

---

**RISIKO-GEBASEERDE BESLUITONDER-  
STEUNING IN SIVIELE INGENIEURSWESE:  
'N METODOLOGIESE BENADERING TOT  
VERBETERENDE INLIGTINGVERSAMELING  
EN BENUTTING**

DEUR

ANDRÉ BESTER

PROEFSKRIF INGELEWER VIR DIE GRAAD DOKTOR IN DIE  
WYSBEGEERTE IN INGENIEURSWESE AAN DIE  
UNIVERSITEIT VAN STELLENBOSCH

PROMOTOR

PROF. J.V. RETIEF



**STELLENBOSCH**

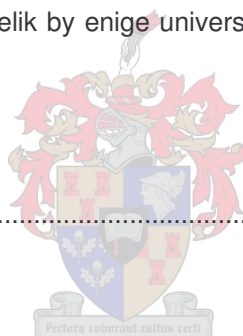
**DESEMBER 2004**

---

---

## VERKLARING

**E**k, die ondergetekende, verklaar hiermee dat die werk in hierdie proefskrif vervat my eie oorspronklike werk is en dat ek dit nie vantevore in die geheel of gedeeltelik by enige universiteit ter verkryging van 'n graad voorgelê het nie.



Geteken:.....

Te: .....

Datum:.....

---



---

## OPSOMMING

---

**B**esluite in konstruksieprojekte word gereeld gekenmerk deur onsekerheid. Een van die bronne van onsekerheid is die eenheidariewe wat tydens detailkosteberekeninge gebruik word. Analitiese tegnieke is beskikbaar om die onsekerhede te modelleer. Inligting word egter benodig om die onsekerhede mee te kwantifiseer. Die voorkeur bron van inligting is historiese data, maar weens onbesikbaarheid en lae akkuraatheid daarvan striem dit die aanwending van die analitiese tegnieke. Veranderde omstandighede in die konstruksiebedryf noodsaak egter die verbetering van databronne en aanwending daarvan op 'n gestruktureerde, omvattende en geïntegreerde wyse.

Die doel van hierdie studie is om te bepaal hoe besluitneming verbeter kan word deur inligtingversameling en benutting van relevante data vir konstruksiekosteberekeninge te verbeter. Dit word gedoen deur die onderskeie aspekte wat deel vorm van 'n inligtingstelsel, nl. die datavloei, datastoring en aanwending te ondersoek en te toon hoe geïntegreerde ontwikkeling van elk van hierdie aspekte met behulp van prosesmodellering, databasisontwikkeling en kwantitatiewe risiko-analise, 'n bydrae kan maak tot verbeterde besluitneming. Dit word gedoen deur twee verskillende omgewings, nl. padbou en dambou, as moontlike toepassingsomgewings te gebruik vir onderskeidelik die herontwikkeling en ontwikkeling van so 'n inligtingstelsel vir besluitneming.

Tydens die ondersoeke is dit gevind dat prosesmodellering 'n waardevolle bydrae kan lewer tot die verbetering van datavloei in die gefragmenteerde konstruksie-omgewing, maar dat tekortkominge in die modelleringsprogramme bestaan. 'n Metode, gebaseer op versameling- en grafiekteorie, is ontwikkel om hierdie aspekte aan te spreek. Dit is ook gevind dat databasisontwikkeling die geleentheid bied om die kwaliteit van elektronies gestoorde databronne te

---

verbeter. 'n Datamodel wat omvattend genoeg is om al die inligting wat met konstruksiekoste, spesifisering en regsdokumente verbandhou te stoor, is ontwikkel met die nodige fout-onderskeppingsmeganismes. Die nut van probabilistiese tegnieke soos Monte Carlo-simulasie en die Limietstaat Kostefunksie om die globale onsekerheid van projekte se kosteberamings te modelleer en die effek van korrelasie op die resultate in ag te neem, is ook getoon.

Die inligting afkomstig van historiese data-ontleding, naas die aanwending in die probabilistiese tegnieke, is gebruik om die eienskappe soos impak, veranderlikheid en korrelasiemeganismes, te bepaal. Dit is in die studie gevind dat 'n tekort aan beskrywende inligting, hetsy op projek- of itemvlak, die oorwegende bron van veranderlikheid is nadat foutiewe data verwyder is. Om egter die beskrywende inligting te bepaal wat 'n bydrae kan maak tot die verlaging van hierdie veranderlikheid is 'n metode ontwikkel met behulp van invloedfaktore en eenheidtariefkorrelasiemeganismes om die mees omvangryke faktore te bepaal vir onderskepping. Dit het aanleiding gegee tot die ontwikkeling van 'n ideale data-aanwendingmodel waarin al die relevante inligting vir die verskillende fases van kosteberekening, nl. langtermyn, konseptueel en detail, onderskep en gebruik word.

Gebaseer op dit wat gevind is, is 'n generiese inligtingstelsel ontwikkeling-model voorgestel wat die interaksie van al die aspekte wat ondersoek is ten opsigte van mekaar toon. Twee van die hoofkenmerke van hierdie model is die selfvoerende proses en die inkrementele ontwikkeling van die stelsel. Op hierdie ontwikkelingmodel is 'n ontwikkelingproses gebou wat gebruik kan word vir die gestruktureerde, omvattende en geïntegreerde ontwikkeling/herontwikkeling van 'n inligtingstelsel vir kosteberekeninge.

---

---

## ABSTRACT

---

**D**ecisions affecting construction projects are often characterised by uncertainty. One of the sources of this uncertainty is the unit costs used in detail cost estimates. Analytical techniques are available to model these uncertainties, but information is required to quantify it. Historical data is the preferred source of information, but due to unavailability and unreliability it hampers the application of the analytical techniques. Changes in the construction environment necessitate the improvement of data sources and the utilisation thereof in a structured, comprehensive and integrated manner.

The purpose of this study is to determine how decision-making can be improved by enhanced information capturing and usage of relevant data for improved cost calculations. This is achieved by investigating the various components of an information system, viz. dataflow, data source, data application, and shown how an integrated development of each of these components with processmodelling, database development and quantitative risk analysis can lead to improved decision-making. This is achieved by the conceptual redevelopment and development of information systems for decision support in two different construction environments, viz. road and dam construction, respectively.

During the research it was found that processmodelling can contribute to the improvement of dataflow in the fragmented construction environment, but that shortcomings exist in the processmodelling tools. A method, based on set and graph theory, was developed to address it. It was also found that database development provides the opportunity to improve the quality of data stored in electronic format. A comprehensive database model to store all information relevant to construction costs, specifications and legal documents with the

---

required error checking mechanisms was developed. The benefit of modelling the global uncertainty of project cost estimates with probabilistic techniques such as Monte Carlo simulation and the Limit State Cost Function, whilst including the effects of correlation, is also shown.

The information gained from historical data-analysis, besides the application in the probabilistic techniques, was used to identify properties such as impact, variability and correlation. It was found in the study that a shortage of descriptive information, whether it be at project or item level, is predominately the source of variability in unit cost data after errors have been removed. In order to identify the information that would contribute to the lowering of variability a method was developed by using influence factors and unit cost correlation mechanisms to identify the most influential factors for data capturing. This led to the development of an ideal data-application model in which all the relevant data used in the different cost-estimation phases, viz. long term, conceptual and detail, would be captured and used.

Based on what was found, a generic information system development model was proposed that indicated the interaction of all the aspects that were investigated. Two of the model's foremost properties are its self-triggering process and incremental development (improvement). On top of this development model a development process was designed that can be used for the structured, comprehensive and integrated development/redevelopment of an information system for cost calculation.



---

## DIE BEITELTJIE

---

Ek kry 'n klein klein beitelkje,  
 ek tik hom en hy klink;  
 toe slyp ek en ek slyp hom  
 totdat hy klink en blink.

die donker naat loop deur my land  
 en kloof hom wortel toe –  
 só moet 'n beitel slaan  
 wat beitel is, of hoé?

