 2 \lambda_5 \geq 2$
 $\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 = 1$
 $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5 \geq 0$



Figuur 13 - DMU vergelijking (Zhu, 2009)

In het voorbeeld wordt DMU5 geëvalueerd. Het oplossen van stelsel vergelijkingen geeft de volgende unieke oplossing: $\theta^* = 0,5$; $\lambda_2^* = 1$ en $\lambda_j^* = 0$ ($j \neq 2$) (Zhu, 2009). Dit betekent dat DMU5's efficiency 0,5 is. De benchmark van DMU5 is DMU2, deze is het meest efficiënt. Daarnaast is er geen andere DMU die eventueel ook efficiënt is zoals DMU2. DMU5 moet zijn cost en response time verlagen naar het niveau van DMU2 om efficiënt te zijn. Dit is grafisch weergegeven in Figuur 13.

Om ook slack meer tastbaar te maken wordt er nu een voorbeeld gegeven met slack met de evaluatie van DMU4 in plaats van DMU5. De DMU's in Tabel 9 worden wederom gebruikt. DMU4 heeft de volgende oplossing: $\theta^* = 1$; $\lambda_4^* = 1$ en $\lambda_j^* = 0$ ($j \neq 4$) (Zhu, 2009). Zoals te zien is in Figuur 13 liggen zowel DMU3 als DMU4 op efficiency frontier. Dit houdt in dat zowel DMU3 als DMU4 efficiënt zijn. Echter is ook te zien dat DMU4 een supply chain cost heeft van 6, terwijl bij DMU3 dit maar 4 is. Dit houdt in dat ondanks DMU4 efficiënt is, er nog wel verbetering mogelijk is. Deze verbetering van 2 is de zogenaamde slack en deze kan verlaagd worden naar het niveau van DMU3. De input slack is dus 2, of $s_1^- = 2$. DMU4 is op dit moment zwak efficiënt, terwijl DMU3 efficiënt is.

Maar deze keer wordt DMU4 geëvalueerd. Dit geeft de volgende oplossing: $\theta^* = 1$; $\lambda_4^* = 1$ en $\lambda_j^* = 0$ ($j \neq 4$). Zoals te zien in Figuur 13 kan DMU4 zijn response time nog verlagen met 2 naar het niveau van DMU2. De input slack is dus 2, of $s_1^- = 2$. DMU4 is zwak efficiënt, omdat ondanks het efficiency niveau van 1. Dit komt omdat er nog verbeter mogelijkheden zijn, gezien het feit dat er nog 'niet-nul' slacks zijn.

2.5.3.4 Voorwaarden aan de data

Bij de vergelijking met DEA wordt er elke keer een DMU vergeleken met de rest. Hierbij kan een keuze worden gemaakt welke input en output er vergeleken wordt. Verder is het van belang dat er rekening gehouden wordt met dat het DEA model uit gaat van hoe hoger de score hoe beter. Als data een opbouw hebben van 'hoger is slechter' moet deze score geïnverteerd worden om te gebruiken. Dit gebeurt met behulp van de volgende formule:

$$x_i^* = 1/x_i$$

DEA vereist dat alle waarden van de gebruikte data positief zijn. Het kan echter zijn dat de dataset zowel positieve als negatieve waarden bevat. Om deze data om te zetten in een dataset die gebruikt kan worden door DEA Frontier Free™, wordt gebruik gemaakt van de volgende indexeringsformule (Roberts, 2012):

$$x_i^* = \text{logit}^{-1}(a_i) = \frac{1}{1 + e^{-a_i}}$$

waarbij a_i de waarde van x is, uitgedrukt in komma getal (voorbeeld: $-87 = -0,87$; $100 = 1,0$).

Omdat de software van Zhu (2009) wordt gebruikt zijn er een aantal praktische zaken waar rekening mee gehouden moet worden. Dit maakt het keuze schema in Figuur 11 eenvoudiger. Dit komt omdat de DEA Frontier Free™ die wordt gebruikt, alleen de input oriëntatie kan gebruiken. Met de betaalde versie kan wel gebruik gemaakt worden van een output oriëntatie, maar die is voor dit onderzoek niet beschikbaar. Outputs worden als constant beschouwd en kunnen niet worden geëvalueerd.

2.5.3.5 Variabelen

Bij het gebruik van DEA voor benchmarking wordt er gebruik gemaakt van input en output variabelen. Deze variabelen kunnen een functionele relatie hebben, maar dit is voor DEA niet nodig. Omdat DEA alles met alles kan vergelijken, is zorgvuldigheid bij de keuze van de variabelen wel noodzakelijk. Dit kan deels worden ondervangen door de keuze van de variabelen te laten goedkeuren door de stakeholders. Dit zorgt voor een extra controle en zorgt ervoor dat alle belangen worden meegenomen. Verder is het verstandig rekening te houden met de eigenschappen van de variabelen. Deze eigenschappen komen overeen met de SMART criteria van KPI's uit het Ex Ante Evaluation van de

Europese Commissie (*Ex Anta Evaluation; A Practical Guide for Preparing Proposals for Expenditure Programmes*, 2001):

- **Specific.** De gekozen inputs en outputs moeten concreet en operationeel zijn.
- **Measurable.** De inputs en outputs moeten meetbaar zijn.
- **Accepted.** De inputs en outputs moeten geaccepteerd zijn door de stakeholders.
- **Realistic.** De gekozen inputs en outputs moeten realistisch zijn. Dit houdt ook in dat de gekozen inputs en outputs gebruikt kunnen worden als doel. Dit hangt samen met dat DEA die er van uit gaat dat er een verbetering mogelijk is.
- **Time-dependent.** Een input of output kan tijdsgebonden zijn, dit hoeft echter niet.

Daarnaast is het van belang om rekening te houden met welke variabelen wel en niet beïnvloedbaar zijn. Het komt voor dat variabelen wel gebruikt worden, maar dat deze niet beïnvloedbaar zijn door de gebruikende partij. Deze variabelen zijn vaak omgevingsvariabelen zoals een gegeven populatie of de lengte van een spoornetwerk. Om meer structuur aan te brengen in de classificatie van deze variabelen wordt het volgende classificatie schema gebruikt (Isoraite, 2005):

Tabel 10 - Classificatie Variabelen (Isoraite, 2005)

Variabele	Beïnvloedbaar
Resource/Input Indicator	Ja
Output Indicator	Ja
Result Indicator (Output)	Nee
Impact Indicator (Input)	Nee

Omdat de railsector ook op een andere manier wordt beoordeeld dan op de behaalde economische resultaten, is het van belang om bij de keuze van input en output hiermee rekening te houden. De railsector is een service sector en wordt uiteindelijk ook afgerekend op het zo goed mogelijk vervoeren van A naar B van goederen of passagiers.

Voor inputvariabelen wordt er in de spoorsector vaak gebruik gemaakt van arbeid of kapitaal gerelateerde variabelen (Merkert et al., 2010). De eigenschappen van elk land qua spoorwezen zijn verschillend, daarom is het van belang om deze toch vergelijkbaar te maken. Dit kan door het uitdrukken in per km of in passagier kilometer (Hansen, 2013).

2.6 Stakeholders

In 2.2.4 is als voorwaarde naar voren gekomen dat medewerking en betrokkenheid van alle benchmarkpartners essentieel is voor een succesvolle benchmark. Daarnaast is het van belang om gekozen variabelen te laten keuren door de stakeholders. Een stakeholder in benchmarking is een betrokken partner bij de vergelijking. Om deze goedkeuring goed te laten verlopen is het van belang dat er bekend is hoe belangrijk elke stakeholder is. Dit kan worden bepaald door gebruik te maken van de stakeholder theory (Mitchell et al., 1997). Het belang van de stakeholder wordt verder toegelicht.

2.6.1 Stakeholder Theorie

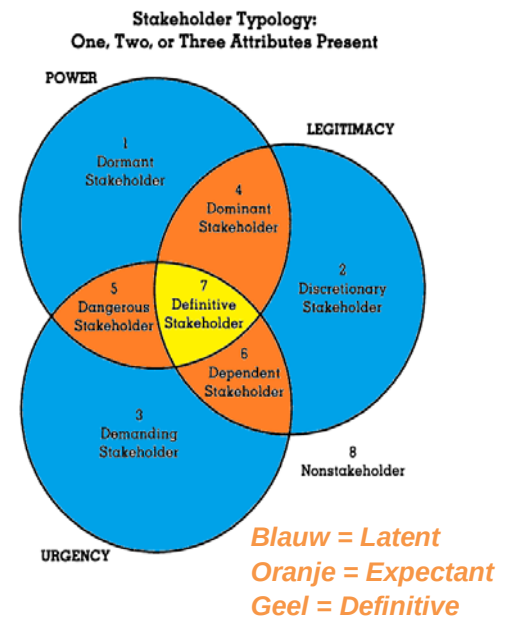
Om voor een goede commitment te zorgen is moeten de belangen van alle partners op juiste manier worden meegewogen. Daarnaast maakt DEA het mogelijk om alles met alles te vergelijken. Dit is zowel een voordeel als een nadeel, omdat het een overvloed aan resultaten kan opleveren. Het is echter niet altijd duidelijk welke nu echt relevant zijn (Roberts, 2012). Om dit te ondervangen moet een stakeholder analyse worden uitgevoerd, zodat bepaald wordt welke stakeholders en daardoor welke variabelen echt relevant zijn. Dit zorgt voor een stakeholder approach van benchmarking. Om structuur aan te brengen

in de analyse, wordt gebruik gemaakt van de Stakeholder Theory (Mitchell et al., 1997). Dit om op een eenduidige manier de stakeholders te kunnen analyseren en te zorgen dat deze dekkend is.

Zoals in Tabel 11 is te zien dat stakeholders 3 attributen hebben waar naar gekeken moet worden. Dit zijn Power, Legitimacy, Urgency. Het attribuut Saliency⁴ volgt uit de relatie tussen de drie attributen. De attributen die een stakeholder heeft, hebben een direct verband met de relatie tussen de verschillende stakeholders. Een overzicht van deze relatie zijn te zien in Tabel 11.

Tabel 11 - Stakeholder Attributen (Mitchell, Aale, & Wood, 1997).
Key Constructs in the Theory of Stakeholder Identification and Saliency

Construct	Definition	Sources
Stakeholder	Any group or individual who can affect or is affected by the achievement of the organization's objectives	Freeman, 1984; Jones, 1995; Kreiner & Bhambri, 1988
Power	A relationship among social actors in which one social actor, A, can get another social actor, B, to do something that B would not have otherwise done	Dahl, 1957; Pfeffer, 1981; Weber, 1947
Bases	Coercive—force/threat Utilitarian—material/incentives Normative—symbolic influences	Etzioni, 1964
Legitimacy	A generalized perception or assumption that the actions of an entity are desirable, proper, or appropriate within some socially constructed system of norms, values, beliefs, definitions	Suchman, 1995; Weber, 1947
Bases	Individual Organizational Societal	Wood, 1991
Urgency	The degree to which stakeholder claims call for immediate attention	Original—builds on the definition from the Merriam-Webster Dictionary Eystone, 1978; Wartick & Mahon, 1994
Bases	Time sensitivity—the degree to which managerial delay in attending to the claim or relationship is unacceptable to the stakeholder Criticality—the importance of the claim or the relationship to the stakeholder	Original—asset specificity from Hill & Jones, 1992; Williamson, 1985
Saliency	The degree to which managers give priority to competing stakeholder claims	Original—builds on the definition from the Merriam-Webster Dictionary



Figuur 14 - Stakeholder Typologie,
Aangepast van Mitchell et al. (1997)

De onderverdeling van de verschillende stakeholders is gegeven Figuur 14. Elk type stakeholder heeft andere belangen waar rekening mee gehouden moet worden bij de analyse van deze stakeholders. Vanuit deze indeling wordt vervolgens gekeken naar de invloed van elke stakeholder op het benchmark proces.

Als een stakeholder maar 1 attribuut heeft dan is het een zogenaamde 'latent stakeholder'. Deze hebben weinig tot geen invloed op het proces en daardoor een passieve houding. Als een stakeholder 2 attributen heeft dan is dit een 'expectant stakeholder'. Deze stakeholder heeft een verwachting van diegene die de analyse uitvoert en heeft een actieve houding en is redelijk belangrijk. De 'definitive stakeholder' heeft alle 3 de attributen en is daardoor het belangrijkste en heeft de hoogste Saliency. Hoe meer Saliency, hoe belangrijker hun keuze is van de KPI's.

2.6.2 Belang stakeholder

De stakeholder heeft verschillende belangen. Ten eerste is het van belang om de gevoeligheid van informatie altijd in de gaten te houden. Vanuit de documentatie van voormalig onderdelen van LR komt meerdere malen naar voren dat het van het grootste belang is dat de informatie intern blijft en niet te herleiden is naar de partijen die deze informatie hebben verstrekt. Als onafhankelijke partij is er vaak een dataverzameling van bedrijven die elkaars directe concurrenten zijn en niet willen dat de andere partij deze data hebben.

⁴ Saliency = Belang

Daarnaast is bij het gebruik van DEA belangrijk dat er rekening gehouden wordt met het feit dat DEA geen functionele relatie nodig heeft tussen input en output variabelen. Hierdoor kan het zijn dat er inputs en outputs gebruikt worden die voor een stakeholder niet van belang zijn of niet beïnvloedbaar zijn. Dit komt de bruikbaarheid niet ten goede. Door samen met de belangrijke stakeholders, zoals die naar voren komen bij het volgen van 2.6.1, kan er gekozen worden welke in- en outputs belangrijk zijn en dus gebruikt moeten worden.

2.7 Conclusie

Het theoretisch kader geeft een antwoord op de deelvragen wat is benchmarking en welke lering kan er getrokken worden uit de documentatie van voormalig onderdelen van LR.

Om antwoord te geven op de eerste deelvraag worden er verschillende onderwerpen behandeld die samen benchmarking beschrijven vanuit de literatuur. Hierbij wordt de definitie als leidraad genomen:

“Benchmarking is een methode waar op een continue en eenduidige manier een objectieve vergelijking gemaakt wordt, met als doel om de spelers te identificeren die de best-practices beheersen en te stimuleren om van hen te leren.”

Er zijn verschillende typen benchmarks die uitgevoerd kunnen worden afhankelijk van het te benchmarken proces alsmede het beoogde niveau binnen de organisatie. Het te kiezen proces varieert van interne, concurrerende, functionele tot generieke benchmark waarbij het niveau oploopt van operationeel tot strategisch niveau. De gemiddelde benchmark duurt zo'n 2 tot 6 maanden en wordt vaak uitgevoerd door een team van 4 personen of minder.

De volgende voordelen kenmerken benchmarking: benchmarking zoekt naar de best presterende werkwijzen, benchmarking geeft een objectief beeld van een organisatie ten opzichte van anderen, benchmarking stelt een organisatie in staat om haar zwakke plekken te identificeren en deze te verbeteren, benchmarking maakt het mogelijk om externe best-practices te vinden, besluitvorming op basis van feiten, benchmarking maakt het mogelijk beter in te spelen op de wensen van de klant en benchmarking dwingt een organisatie om haar eigen processen in kaart te brengen. Tegenover de voordelen staan de volgende valkuilen van benchmarking, het pareren van deze valkuilen wordt gezien als voorwaarden voor een succesvolle benchmark exercitie: betrokkenheid van alle partijen, waarborging van de data kwaliteit, continue verbetering, participatie zonder naming, blaming en shaming en de waarborging van de bescherming van bedrijfsgevoelige informatie.

Vanuit de literatuur zijn 7 modellen bekeken en vanuit de documentatie van LR zijn hier nog 2 modellen bijgevoegd. Deze modellen zijn gebruikt om te kijken welke stappen vaker voorkomen en deze worden dan ook als belangrijk bestempeld. Deze stappen zijn: keuze proces, keuze partners, prestatie indicator ontwikkeling, data verzameling, resultaten, analyse en implementatie. Daarnaast is er behoefte om het belang van de stakeholder te waarborgen. Hierdoor is er gekozen voor twee acceptatie stappen door de stakeholders. Echter biedt geen van de 9 modellen deze acceptatie stappen, wat resulteert dat er in het volgende hoofdstuk een stappenplan wordt gepresenteerd met twee acceptatie momenten.

Als methode om de prestaties tussen verschillende partijen te evalueren is er gekozen voor Data Envelopment Analysis. Vanuit de literatuur kwam naar voren dat DEA vaker wordt gebruikt in de spoorsector vanwege de eigenschappen dat het goed om kan gaan met niet economische variabelen, niet gerelateerde variabelen, externe niet beïnvloedbare invloeden en het gebrek aan vergelijkingspartners.

Om er voor te zorgen dat de belangen van de stakeholder niet uit het oog verloren worden, wordt er gebruik gemaakt van de stakeholder theorie van Mitchell et al. (1997). Door de stakeholders te beoordelen op hun 'power', 'legitimacy' en 'urgency', kan er gekeken worden naar het belang van elke stakeholder en de invloed die hieruit volgt.

3 Benchmark methodologie railsector

3.1 Leeswijzer

Dit hoofdstuk heeft als doel een handleiding te vormen voor het opzetten en uitvoeren van een benchmark. Op basis van de gepresenteerde informatie in hoofdstuk 2, het theoretisch kader, is er een stappenplan ontwikkeld dat toegespitst is op benchmarking in de railsector. Deze methodologie kan gebruikt worden als leidraad bij benchmarking en zorgt voor een gestructureerde en eenduidige manier van benchmarking. Bij de methodologie wordt gebruik gemaakt van een Excel bestand (DEA in Excel). Dit Excel bestand wordt gebruikt voor het berekenen van de efficiency scores met behulp van DEA Frontier™ plug-in voor Excel, ontwikkeld door Zhu (2009).

In § 3.2 wordt het model gepresenteerd en wordt beschreven hoe het model tot stand is gekomen. In § 3.3 staat de daadwerkelijke methodologie met uitleg hoe elke stap moet worden geïnterpreteerd. Ten slotte wordt in § 3.4 verwezen naar de Excel handleiding die gebruikt moet worden bij de methodologie.

3.1.1 Gebruik van het theoretisch kader

Op basis van hoofdstuk 2 is er een methodologie ontwikkeld die rekening houdt met de karakteristieken van de railsector. Waar nodig wordt teruggegrepen op de theorie om extra te verduidelijken waarom of hoe stappen gebruikt dienen te worden.

3.2 Het model

Het gegeven overzicht in 2.3.2 wordt gebruikt om te kijken welke stappen vaker voorkomen en derhalve als belangrijk verondersteld kunnen worden.

Het model bestaat uit 8 stappen en is gebaseerd op Figuur 9 en de twee modellen uit de documentatie van LR. Op basis van het theoretisch kader, de bevindingen en de karakteristieken van de railsector zoals deze naar voren zijn gekomen, is er gekozen voor 8 stappen.

1. Keuze benchmark proces
2. Keuze benchmark partners en identificatie stakeholders
3. Ontwikkeling variabelen voor DEA
4. Verzamelen data
5. Gebruik DEA
6. Analyse resultaten
7. Opstellen verbeterplan
8. Implementatie

De eerste stap betreft het afbakenen van wat er gebenchmarkt moet worden. Deze afbakening moet als eerste gebeuren om te bepalen wat er vergeleken wordt en hier de rest van de benchmark op te baseren. De tweede stap is de keuze van benchmark partners en het maken van een stakeholder analyse. Door de partners te kiezen wordt het speelveld van de benchmark bepaald. Door een stakeholder analyse te doen, wordt het duidelijk wat de rollen zijn van de verschillende partners en kan hier vervolgens rekening gehouden mee worden in de volgende stappen. Zoals aangegeven in 2.6 is het goed om te weten wat de rolverdelingen zijn. Zoals aangegeven in 2.2.4 is de betrokkenheid van de partners belangrijk en dit kan gewaarborgd worden door de stakeholder analyse. De derde stap is het ontwikkelen van variabelen voor DEA. Door de ontwikkeling van deze variabelen wordt duidelijk hoe de partners vergeleken gaan worden. Deze variabelen dienen goed gekeurd te worden door de stakeholders zodat hun belang is gewaarborgd. De vierde stap is het verzamelen van de data om de

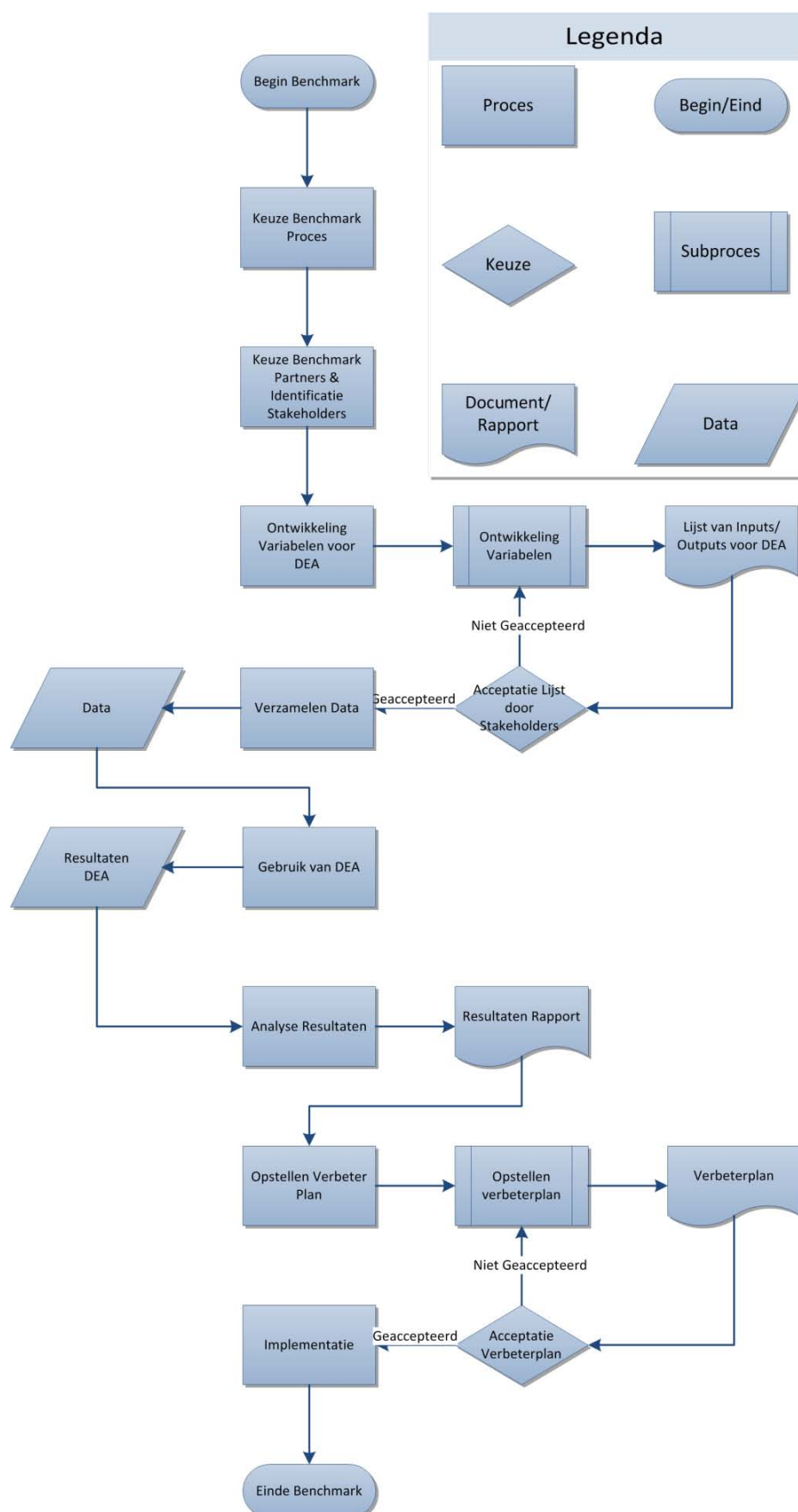
daadwerkelijke vergelijking te kunnen doen. De vijfde stap is het uitvoeren van DEA op basis van de variabelen en de vergaarde data. Bij het uitvoeren van DEA wordt gebruikt gemaakt van de DEA Free Frontier™ software. Dit minimaliseert de kans op fouten en verhoogt het gebruiksgemak. De zesde stap is het analyseren van de resultaten. Door te kijken welke partner bij de benchmarks de BIC is en vervolgens te kijken waarom, kan er voor de andere partners een doel gesteld worden waar ze naar toe kunnen werken en kunnen de voordelen, zoals beschreven in 2.2.3, tot hun recht komen. De zevende stap is het omzetten van de resultaten in een verbeterplan waardoor de benchmark partner het niveau omhoog kan brengen. De achtste stap behelst de implementatie van het verbeterplan.

Zoals in Figuur 9 te zien is, zijn de stappen proces keuze, partner keuze, identificatie van prestatie indicatoren, data collectie, resultaten, analyse en implementatie stappen die vaak terugkomen en hierdoor kan geconcludeerd worden dat dit belangrijke stappen zijn. Uit de literatuur komt naar voren dat de stakeholder belangrijk is en dat waarborging van de stakeholder belangen bijdraagt aan het succes van een benchmark. Om deze belangen te waarborgen wordt er gekozen om deze stap toe te voegen en tevens de goedkeuring van zowel de variabelen, als het implementatieplan door de stakeholders expliciet te noemen. In het model van Van Hoorn (2005), wordt deze stap ook expliciet genoemd.

De 8 stappen kunnen nog worden onderverdeeld in 4 fasen die de benchmark doorloopt. Stap 1 tot en met 3 is de vooronderzoek fase, de eerste fase. Stap 4 en 5 is de data verzamelen en meten fase en is fase twee. Stap 6 is de resultaat fase, fase drie. Fase vier bevat stap 7 en 8 en is de oplossing en implementatie fase. Elke fase resulteert in een document dan wel rapport. Deze 4 fasen kunnen dus ook gebruikt worden om de voortgang in de gaten te houden en de rapporten kunnen als milestones gebruikt worden waarmee de planning in de gaten gehouden kan worden.

3.2.1 Benchmark model

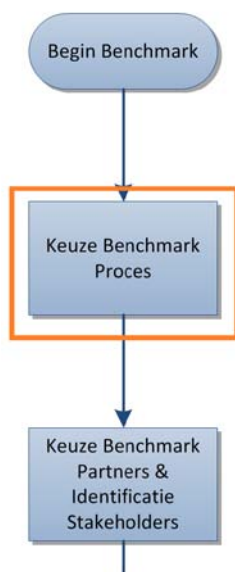
De visuele weergave van het model geeft het proces aan wat doorlopen moet worden bij de benchmark. Tevens is er te zien welke documenten en rapporten er gemaakt moeten worden tijdens de benchmark. Hierdoor komt er inzicht hoe het proces ervoor staat en wat de tussenstand is.



Figuur 15 - Benchmark Model

3.3 Benchmark methodologie

In deze paragraaf wordt de methodologie beschreven hoe een benchmark uitgevoerd dient te worden bij de keuze van het 8 stappenplan zoals in Figuur 15 is weergegeven. Ter ondersteuning van het overzicht is in de linker kolom van de bladzijde een visuele weergave van de huidige stap te zien. Dit geeft eenvoudig weer wat er voor/na de huidige stap komt. In de rechterkolom staat de uitwerkingen van de desbetreffende stap.



3.3.1 Stap 1 – keuze benchmark proces

De eerste stap die genomen wordt is het kiezen van het benchmark proces. De keuze wordt vaak al gemaakt door de klant zelf. Vervolgens is het belangrijk om te kijken naar wat voor soort benchmark er wordt uitgevoerd. Dit heeft betrekking op het niveau binnen de organisatie. Zoals te zien in Tabel 12 kan er gekozen worden uit 3 niveaus binnen een organisatie. Vervolgens wordt er in overleg met de klant gekeken op wat voor manier er vergeleken wordt. Leidraad is dat de manier een hoge relevantie heeft met het niveau binnen de organisatie. Dit om te waarborgen dat de benchmark relevant is.

Tabel 12 - Benchmark Relevantie (Moriarty & Smallman, 2009).

	Intern		Extern	
		Concurrerend	Functioneel	Generiek
Strategisch	Laag	Hoog	Laag	Hoog
Proces	Gemiddeld	Laag	Hoog	Hoog
Prestatie	Gemiddeld	Hoog	Gemiddeld	Laag



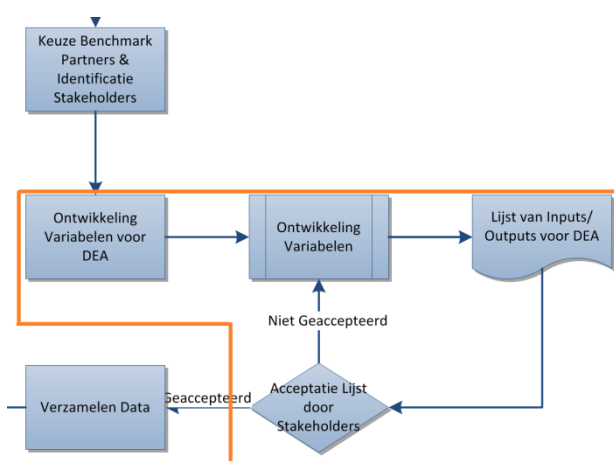
3.3.2 Stap 2 – keuze benchmark partners en stakeholder analyse

Nu de keuze is gemaakt wat voor benchmark er wordt gedaan, is de selectie van partners van belang. Bij de selectie van partners moet er rekening gehouden worden met het type benchmark dat er gedaan wordt.

Wanneer de partners zijn bepaald, is het belangrijk om in kaart te brengen tot welk type stakeholders de verschillende partners behoren. Het doel hiervan is om het belang, of Saliency, in kaart te brengen van elke partner zoals te zien is in Figuur 14. Dit gebeurt op basis van de drie attributen die een stakeholder toegewezen krijgt. In Figuur 14 en Tabel 11 is te zien wat elk attribuut betekent en in welke categorie de stakeholder zich bevindt. Dit vormt de basis van de Saliency van een stakeholder en daarmee de invloed die deze uitoefent op het selectieproces van de in- en output variabelen.

3.3.3 Stap 3 – ontwikkeling van variabelen voor DEA

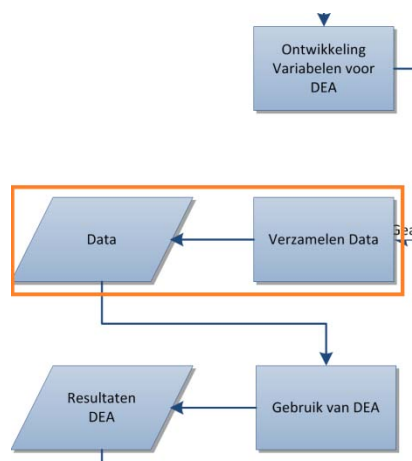
Nu het proces en de partners gekozen zijn moet er gekeken worden naar hoe het proces gemeten wordt. Aan de hand van welke variabelen wordt het proces weergegeven. Door gebruik te maken van DEA hoeft er geen functionele relatie te bestaan tussen een input en een output. Dit heeft als voordeel dat er gekeken kan worden naar welke in- en output er belangrijk worden gevonden, zonder rekening te houden of deze op elkaar aansluiten. Bij het ontwikkelen van deze variabelen moet rekening gehouden worden met de SMART zoals beschreven in 2.5.3.5.



Er wordt een lijst aangelegd waarin inputs en outputs worden genoemd die gemeten moet en gaan worden. Wanneer de lijst klaar is, moet deze lijst worden goedgekeurd door de stakeholders. Dit om de kwaliteit te waarborgen en er voor te zorgen dat er geen input of output vergeten wordt. De rol die de stakeholders in dit proces hebben komt voort uit de belangrijkheid die de stakeholder heeft volgens de stakeholder theory in 2.6.1. De uiteindelijke keuze kan worden gemaakt door de uitvoerder van de benchmark of in overleg met de stakeholders.

3.3.4 Stap 4 – verzamelen data

Het verzamelen van data is de volgende stap. Gevraagd wordt aan de partners om de benodigde data aan te leveren. Wat er precies nodig is moet voortkomen uit de lijst met in- en outputs. De ruwe data die de partners aanleveren moeten vervolgens worden opgeschoond en ontdaan van eventuele fouten. Denk hierbij aan negatieve waarden, waar dit niet mogelijk is, en eventueel andere (grote) uitschieters. Alle data worden verzameld in het Excel bestand 'DEA in Excel' in het tabblad 'Ruwe Data' (kleurcode blauw).



Uit deze ruwe data kunnen vervolgens de benodigde input en output variabelen uit worden onttrokken. Vaak kan dit gedaan worden door een sommatie of gemiddelde te nemen vanuit de ruwe data.