Ek sit 'n klippie op 'n rots:  
 - mens moet jou vergewis:  
 'n beitel moet kan klip breek  
 as hy 'n beitel is –



Dan, met twee goue afgronde  
 val die planeet aan twee  
 en oor die kranse, kokend,  
 verdwyn die vlak groen see

ek slaat hom met my beitelkje  
 en dié was sterk genoeg;  
 daar spring die klippie stukkend  
 so skoon soos langs 'n voeg:

en op die dag sien ek die nag  
 daar anderkant gaan oop  
 met 'n bars wat van my beitel af  
 dwarsdeur die sterre loop.

toe, onder my tien vingers bars  
 die grys rots middeldeur  
 en langs my voete voel ek  
 die sagte aarde skeur,

---

- NP van Wyk Louw



## BEDANKINGS

---

In enige studie soos hierdie is daar 'n groot groep mense wie op 'n direkte of indirekte wyse 'n bydrae tot dit gemaak het. Op die akademiese gebied wil ek graag my dank betuig aan:

- Prof Johan Retief, my promotor, wie se kennis en insig, maar veral sy geduld my na ses jaar nog steeds verstom. Sonder sy insette sou hierdie studie nie moontlik gewees het nie;
- Prof Peter Dunaiski, my departementshoof, wie 'n leidende rol gespeel het in die vestiging van Siviele Ingenieursinformatika aan die Universiteit van Stellenbosch;
- Prof. Dr.-Ing. Wolfgang Huhnt, wie my aan die wêreld van prosesmodellering bekend gestel het.

In my persoonlike lewe het die volgende persone 'n groot bydrae gelewer deur hulle ondersteuning vir my om met hierdie studie te kon begin en af te handel. My innige danke aan hulle:

- Koot en Helené, my ouers, wie se onbaatsugtige opofferinge vir my en stimulerende blootstelling wat hulle vir my gegee het, geleenthede gebied het wat hulle nie gehad het nie;
  - Hans, my lewensmaat, wat met sy aanmoediging en vertoue in my verseker het dat hierdie studie afgehandel is;
  - Dr. Chris Oosthuizen, my “baas” en soms “tweede pa”, wie 'n omgewing geskep het waarbinne ek kon werk en studeer;
-

- Prof. Albert en Mev. Elbie Rooseboom, vir hulle ondersteuning en besorgdheid;
- Al my vriende en vriendinne wie ek in die afgelope ses jaar gereeld afgeskeep het.

Hierdie studie sou ook nie moontlik gewees sonder die finansiële ondersteuning wat ek ontvang het nie. My dank aan:

- Die Charl van der Merwe-stigting;
- Die Afdeling Struktuur Ingenieurswese en Siviele Ingenieursinformatika, Universiteit van Stellenbosch;
- Die Bauhaus Universiteit, Weimar;
- Die Departement van Waterwese en Bosbou.

Laastens wil ek graag die personeel van die Departement Siviele Ingenieurswese bedank vir die bydrae wat hulle gemaak het om 'n aangename omgewing te skep waarbinne ek kon werk en studeer. Die Universiteit van Stellenbosch sal altyd vir my 'n sagte plekkie hê danksy hulle.



---

## INHOUDSOPGAWE

<b>VERKLARING</b> .....	<b>I</b>
<b>OPSOMMING</b> .....	<b>III</b>
<b>ABSTRACT</b> .....	<b>V</b>
<b>BEDANKINGS</b> .....	<b>VIII</b>
<b>LYS VAN FIGURE</b> .....	<b>XVI</b>
<b>LYS VAN TABELLE</b> .....	<b>XXI</b>
<b>LYS VAN AFKORTINGS</b> .....	<b>XXII</b>
<b>HOOFSTUK 1</b>	
<b>INLEIDING</b> .....	<b>1</b>
1.1 <b>BESLUITNEMING</b> .....	<b>1</b>
1.2 <b>BEHOEFTE AAN HISTORIESE DATA</b> .....	<b>3</b>
1.3 <b>INLIGTINGSTELSEL</b> .....	<b>4</b>
1.4 <b>DOEL VAN STUDIE</b> .....	<b>5</b>
1.5 <b>METODOLOGIE</b> .....	<b>5</b>
1.5.1 <b>Bestaande Omgewing</b> .....	<b>6</b>
1.5.2 <b>Nuwe Omgewing</b> .....	<b>6</b>
1.5.3 <b>Geïntegreerde Besluitsteunstelsel</b> .....	<b>6</b>
1.6 <b>UITLEG VAN STUDIE</b> .....	<b>7</b>
<b>HOOFSTUK 2</b>	
<b>PROSESMODELONTWIKKELING</b> .....	<b>10</b>
2.1 <b>PROSESMODELLERING</b> .....	<b>10</b>
2.1.1 <b>Doel van Prosesmodelle</b> .....	<b>11</b>
2.1.2 <b>Prosesmodellering in Siviele Ingenieurswese</b> .....	<b>11</b>
2.1.3 <b>Kwantitatiewe Risikobestuur in Prosesmodellering</b> .....	<b>13</b>
2.1.4 <b>Ontwikkeling van 'n Prosesmodel</b> .....	<b>15</b>

---

2.1.5	Tekortkomings van Modelleringsgereedskap .....	18
2.2	ONTWIKKELING VAN 'N ANALISERINGSPROGRAM .....	19
2.2.1	Wiskundige Agtergrond.....	20
2.2.2	Implementering .....	29
2.3	AANWENDING VAN ANALISERINGSPROGRAM .....	36
<b>HOOFSTUK 3</b>		
	<b>DATABASISONTWIKKELING .....</b>	<b>38</b>
3.1	INLEIDING .....	38
3.1.1	Databasis .....	39
3.1.2	Databasis Bestuurstelsel .....	39
3.1.3	Databasis Stelsel .....	40
3.2	DATABASIS LEWENSIKLUS .....	41
3.3	TEGNIESE AGTERGROND .....	44
3.3.1	Basiese Begrippe .....	44
3.3.2	Sleutels .....	45
3.3.3	EV-diagramme .....	46
3.3.4	Normaalvorms.....	48
3.3.5	Sequel .....	51
3.4	EVALUASIE .....	51
3.4.1	Spesifisering.....	52
3.4.2	Data.....	53
3.4.3	Databasismodel .....	54
3.5	ONTWIKKELING VAN NUWE DATAMODEL .....	55
3.5.1	Konseptuele Ontwikkeling.....	55
3.5.2	Normalisering van Huidige Model.....	61
3.5.3	Logiese Ontwerp.....	63
3.6	OORDRAG VAN DATA NA NUWE MODEL .....	67
3.7	TOETSING VAN OORGEDRAGTE DATA.....	68
3.8	UITBREIDING VAN DIE NUWE MODEL.....	69
3.9	VOORDELE VAN NUWE ONTWERP .....	71
3.9.1	Optimale Ruimte Benutting .....	72
3.9.2	Integriteit .....	72
3.9.3	Uitbreiding Vergemaklik .....	72
<b>HOOFSTUK 4</b>		
	<b>KWANTITATIEWE RISIKO-ANALISE .....</b>	<b>73</b>
4.1	RISIKO EN ONSEKERHEID .....	73
4.2	RISIKOBESTUUR.....	75
4.3	KOSTEBERAMING .....	76
4.4	MONTE CARLO-SIMULASIE TEGNIEK .....	79
4.4.1	Ontwikkeling van 'n Monte Carlo-Simulasiemodel.....	80
4.4.2	Kommentaar oor Monte Carlo-Simulasie.....	82
4.5	LIMIETSTAAT KOSTEFUNKSIE .....	87

4.5.1	Limietstaatteorie.....	87
4.5.2	Vorms van Limietstaat Kostefunksie.....	90
4.5.3	Kommentaar oor LSKF-Tegniek .....	94
4.6	SKYNINTELLIGENSIE .....	96
4.6.1	Ontwikkeling van 'n Neurale netwerk.....	97
4.6.2	Kommentaar oor Neurale Netwerk .....	100
4.7	VERSKILLE EN INTEGRERING.....	101
<b>HOOFSTUK 5</b>		
	<b>DATA-ONTLEDING .....</b>	<b>103</b>
5.1	IDENTIFISERING VAN BEDUIDENDE ITEMS.....	103
5.2	FOUTE IN DATA .....	110
5.3	TYDVERANDERLIKHEID VAN DATA.....	114
5.4	BENADERDE STATISTIESE BESKRYWINGS.....	120
5.5	KORRELASIE .....	125
5.6	EIENSKAPPE VAN DATA .....	131
5.6.1	Noemenswaardige Bydrae.....	131
5.6.2	Veranderlikheid .....	132
5.6.3	Korrelasie .....	133
5.7	VOORBEELD .....	134
<b>HOOFSTUK 6</b>		
	<b>UITBREIDING VAN DATA-ONDERSKEPPING .....</b>	<b>135</b>
6.1	PROJEKFAKTORE EN ITEMSPESIFIKASIES.....	135
6.2	MODELLE VAN DATA-AANWENDINGSG.....	137
6.2.1	Bestaande Model .....	137
6.2.2	Ideale Model.....	140
6.3	INVLOEDSDIAGRAMME .....	145
6.4	KOSTEBEREKENINGMODELLE.....	146
6.4.1	Langtermynkostemodel.....	147
6.4.2	Konseptuele Kostemodel .....	148
6.4.3	Detailkostemodel .....	149
6.4.4	Uitkomskostemodel.....	151
6.5	LANGTERMYNKOSTEMODEL INVLOEDFAKTORE.....	151
6.6	KONSEPTUELE KOSTEMODEL INVLOEDFAKTORE .....	154
6.6.1	Kwantitatiewe Faktore.....	155
6.6.2	Kwalitatiewe Faktore.....	156
6.7	DETAILKOSTEMODEL INVLOEDFAKTORE.....	158
6.7.1	Kontrakteurperspektief.....	158
6.7.2	Raadgewende Ingenieurperspektief .....	161
6.8	INVLOEDFAKTORE .....	163
6.8.1	Identifisering.....	163
6.8.2	Bepaling van Beduidendste Invloedfaktore .....	165
6.8.3	Nuwe Projektoepassing .....	170



6.8.4 Voorstelling en Uitbreiding van Databasismodel .....170  
 6.8.5 Aanwending .....173

**HOOFSTUK 7**

**ONTWIKKELING VAN 'N NUWE DATABRON ..... 176**

7.1 ONTWIKKELINGSVLAK .....176  
 7.1.1 Stapsgewyse Ontwikkeling .....176  
 7.1.2 Struktuur van Omgewing .....178  
 7.2 ONTWIKKELINGSTRUKTUUR .....179  
 7.3 BEHOEFTES .....179  
 7.3.1 Projektipies .....180  
 7.3.2 Vlak van Detail en Volledigheid .....180  
 7.3.3 Kwantifisering van Risiko .....181  
 7.3.4 Betroubare Bron.....182  
 7.4 ONTWIKKELINGASPEKTE .....183  
 7.4.1 Datavloei .....183  
 7.4.2 Databasis .....184  
 7.4.3 Karakterisering van Omgewing.....186  
 7.5 GEBREK AAN STANDAARDITEMS.....187  
 7.5.1 Items teenoor Aktiwiteite .....187  
 7.5.2 Generiese Items.....189  
 7.5.3 Hoëvlak Kosteberekeninge .....190  
 7.6 GROEI VAN DATABRON .....191  
 7.6.1 Beginfase van Databrongroei .....191  
 7.6.2 Selfvoerende Fase van Databrongroei .....192  
 7.7 VERDERE INKREMENTELE ONTWIKKELING .....192

**HOOFSTUK 8**

**ONTWIKKELING VAN 'N INLIGTINGSTELSEL ..... 194**

8.1 BEHOEFTE AAN VERBETERDE BESLUITNEMING .....194  
 8.2 GENERIESE INLIGTINGSTELSEL ONTWIKKELINGMODEL .....195  
 8.3 KONSTRUKSIEKOSTE-INLIGTING .....197  
 8.3.1 Probabilistiese Tegnieke .....197  
 8.3.2 Skynintelligensietegnieke.....197  
 8.4 DATABRON .....197  
 8.4.1 Data-ontleding.....198  
 8.4.2 Selfvoerende Proses.....198  
 8.4.3 Inkrementele Verbetering.....199  
 8.5 BETROUBAARHEID VAN BRON .....200  
 8.5.1 Datavloei .....200  
 8.5.2 Datastoor.....201  
 8.5.3 Beskrywende Inligting .....201  
 8.5.4 Spesifisering.....201  
 8.6 GEREEDSKAP BESKIKBAAR .....201

8.6.1	Prosesmodellering .....	201
8.6.2	Databasisontwikkeling .....	202
8.6.3	Invloedfaktore.....	202
8.6.4	Generiese Items.....	202
8.6.5	Geïntegreerde Ontwikkeling .....	203
8.7	KOSTE VAN 'N VERBETERDE INLIGTINGSTELSEL .....	203
8.8	ONTWIKKELINGPROSES .....	204
8.8.1	Behoeftebepaling .....	205
8.8.2	Oorhoofse Prosesmodel .....	205
8.8.3	Opstel van Aanwendingmodel .....	206
8.8.4	Ontleding van Databronne .....	206
8.8.5	Invloedfaktore.....	206
8.8.6	Opdatering van Prosesmodel .....	207
8.8.7	Databasisontwikkeling .....	207
8.8.8	Aanwendingstegnieke.....	208
8.8.9	Herevaluering van Stelsel .....	208
<b>HOOFSTUK 9</b>		
<b>SAMEVATTING EN GEVOLGTREKKINGS .....</b>		<b>209</b>
9.1	BESTAANDE DATABASIS .....	209
9.1.1	Prosesmodellering .....	210
9.1.2	Databasisontwikkeling .....	210
9.1.3	Kwantitatiewe Risiko-analise.....	211
9.1.4	Data-ontleding.....	212
9.1.5	Uitbreiding van Data-onderskepping .....	213
9.2	NUWE OMGEWING .....	213
9.2.1	Damkonstruksie .....	214
9.2.2	Generiese Ontwikkelingmodel .....	215
9.3	GEVOLGTREKKINGS .....	215
9.4	VERDER STUDIE .....	216
<b>HOOFSTUK 10</b>		
<b>VERWYSINGS.....</b>		<b>218</b>
<b>AANHANGSEL A</b>		
<b>PADKONSTRUKSIEPROSEMODEL .....</b>		<b>A-1</b>
<b>AANHANGSEL B</b>		
<b>IMPLEMENTERING VAN TOETSPROGRAM .....</b>		<b>B-1</b>
<b>AANHANGSEL C</b>		
<b>DATA-OORDRAG.....</b>		<b>C-1</b>
<b>AANHANGSEL D</b>		
<b>TOETSING VAN OORGEDRAGDE DATA .....</b>		<b>D-1</b>
<b>AANHANGSEL E</b>		
<b>VERDELINGS .....</b>		<b>E-1</b>
<b>AANHANGSEL F</b>		

<b>MCS-TEGNIK</b> .....	<b>F-1</b>
<b>AANHANGSEL G</b>	
<b>LSKF-TEGNIK</b> .....	<b>G-1</b>
<b>AANHANGSEL H</b>	
<b>NEURALE NETWERK</b> .....	<b>H-1</b>
<b>AANHANGSEL I</b>	
<b>KUMULATIEWE BYDRAE TOT PROJEKKOSTE</b> .....	<b>I-1</b>
<b>AANHANGSEL J</b>	
<b>IDENTIFISERING VAN BEDUIDENDE ITEMS</b> .....	<b>J-1</b>
<b>AANHANGSEL K</b>	
<b>BEPALING VAN UITSKIETER PROJEKTE</b> .....	<b>K-1</b>
<b>AANHANGSEL L</b>	
<b>INFLASIE-EFFEK</b> .....	<b>L-1</b>
<b>AANHANGSEL M</b>	
<b>INFLASIE-AANPASSINGS</b> .....	<b>M-1</b>
<b>AANHANGSEL N</b>	
<b>VERWERKING VAN RESULTATE</b> .....	<b>N-1</b>
<b>AANHANGSEL O</b>	
<b>KORRELASIE BEREKENINGE</b> .....	<b>O-1</b>
<b>AANHANGSEL P</b>	
<b>INVLOEDDIAGRAMME</b> .....	<b>P-1</b>
<b>AANHANGSEL Q</b>	
<b>BEDUIDENDSTE INVLOEDFAKTORE</b> .....	<b>Q-1</b>
<b>AANHANGSEL R</b>	
<b>KOMPAKTE SKYF</b> .....	<b>R-1</b>