Als deze variabelen bekend zijn, moet er nogmaals nagelopen worden of er geen fouten of uitschieters in de data zitten. Deze extra check is nodig omdat DEA als evaluatie techniek gevoelig is voor uitschieters in de data.

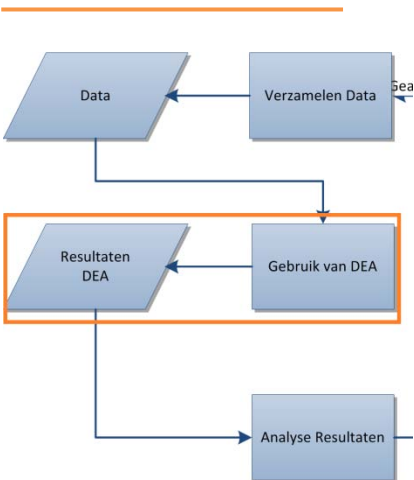
DEA Frontier Free werkt met de insteek hoger is beter en alle getallen moeten positief zijn. Mocht dit niet het geval zijn kunnen met behulp van de inverse formule en eventueel indexatie de cijfers gecorrigeerd worden (verwijzing).

Als laatste moeten de variabelen geordend worden. Dit is om duidelijke aan te geven of een variabele een input of een output is. De ordening is te zien in Tabel 13. Deze lijst met variabelen geordend naar partner (DMU) wordt hierna gebruikt door de DEA software. Door het aanhouden van deze indeling, zijn de data eenvoudig te gebruiken in de volgende stap. Het Excel bestand is dusdanig ingericht dat het invoeren zichzelf zoveel mogelijk wijst.

Tabel 13 - Variabelen Format

DMU	Inputs			Lege Kolom	Outputs		
	Input 1	Input 2	Input n		Output 1	Output 2	Output n
DMU1							
DMU2							
DMUi							

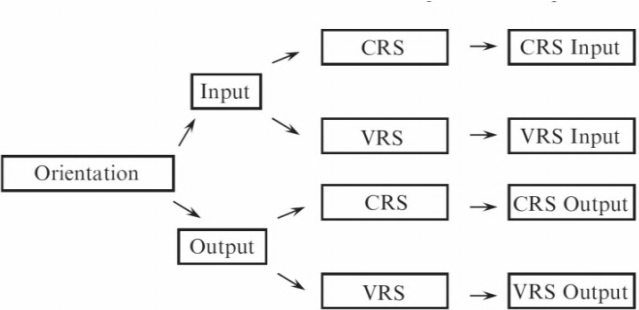
Het resultaat van deze stap is een verzameling van ruwe data die gebruikt kunnen worden.



3.3.5 **Stap 5 – gebruik van DEA**

Het daadwerkelijke gebruik van DEA en de vergelijking die hier uit volgt is de volgende stap. Elke DMU heeft een aantal inputs en outputs en met behulp van de software wordt de relatieve efficiency uitgerekend van elke DMU. Deze in- en outputs moeten in het tabblad “Data” (kleurcode geel) gekopieerd worden.

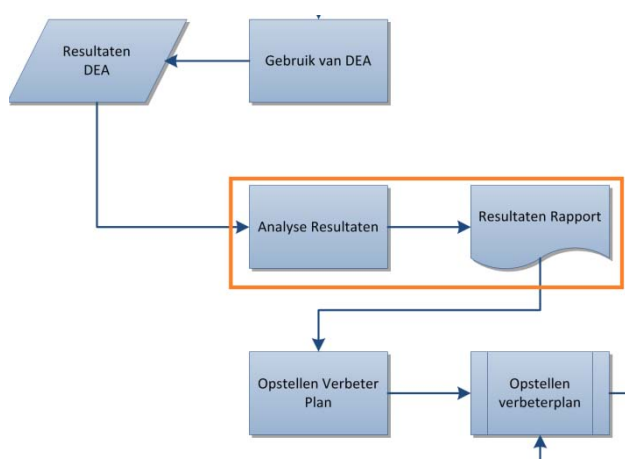
Daarna moet er gekozen worden voor een input of een output benadering. Vervolgens of er een CRS of VRS model gehanteerd wordt. Beide modellen worden gebruikt en als blijkt dat er meer dan één optimale oplossing mogelijk is, moet het VRS model gehanteerd worden.



Figuur 16 - Basic Model Orientation (Ozcan, 2008)

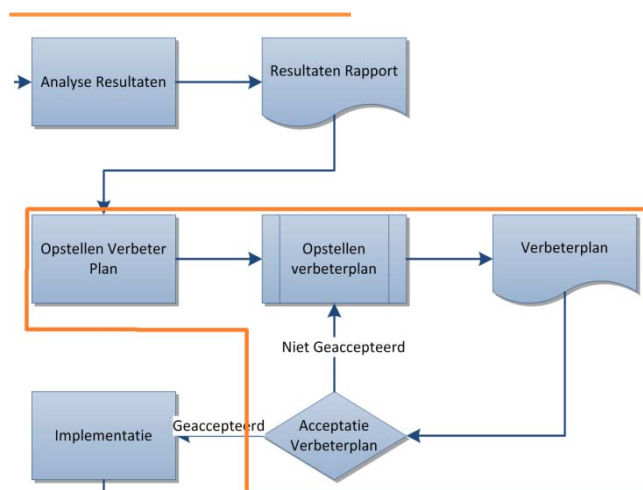
DEA Frontier Free™ wordt gebruikt om de efficiency scores uit te rekenen. Binnen de gratis versie, die in dit rapport wordt gebruikt, zijn er twee typen modellen die gebruikt kunnen worden. In Appendix B.2 wordt de input oriëntatie uitgewerkt, alsmede de modellen die de software biedt. Voor elke combinatie van input en output moet de software opnieuw gebruikt worden. Dus wanneer er eerst 1 input gebruikt wordt en daarna 2 inputs moet de software 2 keer gebruikt worden. Belangrijk hierbij is dat elk gebruik apart wordt opgeslagen, anders verdwijnen de resultaten.

Deze stap resulteert in Excel bestanden met elke de verschillende resultaten van DEA.



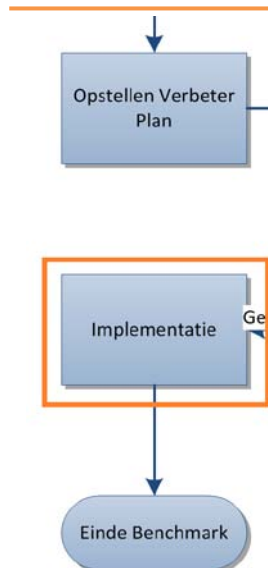
3.3.6 Stap 6 – analyse van resultaten

Wanneer alle verschillende combinaties zijn gedaan, moeten de resultaten worden geanalyseerd worden. Van elke berekende combinatie kan worden gekeken wie er efficiënt is en wie er achter lopen. Belangrijk is om in acht te nemen, is dat de efficiency cijfers relatief zijn. Verder kan er uitspraak gedaan worden over wat van de verschillende DMU's de hypothetische efficiency is. Op basis hiervan kan een rapport geschreven worden welke DMU's efficiënt zijn en de aanbevelingen die hierbij horen.



3.3.7 Stap 7 – opstellen verbeterplan

Nadat de analyse van de resultaten is geweest, kan er gekeken worden naar hoe de DMU's die niet efficiënt zijn efficiënter kunnen worden. Het verbeterplan is een stap die geheel op zichzelf kan staan. Het is het vervolgonderzoek op de eerste 6 stappen. Er moet onderzocht worden waarom de een beter presteert dan de ander. Dit onderzoek is feitelijk een nieuw onderzoek. Het onderzoek wordt gebaseerd op de bevindingen uit stap 6, waarbij er gekeken wordt naar welke DMU efficiënt is en wordt er gekeken waarom. Dit wordt vervolgens omgevormd tot een verbeterplan waarin wordt beschreven hoe de DMU zichzelf zou kunnen verbeteren.



3.3.8 Stap 8 – implementatie

De laatste stap is het implementeren van het verbeterplan. Zoals bij elke implementatie is het samenspel tussen de verschillende partijen belangrijk. Rekening gehouden moet worden met het gegeven dat alle partijen verschillend zijn en dat werkwijzen niet een op een overgenomen kunnen worden, maar moet worden aangepast naar de DMU in kwestie.

Het belangrijkste aan het gehele proces dat doorlopen is, is dat het proces van benchmarking niet stopt. Zoals in 2.2.3 is uitgelegd, is benchmarking een continu proces van verbetering. Om de gevolgen van de implementatie goed te kunnen beoordelen is het van belang dat er wordt getoetst of er verbetering plaats vindt. Daarnaast wordt er continu geïnnoveerd en verbeterd, dit maakt het belang van benchmarking nog groter. Want er worden altijd best-practices nieuwe manieren en oplossingen bedacht voor bestaande en nog niet bekende problemen. Na stap 8, begint het proces dus feitelijk opnieuw.

3.4 Excel Handleiding

In Appendix B is de handleiding te vinden voor het gebruik van de ontwikkelde Excel sheet in combinatie met DEA Frontier Free™. Deze handige software kan gebruikt worden om het kwantitatieve gedeelte van het vergelijken te doen. Het is belangrijk dat bij het uitvoeren van een benchmark de handleiding gevolgd wordt om de kans op fouten te minimaliseren.

4 Benchmark casus

Om het model te testen en de toepasbaarheid te testen in de praktijk, wordt in overleg met LRRE een testcasus uit de praktijk gebruikt. De test casus is gekozen door LRRE omdat het een interessante casus is voor hen om te onderzoeken. Het beidt de mogelijkheid om de methodologie voor de eerste keer te volgen en te kijken hoe het gebruik ervan is. Dit wordt vervolgens gebruikt om uitspraak te doen over de bruikbaarheid en om aanbevelingen te geven voor een verbetering van de methodologie.

4.1.1 Korte leeswijzer

In § 4.2 wordt beschreven hoe de casus is gekozen. In § 4.3 worden de stappen van de methodologie doorlopen zoals beschreven in het vorige hoofdstuk. Hierin worden verschillende scenario's getest om een idee te krijgen hoe de situatie is in de gekozen casus. Ten slotte worden in § 4.4 de bevindingen besproken van de casus zelf, alsmede de toepasbaarheid en wat LRRE hier van vindt. De paragraaf wordt afgesloten met de conclusie en aanbevelingen op de casus.

4.2 Keuze casus

De casus moet aan verschillende voorwaarden voldoen om de toepasbaarheid te kunnen testen. De eerste voorwaarde is dat de testcasus een casus is in de railsector. De tweede voorwaarde is dat er data beschikbaar moeten zijn. De derde voorwaarde is dat de casus van toegevoegde waarde kan zijn voor LRRE.

De casus die geselecteerd is, is de casus van het klein onderhoud aan het treintype Sprinter Light Train (SLT), de nieuwe sprinter van de Nederlandse Spoorwegen. Dit gebeurt door verschillende Service Bedrijven (SB) van NTR. De casus is geselecteerd omdat het voldoet aan alle drie de gestelde voorwaarden. De casus gaat om het onderhoud aan de SLT, wat betekent dat de casus over de railsector gaat, voorwaarde één. Omdat LRRE sinds de ontwikkeling bij SLT betrokken is, is er toegang tot bepaalde data. Deze data kunnen dus gebruikt worden en hiermee voldoet de casus aan voorwaarde twee. De laatste voorwaarde is dat de casus een meerwaarde kan zijn voor LRRE. Hier wordt aan voldaan, omdat de resultaten van de testcasus (nieuwe) inzichten kunnen geven over hoe het onderhoud wordt uitgevoerd wanneer er naar de cijfers gekeken wordt. De resultaten kunnen de aanleiding zijn om verder onderzoek te doen naar waar de resultaten vandaan komen. Dit kan voor zowel LRRE als voor NTR van toegevoegde waarde zijn.

Belangrijk blijft om te realiseren dat de benchmark casus geen antwoord geeft op de vraag waarom er verschil zit in prestaties van SB's, maar inzichten waar gezocht moet worden voor best practices op het gebied van onderhoud.

4.3 Casus

4.3.1 Stap 1 – keuze benchmark proces

Zoals al eerder beschreven betreft de casus het klein onderhoud van de SLT door verschillende SB's van NTR. Er wordt gekeken naar de doorlooptijd en welke factoren hierop van invloed. Op basis van de uitkomsten van de benchmark kan een uitspraak gedaan worden bij welke van de verschillende SB's mogelijk best practices aanwezig zijn. Dit kan worden afgeleid uit het feit dat zij op bepaalde aspecten efficiënter zijn.

4.3.2 Stap 2 – keuze benchmark partners en identificatie stakeholders

Bij elke benchmark zijn meerdere partijen betrokken. Om alle belangen goed te vertegenwoordigen is het van belang dat er gekeken wordt welke partijen betrokken zijn bij de benchmark en wat voor stakeholder ze zijn. Er zijn bij deze testcasus een aantal partijen direct, dan wel indirect betrokken. De eerste is LRRE. Zij zijn als onderzoekende partij betrokken bij de directe uitvoer. De tweede betrokken

partij is NTR. Zij voeren het onderhoud uit via de verschillende service bedrijven, aangezien zij de uitvoerende organisatie zijn van de SB's wordt NTR als geheel als stakeholder aangestipt. De derde partij is de NS. Dit omdat de NS de treinen gebruikt.

Om inzicht te krijgen in de verschillende belangen van de stakeholders is het belangrijk om te kijken wat de positie is van elke partij in de benchmark. Dit wordt gedaan aan de hand van de stakeholder theorie zoals eerder uitgewerkt in 2.6.1.

Belangrijk om te vermelden is dat binnen het tijdsbestek van de opdracht het niet mogelijk is geweest om alle betrokken partijen bij het stakeholder proces te betrekken. Dit houdt in dat de variabelen die verder zijn ontwikkeld voornamelijk door LRRE zijn vastgesteld.

4.3.2.1 Lloyd's Register Rail Europe B.V.

In dit onderzoek is LRRE de 'definitive stakeholder' en daarmee de meest belangrijke stakeholder. Dit volgt uit het feit dat LRRE de opdrachtgever is van het onderzoek. Daardoor hebben ze als stakeholder zowel power, legitimacy en urgency bij het onderzoek. Power volgt uit het feit dat de ondervindingen uit het onderzoek belangrijk kunnen zijn voor bijvoorbeeld NTR en de NS. Legitimacy volgt uit het feit dat het onderzoek gewenst is om inzichten te verkrijgen in de stand van zaken op het gebied van onderhoud aan de SLT door NTR. Bovendien is het LRRE ook aan gelegen om de methodologie te testen doormiddel van een casus. Urgency volgt uit dat LRRE als opdrachtgever van het benchmarkonderzoek binnen een redelijke termijn resultaten wil hebben, voor het mogelijk opzetten van een vervolg onderzoek.

4.3.2.2 NedTrain B.V.

NTR is als uitvoerder van het onderhoud een belangrijke stakeholder. Binnen de scope van het onderzoek zijn ze in dit geval een 'dependant stakeholder'. Legitimacy komt voort uit dat het onderzoek voor NTR gewenst kan zijn. Het vinden van de plekken waar best-practices voorkomen is goed voor de uitvoering van het onderhoud. Urgency komt voort uit dat NTR de uitvoerder is van het onderhoud en heeft daardoor baat bij resultaten op een redelijke termijn.

Er is binnen NTR nog een onderverdeling tussen de organisatie NTR en de diverse SB's. De diverse SB's zelf 'dependant stakeholder' van NTR. Voor NTR zijn de SB's onmisbaar en daarom hebben zij de Legitimacy om belangrijk te zijn voor NTR, daarnaast is de relatie tussen SB en NTR essentieel voor het functioneren van beiden. Hier ontleen de SB's de Urgency aan.

4.3.2.3 Nederlandse Spoorwegen N.V.

De laatste stakeholder is de NS. Zij zijn als eigenaar van de treinen afhankelijk van NTR voor het onderhoud. Zij hebben dus baat bij de implementatie van de best-practices. Dit geeft ze de Legitimacy. Verder heeft de NS geen andere attributen en is dus een 'discretionary stakeholder'.

Als eigenaar van de treinen die onderhouden worden door NTR en als eigenaar van NTR heeft de NS baat bij een goede uitvoering van het onderhoud.