## LYS VAN FIGURE

FIGUUR 1-1: UITLEG VAN EERSTE GEDEELTE VAN STUDIE .....	7
FIGUUR 2-1: INTERAKSIES TUSSEN VERSKILLENDE GROEPE INLIGTING .....	14
FIGUUR 2-2: UITTREKSEL VAN PROSESMODEL .....	16
FIGUUR 2-3: GENERIESE ONTWERP EN KONSTRUKSIE PROSESPROTOKOL .....	17
FIGUUR 2-4: ONGERIGTE GRAFIEK .....	24
FIGUUR 2-5: GERIGTE GRAFIEK.....	24
FIGUUR 2-6: GERIGTE GRAFIEK MET 'N LUS BY $E_3$ EN PARALLELE BOË $E_8$ EN $E_9$ .....	25
FIGUUR 2-7: GERIGTE GRAFIEK MET TWEE SIKLIESE KOMPONENTE .....	25
FIGUUR 2-8: BIPARTIETE GRAFIEK MET VERSAMELINGS $V_1$ EN $V_2$ .....	26
FIGUUR 2-9: GRAFIEK MET STERK SAMEHANGENDE KOMPONENT $\{V_2, V_3, V_4, V_5\}$ .....	26
FIGUUR 2-10: GRAFIESE AFVOER: AANTAL GROEPE BETROKKE .....	30
FIGUUR 2-11: GRAFIESE AFVOER: GROEP BETROKKENHEID.....	30
FIGUUR 2-12: GRAFIESE AFVOER: UITVOERING VAN 'N AKTIWITEIT .....	30
FIGUUR 2-13: GRAFIESE AFVOER: WATTER AKTIWITEITE 'N DATA GEBRUIK .....	32
FIGUUR 2-14: GRAFIESE AFVOER: DATA DEUR 'N AKTIWITEIT GEBRUIK .....	32
FIGUUR 2-15: GRAFIESE AFVOER: AANTAL DATA BETROKKE BY AKTIWITEITE.....	32
FIGUUR 2-16: GRAFIESE AFVOER: WIE DATA GEBRUIK MET TOEGANG .....	34
FIGUUR 2-17: GRAFIESE AFVOER: DATA DEUR 'N GROEP GEBRUIK MET TOEGANG ..	35
FIGUUR 2-18: GRAFIESE AFVOER: SIKLUSSE EN PAAIE .....	36
FIGUUR 3-1: VOORSTELLING VAN 'N DB .....	39
FIGUUR 3-2: VOORSTELLING VAN 'N DBBS .....	40
FIGUUR 3-3: VOORSTELLING VAN 'N DBS .....	41
FIGUUR 3-4: DATABASIS LEWENSIKLUS.....	42
FIGUUR 3-5: PROJEK RELASIE .....	45
FIGUUR 3-6: VERWANTSKAP TUSSEN TWEE TABELLE .....	46
FIGUUR 3-7: VOORSTELLING VAN 'N ENTITEIT MET ATTRIBUTE .....	47
FIGUUR 3-8: VOORSTELLING VAN VERWANTSKAPPE TUSSEN ENTITEITE .....	47

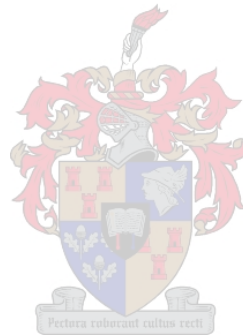
FIGUUR 3-9: KONNEKTEERDHEID TUSSEN TWEE TABELLE.....	48
FIGUUR 3-10: TABEL KONTRAKTEUR VERANDER NA 1NV .....	49
FIGUUR 3-11: DRIE VOLLEDIGE FUNKSIONELE AFHANKLIKHED E IN EEN RELASIE .....	49
FIGUUR 3-12: RELASIES VERDEEL OM TE VOLDOEN AAN 2NV.....	50
FIGUUR 3-13: RELASIE VERDEEL OM TE VOLDOEN AAN 3NV.....	50
FIGUUR 3-14: BESTAANDE DATABASISMODEL .....	54
FIGUUR 3-15: PROJEK-ENTITEIT MET ATTRIBUTE .....	55
FIGUUR 3-16: ITEM-ENTITEIT MET ATTRIBUTE.....	56
FIGUUR 3-17: SWAK ENTITEIT TENDERITEM MET ATTRIBUTE .....	56
FIGUUR 3-18: KONTRAKTEUR- EN RAADGEWEND-ENTITEIT MET ATTRIBUTE...	56
FIGUUR 3-19: VERWANTSAP TUSSEN PROJEK EN ITEM .....	57
FIGUUR 3-20: PROJEK EN ITEM VERWANTSAP IS M:N.....	57
FIGUUR 3-21: VERWANTSAP TUSSEN KONTRAKTEUR EN PROJEK .....	57
FIGUUR 3-22: VERWANTSAP TUSSEN RAADGEWEND EN PROJEK .....	57
FIGUUR 3-23: VERWANTSAP TUSSEN KONTRAKTEUR, PROJEK, ITEM EN TENDERITEM .....	58
FIGUUR 3-24: SAAMGESTELDE ENTITEIT TENDER EN L_V_H .....	59
FIGUUR 3-25: PROJEK EN ITEM MET SAMEGESTELDE ENTITEIT L_V_H .....	59
FIGUUR 3-26: VERBETERDE FORMULERING VAN DIE VERWANTSAPPE.....	60
FIGUUR 3-27: L_V_H-ENTITEIT MET ATTRIBUTE .....	60
FIGUUR 3-28: TENDER-ENTITEIT MET ATTRIBUTE .....	61
FIGUUR 3-29: OPSOEK TABELLE VIR L_V_H, ITEM EN PROJEK.....	61
FIGUUR 3-30: RELASIE VERDEEL OM TE VOLDOEN AAN 2NV.....	62
FIGUUR 3-31: VOORGESTELDE DATAMODEL IN BCNF .....	63
FIGUUR 3-32: VERWANTSAPMODEL VAN VOORGESTELDE DATABASIS.....	67
FIGUUR 3-33: VERDELING VAN CONTRACT_No NA TWEE TABELLE .....	68
FIGUUR 3-34: KONSEPTUELE UITGEBREIDE MODEL.....	71
FIGUUR 4-1: RISIKO WEENS ONSEKERHEID .....	74
FIGUUR 4-2: STAPPE IN DIE ONTWIKKELING VAN 'N MCS-MODEL .....	80
FIGUUR 4-3: EFFEK VAN VERSKILLENDE VERDELINGS OP RESULTATE .....	84
FIGUUR 4-4: EFFEK VAN KORRELASIE OP RESULTATE .....	87
FIGUUR 4-5: RUIMTE VAN GEREDUSEERDE VERANDERLIKES X'EN Y'.....	89
FIGUUR 4-6: VERSKILLE IN RESULTATE TUSSEN MCS EN LSKF.....	92
FIGUUR 4-7: VERSKILLE IN RESULTATE TUSSEN VERSKILLENDE LSKF-TEGNIKE .....	94
FIGUUR 4-8: KOSTE VIR 'N GEGEWE BETROUBAARHEID MET KORRELASIE .....	95
FIGUUR 4-9: KOSTE VIR 'N GEGEWE BETROUBAARHEID SONDER KORRELASIE .....	96
FIGUUR 4-10: SKEMATIESE ONTWIKKELING VAN 'N NEURALE NETWERK .....	98
FIGUUR 5-1: ONDERSKEPPINGSTRATEGIE MATRIKS.....	104
FIGUUR 5-2: BYDRAE VAN ITEMS TOT PROJEKKOSTE .....	106
FIGUUR 5-3: AANTAL ITEMS SE SENSITIEWEIT VIR VERANDERINGE IN DIE KRITERIA	108
FIGUUR 5-4: ITEM 16.02 MET FOUTIEWE DATAPUNT.....	110
FIGUUR 5-5: KOEFFISIËNT VAN VARIASIE VERLAGING DEUR FOUT VERWYDERING.	111

FIGUUR 5-6: GRENSE VAN UITSKIETER GEBIEDE .....	112
FIGUUR 5-7: INFLASIE-EFFEK OP UNIFORMVERDEELDE DATA .....	115
FIGUUR 5-8: INFLASIE-EFFEK OP NORMAALVERDEELDE DATA .....	115
FIGUUR 5-9: OORSPRONKLIKE EN INDEKSAANGEPASTE DATA - ITEM 33.01(A)(I)....	117
FIGUUR 5-10: OORSPRONKLIKE EN INDEKSAANGEPASTE DATA - ITEM 45.01(B)(IV)	118
FIGUUR 5-11: TENDENSAANGEPASTE DATA VAN ITEM 33.01(A)(I) .....	119
FIGUUR 5-12: TENDENSAANGEPASTE DATA VAN ITEM 45.01(B)(IV) .....	120
FIGUUR 5-13: GEMIDDELDE JAARLIKSE INFLASIE VAN DIE BEDUIDENDE ITEMS .....	120
FIGUUR 5-14: TOENAME IN EENHEIDTARIEWE VAN BEDUIDENDE ITEMS OOR 'N 10- JAAR PERIODE .....	123
FIGUUR 5-15: ITEM 33.03(B) SE KOSTE AANPASSINGS .....	124
FIGUUR 5-16: KOËFFISIËNTE VAN VARIASIE .....	124
FIGUUR 5-17: KORRELASIEMEGANISMES .....	127
FIGUUR 5-18: KORRELASIEMATRIKS VAN EENHEIDTARIEF .....	128
FIGUUR 5-19: KORRELASIEMATRIKS VAN HOEVEELHEID .....	128
FIGUUR 5-20: KORRELASIEMATRIKS VAN HOEVEELHEID EN EENHEIDTARIEF .....	129
FIGUUR 6-1: BESTAANDE DATA-AANWENDINGSMODEL .....	138
FIGUUR 6-2: IDEALE DATA-AANWENDINGSMODEL .....	140
FIGUUR 6-3: EENVOUDIGE VOORBEELD VAN 'N PADOPGRADERING .....	146
FIGUUR 6-4: INVLOEDFAKTORE VIR LANGTERMYN KONSTRUKSIEKOSTES .....	148
FIGUUR 6-5: INVLOEDFAKTORE VIR KONSEPTUELE MODEL .....	149
FIGUUR 6-6: INVLOEDFAKTORE VIR DIE DETAIL MODEL VOOR TENDER .....	150
FIGUUR 6-7: INVLOEDFAKTORE VIR DETAIL MODEL NA TENDER .....	151
FIGUUR 6-8: INFLASIE-INVLOEDFAKTORE VAN LANGTERMYNKOSTEMODEL .....	152
FIGUUR 6-9: PROJEKEIENSKAP INVLOEDFAKTORE VAN LANGTERMYNKOSTEMODEL	153
FIGUUR 6-10: KONSTRUKSIE-OMGEWING INVLOEDFAKTORE VAN LANGTERMYNKOSTEMODEL .....	154
FIGUUR 6-11: KWALITATIEWE INVLOEDFAKTORE VAN KONSEPTUELE KOSTEMODEL	156
FIGUUR 6-12: KWALITATIEWE INVLOEDFAKTORE VAN KONSEPTUELE MODEL .....	158
FIGUUR 6-13: GENERIESE INVLOEDSMODEL VAN KONTRAKTEUR .....	159
FIGUUR 6-14: INVLOEDSMODEL: TIPIESE ITEM VANUIT KONTRAKTEURPERSPEKTIEF	161
FIGUUR 6-15: GENERIESE INVLOEDSMODEL VAN RAADGEWENDE INGENIEUR .....	162
FIGUUR 6-16: INVLOEDSMODEL: ITEM 31.01 VANUIT RAADGEWENDE INGENIEURPERSPEKTIEF .....	163
FIGUUR 6-17: BEPALING VAN AKTIWITEITE SE KOSTEBYDRAE PER ITEM .....	167
FIGUUR 6-18: GESKATTE INVLOEDFAKTORSTERKTE OP AKTIWITEITE .....	168
FIGUUR 6-19: INVLOEDFAKTORSTERKTE OP ITEMS .....	168
FIGUUR 6-20: BEREKENDE STERKTE-VAN-VERBAND .....	169
FIGUUR 6-21: VERSKIL TUSSEN BEREKENDE VERBANDSTERKTE EN KORRELASIE ..	169
FIGUUR 6-22: SOMPRODUK VAN ITEMKOSTES EN INVLOEDFAKTORSTERKTE .....	170
FIGUUR 6-23: RELATIEWE IMPAKTE VAN INVLOEDFAKTORE .....	170
FIGUUR 6-24: AKTIWITEIT- EN INVLOED-ENTITEIT MET ATTRIBUTE .....	171

FIGUUR 6-25: SAAMGESTELDE ENTITEITE BESTAAN UIT EN BEÏNVLOED .....	171
FIGUUR 6-26: OPSOEKTABLE VIR BESTAAN UIT EN BEÏNVLOED .....	172
FIGUUR 6-27: UITGEBREIDE DATAMODEL .....	173
FIGUUR 7-1: ONTWIKKELINGSMODEL VIR 'N NUWE OMGEWING .....	179
FIGUUR 7-2: TIPIESE ITEM IN 'N SPESIFIKASIE .....	187
FIGUUR 7-3: TIPIESE AKTIWITEIT SE KOSTES .....	188
FIGUUR 7-4: EENHEIDTARIEF VARIASIE WEENS VERSKIL IN SPESIFIKASIE .....	189
FIGUUR 7-5: INVLOEDSMODEL VAN 'N GENERIESE ITEM .....	190
FIGUUR 8-1: GENERIESE INLIGTINGSTELSEL ONTWIKKELINGSMODEL .....	196
FIGUUR 8-2: TOEPASSING VAN DATA-ONTLEDINGS .....	198
FIGUUR 8-3: SELFVOERENDE PROSES TUSSEN DATABRON EN AANWENDING .....	199
FIGUUR 8-4: INKREMENTELE VERBETERING VAN DATABRON .....	200
FIGUUR 8-5: OPTIMALE VERBETERING BESLUIT .....	204
FIGUUR 8-6: ONTWIKKELINGPROSES VAN INLIGTINGSTELSEL .....	205
FIGUUR A-1: TIPOGRAMME IN 'N PROSESMODEL .....	A-1
FIGUUR A-2: PROSESMODEL .....	A-2
FIGUUR A-3: PROSESMODEL (VERVOLG) .....	A-3
FIGUUR A-4: PROSESMODEL (VERVOLG) .....	A-4
FIGUUR A-5: PROSESMODEL (VERVOLG) .....	A-5
FIGUUR B-1: KLASDIAGRAM VAN DIE PROGRAM CORRECTNESS .....	B-2
FIGUUR E-1: GUMBELVERDELING .....	E-2
FIGUUR E-2: LOG-NORMAALVERDELING .....	E-3
FIGUUR E-3: DRIEHOEKVERDELING .....	E-4
FIGUUR E-4: GAMMAVERDELING .....	E-5
FIGUUR E-5: BETAVERDELING .....	E-6
FIGUUR E-6: CHI-KWADRAATVERDELING .....	E-7
FIGUUR P-1: ITEM 31.01 - OORMAAT BOLAAG .....	P-1
FIGUUR P-2: ITEM 33.01(A)(I) - OPVULLINGS < 200MM UIT DEURGRAWINGS .....	P-1
FIGUUR P-3: ITEM 33.01(B)(I) - OPVULLINGS 200 - 500MM UIT DEURGRAWINGS .....	P-2
FIGUUR P-4: ITEM 33.01(C) - ROTSOPVULLING .....	P-2
FIGUUR P-5: ITEM 33.03(B) - EKSTRA OOR 33.01 HARD .....	P-2
FIGUUR P-6: ITEM 33.04(A) - WEGGOOI VAN DEURGRAWINGSMATERIAAL .....	P-2
FIGUUR P-7: ITEM 34.01(A) - GRUIS KEURLAAG 90% VAN GEWYSIGDE AASHTO- DIGTHEID .....	P-2
FIGUUR P-8: ITEM 34.01(C) GRUIS STUTLAAG 95% VAN GEWYSIGDE AASHTO- DIGTHEID .....	P-3
FIGUUR P-9: ITEM 34.01(E) GRUIS KROONLAAG 97% VAN GEWYSIGDE AASHTO- DIGTHEID .....	P-3
FIGUUR P-10: ITEM 34.01(F) - GRUIS SKOUERS 93% VAN GEWYSIGDE AASHTO- DIGTHEID .....	P-3
FIGUUR P-11: ITEM 34.01(G) GRUIS SLYTLAAG 93% VAN GEWYSIGDE AASHTO- DIGTHEID .....	P-4

FIGUUR P-12: ITEM 35.02(A) - STABILISERINGSMIDDEL PORTLANDSEMENT ..... P-4  
FIGUUR P-13: ITEM 35.02(B) - STABILISERINGSMIDDEL HOOGOONDSLAKSEMENT. P-4  
FIGUUR P-14: ITEM 35.02(C) - STABILISERINGSMIDDEL GEBLUSTE KALK ..... P-4  
FIGUUR P-15: ITEM 35.07(A) - ANIONIESE BITUMINEUSE STABILISERINGSMIDDEL. P-4  
FIGUUR P-16: ITEM 36.01(A) - G1 GEBREEKTE KLIP KROONLAAG KOMMERSIEEL P-5  
FIGUUR P-17: ITEM 36.01(C) - G2 GEBREEKTE KLIP KROONLAAG KOMMERSIEEL P-5