4.3.3 Stap 3 – ontwikkeling variabelen voor DEA

In de methodologie wordt gebruik gemaakt van DEA. Voor het gebruik van DEA moeten er input en output variabelen ontwikkeld worden. Er moet bepaald worden wat de DMU's zijn. De DMU's zijn de verschillende SB's waar klein onderhoud verricht wordt aan de SLT. Met behulp van deze variabelen kan de DEA uitgevoerd worden. De variabelen worden gerangschikt naar of het input of output variabelen zijn. De variabelen zijn ontwikkeld uit het gegevens bestand 'Uitgifte Herstelbestand 2010+'. In dit bestand is gekeken welke informatie beschikbaar is van het klein onderhoud aan de SLT. Daarnaast is ook nog een schatting gemaakt van het aantal monteurs in elk van de SB's (in overleg met LRRE). Daarnaast wordt het uitvoeren het aantal treinstellen waarop een reset van de software wordt uitgevoerd meegenomen. Het resetten van de software is een standaard procedure waarmee sommige

defecten verholpen kunnen worden. Als laatste is het aantal treinstellen waarbij een onderdeel wordt vervangen meegenomen. Door een defect onderdeel te vervangen, hoeft dit onderdeel niet te worden gerepareerd en zou dus, in theorie, voor een kortere onttrekkingstijd zorgen.

In Tabel 14 is te zien welke DMU's er zijn. Dit zijn de verschillende locaties waar klein onderhoud wordt uitgevoerd aan de SLT. De afkortingen staan voor de verschillende locaties. De volledige namen van de locaties staan in de afkortingen lijst in Appendix A.

De verschillende input en output variabelen zijn geclassificeerd. Hierdoor kan eenvoudig rekening gehouden worden met of een variabele wel of niet beïnvloedbaar is, zoals beschreven in 2.5.3.5. De verschillende variabelen die ontwikkeld zijn, hebben als doel om een beeld te vormen over welke potentiële invloeden effect kunnen hebben op de gemiddelde doorloop tijd. In Tabel 14 en Tabel 15 zijn de ontwikkelde variabelen weergegeven met hun classificatie.

Tabel 14 – DMU's

Locatie
OLDD
SAMR
SBKH
SRSD
SRTD
SUTC
SUTZ
SWGM

Tabel 15 - Input en output classificatie

Input	Classificatie	Output	Classificatie
Aantal Monteurs	Input Indicator	Gem. Duur Onttrekking	Result Indicator
Aantal Monteurs naar verhouding SLT/Overige	Input Indicator	Gem. Duur R; GV	Result Indicator
Aantal Onttrekkingen	Impact Indicator	Gem. Duur GR; V	Result Indicator
Aantal GR; V	Input Indicator		
Aantal R; GV	Input Indicator		

- GR = Geen Reset.
- R = Reset
- GV = Geen vervanging
- V = Vervanging

4.3.4 Stap 4 – data verzamelen

Voor het uitvoeren van DEA zijn data nodig. De data die worden gebruikt, komen uit datasets die gebruikt worden voor het analyseren van het onderhoud bij de diverse SB's. In deze logfiles wordt bijgehouden welk treinstel waar binnen komt, hoe laat, wat ermee gebeurd is, et cetera. Deze data worden gebruikt door zowel NedTrain, de Nederlandse Spoorwegen en LRRE. Door gebruik te maken

van Excel zoekfuncties kan er informatie worden onttrokken aan deze data over bijvoorbeeld de gemiddelde onttrekkingduur, het aantal vervangingen dat plaats vindt. Deze worden uitgelezen middels een draaitabel en genoteerd in het tabblad 'Ruwe Data' in het Excel bestand 'DEA in Excel'. Exacte gegevens over het aantal fte's is niet bekend, maar er is een inschatting gemaakt van het aantal monteurs in overleg met LRRE. Alle data worden verzameld in het Excel bestand, zodat het eenvoudig gebruikt kan worden in stap 5.

Belangrijk om te vermelden is dat de doorlooptijd geïnverteerd is. Dit is noodzakelijk omdat de output in DEA gewaardeerd wordt volgens het 'hoger = beter' principe. De cijfers zijn daarna vermenigvuldigd met 100, om de cijfers meer leesbaar te maken.

4.3.5 Stap 5 – gebruik van DEA

De data die verzameld zijn in stap 4 worden nu gebruikt voor het uitvoeren van DEA. Voor een goed beeld moeten er meerdere scenario's worden gekozen die doorgerekend worden in Excel. Dit om te kijken hoe verschillende inputs en outputs de resultaten beïnvloeden. In samenspraak met LRRE zijn er 6 scenario's opgesteld om de invloed van de variabelen te kunnen bepalen.

Deze 6 scenario's zijn bedoeld om inzicht te krijgen in hoe verschillende SB's scoren waar de SLT wordt onderhouden. De mogelijkheden tot vergelijken zijn zeer divers en groot en daarom is het aantal begrensd op 6. Hiermee wordt gepoogd een eerste inzicht te verwerven in de verschillende SB's en hoe deze zich onderling verhouden en welke variabelen mogelijk invloed hierop kunnen hebben. In elk scenario worden de DMU's, inputs, outputs en beoogde invloeden beschreven. De uitkomsten worden pas in stap 5 geanalyseerd, er wordt dus nog geen uitspraak gedaan over de gevonden (in)efficiëntie van DMU's.

Van de 6 scenario's wordt alleen scenario 1 uitgewerkt weergegeven. De uitwerking van de overige 5 scenario's wordt gegeven in Appendix D. Door scenario 1 uit te werken krijgt de lezer een idee hoe een scenario wordt aangepakt.

4.3.5.1 Scenario's

In alle scenario's worden alle DMU's meegenomen in de vergelijking. In Tabel 16 Van elk scenario wordt weergegeven welke variabelen er mee genomen worden.

Tabel 16 - Scenario variabelen overzicht

Scenario:	1	2	3	4	5	6
Input:	- Aantal monteurs - Aantal onttrekkingen	- Aantal monteurs naar verhouding SLT/Overig - Aantal onttrekkingen	- Aantal monteurs naar verhouding SLT/Overig - Aantal onttrekkingen - Aantal R; GV	- Aantal monteurs naar verhouding SLT/Overig - Aantal onttrekkingen - Aantal GR; V	- Aantal monteurs naar verhouding SLT/Overig - Aantal onttrekkingen	- Aantal monteurs naar verhouding SLT/Overig - Aantal onttrekkingen
Output:	Gem. Duur Onttrekking	Gem. Duur Onttrekking	Gem. Duur Onttrekking R; GV	Gem. Duur Onttrekking GR;V	Gem. Duur Onttrekking	Gem. Duur Onttrekking

4.3.5.2 Scenario 1 – basis

4.3.5.2.1 Beschrijving

Het eerste scenario is de uitgangspositie van de SB's. Doel van dit scenario is om een eerste beeld te geven hoe de situatie is bij het meenemen van alle materieelsoorten op de locatie waar SLT's worden onderhouden. Dit scenario vormt als het ware de basis waar de andere scenario's op voortborduren. Het basis scenario heeft de volgende inputs: het aantal monteurs en het aantal onttrekkingen. Beide zijn per SB. De output variabele is de gemiddelde onttrekkingduur. Omdat de data uitschieters bevatten, zijn de data geschoond op onttrekkingduur waarden van minder dan 0 uur en meer dan 240

uur. Wanneer een trein meer dan 240 uur onttrokken is uit de dienst wordt dit niet meer gezien als klein onderhoud en vandaar dat de grens gezet is op 240 uur. Daarnaast worden er door deze filtering ook foutieve waarden eruit gefilterd die foutief zijn ingevoerd. Denk hieraan negatieve onttrekkingstijden of onttrekkingen die meer dan 3000 uur duren. De periode die is bekeken is de van januari 2010 tot juni 2013.

4.3.5.2.2 Variabelen

Tabel 17 - Variabelen Scenario 1

Locatie	Input	Classificatie	Output	Classificatie
OLDD	Aantal Monteurs	Input Indicator	Gem. Duur Onttrekking	Result Indicator
SAMR	Aantal Onttrekkingen	Impact Indicator		
SBKH				
SRSD				
SRTD				
SUTC				
SUTZ				
SWGM				

4.3.5.2.3 Uitwerking

Het model dat gebruikt is, is Measure Specific waarbij het aantal onttrekkingen een gegeven is waardoor het constant en niet beïnvloedbaar is. Als model is er gekozen voor een VRS model, omdat er meerdere partijen efficiënt zijn.

Tabel 18 - Scenario 1 Measure Specific VRS - Tabblad 'Efficiency'

DMU No.	DMU Name	Input-Oriented	Optimal Lambdas with Benchmarks	
		VRS Measure-Specific Efficiency		
1	OLDD	0,04192	1,000	SRSD
2	SAMR	1,00000	1,000	SAMR
3	SBKH	0,19337	1,000	SRSD
4	SRSD	1,00000	1,000	SRSD
5	SRTD	0,37102	1,000	SRSD
6	SUTC	0,44492	1,000	SRSD
7	SUTZ	0,36585	1,000	SRSD
8	SWGM	0,33981	1,000	SRSD

In Tabel 18 is de uitkomst gegeven zoals DEA Frontier Free™ deze genereert. In de linkerkolom staat het nummer van de DMU. De tweede kolom staat de naam van de DMU. In de derde kolom staat de efficiency score van de DMU, waarbij 1 efficiënt is. In de kolom optimal lambdas staat in welke mate de DMU in kwestie zich moet richten op de Best-In-Class, de DMU die in de meest rechtse kolom staat.

In Tabel 18 moeten de DMU's zich maar op een BIC richten. Het kan echter zo zijn dat een DMU zich op meer dan een moet richten.

4.3.6 Stap 6 – analyse resultaten

In stap 5 is scenario 1 uitgewerkt, in Appendix D staat de uitwerking van de overige 5 scenario's. Op basis van de uitkomsten van de 6 scenario's kan uitspraak gedaan worden over welke zaken opvallen in de uitkomsten. Door het maken van een analyse kan gezegd worden welke DMU's efficiënt zijn,

welke variabelen hier invloed op hebben, wat de BIC is voor deze DMU's en eventueel nog een verklaring waar de mogelijke verschillen in prestatie vandaan komen tussen de DMU's in de verschillende situaties.

4.3.6.1 Scenario 1

In het basis scenario wordt gekeken naar de huidige situatie van alle binnenkomsten op de SB's waar de SLT wordt onderhouden. De variabelen zijn het aantal monteurs, het totaal aantal binnenkomsten en de gemiddelde duur van de onttrekking.

Te zien is in Tabel 18 is dat zowel SAMR (Amersfoort) als SRSD (Roosendaal) het meest efficiënt zijn volgens DEA. OLDD (Onderhoudsbedrijf Leidschendam) heeft de laagste efficiency. Met het aantal onttrekkingen als een constante en het aantal monteurs als variabele blijkt dat deze twee het meest efficiënt zijn om dit te vertalen in een lage gemiddelde duur van de onttrekking. Deze twee SB's (SAMR, SRSD) zijn tevens ook de kleinste. Wat vooral opvalt, is dat AMR behoorlijk veel binnenkomsten heeft, maar een lage onttrekkingduur. Een snelle blik in de databestanden laat zien dat er weinig grote problemen binnenkomen in SAMR en dat wanneer deze binnenkomen ze snel worden doorverwezen naar de grote onderhoudsbedrijven. Daarnaast is Roosendaal volgens DEA het streef SB. Dit wil zeggen dat de targets voor de reductie van het aantal monteurs die benodigd zijn worden bepaald door DEA op basis van SRSD.

4.3.6.2 Scenario 2

In scenario 2 wordt er gekeken naar de SLT. Hierbij wordt het aantal monteurs geschaald naar het aantal binnenkomsten van de SLT. De gemiddelde onttrekkingstijd is ook van de SLT. Hierdoor komt er een beeld over hoe de situatie er uit ziet wanneer er gekeken wordt naar de SLT. Interessant is om te kijken of er verschillen optreden in vergelijking met scenario 1.

Zoals blijkt uit Tabel 22 zijn SAMR en SRSD wederom het meest efficiënt. Deze SB's zijn ook bij de SLT in staat om met de minste aantal monteurs, gegeven het aantal binnenkomsten, de laagste gemiddelde onttrekkingduur te realiseren. SAMR is nu de target DMU voor de andere DMU's. Wat verder nog opvalt, is dat SWGM een stuk beter presteert bij de SLT dan wanneer er wordt gekeken naar alle treinen.

4.3.6.3 Scenario 3

In scenario 3 wordt er gekeken of het uitvoeren van een reset invloed heeft op de onttrekking. Zoals te zien is in Tabel 24 zijn er 3 DMU's efficiënt. Dit zijn OLDD, SAMR en SBKH (Brinkhorst). De rest van de DMU's zijn niet efficiënt. Het effect van het aantal resetten is hier terug te vinden in de efficiency score. Bij de andere partijen is er volgens DEA minder positieve invloed van terug te vinden.

Wat verder opvalt, is dat in vergelijking met scenario 2, is dat er beter wordt gescoord op gemiddelde onttrekkingduur. Dat wil zeggen dat het uitvoeren van een reset een positieve invloed heeft. De gemiddelde tijden liggen lager.

4.3.6.4 Scenario 4

In scenario 4 wordt gekeken of het aantal vervangingen invloed heeft op de gemiddelde doorlooptijd. Te zien is in Tabel 26 dat alleen SRSD het goed doet en dat de overige SB's het allemaal zeer slecht doen. Dit kan aan twee zaken liggen. Of ze doen het slecht, of de data zijn niet van voldoende niveau. Wanneer er wordt gekeken naar de data, dan is te zien SRSD een hogere output heeft. Dit duidt op een lage onttrekkingduur. Dit komt echter omdat er maar vervanging heeft plaatsgevonden. Dit vertekent het beeld behoorlijk, omdat het ver af ligt van wat er gemiddeld wordt gedaan en zoals eerder genoemd is DEA gevoelig voor uitschieters.

Wanneer er wordt gekeken naar SB's waar vaker vervangingen plaatsvinden, zoals OLDD, SRTD en SUTC, dan is te zien dat de waarden al meer richting de 1,5 gaan. Dit ligt al veel meer in lijn met wat de

gemiddelde duur is van een onttrekking. Over deze DMU's kan worden gezegd dat OLDD en SRTD een lagere onttrekkingduur hebben en SUTC (Utrecht Cartesius) een hogere. Of dit statistisch significant is ligt buiten de scope.

4.3.6.5 Scenario 5

Scenario 5 en 6 kijken of er verschillen te vinden zijn tussen een rustige en een drukke periode. Om een vergelijking te maken is het van belang dat de verhouding SLT/overige treinen ongeveer gelijk is. Scenario 5 behandelt de rustige week. In Tabel 28 is te zien dat SRTD (Rotterdam) en SUTC (Utrecht Cartesius) het het beste doen in de rustige periode. De doorloop tijd bij deze SB's is een stuk lager dan bij de andere 2 SB's. Dit houdt in dat zij in met het aantal monteurs wat zij tot hun beschikking hebben, het beste presteren. Hierbij moet wel gezegd worden dat het aantal behandelingen zeer laag is in deze week (OLDD 3;SRTD 1;SUTC 1;SUTZ 2). Dit maakt de gevoeligheid met betrekking tot uitschieters hoog.

Er zijn 4 SB's die in deze week geen treinstel behandelen, deze zijn derhalve niet meegenomen in de resultaten.

4.3.6.6 Scenario 6

In scenario 6 wordt er gekeken of drukte invloed heeft op de prestaties. Uit Tabel 30 is af te leiden dat SWGM (Watergraafsmeer) het meest efficiënt is. Hier moet echter wel worden aangegeven, dat SWGM maar 1 SLT heeft behandeld. Verder scoren de overige DMU's niet hoog. Dit komt waarschijnlijk door de uitschieter van SWGM. Daarnaast scoren zowel SRTD als SUTC minder goed dan in het rustige scenario. Dit kan impliceren dat de prestatie bij drukte terugloopt.

Daarnaast valt op bij de vergelijking met scenario 5 dat OLDD het weer een stuk minder doet. Deze scoort door alle scenario's heen constant laag. Daarnaast scoort SUTZ (Utrecht Zuid) nog steeds niet goed.

4.3.7 Stap 7 – opstellen verbeterplan

Zoals eerder aangegeven wordt het opstellen van het verbeterplan niet in dit rapport uitgewerkt. Dit komt omdat de vraag waardoor de verschillen veroorzaakt worden een vraag op zichzelf is. Wat belangrijk is bij het opstellen van het verbeterplan is dat er gekeken wordt welke SB's goed presteren en in wat voor soort omgeving dit gebeurt. Er kan verondersteld worden dat als een SB goed presteert in een drukke periode, dat er binnen dit SB op een manier gewerkt wordt die mogelijk als best practise beschouwd kan worden. Het verbeterplan moet zich dan ook richten op het vinden van het antwoord op de vraag waarom presteert het ene SB beter presteert dan de andere en hoe dit kan worden omgezet in een plan om hiermee andere SB te verbeteren.