---



## LYS VAN TABELLE

TABEL 3–1: NORMAALVORMS VAN DIE TABELLE.....	62
TABEL 3–2: LOGIESE ONTWERP VIR PROJECTS .....	65
TABEL 3–3: LOGIESE ONTWERP VIR ITEMS.....	65
TABEL 3–4: LOGIESE ONTWERP VIR BILLOFQUANTITIES .....	66
TABEL 3–5: LOGIESE ONTWERP VIR TENDER.....	66
TABEL 3–6: LOGIESE ONTWERP VIR FIXED_PRICES .....	66
TABEL 3–7: LOGIESE ONTWERP VIR CONTRACTORS.....	66
TABEL 4–1: KLASSE VAN ONSEKERHEID .....	77
TABEL 4–2: VERSKILLENDE Tipes KONTINUE VERDELINGS .....	83
TABEL 4–3: VOORKOMS VAN VERDELINGS .....	84
TABEL 5–1: LYS VAN DIE CSRA-SERIES.....	106
TABEL 5–2: VOORKOMS VAN ITEMS MET VERSKILLENDE KRITERIA .....	107
TABEL 5–3: BEDUIDENDE ITEMS MET BESKRYWINGS .....	109
TABEL 5–4: VOORKOMS VAN UITSKIETER ITEMS EN PROJEKTE .....	113
TABEL 5–5: UITSKIETER PROJEKTE .....	113
TABEL 5–6: AANGEPASTE STATISTIEKE VAN UNIFORMVERDEELDE DATA .....	116
TABEL 5–7: AANGEPASTE STATISTIEKE VAN NORMAALVERDEELDE DATA .....	116
TABEL 5–8: EERSTE EN TWEDE MOMENTE VAN BEDUIDENDE ITEMS.....	122
TABEL 5–9: VOORKOMS VAN KORRELASIES TUSSEN BEDUIDENDE ITEMS (OORSPRONKLIK) .....	130
TABEL 5–10: VOORKOMS VAN KORRELASIES TUSSEN BEDUIDENDE ITEMS (TENDENSAANGEPAS).....	130
TABEL 6–1: KOSTEBYDRAE SKAAL.....	166
TABEL 6–2: INVLOEDFAKTORSTERKTE SKAAL.....	167
TABEL 6–3: INVOERDATA VIR NEURALE NETWERK.....	174
TABEL 6–4: AFVOER VAN NEURALE NETWERK .....	174

## LYS VAN AFKORTINGS

---

ANSI .....	AMERIKAANSE NASIONALE STANDAARDE INSTITUUT
BCNV .....	BOYCE-CODD NORMAALVORM
CSRA.....	KOMITEE VAN STAATS PADOWERHEDE
CIC.....	REKENAAR GEÏNTEGREERDE KONSTRUKSIE
DB.....	DATABASIS
DBBS .....	DATABASIS BESTUURSTELSEL
DBS .....	DATABASIS STELSEL
DWB .....	DEPARTEMENT VAN WATERWESE EN BOSBOU
EV .....	ENTITEITVERWANTSKAP
FA .....	FUNKSIONELE AFHANKLIKHEID
GB.....	GROSBEDRAG
GOKPP .....	GENERIESE ONTWERP EN KONSTRUKSIE PROSESPROKOL
IT.....	INFORMASIEGNOLOGIE
KS .....	KOMPAKTE SKYF
LSKF .....	LIMIETSTAAT KOSTEFUNKSIE
MCS.....	MONTE CARLO-SIMULASIE
NV .....	NORMAALVORM
SABS .....	SUID-AFRIKAANSE BURO VAN STANDAARDE
SANRAL .....	SUID-AFRIKAANSE NASIONALE PADAGENTSAP BEPERK

---

---

## HOOFSTUK 1

---

### INLEIDING

---

**B**esluite in konstruksieprojekte word gereeld gekenmerk deur groot onsekerhede. Analitiese tegnieke is beskikbaar om hierdie onsekerhede mee te modelleer, maar inligting word benodig om die onsekerhede te kwantifiseer. Beperkings, soos onbeskikbaarheid en lae akkuraatheid van inligting, striem egter die aanwending van hierdie tegnieke. Informasietegnologie hulpmiddels bied die geleentheid om hierdie beperkings aan te spreek.

In hierdie hoofstuk word die agtergrond van die studie geskets, nl. besluitneming tydens onsekerheid en die behoefte aan betroubare inligting om dit te kan ondersteun. Aangesien hierdie twee komponente afhanklik is van mekaar, word die interaksie tussen hierdie komponente ondersoek, asook die impak wat 'n veranderende informasietegnologie omgewing daarop het. In die laaste gedeelte van die hoofstuk word die doel van hierdie studie gestel en die metodologie wat gevolg is om dit te bereik word verduidelik.

---

#### 1.1 BESLUITNEMING

---

Gedurende 'n konstruksieprojek word 'n legio van besluite deur 'n verskeidenheid van persone met verskillende perspektiewe geneem. Hierdie besluite kan wissel van die kliënt wat 'n gegewe model aanvaar om die uitkoms, hetsy finansieel of maatskaplik, van 'n projek te modelleer, of die ontwerper wat moet besluit op 'n ontwerpfilosofie, of die kontrakteur wat 'n konstruksiemetode moet kies.

---



Alhoewel die aard en die effek van besluite dramaties van die een tot die ander wissel, kan dit essensieel beskryf word as 'n punt waar daar alternatiewe voorkom en 'n keuse gemaak moet word. Hierdie alternatiewe kan diskreet wees, byvoorbeeld of daar met 'n projek voortgegaan moet word en of dit gestaak moet word, of dit kan kontinuu wees soos wanneer 'n besluit geneem moet word oor die bekostigbare grootte van 'n kapitaaluitgawe wat aangegaan moet word.

Die verskil tussen 'n goeie en 'n slegte besluit wat by so 'n keuse gemaak word, kan as volg gedefinieer word: 'n Goeie besluit is een waar logika toegepas word, al die inligting gebruik en alternatiewe beskou word en kwantitatiewe tegnieke toegepas word. Soms kan 'n goeie besluit egter 'n onverwagte of ongunstige uitkoms hê, maar desnieteenstaande is dit dan nog steeds 'n goeie besluit wat geneem is. 'n Slegte besluit is een wat verwyder is van logika, nie al die alternatiewe in ag neem nie en ook nie van al die inligting gebruik maak nie. Indien 'n slegte besluit 'n gunstige uitkoms het, is dit nog steeds 'n slegte besluit. Goeie besluite sal op die langduur altyd 'n beter kans op sukses bied as slegte besluite. [Render (2000)]

Besluite wat geneem word tydens die verloop van 'n konstruksieprojek, beginnend by die konseptuele ontwerp tot die buitediensstelling daarvan, bepaal oorwegend of 'n projek 'n sukses gaan wees in terme van die vereistes wat deur die kliënt gestel word. Korrekte besluite wat vroeg in die projek se leeftyd geneem word, terwyl die ontwerpe nog aanpasbaar is, het 'n groter kans om te verseker dat die uitkoms van die projek 'n sukses gaan wees, terwyl verkeerde besluite vroeg in 'n projek kan veroorsaak dat die potensiële voordeel wat deur die kliënt voorsien is, nooit bereik word nie.

Besluite wat egter vroeg in 'n projek geneem word, word gewoonlik gedoen wanneer daar onsekerheid bestaan oor 'n verskeidenheid van aspekte as gevolg van die min en vae inligting wat beskikbaar is om die besluite op te baseer. Hierdie onsekerhede kan bestaan by aspekte wat tegnies van aard is, soos byvoorbeeld fondasietoestande, of ekonomies van aard, soos byvoorbeeld wat die koste van items gaan wees, of omgewingsaspekte, soos sosiale impakte en nog vele meer. 'n Verskeidenheid van metodes om op 'n logiese en gestruktureerde wyse alternatiewe te identifiseer en tegnieke om die alternatiewe mee te ontleed, is beskikbaar om besluitneming te ondersteun. Dit is egter nie voldoende om te verseker dat die beste besluit geneem word nie. Meeste van hierdie tegnieke is gebaseer op kwantitatiewe metodes waar historiese inligting of die insette van spesialiste benodig word. Hierdie inligting of insette is nie noodwendig altyd beskikbaar nie.

---

---

## 1.2 BEHOEFTE AAN HISTORIESE DATA

---

Die twee tradisionele bronne van besluitsteuninligting is die ervaring van ingenieurs en data van historiese projekte. Historiese data behoort die mees betroubare bron van inligting te wees van die twee bronne, omdat dit 'n objektiewe weerspieëling van historiese tendense bied. Ervaring van ingenieurs, alhoewel belangrik, word beperk deur beskikbaarheid en word gesien as subjektief, omdat tydsveranderlike tendense en korrelasiemeganismes wat data mag hê nie altyd in ag geneem word nie [Boles (2000)]. Historiese data as bron van inligting word egter nie benut nie weens 'n verskeidenheid van redes.

Die redes waarom data as bron van inligting nie gebruik word nie kan wissel, afhangend van databronne wat:

- nie bestaan nie, omdat daar voorheen nie 'n behoefte vir die onderskepte data was nie, dit verlore geraak het, omdat dit as nutteloos gesien is of data-onderskepping finansieel nie lonend was nie;
- onbetroubaar voorkom, omdat die data statisties groot veranderlikheid toon, die data verouderd blyk te wees, gapings in die onderskepte data voorkom of die gebruikers nie vertrou is met die omgewing waarbinne die data ontstaan het nie;
- in 'n onhanteerbare formaat is, omdat dit moeilik aanpasbaar is by die formaat waarin dit benodig word, onekonomies is om die data te transformeer na 'n meer geskikte formaat of die gebruikers nie oor die nodige vaardighede beskik om die transformasies uit te voer nie.

Weens veranderende omstandighede wat intern sowel as ekstern tot die konstruksiebedryf voorkom, word verbeterde data-onderskepping en verwerking al hoe meer moontlik en ook kritiek. Intern tot die industrie striem verouderde informasiebestuur die produktiwiteit van die datagebruikers, omdat data nie maklik bekombaar en aanpasbaar is nie [Rasdorf (2000)]. Kompetisie in die bedryf, weens 'n toenemende skaarste aan finansiële bronne vir ontwikkeling, noodsaak ook dat meer effektiewe tegnieke geïmplementeer word. Hierdie verbeterde tegnieke is nie net beperk tot die tegniese gebied nie. Ook op die gebied van bestuur waar beter beramings gedoen en besluite geneem moet word, gebaseer op betroubare data, bestaan die behoefte vir verbetering [Mitropoulos (2000)].

Ekstern tot die industrie is daar die ontwikkeling wat plaasvind op die gebied van informasietegnologie. As dit reg toegepas word, bied dit die geleentheid om data meer koste effektief te onderskep, te bestuur, te manipuleer en te analiseer. Veranderde omstandighede noodsaak nie net nie, maar bied ook die

---

geleentheid om 'n verbetering en 'n vermeerdering in die kwaliteit en die kwantiteit van historiese data wat gebruik kan word vir bestuursdoeleindes teweeg te bring. [Cleveland (1999)]

---

### 1.3 INLIGTINGSTELSEL

---

Gedurende die verskillende aktiwiteite wat plaasvind in die leeftyd van 'n konstruksieprojek, vanaf die konseptuele fase tot die afsluiting van die konstruksiefase en tydens die bedryf van die fasiliteit, word inligting geskep of benodig vir projekadministrasie en besluitneming. 'n Stelsel wat hierdie data kan onderskep, stoor, analiseer en beskikbaar stel staan bekend as 'n inligtingstelsel.

'n Inligtingstelsel het drie hoof komponente, nl. datavloei, databestuur en data-aanwending. Datavloei verwys na die oordrag van data tussen die verskillende aktiwiteite, m.a.w. hoe data vanaf die punt waar dit onderskep word beweeg, in watter vorm ook al, tot waar dit weer aangewend word. Data-aanwending verwys na hoe data gemanipuleer, geanaliseer en inligting daaruit verkry word vir onder andere projekadministrasie en besluitneming. In hierdie studie word kwantitatiewe risikobestuur as data-aanwendingstegniek ondersoek, asook die behoefte wat dit op datakwaliteit en kwantiteit plaas. Dit is ook nodig dat daar 'n tussenfase is waar data gestoor kan word, aangesien die punt waar die data geskep word en waar dit benodig word nie noodwendig mekaar kronologies in die proses volg nie. Die tussenfase kan gesien word as die databestuurgedeelte. [Rob (1997)]

Vir ontwikkelingsdoeleindes is dit nodig om die interaksie wat bestaan tussen hierdie drie komponente te herken. Indien slegs die een komponent ontwikkel word, maar daar word niks aan die ander gedoen nie, gaan die voordele wat verkry kon word in meeste gevalle verlore weens oneffektiwiteit. Hier kan die twee uiterste voorbeelde geskets word waar slegs die data-aanwending of die datavloei ontwikkel word. Vir eersgenoemde geval kan 'n nuwe analise tegniek voordelig wees om die akkuraatheid van kosteberekeninge te verbeter, maar weens geen verbetering in die datavloei nie word die voordele nietig gemaak deur die insette wat dit verg om die nodige data te versamel. So kan die datavloei ook verbeter word sonder inagneming van waarvoor die data benodig word. Weereens kan dit lei tot oneffektiwiteit, omdat nuttelose inligting onderskep kan word wat nie toepaslik is vir die aanwending nie. [Back (2000a) , (2000b)]

Die voordeel wat 'n geïntegreerde stelsel kan bied is voor die hand liggend. Die kwantiteit, kwaliteit en gebruik van relevante data kan verbeter deur nie net die wagtyd op korrekte inligting te verminder nie, maar ook die koste vir die

---

versameling van die benodigde data wat in administrasie en besluitneming gebruik word, te kan besnoei.

---

## 1.4 DOEL VAN STUDIE

---

Die vertrekpunt van hierdie studie is besluitneming wat plaasvind in die siviele ingenieurswese omgewing waar onsekerhede meeste van die tyd voorkom. Gereedskap word benodig om hierdie onsekerhede mee te modelleer. Hierdie gereedskap benodig data waarmee die onsekerhede gekwantifiseer kan word. Massas data word geskep tydens siviele ingenieursprosesse, maar die onderskepping van hierdie data vir latere gebruik word in meeste gevalle nooit gedoen nie. Ontwikkeling op die gebied van informasietegnologie bied die geleentheid om die situasie te verbeter.

Die doel van hierdie studie is om te bepaal hoe besluitneming verbeter kan word deur inligtingversameling en benutting van relevante data vir konstruksiekosteberekeninge te verbeter. Dit word gedoen deur die onderskeie aspekte wat deel vorm van 'n inligtingstelsel, nl. die datavloei, datastoring en aanwending te ondersoek en te toon hoe geïntegreerde ontwikkeling van elk van hierdie aspekte met behulp van prosesmodellering, databasisontwikkeling en kwantitatiewe risiko-analise, 'n bydrae kan maak tot verbeterde besluitneming. Die insigte wat data-ontleding kan bied met die ontwikkeling/herontwikkeling van so 'n stelsel word geïdentifiseer. Dit word aangewend saam met 'n metode wat ontwikkel is gebaseer op invloedfaktore om op 'n geselekteerde wyse data-onderskepping uit te brei. 'n Geïntegreerde ontwikkelingsmodel word sodoende saamgestel wat die interaksie van die onderskeie aspekte met mekaar toon. Hierdie model vorm die basis vir die inkrementele ontwikkelingsproses van so omvattende inligtingstelsel wat in die studie voorgestel word.

---

## 1.5 METODOLOGIE

---

In hierdie studie word die interaksie van vier aspekte ondersoek en geïntegreer, nl. prosesmodellering, databasisontwikkeling, besluitneming en die identifisering en verlaging van onsekerheid weens 'n tekort aan inligting, vir siviele ingenieurtoepassings. Van hierdie aspekte vorm nie deel van die tradisionele siviele ingenieursomgewing nie en word derhalwe vanaf basiese beginsels gedefinieer. Die sistematiese ontwikkeling van die studie wentel om twee toepassingsomgewings van uiteenlopende velde binne siviele ingenieurswese, nl. pad- en dambou, en uiteenlopende vlakke van ontwikkeling binne die spesifieke veld. Dit word veralgemeen in die ontwikkeling van 'n geïntegreerde besluitsteunstelsel.