4.3.8 Stap 8 – implementatie

Net zoals bij Stap 7 wordt de implementatie van het verbeterplan niet in dit rapport uitgewerkt. Wel is van belang bij het uitvoeren van het verbeterplan dat er rekening gehouden wordt met dat werkwijzen bij het ene SB niet zomaar een-op-een kunnen worden ingevoerd bij het andere SB.

Echter het belangrijkste uit het hele verbetertraject is dat er gekeken wordt of de verbetering optreedt. Met andere woorden, nadat het plan geïmplementeerd is, is het noodzakelijk opnieuw een benchmark uit te voeren, zodat er gekeken kan worden naar de verbetering over tijd.

4.4 Bevindingen casus

Op basis van de casus kan er uitspraak gedaan worden over de gevonden resultaten van de casus. Eerst worden de resultaten besproken van alle 6 scenario's bij elkaar, vervolgens worden er conclusies getrokken over de resultaten van de casus gevolgd door de aanbevelingen. Daarna wordt de toepasbaarheid van het model besproken samen met de discussie bij LRRE naar aanleiding van de presentatie van de uitkomsten van de casus die gegeven is bij LRRE. Als laatste wordt er middels de conclusie antwoord gegeven op de gestelde deelvraag in 1.4:

“Wat zijn de resultaten bij het uitvoeren van de benchmark bij NedTrain BV die uitgevoerd wordt op de Service Bedrijven van de Nederlandse Spoorwegen?”

4.4.1 Scenario's

In Figuur 17 is een overzicht gegeven met alle zes scenario's en de scores van de DMU's. Tevens is gegeven welke oriëntatie, CRS of VRS en welk DEA model er is gebruikt. In deze paragraaf worden de resultaten van de casus besproken en nog niet de toepasbaarheid van het model.

Scenario	1 Basis	2 SLT/ Overig	3 Reset	4 Vervanging	5 Rustig	6 Druk
Orientation	Input-Oriented	Input-Oriented	Input-Oriented	Input-Oriented	Input-Oriented	Input-Oriented
CRS / VRS	VRS	VRS	CRS	CRS	CRS	CRS
DEA Model	Measure-Specific Efficiency	Measure-Specific Efficiency	Measure-Specific Efficiency	Measure-Specific Efficiency	Measure-Specific Efficiency	Measure-Specific Efficiency
DMU Name						
OLDD	0,04192	0,03044	1	0,00485	0,15904	0,02429
SAMR	1	1	1	-	-	-
SBKH	0,19337	0,25243	1	0,01847	-	0,05885
SRSD	1	1	0,75794	1	-	-
SRTD	0,37102	0,21311	0,46988	0,03125	1	0,66484
SUTC	0,44492	0,29545	0,8872	0,03372	1	0,33587
SUTZ	0,36585	0,24528	0,37425	0,01962	0,31634	0,23698
SWGM	0,33981	0,74286	0,21949	0,19747	-	1

Figuur 17 - Overzicht Scenario's

Te zien is uit scenario 1 en 2 komen SAMR en SRSD als meest efficiënte SB's naar voren. Ze zijn in staat om met zo min mogelijk aantal monteurs, met het aantal binnenkomsten als constante, de laagste gemiddelde onttrekkingstijd te realiseren.

Wanneer er in scenario 3 gekeken wordt naar de invloed van het uitvoeren van een reset op de gemiddelde onttrekkingduur blijkt dat OLDD, SAMR en SBKH hier het meest profijt van ondervinden. Met name OLDD is verrassend, omdat dit SB in de eerste twee scenario's het slecht deed. Hierbij moet worden aangetekend dat het aantal monteur bij OLDD zeer hoog is. Dit komt door twee redenen. De eerste is dat het een schatting is en ten tweede dat er bij deze schatting geen rekening gehouden is met of de monteur werkte voor het SB gedeelte van OLDD of voor het OB (Onderhoudsbedrijf), waar groot onderhoud wordt verricht aan treinstellen. Dit geeft een vertekend beeld waardoor OLDD systematisch lager scoort, omdat ze meer monteurs hebben. Dit kon in de scope van het onderzoek niet verder worden uitgesplitst en zorgt dus voor een vertekend beeld van OLDD.

In scenario 4 wordt er gekeken naar de invloed van vervangingen op de onttrekkingstijd. Het SB wat het goed doet is SRSD. Hierbij wordt aangetekend dat SRSD maar 1 SLT heeft behandeld. Dit geeft dus een vertekend beeld.

Scenario 5 en 6 zijn bedoeld om te kijken of drukte invloed heeft op de gemiddelde doorlooptijd. Er is gekeken naar wie het beste scoort in een rustige periode en wie het beste scoort in een drukke periode. Helaas komt naar voren dat er onvoldoende gegevens zijn om 1 week in een rustige en drukke periode te bekijken. Er komen te weinig SLT treinen binnen om een uitspraak te doen die niet stoelt op een toevallige binnenkomst.

Tabel 19 - Prestatie over 6 scenario's

DMU	Efficiënt	Niet Efficiënt
OLDD	1	5
SAMR	3	3
SBKH	1	5
SRSD	3	3
SRTD	1	5
SUTC	1	5
SUTZ	0	6
SWGMM	1	5

In Tabel 19 is een overzicht gegeven van hoe vaak een SB naar voren komt als efficiënt. Te zien is dat zowel SAMR als SRSD beide 3 keer als efficiënt naar voren komen. SUTZ komt in alle Er zijn 2 SB's die het beter doen dan de rest, dit zijn SAMR en SRSD. In 3 van de 6 scenario's komen deze als beste naar voren. Dit duidt er op dat deze twee SB's het goed doen qua onttrekkingduur bij het onderhoud van de SLT in verschillende scenario's. De andere SB's blijven duidelijk achter.

De vraag die hier op volgt is uiteraard de vraag waar deze verschillen vandaan komen. Deze

vraag kan nu meer specifiek gesteld worden, omdat er dankzij de uitgevoerde benchmark een indicatie is waar gezocht moet worden naar best-practices. Door te onderzoeken waarom SAMR en SRSD het beter doen, kan er middels stap 7, opstellen verbeterplan, en stap 8, de implementatie, een vervolg worden gegeven aan de eerste 6 stappen die doorlopen zijn in de casus. Uiteindelijk kan dit resulteren in een lagere gemiddelde onttrekkingduur van het onderhoud van de SLT en potentieel ook ander materiaal.

4.4.2 Toepasbaarheid van de methodologie

Het doel van de casus is, naast het analyseren van de uitkomsten, ook om de toepasbaarheid van de methodologie te testen. Om te bepalen of en hoe de methodologie toepasbaar is, wordt er naar hoe het gebruik van de methodologie is, of deze makkelijk hanteerbaar is en of de stappen eenvoudig zijn. Daarnaast wordt er gekeken wat de problemen zijn die naar boven komen.

De stappen laten zich eenvoudig en gestructureerd gebruiken voor het opbouwen van de benchmark casus. Door het doorlopen worden er steeds keuzes gemaakt die de casus steeds meer houvast en richting geven. Door eerst het proces te kiezen, wordt de keuze voor benchmark partners eenvoudiger. Wanneer deze partners bekend zijn en de belangen in kaart zijn gebracht door een stakeholder analyse gestoeld op Mitchell et al. (1997), ligt er een goede basis om variabelen te ontwikkelen waarmee de efficiëntie gemeten kan worden van het gekozen proces. De keuze welke variabelen belangrijk zijn, kan effectief worden gemaakt vanwege het feit dat de belangen van alle stakeholders in kaart zijn gebracht. Dit maakt de overweging makkelijk. In de testcasus was het voor LRRE en gezien het tijdsbestek, een goede kans om de voor hen interessante variabelen te kiezen binnen de testcasus om te kijken wat de uitkomst hiervan was.

Doordat er gebruik wordt gemaakt van DEA Frontier Free™ software kan er in ieder geval worden gezegd dat de vergelijkingsmethode goed is om te gebruiken. Doordat het softwarepakket stevig onderbouwd is door Zhu (2009) kan op een goede manier gebruik worden gemaakt van DEA en kan er vanuit worden gegaan dat de resultaten die door de software gegenereerd worden op basis van de input data valide zijn. Echter zegt dit niet over of de resultaten zelf valide zijn, want zoals aangegeven in 2.5 is DEA gevoelig voor ruis en uitschieters in de data en leveren slechte data, slechte vergelijkingen op (Luu et al., 2008; Merkert et al., 2010). Omdat de data al direct enigszins zijn opgeschoond voor negatieve onttrekkingtijden en enorm hoge onttrekkingtijden, wordt geconcludeerd dat de data niet van dermate hoge kwaliteit zijn dat de ondervonden resultaten een goede weerspiegeling van de werkelijkheid zijn. Aangezien de data voor een deel gefilterd zijn, maar dit niet automatisch betekent dat alle fouten eruit zijn betekent dit dat de data niet van hoge kwaliteit zijn. Dit in combinatie met dat er geen referentiekader beschikbaar is, volgt hieruit dat de methodologie niet gevalideerd kan worden binnen dit kader.

Ondanks de kwaliteit van de data, laat de methodologie zich makkelijk gebruiken en geeft het structuur en houvast voor het uitvoeren van een benchmark. Binnen het gestelde tijds kader kan niet worden uitgezocht hoe de resultaten zich verhouden met de werkelijkheid.

4.4.3 Discussie LRRE

Ondanks dat de data niet van de vereiste kwaliteit zijn, is de methodologie desondanks wel een keer succesvol toegepast in de praktijk middels de casus. De resultaten zijn gepresenteerd bij Lloyd's Register Rail Europe en uit de discussie die volgde was duidelijk te merken dat de problemen de lage kwaliteit van de data een bekend probleem is. Dit werd onderstreept dat door het gebruik van DEA met variabelen die interessant zijn om te gebruiken, er uitkomsten uit zijn gekomen die voor een aantal personen als onverwacht zijn bestempeld en niet in de lijn der verwachtingen met wat zij hadden gedacht en voorspeld op basis van hun jaren lange ervaring in de railsector. Echter werd de methodologie met behulp van DEA wel zeer interessant en goed onderbouwd gevonden, juist omdat de methodologie in combinatie met DEA de weging van meerdere variabelen tegelijk mogelijk maakt en dit nog niet eerder op deze manier was gepresenteerd. Dit biedt volgens hen kansen om op een andere manier zaken te kunnen onderzoeken om te kijken of hier andere resultaten of verbanden uitkomen. De data op hoger niveau krijgen is een lastige zaak, omdat ze LRRE hierbij afhankelijk is van NedTrain en zij er geen werk van maken, omdat ze bezig zijn met het overstappen naar een nieuw systeem. Daarnaast is het niet te doen om zelf data te verzamelen vanwege het feit dat iemand dan in elke SB aanwezig moet zijn en een eigen administratie moet gaan bijhouden van binnenkomsten. Hoe de resultaten zich verhouden naar de werkelijkheid zou een interessant onderzoek zijn voor LRRE.

4.4.4 Aanbevelingen

Op basis van de casus kunnen er een aantal aanbevelingen gegeven worden. Deze vallen uiteen in twee categorieën. De eerste categorie heeft betrekking op de casus zelf, de tweede categorie heeft betrekking op methodologie als geheel. De tweede categorie wordt behandeld in hoofdstuk 5 Conclusie en aanbevelingen.

De eerste aanbeveling die gegeven wordt over de casus zelf is dat het interessant is om onderzoek te doen waarom SARM en SRSD het beter doen in de onderzochte scenario's. Ondanks dat de data niet van hoge kwaliteit zijn, zijn het wel de data die door het gehele NedTrain en NS concern worden gebruikt voor analyse.

De tweede aanbeveling is om een aantal aannames die gedaan worden te controleren bij LRRE zelf of bij NTR. Aannames die gemaakt worden, zijn onder andere dat de verdeling SLT/overige treinen een-op-een kan worden overgenomen bij de verdeling van de monteurs. Daarnaast is de aanname gemaakt dat het uitvoeren van een reset, of vervanging een positieve invloed heeft op de onttrekkingduur.

De derde aanbeveling is dat er veel meer scenario's zijn die uitgevoerd kunnen worden. Om een richting te geven wordt een lijst met scenario's gegeven die potentieel interessant zijn om te bekijken. Dit kan eventueel een nieuwe beeld opleveren over hoe elk SB scoort en of dit nog invloed heeft op welke SB's naar voren komen als zijnde goed of minder goed. Gegeven is een lijst met mogelijke variabelen dan wel opties die mogelijk zijn voor andere scenario's:

- Meerdere rustige en drukke perioden met elkaar vergelijken.
- Aantal behandelde storingen per monteur.
- Gemiddelde Onttrekkingduur per monteur per SB.
- Aantal Resetten per monteur.
- Prestaties van verschillende monteurs binnen een SB.
- Een vergelijking SLT-VIRM. Dit aangezien ze een overeenkomstig intern diagnosesysteem hebben.

- Specifieke defecten vergelijken tussen SB's (tractie installatie, klimaat systeem etc.).
- Invloed van aanwezige hulpmiddelen bij verschillende SB's (outillage, draaibank, put spoor, meest moderne apparatuur etc.) op de gemiddelde doorlooptijd.
- Een bekende casus over het onderhoud bij SB's vergelijken. Dit met als doel om de methodologie en DEA te gebruiken en te vergelijken met de uitkomsten van het andere onderzoek.
- Het uitvoeren van output georiënteerde analyses. Hiervoor is echter wel de betaalde software voor nodig.
- Gealloceerd budget per SB als input.
- Aantal managers ten opzichte van aantal monteurs.

4.4.5 Conclusie

Uit de toepassing van de case blijkt dat de methodologie zich leent om gebruikt te worden voor het uitvoeren van een benchmark en dus toepasbaar is. Er ligt een stevige academische onderbouwing onder de benchmark methodologie waarin wordt behandeld waarom stappen worden uitgevoerd en de reden hierachter. Vanuit de literatuur wordt onderbouwd dat DEA goed te gebruiken is, mits er data van goede kwaliteit voor handen zijn. Echter blijkt dat de data niet van de kwaliteit zijn, die nodig is om DEA goed te kunnen gebruiken. Dit resulteert er in dat de resultaten uit het doorlopen van de benchmark methodologie niet in de lijn der verwachtingen liggen van LRRE en dat er dus niet hard gemaakt kan worden of de resultaten waarheidsgetrouw zijn. Het tijds kader laat het ook niet toe om te onderzoeken hoe de resultaten zich verhouden met de werkelijkheid. Dit leidt tot de conclusie dat de methodologie wel toepasbaar is binnen de rail sector, maar niet gevalideerd kan worden binnen het kader van dit onderzoek. Naast de scenario's die al onderzocht zijn, zijn er nog verschillende scenario's te bedenken die onderzocht kunnen worden om een beter beeld te krijgen van de verschillende SB's. Gegeven in de aanbevelingen is een lijst met variabelen en opties die gebruikt kunnen worden om nieuwe en andere scenario's te onderzoeken.

5 Conclusie en aanbevelingen

5.1 Conclusie

Vanuit Lloyd's Register Rail Europe is de vraag gekomen hoe een benchmark methodologie eruit ziet die gericht is op de railsector. Door middel van dit rapport wordt antwoord gegeven op deze vraag. De probleemstelling van dit rapport is opgebouwd uit een hoofdvraag met vier deelvragen. De vier deelvragen zijn door de verschillende hoofdstukken heen beantwoord en hiermee kan antwoord gegeven worden op de hoofdvraag:

“Hoe ziet een benchmark methodologie eruit die gericht is op de railsector en is deze toepasbaar?”

Aan de hand van een uitgebreid onderzoek in de literatuur is er een benchmark methodologie ontwikkeld. Deze methodologie, zoals gepresenteerd in hoofdstuk 3, is het antwoord op de hoofdvraag. In de ontwikkelde methodologie is rekening gehouden met verschillende unieke karakteristieken van de railsector. Zoals te lezen in 2.4.3 op bladzijde 24 wordt de railsector gekenmerkt door weinig vergelijkingspartners, externe factoren zoals overheden en factoren die niet economische ingestoken zijn. Vanuit de literatuur is naar voren gekomen dat Data Envelopment Analysis (DEA) een geschikte vergelijkingstechniek is om met deze kenmerken om te gaan. Door gebruik te maken van het ontwikkelde Excel bestand in combinatie met DEA Frontier Free™ kan er op een eenvoudige en valide manier gebruik worden gemaakt van DEA. Verder houdt de methodologie rekening met de verschillende stakeholders doormiddel van twee acceptatie momenten. Zoals naar voren is gekomen in 3.2 is het belang van de stakeholder hiermee gewaarborgd.

Doormiddel van een testcasus is in hoofdstuk 4 gekeken naar de toepasbaarheid van de methodologie. Door de methodologie te testen is er gekeken naar hoe de methodologie in gebruik is en welke problemen naar voren komen. De casus is in samenspraak met LRRE. De toepasbaarheid van de methodologie is goed en de benchmark methodologie houdt de definitie zoals gegeven in 2.2.1 aan. Door gebruik te maken van de 8 stappen, kan op een eenvoudige en gestructureerde manier een benchmark worden opgezet. Na het uitvoeren van de benchmark kwam naar voren dat Roosendaal en Amersfoort beide als beste uit de bus kwamen. Dit strookt echter niet met wat LRRE had verwacht met alle kennis die ze hebben van de sector en de resultaten zijn dus onverwacht. Dit kan een aantal redenen hebben. Zoals eerder is aangegeven is een onderzoek zo goed als de data zijn. De data die gebruikt zijn, zijn dezelfde als LRRE, Nedtrain en de NS gebruiken. Echter staat de kwaliteit van deze data al jaren ter discussie. Er is naar voren gekomen dat de data niet van hoge kwaliteit zijn. Verdere implicatie hiervan is dat DEA als methode gevoelig is voor uitschieters in de data die gebruikt worden. Als de data veel uitschieters bevat, wat goed kan met de lage kwaliteit, kan het zijn dat DEA niet goed geschikt is voor de vergelijking. Als laatste kan het ook betekenen dat deze twee SB's goed scoren op de geteste punten. Om te kijken wat de resultaten nu echt in houden, vergeleken met de werkelijkheid, moet de methode gevalideerd worden. Echter is dit binnen de scope van het onderzoek niet mogelijk. Naast dat het doel van het onderzoek gericht was op toepasbaarheid, mist er ook een referentiekader waaraan de methodologie getest kan worden. Zonder dit referentiekader kan niet worden vastgesteld hoe de resultaten zich verhouden.

Hiermee is de hoofdvraag in zijn geheel beantwoord. De benchmark methodologie geeft eenvoudig aan welke partners het meest efficiënt zijn en bij deze moet dan gezocht worden naar de best-practices en sluit hier mee aan op de gegeven definitie in 2.2.1 op bladzijde 11. Rekening moet gehouden worden met het feit dat het model niet gevalideerd is en dat DEA gevoelig kan zijn voor data van lage kwaliteit.

5.2 Aanbevelingen

Omdat de methodologie alleen op toepasbaarheid, maar niet op validiteit getest is, kan er geen uitspraak gedaan worden over hoe de resultaten zich echt verhouden met de werkelijkheid. Omdat dit wel een belangrijk onderdeel is van elke methodologie, is de eerste aanbeveling om de methodologie te testen op de validiteit. Om te testen op validiteit is er een referentiekader nodig. Dit referentiekader kan het beste binnen LRRE gezocht worden, omdat dat het valideren het makkelijkste maakt. Aangezien valideren inhoudt dat er gekeken wordt of de methodologie zich verhoudt naar hoe de werkelijkheid is, is het beste om een casus te pakken uit het verleden. LRRE heeft meerdere casussen waar zaken zijn uitgezocht en vergeleken. Deze manieren zijn destijds als valide gebruikt naar de klant toe en kunnen dus voor waarheid aangenomen worden. Door deze bestaande casus als referentiekader te gebruiken en vervolgens de data uit deze casus te gebruiken bij de benchmark methodologie uit de dit onderzoek kan er gekeken worden naar hoe de resultaten uit het gebruik ervan zich verhouden tot resultaten die eerder ondervonden zijn. Zaken waar op gelet moet worden, zijn wat de kwaliteit van de data zijn en wat de aannames waren waarmee de vergelijking destijds gemaakt is. De uitkomst van het opnieuw doen van de casus kan twee kanten op gaan. De ene kant is dat de benchmark methodologie dezelfde resultaten oplevert als het eerdere onderzoek. In dat geval kan er gezegd worden dat de benchmark methodologie valide is. Echter kan het ook zo zijn dat er een verschillend resultaat uitkomt. Wanneer dit het geval is moet onderzocht worden waar dit verschil vandaan kan komen en daarnaast wat dat voor impact heeft op de methodologie alsmede de uitkomst van het eerdere onderzoek. Door de methodologie te testen met een bestaande casus, moet er een uitspraak gedaan kunnen worden over de validiteit van de benchmark methodologie.

Zoals blijkt is het lastig om te waarborgen dat de kwaliteit van dusdanig niveau is een geavanceerdere vergelijkstechniek zoals DEA goed gebruikt kan worden. De tweede aanbeveling is om te onderzoeken hoe men kan bepalen of de data die worden aangeleverd van goede of van mindere kwaliteit zijn. Hoe dit exact gedaan kan worden valt buiten de scope van dit onderzoek, maar het hebben van een goede procedure om data te controleren op kwaliteit kan niet alleen voor de benchmark methodologie een goede zijn, maar voor menig onderzoek kan het een procedure zijn om te kunnen uitvoeren.

Mocht bij een vervolg onderzoek blijken dat de kwaliteit niet van een dusdanig niveau is dat er gebruik gemaakt kan worden van DEA is het belangrijk dat dit te ondervangen is. Omdat DEA gebaseerd is op extreme waarden en daardoor gevoelig is voor uitschieters is er een alternatief stappenplan. Deze maakt onderscheid tussen hoge kwaliteit data, waarbij DEA gebruikt kan worden, en lage kwaliteit data. Bij data met een lage kwaliteit kan er gebruik worden gemaakt van een eenvoudigere methode zoals Key Performance Indicators (KPI). In Appendix E is een aangepast stappenplan dat bij een lage kwaliteit data aangeeft dat er gebruik gemaakt dient te worden van KPI's, terwijl erbij hoge kwaliteit data er gebruik gemaakt kan worden van DEA. Door deze splitsing te verwerken in het model, blijven alle stappen nagenoeg gelijk, maar vermindert de gevoeligheid op uitschieters door met KPI's te werken als de data van lage kwaliteit zijn.

De laatste aanbeveling is om bij gebruik van de benchmark methodologie in combinatie met DEA altijd het ontwikkelde Excel bestand te gebruiken. Dit ondervangt het grootste gedeelte aan invoer fouten en vergroot daarmee de kwaliteit van de vergelijking. Als het goed werkt is de optie om de volledige versie van DEA Frontier Free™ aan te schaffen een die het overwegen waard is. De volledige versie biedt veel extra mogelijkheden en met name de output oriëntatie kan veel nieuwe inzichten opleveren.

Literatuurlijst

- Adebanjo, D., Abbas, A., & Mann, R. (2010). An Investigation of the Adoption and Implementation of Benchmarking. *International Journal of Operations & Production Management*, 30(11), 1140-1169. doi: 10.1108/01443571011087369
- Anderson, K., & McAdam, R. (2005). An Empirical Analysis of Lead Benchmarking and Performance Measurement: Guidance for Qualitative Research. *International Journal of Quality & Reliability Management*, 22(4), 354-375. doi: 10.1108/02656710510591200
- Anderson, R. J. H., Robin C;Trompet, Mark;Adeney, William E. (2003). Developing Benchmark Methodologies for Railway Infrastructure Management Companies: Imperial College London.
- Banker, R. D. C., A.;Cooper, W. W. (1984). Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. *Management Science*, 30(9), 1078-1092.
- Bottasso, A. C., Maurizio. (2011). Quantitative Techniques for Regulatory Benchmarking. Brussels: Centre on Regulation in Europe.
- Braadbaart, O., & Yusnandarshah, B. (2008). Public Sector Benchmarking: a Survey of Scientific Articles, 1990--2005. *International Review of Administrative Sciences*, 74(3), 421-433. doi: 10.1177/0020852308095311
- Caldas, M. A. F. C., R. L.; Gabriele P. D.; Ramos, T. G. (2013). *The Efficiency of Freight Rail Transport; An Analysis from Brazil and United States*. Paper presented at the 13th WCTR, Rio de Janeiro, Brazil.
- Camp, R. (1989). *Benchmarking: The Search for Industry Best Practices that Lead to Superior Performance*. Milwaukee, WI: ASQC Quality Press.
- Charnes, A. C., W.W.;Rhodes, E. (1978). Measuring the Efficiency of Decision Making Units. *European Journal of Operational Research*, 2(4), 429-444.
- . *Ex Anta Evaluation; A Practical Guide for Preparing Proposals for Expenditure Programmes*. (2001).
- Finnigan, J. P. (1996). *The Manager's Guide to Benchmarking: Essential Skills for the New Competitive-Cooperative Economy* (1st ed.). San Fransisco, California: Jossey-Bass.
- Growitsch, C., & Wetzels, H. (2009). Testing for Economies of Scope in European Railways. *Journal of Transport Economics and Policy*, 43(1), 1-24.
- Hansen, I. A. W., Paul B. L.; Wolff, Jeroen W. (2013). Benchmark Analysis of Railway Networks and Undertakings: UIC.
- Isoraite, M. (2004). Benchmarking Methodology in a Transport Sector. *Transport*, 19(6), 269-275. doi: 10.1080/16484142.2004.9637986
- Isoraite, M. (2005). Analysis of Transport Performance Indicators. *Transport*, 20(3), 111-116. doi: 10.1080/16484142.2005.9638006
- Joo, S.-J., Nixon, D., & Stoeberl, P. A. (2011). Benchmarking with Data Envelopment Analysis: a Return on Asset Perspective. *Benchmarking: An International Journal*, 18(4), 529-542. doi: 10.1108/14635771111147623
- Keyley, P. M.-B. S. A. (1997). Can Benchmarking for Best Practices Work for Government? *Quality Progress*, March, 75-80.
- Laird, K. Z., Joffery;Antoniazzi, Federico;Focsaneanu, Emil;Israel, Sabine;Ionnidou, Anna-Maria. (2011). Report On Technical Benchmarking of European Railways: European Railway Agency.

- Luu, V. T., Kim, S.-Y., & Huynh, T.-A. (2008). Improving Project Management Performance of Large Contractors Using Benchmarking Approach. *International Journal of Project Management*, 26(7), 758-769. doi: 10.1016/j.ijproman.2007.10.002
- McAdam, R., Hazlett, S.-A., & Anderson-Gillespie, K. (2008). Developing a Conceptual Model of Lead Performance Measurement and Benchmarking: A Multiple Case Analysis. *International Journal of Operations & Production Management*, 28(12), 1153-1185. doi: 10.1108/01443570810919350
- McCoullough, G. J. (2007). *US Railroad Efficiency: A Brief Economic Overview*. Department of Applied Economics, University of Minnesota.
- Merkert, R., Smith, A. S. J., & Nash, C. A. (2010). Benchmarking of Train Operating Firms – a Transaction Cost Efficiency Analysis. *Transportation Planning and Technology*, 33(1), 35-53. doi: 10.1080/03081060903429330
- Mitchell, R. K., Agle, B. R., & Wood, D. J. (1997). Toward a Theory of Stakeholder Identification and Salience: Defining the Principle of Who and What Really Counts. *The Academy of Management Review*, 22(4), 853-886.
- Moffett, S., Anderson-Gillespie, K., & McAdam, R. (2008). Benchmarking and Performance Measurement: a Statistical Analysis. *Benchmarking: An International Journal*, 15(4), 368-381. doi: 10.1108/14635770810887203
- Moriarty, J. P. (2011). A Theory of Benchmarking. *Benchmarking: An International Journal*, 18(4), 588-611. doi: 10.1108/14635771111147650
- Moriarty, J. P., & Smallman, C. (2009). En Route to a Theory of Benchmarking. *Benchmarking: An International Journal*, 16(4), 484-503. doi: 10.1108/14635770910972423
- Ozcan, Y. A. (2008). *Health Care Benchmarking and Performance Evaluation*: Springer-Science+Business Media, LCC.
- Roberts, G. (2012). *GB Rail Efficiency & Benchmarking*. (Masters Master Thesis), University of Strathclyde.
- Spendolini, M. J. (1992). *The Benchmarking Book*: AMACON, American Management Association.
- Training, N. T. R. a. (2003). BOB Railway Case Benchmarking Passenger Transport in Railways. Rijswijk, The Netherlands: European Commission.
- Underground, L. (2012). TfL Rail and Underground Annual Benchmarking Report.
- van Hoorn, A. (2005). Benchmarking OK: Van appels met peren naar Elstar met Jonagold (Vol. Deel I: Doel, Visie, Methode & Systemenkenmerken): NFU Projectgroep Benchmarking OK.
- Vaziri, M. M. (1992). Using Competitive Benchmarking to Set Goals. *Quality Progress*, 25(10), 81-85.
- Winston, W. L. (2004). *Operations Research* (Fourth ed.): Brooks/Cole - Thomson Learning.
- Wu, W.-W. (2012). An Integrated Solution for Benchmarking Using DEA, Gray Entropy, and Borda Count. *The Service Industries Journal*, 32(2), 321-335. doi: 10.1080/02642069.2011.517833
- Yasin, M. M. (2002). The Theory and Practise of Benchmarking; Then and Now. *Benchmarking: An International Journal*, 9(3), 217-243.
- Zhu, J. (2009). *Quantitative Models for Performance Evaluation and Benchmarking* (Second ed.): Springer.

Appendix A Begrippenlijst

BIC	Best In Class
DMU	Decision Making Unit
DEA	Data Envelopment Analysis
LSR	Least Square Regression
TFP	Total Factor Productivity
SFA	Stochastic Frontier Analysis
LP	Lineair Programmeren
CRS	Constant Return to Scale
VRS	Variable Return to Scale
NoBo	Notified Body
NS	Nederlandse Spoorwegen
LRRE	Lloyd's Register Rail Europe BV
NTR	NedTrain B.V.
SLT	Sprinter Light Train
UIC	International Union of Railways
ERA	European Railway Agency
EC	Europese Commissie
SB	Service Bedrijf van NedTrain
OB	Onderhoudsbedrijf van NedTrain
OLDD	Onderhoudsbedrijf Leidschendam
SAMR	Servicebedrijf Amersfoort
SBKH	Servicebedrijf Den Haag Binckhorst
SRSD	Servicebedrijf Roosendaal
SRTD	Servicebedrijf Rotterdam
SUTC	Servicebedrijf Utrecht Cartersiusweg
SUTZ	Servicebedrijf Utrecht Zuid
SWGM	Servicebedrijf Watergraafsmeer

Appendix B Excel Handleiding

Door gebruik te maken van Microsoft Office Excel in combinatie met de Excel plug-in DEA Frontier Free™, ontwikkeld door [ENREF_37](#)Zhu (2009), kunnen er eenvoudig DEA analyses uitgevoerd worden. Het gebruik van de plug-in vereist wel wat dat handelingen op de juiste manier moeten worden uitgevoerd, anders werkt het programma niet. In deze handleiding worden de juiste stappen uitgelegd. In combinatie met de het Excel bestand 'DEA in Excel' moet het gebruik eenvoudig, duidelijk en vooral goed zijn. In Appendix C is een overzicht gegeven van hoe de eerste 3 tabbladen eruit zien.

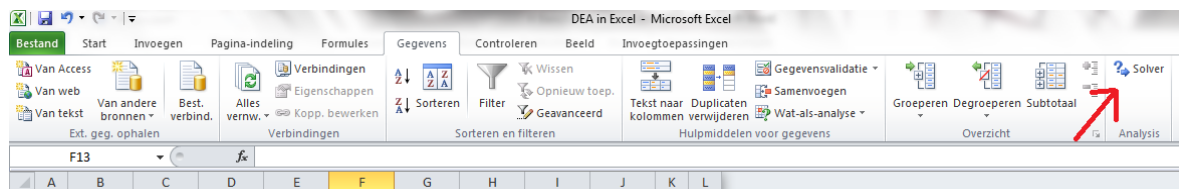
B.1 Belangrijk

Er zijn twee zaken essentieel bij het gebruik van DEA Frontier Free™. Deze twee zaken moeten altijd gecheckt worden voordat gebruik wordt gemaakt van de software. Deze twee zaken zijn dat Excel Solver actief is in Excel en dat DEA Frontier Free™ ingeladen is.

B.1.1 Excel Solver

De eerste check die gedaan moet worden, is de controle dat Excel Solver actief is in Excel. Dit houdt in dat Excel Solver geïnstalleerd en actief moet zijn. Dit kan gecheckt worden door in de werkbalk bovenaan in Excel, bij het kopje 'gegevens' te kijken. Hier moet helemaal rechts bij 'Analysis' het Solver pictogram staan, zoals te zien bij de rode pijl in Figuur 18. Mocht deze niet actief zijn, dan is het volgen van de stappen op de site van Microsoft Office voor het activeren van de Excel Solver hetgene dat gedaan dient te worden. Deze stappen kunnen op de volgende site gevonden worden:

<http://office.microsoft.com/nl-nl/excel-help/snel-aan-de-slag-een-invoegtoepassing-activeren-en-gebruiken-HA010370161.aspx>



Figuur 18 - Solver in Excel

B.1.1.1 DEA Frontier Free™

De tweede stap is dat de plug-in DEA Frontier Free™ wordt geladen. Doe dit pas nadat de Excel solver actief is, anders kan de plug-in niet goed werken en komt er een foutmelding bij gebruik.

B.2 Tabbladen

Er zijn een aantal tabbladen in de Excel file. Deze worden hier uitgelegd zodat de werking hiervan duidelijk is. In Appendix C is te zien hoe de tabbladen eruit zien.

B.2.1 Voorblad

Het tabblad 'voorblad' (kleurcode groen) dient als navigatie blad, zie Figuur 19. Er staat een korte uitleg en er wordt verwezen naar de verschillende andere tabbladen. Het voorblad kan niet worden aangepast.

B.2.2 Uitleg

Voor extra uitleg binnen Excel over het gebruik van DEA Frontier Free™ is in het tabblad 'Uitleg' (kleurcode groen), zie Figuur 19, extra informatie te vinden over hoe er op een correcte manier gebruik gemaakt kan worden van de software. De uitleg kan niet worden aangepast.

B.2.3 Formules

Omdat gegevens op de juiste manier gepresenteerd moeten zijn voor het gebruik, staan in het tabblad 'formules' (kleurcode groen) formules om waarden te inverteren dan wel indexeren.

B.2.4 Data

In het tabblad 'Data' (kleurcode geel), zie Figuur 19, moeten de data worden ingevoerd die gebruikt gaan worden voor DEA. Dit dient op de juiste manier te gebeuren, want dit tabblad wordt gebruikt door de software om de DEA analyse uit te voeren. Een aantal aanpassingsmogelijkheden binnen dit tabblad zijn niet mogelijk en er zitten verscheidene checks in om te zorgen dat er zo min mogelijk fouten worden gemaakt bij het invoeren van de benodigde data.

B.2.5 Ruwe Data

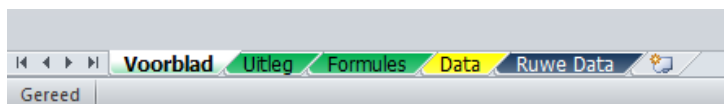
Het tabblad 'Ruwe Data' (kleurcode blauw), zie Figuur 19, kunnen de verzamelde data van de verschillende DMU verzameld worden. Dit tabblad dient puur voor het eigen gebruik en is bedoeld om alle benodigde gegevens in een Excel bestand te verzamelen. Behalve het tabblad 'Ruwe Data' kunnen er ook nog eigen tabbladen worden toegevoegd die dezelfde functie hebben.

B.2.6 Overige tabbladen

De plug-in maakt bij het gebruik een aantal nieuwe tabbladen aan. Deze tabbladen zijn de output van de DEA analyse. Hierin staat wat de verschillende scores zoals die berekend zijn. De twee modellen die gebruikt worden genereren een drietal nieuwe tabbladen. Deze tabbladen zijn: 'Target', 'Slack', en 'Efficiency'.

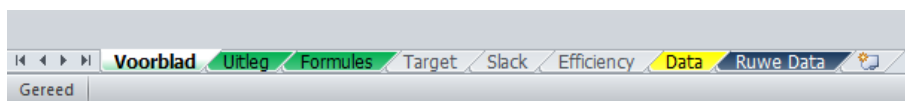
B.3 Overzicht

Bij het openen van 'DEA in Excel' ziet de opbouw van tabbladen er als volgt uit, zie Figuur 19.



Figuur 19 – Tabbladen DEA in Excel, uitgangspositie

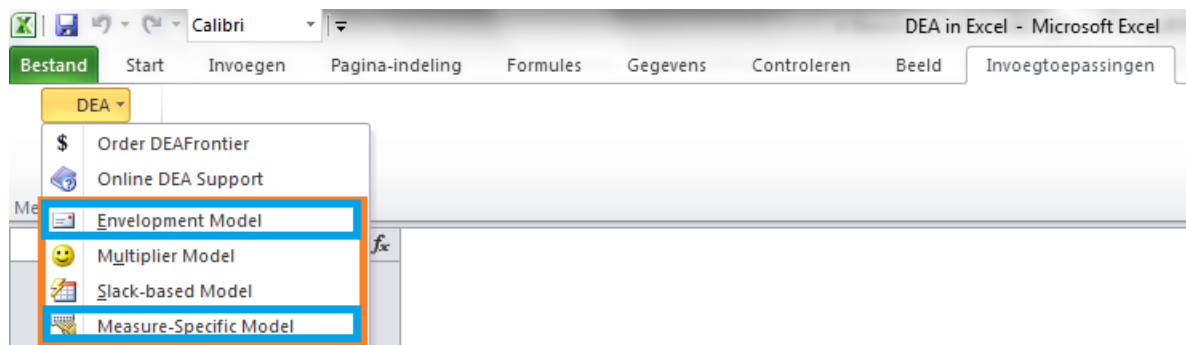
Door het gebruik van DEA Frontier komen er een aantal tabbladen bij. Dit ziet er als volgt uit wat te zien is in Figuur 20.



Figuur 20 – Tabbladen DEA in Excel, na gebruik DEA Frontier Free™

B.4 Modellen beschikbaar in DEA Frontier Free™

In DEA Frontier Free™ zijn vier modellen beschikbaar, namelijk: Envelopment Model, Multiplier Model, Slack Based Model en Measure Specific Model, zoals te zien in Figuur 21. In dit onderzoek worden 2 modellen gebruikt, de andere twee modellen worden niet gebruikt. De twee modellen die niet worden gebruikt zijn het Slack-Based Model en het Multiplier Model. Deze twee modellen hebben alleen functionaliteit als de betaalde versie wordt gebruikt. Hierdoor blijven er twee modellen over die gebruikt kunnen worden. Het Envelopment Model en het Measure-Specific Model.



Figuur 21 - DEA Modellen in DEA Frontier Free™

B.4.1 Envelopment model

Het Envelopment Model is het basis model van DEA. Met de geselecteerde inputs en outputs wordt de situatie weergegeven op het gebied van efficiency. De gegenereerde tabbladen zijn 'Target', 'Slack' en 'Efficiency'. In het tabblad 'Target' geeft weer wat het doel is voor elke DMU. Van elke input is te zien wat het niveau moet zijn om efficiënt te zijn. 'Slack' geeft aan hoe groot de reductie van input moet zijn om efficiënt te zijn. Het tabblad 'Efficiency' geeft aan wat de huidige scores zijn en tevens ook welke DMU het meest efficiënt is. Een DMU is efficiënt als de efficiency score in kolom 3, Input-Oriented CRS Efficiency, de waarde 1,000 heeft zoals te zien in Tabel 20.

Tabel 20 - Tabblad Efficiency

DMU No.	DMU Name	Input-Oriented CRS Efficiency	Optimal Lambdas			
			Sum of lambdas	RTS	with Benchmarks	
1	Seattle	1,00000	1,000	Constant	1,000	Seattle
2	Denver	1,00000	1,000	Constant	1,000	Denver
3	Philadelphia	1,00000	1,000	Constant	1,000	Philadelphia
4	Minneapolis	0,98238	0,987	Increasing	0,737	Philadelphia
5	Raleigh-Durham	1,00000	1,000	Constant	1,000	Raleigh-Durham
6	St. Louis	1,00000	1,000	Constant	1,000	St. Louis
7	Cincinnati	0,80770	0,763	Increasing	0,095	Seattle
8	Washington	1,00000	1,000	Constant	1,000	Washington
9	Pittsburgh	0,89809	0,790	Increasing	0,145	Seattle
10	Dallas-Fort Worth	0,86440	0,939	Increasing	0,655	Philadelphia
11	Atlanta	0,82216	1,009	Decreasing	1,009	Denver
12	Baltimore	1,00000	1,000	Constant	1,000	Baltimore
13	Boston	1,00000	1,000	Constant	1,000	Boston
14	Milwaukee	0,81208	1,005	Decreasing	0,571	Philadelphia
15	Nashville	0,78108	0,944	Increasing	0,034	Raleigh-Durham

In Tabel 20 is in de laatste kolom te zien welke DMU het meest efficiënt is voor de DMU die geëvalueerd wordt.

B.4.2 Measure specific model

Het 'Measure Specific Model' kan worden gebruikt als een of meerdere inputs een vaste waarde hebben. Dit kan komen doordat de waarde vast is, bijvoorbeeld aantal kilometers spoor, of als de DMU geen invloed heeft op de variabele, bijvoorbeeld toebedeeld budget. Tevens kan het Measure Specific Model worden toegepast wanneer men de invloed van één variabele wilt weten. Dan kunnen de andere variabelen als vaste waarden worden ingesteld. Doordat de variabelen moeten worden geclassificeerd, is het eenvoudig te zien wanneer er een Measure Specific model gebruikt dient te worden, dit is namelijk het geval zodra er een variabele een impact/result variabele is. Deze moet dan ook worden geselecteerd als zijnde constant.

Om de ontwikkelde methodologie te testen is het gewenst om een test casus uit te voeren. In de casus wordt de methodologie doorlopen en wordt gekeken welke resultaten dit oplevert. Op basis van deze

bevindingen kan uitspraak worden gedaan over de methodologie en kunnen tekortkomingen worden ontdekt en aanpassingen worden voorgedragen. Dit geeft antwoord op deelvraag 4. Belangrijk hierbij is dat het stappenplan tot stap 6 wordt uitgevoerd. Stap 7 en stap 8 onderdeel zijn van een eventuele vervolgstudie naar aanleiding van de resultaten. Stap 7 en 8 worden daarom wel kort toegelicht, maar niet in volledigheid uitgewerkt.

Appendix C DEA in Excel

C.1 Voorblad

Lloyd's
Register

Made by: T.W. Duffhues

DEA In Excel

Versie: 1.0

LIFE MATTERS

Voor het juiste gebruik van DEA in Excel is het van belang om er voor te zorgen dat alle variabelen goed worden ingevoerd. Daarnaast moet de DEA Frontier Free zijn ingeladen. Het Excelbestand is dusdanig ontworpen dat een foutieve invoer niet logisch dan wel mogelijk is. Tezamen met de Excel uitleg in het rapport zorgt het voor een feilloze werking.

Tabbladen

[Voorblad](#)

Dit tabblad geeft overzicht over wat de inhoud is van DEA in Excel.

[Uitleg](#)

In dit tabblad staat uitgelegd hoe DEA in Excel gebruikt moet worden.

[Formules](#)

In dit tabblad staat uitgelegd hoe data kan worden geïnterteerd of geïndexeerd. Dit om het gebruik in DEA goed te laten verlopen.

[Data](#)

In dit tabblad wordt de Data ingevoerd voor het toepassen van DEA. Volg voor het juiste gebruik de handleiding in het rapport en de punten zoals genoemd in DEA Gebruik.

[Ruwe Data](#)

Hier kan eigen Ruwe Data in worden gekopieerd voor makkelijk gebruik.

Overige Opmerkingen

Er worden door de DEA Frontier Plugin verschillende tabbladen aangemaakt tijdens gebruik. Hierin staan de verschillende resultaten. Voor wat deze betekenen wordt u verwezen naar de Excel handleiding het rapport.

1

C.2 Uitleg

Lloyd's Register DEA In Excel

Made by: T.W. Duffhues

LIFE MATTERS

Versie: 1.0

Uitleg

DEA Gebruik

Door gebruik te maken van de DEA Frontier Excel plug-in van Zhu kan er eenvoudig gebruik gemaakt worden van DEA in Excel. Hieronder staat een uitleg hoe het format moet worden gebruikt voor een goede werking. Dit in combinatie met de Excel uitleg in het rapport moet een goede werking bewerkstelligen.

DMU

In de kolom DMU moeten alle DMU's ingevuld worden. De DMU's zijn de betrokken partijen. Er kan maar 1 kolom met DMU's zijn. Eventueel kan nog voor het overzicht DMU als naam verandert worden in de naam van de variabelen (Spoorwegmaatschappijen, Beheerders etc.) Het maximum aan DMU in de DEA Frontier Free versie is 20. Meer kunnen er ook niet worden ingevuld.

Input

De kolom input staat voor de inputs die in het model gebruikt worden. Dit kan er een zijn, maar ook meerdere. Geef de kolom 'input' de naam van de variabele. Bij een 2e input dient men een nieuwe kolom aan te maken. Dit doet men door op de kolom input te klikken met de rechtermuisknop en dan invoegen te klikken.

Output

De kolom outputs moet altijd gescheiden zijn door een lege kolom voor de juiste werking van DEA Frontier. Geef de kolom 'output' de naam van de output variabele. Nieuwe output variabelen kunnen rechts van de output kolom worden toegevoegd.

Voorbeeld Format (Zhu)

Format for Data sheet

	A	B	J	R	S	W	Z
1	DMU Name	Input Measure	Input Measure		Output Measure	Output Measure	
2							
3							
4							
5	In column A (cell A2), enter the DMUs	Starting with column B (cell B2), enter data for inputs , no blank columns are allowed			enter data for outputs , no blank columns are allowed		
6							
7							
8							
9							
10							
11							
12							
13							
14							
15		Multiple inputs are allowed, although only two are shown			Multiple outputs are allowed, although only two are shown here		
16							
17							
18							
19							
20							
21							

Name the sheet containing the data on DMUs as Data

Leave one blank column between the last input measure and the first output measure



For example,

	A	B	C	D	E	F	G
1	Company	Assets	Equity	Employees		Revenue	Profit
2	Mitsubishi	91920.6	10650	36000		184365.2	346.2
3	Mitsui	68770.9	5553.9	80000		181518.7	314.8
4	Itochu	65708.9	4271.1	7182		169164.6	121.2
5	General Motors	217123.4	23345.5	709000		168828.6	6880.7
6	Sumitomo	50268.9	8681	6193		167530.7	210.5
7	Marubeni	71439.3	5239.1	6702		161057.4	156.6
8	Ford Motor	243283	24547	346990		137137	4139
9	Totota Motor	106004.2	49691.6	146855		111052	2662.4
10	Exxon	91296	40436	82000		110008	6470
11	Royal Dutch/Shell Group	118011.6	58986.4	104000		109633.7	6904.6
12	Wal-Mart	37871	14762	675000		93627	2740
13	Hitachi	91620.9	29907.2	331852		84167.1	1468.8
14	Nippon Life Insurance	364762.5	2241.9	89690		83206.7	2426.6
15	Nippon Telegraph & Telephone	127077.3	42240.1	231400		81937.2	2209.1
16	AT&T	88884	17274	296300		79809	139

DMUs

Inputs

Outputs

2

C.3 Formules



DEA In Excel

LIFE MATTERS

Made by: T.W. Duffhues

Versie: 1.0

Wanneer waarde moeten worden omgezet 'van-hoog-naar-laag' tot 'van-laag-naar-hoog' kan de inverteer formule gebruikt worden.

Formule voor Inverteren

#DEEL/0!

Formule: =1/doelcel

Voorbeeld:

Geinverteerde waarde	Orginele waarde
0,1000	10,00
0,2000	5,00
0,7463	1,34

Wanneer er negatieve waarden in de data zitten, kunnen deze met behulp van de indexeer formule worden omgerekend.

Formule voor Indexeren

0,5

Formule: =1/(1+EXP((-1)*doelcel))

Voorbeeld:

Geindexeerde waarde	Orginele waarde
0,57444	0,30
0,43045	-0,28
0,32082	-0,75

3

Appendix D Casus scenario's

D.1 Scenario 2 – basis, genormeerd naar SLT/Overige treinen

D.1.1 Beschrijving

Het tweede scenario bekijkt de situatie van de SLT's. Er wordt gekeken naar de aantallen onttrekkingen en de gemiddelde duur van de onttrekking, tevens is het aantal monteurs genormeerd naar het aantal SLT ten opzichte van het overige materiaal. De aanname is dat de verdeling van de verdeling van monteurs gelijk is aan de verdeling van de treinsoorten. Dit wordt onder andere afgeleid van het feit dat een monteur in principe in zijn eentje aan een storing werkt.

De inputs zijn: Aantal monteurs naar verhouding SLT/Overige, aantal onttrekkingen van de SLT en als output de gemiddelde duur van de onttrekking van de SLT. De data zijn ontdaan van uitschieters, zoals uitgelegd in 4.3.5.2.1. De periode die is bekeken is de van januari 2010 tot juni 2013.

D.1.2 Variabelen

Tabel 21 - Variabelen Scenario 2

Locatie	Input	Classificatie	Output	Classificatie
OLDD	Aantal Monteurs naar verhouding SLT/Overige	Input Indicator	Gem. Duur Onttrekking	Result Indicator
SAMR	Aantal Onttrekkingen	Impact Indicator		
SBKH				
SRSD				
SRTD				
SUTC				
SUTZ				
SWGM				

D.1.3 Uitwerking

Het model dat gebruikt is, is Measure Specific, waarbij het aantal onttrekkingen als constant wordt beschouwd, deze is dus niet beïnvloedbaar. Als model is er gekozen voor een VRS model, omdat er meerdere partijen efficiënt zijn.

Tabel 22 - Scenario 2 Measure Specific VRS Tabblad 'Efficiency'

DMU No.	DMU Name	Input-Oriented VRS Measure-Specific Efficiency	Optimal Lambdas with Benchmarks	
1	OLDD	0,03044	1,000	SAMR
2	SAMR	1,00000	1,000	SAMR
3	SBKH	0,25243	1,000	SAMR
4	SRSD	1,00000	1,000	SRSD
5	SRTD	0,21311	1,000	SAMR
6	SUTC	0,29545	1,000	SAMR
7	SUTZ	0,24528	1,000	SAMR
8	SWGGM	0,74286	1,000	SAMR

D.2 Scenario 3 – invloed reset

D.2.1 Beschrijving

Het derde scenario kijkt of het uitvoeren van een reset invloed heeft op de gemiddelde onttrekkingstijd. Een reset is in dit geval het uitvoeren van een software reset waarmee getracht wordt om problemen op te lossen. Vaak als een systeem niet meer reageert, kan een reset uitkomst bieden. De aanname die gemaakt wordt gemaakt bij dit scenario is dat er een positieve relatie is tussen het aantal resetten en de onttrekkingduur. Omdat er een positieve relatie wordt verondersteld dat het uitvoeren van een reset een positieve invloed heeft op de onttrekkingduur (die dus afneemt), moeten de waarden van het aantal resetten geïnverteerd worden. Omdat DEA werkt volgens het principe dat een input verlaging positief is en het in dit geval andersom moet zijn. De periode die wordt bekeken is van januari 2010 tot juni 2013.

D.2.2 Variabelen

Tabel 23 - Variabelen Scenario 3

Locatie	Input	Classificatie	Output	Classificatie
OLDD	Aantal Monteurs naar verhouding SLT/Overige	Input Indicator	Gem. Duur R; GV	Result Indicator
SAMR	Aantal Onttrekkingen	Impact Indicator		
SBKH	Aantal R; GV	Input Indicator		
SRSD				
SRTD				
SUTC				
SUTZ				
SWGGM				

D.2.3 Uitwerking

Het model dat gebuikt is, is het Measure Specific. Hierbij is het aantal onttrekkingen constant gehouden en wordt in dit model gevarieerd worden met de aantallen monteurs en de aantallen resetten. De

resetten hebben in dit geval geen vervanging van een onderdeel tot gevolg. Er is voor een CRS model gekozen, omdat het VRS model bijna alleen maar efficiënte DMU's oplevert, waardoor de betekenis van het uitvoeren van het scenario volledig wegvalt.

Tabel 24 - Scenario 3 Measure Specific CRS Tabblad 'Efficiency'

DMU No.	DMU Name	Input-Oriented CRS Measure-Specific Efficiency	Optimal Lambdas with Benchmark s					
			Sum of lambdas		RTS			
1	OLDD	1,00000	1,000	Constant	1,000	OLDD		
2	SAMR	1,00000	1,000	Constant	1,000	SAMR		
3	SBKH	1,00000	1,000	Constant	1,000	SBKH		
4	SRSD	0,75794	1,038	Decreasing	0,995	SAMR	0,043	SBKH
5	SRTD	0,46988	0,586	Increasing	0,039	SAMR	0,547	SBKH
6	SUTC	0,88720	0,950	Increasing	0,256	SAMR	0,693	SBKH
7	SUTZ	0,37425	0,480	Increasing	0,127	SAMR	0,353	SBKH
8	SWGM	0,21949	0,295	Increasing	0,295	SAMR		

D.3 Scenario 4 – invloed vervanging

D.3.1 Beschrijving

Het vierde scenario kijkt of het vervangen van een onderdeel effect heeft op de gemiddelde doorlooptijd. Dit kan bijdragen aan het onderzoek naar of preventief vervangen qua reparatietijd voordeliger is dan alleen verhelpen van een storing zonder een eventuele preventieve vervanging. Daarnaast kan er gekeken worden of er bij bepaalde DMU's vaker onderdelen worden vervangen dan bij andere. De periode die wordt bekeken is van januari 2010 tot juni 2013 en geeft dus een totaal beeld over een langere periode.

D.3.2 Variabelen

Tabel 25 - Variabelen Scenario 4

Locatie	Input	Classificatie	Output	Classificatie
OLDD	Aantal Monteurs naar verhouding SLT/Overige	Input Indicator	Gem. Duur GR; V	Result Indicator
SAMR	Aantal Onttrekkingen	Impact Indicator		
SBKH	Aantal GR; V	Input Indicator		
SRSD				
SRTD				
SUTC				
SUTZ				
SWGM				

D.3.3 Uitwerking

Het model dat gebruikt is, is het Measure Specific. Hierbij is het aantal onttrekkingen als constant verondersteld en wordt in het model gevarieerd met het aantal monteurs en het aantal vervangingen. Er is voor een CRS model gekozen, omdat bij het VRS model er uit kwam dat SAMR efficiënt is, terwijl deze geen vervangingen heeft uitgevoerd.

Tabel 26 - Scenario 4 Measure Specific CRS Tabblad 'Efficiency'

Input-Oriented			Optimal Lambdas			
DMU No.	DMU Name	CRS Measure- Specific Efficiency	Sum of lambdas	RTS	with Benchmarks	
1	OLDD	0,00485	0,104	Increasing	0,104	SRSD
2	SAMR	0,00000	0,000	Increasing		
3	SBKH	0,01847	0,048	Increasing	0,048	SRSD
4	SRSD	1,00000	1,000	Constant	1,000	SRSD
5	SRTD	0,03125	0,095	Increasing	0,095	SRSD
6	SUTC	0,03372	0,074	Increasing	0,074	SRSD
7	SUTZ	0,01962	0,052	Increasing	0,052	SRSD
8	SWGM	0,19747	0,173	Increasing	0,173	SRSD

D.4 Scenario 5 – invloed drukte I, rustige periode

D.4.1 Beschrijving

Scenario 5 en 6 zijn bedoeld om te kijken of drukte in een SB effect heeft op de gemiddelde doorlooptijd. Drukke wordt hier gedefinieerd als een hoog aanbod van het aantal treinen wat binnenkomt met een storting. De grens van wanneer het druk is, is bepaald op meer dan 30 treinen in de week die binnenkomen met een storting. De drukte wordt gemeten door te kijken naar alle treinen.

Om te kijken wat voor effect drukte heeft op de gemiddelde doorlooptijd van de SLT, zijn er twee weken genomen waarin de verhouding SLT/Overige treinen rond de 40% ligt. Dit wil zeggen dat van alle treinen die binnenkomen er 40% SLT is. Hierdoor is het mogelijk om een te kijken naar de invloed van drukte op de doorlooptijd bij de verschillende SB's.

De verhouding van SLT/overige treinen ligt ongeveer op 40%. De twee periodes die zijn week 32 in 2012 en week 49 in 2010. Hierin is week 32 de rustige periode, deze wordt in dit scenario gebruikt, en week 49 de drukke periode, deze wordt in scenario 6 gebruikt.

D.4.2 Variabelen

Tabel 27 - Variabelen Scenario 5

Locatie	Input	Classificatie	Output	Classificatie
OLDD	Aantal Monteurs naar verhouding SLT/Overige	Input Indicator	Gem. Duur Onttrekking	Result Indicator
SAMR	Aantal Onttrekkingen	Impact Indicator		
SBKH				
SRSD				
SRTD				
SUTC				
SUTZ				
SWGM				

De verhouding van het aantal monteurs is wederom genormeerd naar de verhouding SLT/Overige treinen zoals deze zich verhoudt in week 32 in 2012. Daarnaast zijn het aantal onttrekkingen en de gemiddelde duur van de onttrekking van de SLT.

D.4.3 Uitwerking

Het model dat gekozen is, is Measure Specific. Hierbij is het aantal onttrekkingen een constante en kan er gevarieerd worden met het aantal monteurs. Er is voor een CRS model gekozen, omdat het VRS model zou opleveren dat bijna alle partijen waar onderhoud wordt uitgevoerd efficiënt zouden zijn.

Tabel 28 - Scenario 5 Measure Specific CRS Tabblad 'Efficiency'

DMU No.	DMU Name	Input-Oriented CRS Measure-Specific Efficiency	Optimal			
			Sum of	Lambdas		
			lambdas	RTS	with Benchmarks	
1	OLDD	0,15904	0,477	Increasing	0,477	SRTD
2	SAMR	0,00000	0,000	Increasing		
3	SBKH	0,00000	0,000	Increasing		
4	SRSD	0,00000	0,000	Increasing		
5	SRTD	1,00000	1,000	Constant	1,000	SRTD
6	SUTC	1,00000	1,000	Constant	1,000	SUTC
7	SUTZ	0,31634	0,633	Increasing	0,216	SRTD 0,417 SUTC
8	SWGM	0,00000	0,000	Increasing		

Er zijn vier DMU's die een score van 0 hebben. Bij deze DMU's is er geen onderhoud aan de SLT uitgevoerd in week 32 en deze kunnen derhalve ook niet worden meegenomen in de evaluatie.

D.5 Scenario 6 – invloed drukte II, drukke periode

D.5.1 Beschrijving

Zoals eerder genoemd is scenario 6 bedoeld om te kijken of drukte effect heeft op de gemiddelde doorlooptijd. De periode waarnaar gekeken wordt is qua verhouding SLT/overige treinen ongeveer

gelijk ten opzichte van de drukke periode. Hierdoor kan gekeken worden naar de invloed van drukte op de gemiddelde doorlooptijd en of hier verschillen zijn tussen de verschillende SB's. De periode die wordt bekeken is week 49 in 2010. Dit is de drukke periode.

D.5.2 Variabelen

Tabel 29 - Variabelen Scenario 6

Locatie	Input	Classificatie	Output	Classificatie
OLDD	Aantal Monteurs naar verhouding SLT/Overige	Input Indicator	Gem. Duur Onttrekking	Result Indicator
SAMR	Aantal Onttrekkingen	Impact Indicator		
SBKH				
SRSD				
SRTD				
SUTC				
SUTZ				
SWGM				

De verhouding van het aantal monteurs is wederom genormeerd naar de verhouding SLT/Overige treinen zoals deze zich verhoudt in week 49 in 2012. Daarnaast zijn het aantal onttrekkingen en de gemiddelde duur van de onttrekking van de SLT.

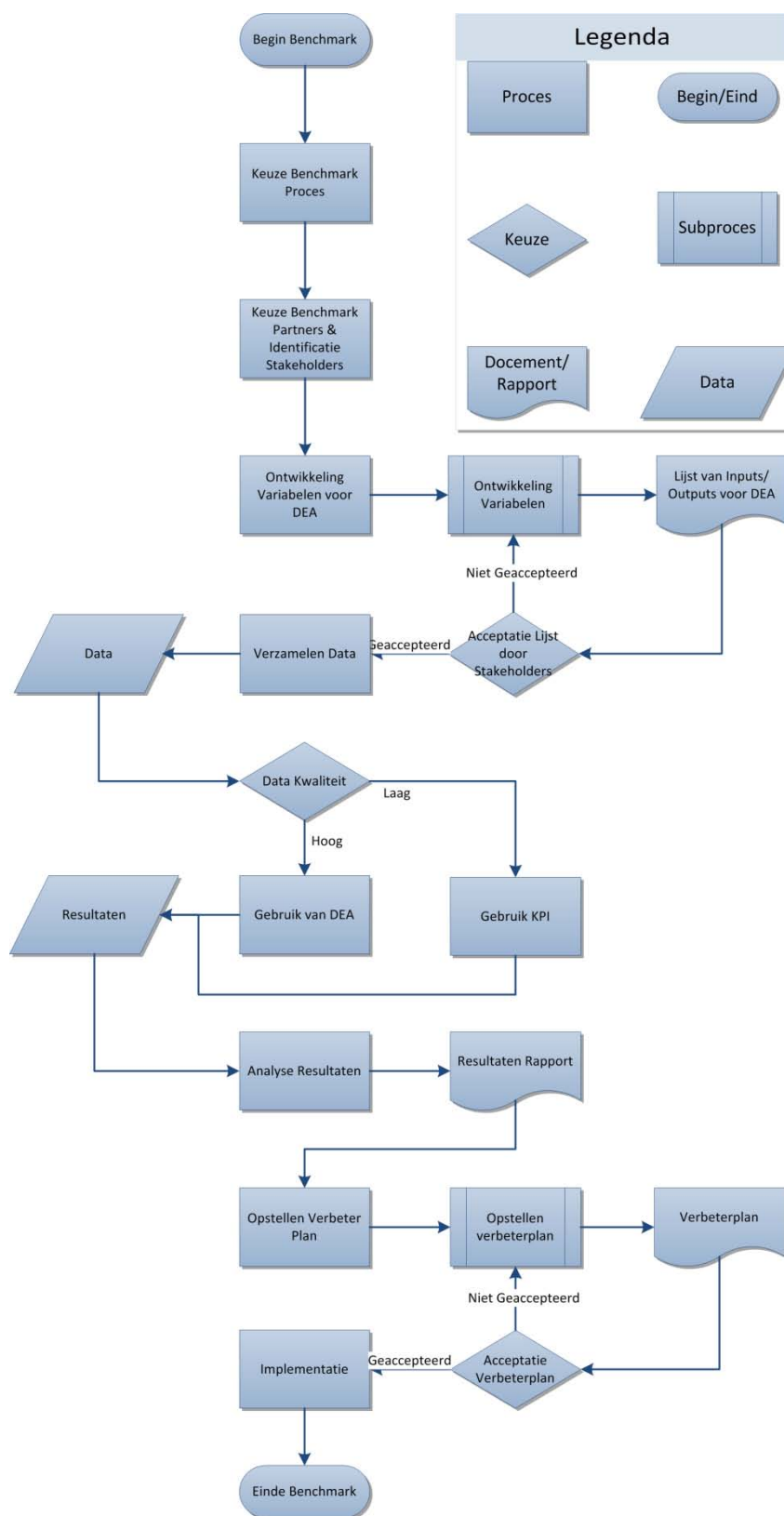
D.5.3 Uitwerking

Tabel 30 - Scenario 6 Measure Specific CRS Tabblad 'Efficiency'

Input-Oriented						
DMU No.	DMU Name	CRS Measure-Specific Efficiency	<i>Sum of</i>		<i>Optimal</i>	
			<i>lambdas</i>	<i>RTS</i>	<i>Lambdas with</i>	<i>Benchmarks</i>
1	OLDD	0,02429	0,291	Increasing	0,291	SWGM
2	SAMR	0,00000	0,000	Increasing		
3	SBKH	0,05885	0,118	Increasing	0,118	SWGM
4	SRSD	0,00000	0,000	Increasing		
5	SRTD	0,66484	1,995	Decreasing	1,995	SWGM
6	SUTC	0,33587	1,008	Decreasing	1,008	SWGM
7	SUTZ	0,23698	0,237	Increasing	0,237	SWGM
8	SWGM	1,00000	1,000	Constant	1,000	SWGM

Er zijn twee DMU's die een score van 0 hebben. Bij deze DMU's is er geen onderhoud aan de SLT uitgevoerd in week 49 en deze kunnen derhalve ook niet worden meegenomen in de evaluatie.

Appendix E Aangepast model



Figuur 22 - Benchmark Model aangepast