---

### **1.5.1 BESTAANDE OMGEWING**

---

In padkonstruksie, soos in enige ander veld van konstruksie, word spesifikasies en tekeninge gebruik om die gedagtes van die ontwerper oor te dra aan die kontrakteur. Indien hierdie spesifikasies gestandaardiseer is kan dit gebruik word as die basis vir die ontwikkeling van hoeveelheidslyste met standaarditems. Die ontstaan van die “Committee of State Road Authorities” (CSRA) se databasis van tenderdata (CSRA-databasis) is oorwegend hierdeur teweeggebring toe standaardspesifikasies vir die konstruksie van paaie en brûe in 1987 ontwikkel is. [Nel (1987)]

Die CSRA-databasis se data word gebruik vir kosteberekeninge op verskillende vlakke van detail en fases in die konstruksieproses. Inligting aangaande veralgemeende konstruksiekostes wat uit die databasis verkry word, kan gebruik word vir beplanning- en vooruitskattingsdoeleindes en kan ook gebruik word in kosteberekeninge tydens die lewensvatbaarheidstudie van padprojekte. Verder kan spesifieke data oor eenheidtariewe van items wat tydens kosteberekeninge in die ontwerpfase en ook tydens die finale kosteberekening benodig word, uit die databasis verkry word vir detailkosteberekeninge.

In die eerste ses hoofstukke van hierdie studie word die padkonstruksie omgewing, spesifiek die omgewing waarbinne die CSRA-databasis funksioneer, as basis gebruik om die konsepte mee te verduidelik wat gebruik kan word vir die ontwikkeling of verbetering van ’n bestaande inligtingstelsel. Hierdie konsepte word gedurende die verduideliking ook aangewend om potensiële verbeterings aan die bestaande stelsel te toon en word in sommige gevalle toegepas.

### **1.5.2 NUWE OMGEWING**

---

Die dambou omgewing in Suid-Afrika word gebruik as voorbeeld van ’n omgewing waar daar nie so ’n databasis bestaan nie, alhoewel historiese data voorheen onderskep is, maar egter verlore geraak het. Veranderde omstandighede noodsaak egter dat daar in hierdie omgewing ondersoek ingestel word na die ontwikkeling van ’n sentrale bron van betroubare koste-inligting.

Die insig wat ontwikkel word tydens die ontleding van die padkonstruksie databasis word in hierdie omgewing gebruik om die onderskeie aspekte wat ondersoek is logies teenoor mekaar op te stel. Aspekte wat verskil tussen die twee omgewings word geïdentifiseer en metodes word voorgestel en ontwikkel om dit te kan akkommodeer.

### **1.5.3 GEÏNTEGREERDE BESLUITSTEUNSTELSEL**

---

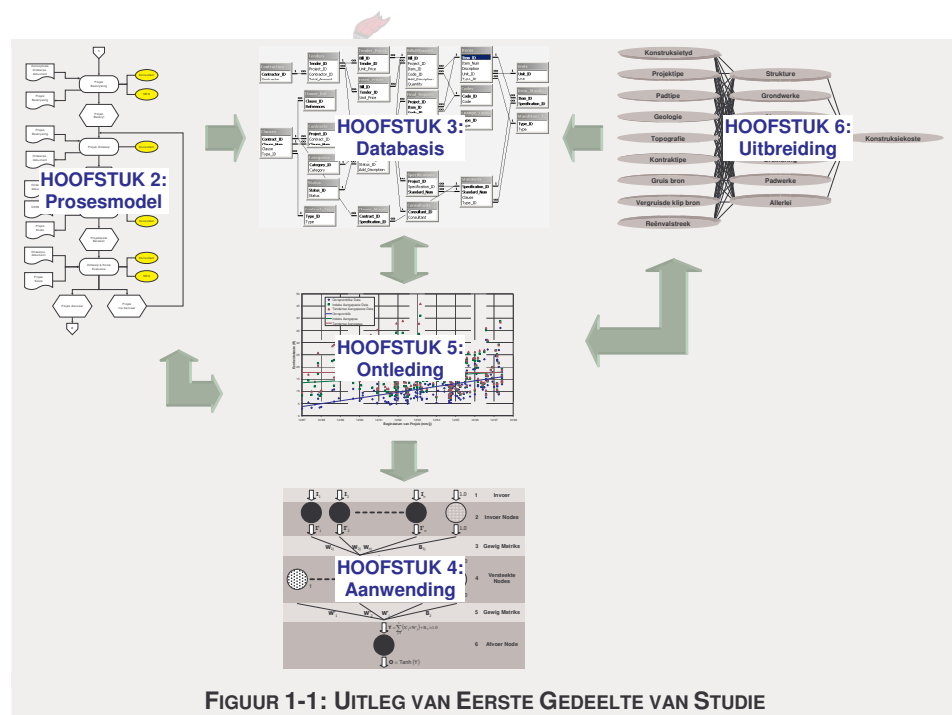
In die laaste gedeelte van die studie word die verskillende aspekte wat gedurende die konseptuele ontwikkeling van ’n inligtingstelsel in die twee omgewings

---

gevind is veralgemeen in die vorm van 'n generiese ontwikkelingmodel. In hierdie model word die onderskeie aspekte logies ten opsigte van mekaar gerangskik om die interaksie van die verskillende komponente van 'n geïntegreerde besluitsteunstelsel te toon. Die stapsgewyse proses hoedat so 'n stelsel sistematies ontwikkel kan word vir 'n bestaande of nuwe omgewing word aan die hand van hierdie generiese ontwikkelingmodel verduidelik.

## 1.6 UITLEG VAN STUDIE

Die resultate van die studie om 'n geïntegreerde besluitsteunstelsel te ontwikkel, is saamgevat in die proefskrif wat uit nege hoofstukke bestaan. Hierdie inleidende hoofstuk skets kortweg die verskillende komponente wat in hierdie studie ondersoek gaan word en die interaksie wat tussen hierdie komponente bestaan, asook die veranderende IT-omgewing wat die geleentheid bied vir die ontwikkeling van so 'n stelsel. Die metodologie wat in hierdie studie gevolg is, word ook hier beskryf. In *Figuur 1-1* word die uitleg van die eerste vyf (hoof) hoofstukke van die studie skematies getoon waarop die daaropvolgende twee hoofstukke gebaseer is.



FIGUUR 1-1: UITLEG VAN EERSTE GEDEELTE VAN STUDIE

In *Hoofstuk 2* word prosesmodellering bespreek. Onderzoek word ingestel na die historiese ontwikkeling hiervan, hoe dit verskil van tradisionele skedule- ring, die rol wat dit speel by die ontwikkeling van die datavloeiemodel en hoe dit van voordeel kan wees tot die verbetering van databronne. 'n Generiese model wat op feitlik enige tipe konstruksieprojek toegepas kan word, word voorgestel met die fokus op die aanwending en onderskepping van koste-inligting. Kwali- teitsbeheer word aangespreek deur 'n metode te ontwikkel wat gebruik kan word om die beperkte funksionaliteit van prosesmodelleringsgereedskap te

oorkom. *Hierdie metode bied die geleentheid om vir volledigheid en korrektheid van prosesmodelle te toets om sodoende 'n meer realistiese model te skep wat verbeterde datavloei te weeg kan bring.*

Databasisse moet ontwerp word volgens streng reëls wat gebaseer is op relasionele algebra. Die doel hiervan is om te verseker dat daar nie foute in die data weens onnodige duplisering tydens toevoeging en verwydering van data, ontstaan nie. Die herontwerp van 'n bestaande databasis is as basis gebruik om 'n voorgestelde datamodel vir kosteberamings te ontwikkel. *In Hoofstuk 3, word die basiese databasis reëls bespreek en toegepas om hierdie generiese databasismodel vir tenderinligting daar te stel wat maklik uitbreibaar is.* Die moontlikheid om die model uit te brei om addisionele inligting wat verbandhou met kosteberamings, nl. spesifikasies en regskontrakte, te stoor, word ook ondersoek en daar word getoon hoe verdere projekverwante inligting ook in die voorgestelde databasismodel geakkommodeer kan word vir latere besluitnemingsdoeleindes.

Die spil waarom hierdie studie draai is besluitsteuntegnieke wat aangewend kan word tydens onsekerheid. *In Hoofstuk 4 word daar gefokus op die kwantitatiewe gereedskap wat beskikbaar is vir besluitneming. Monte Carlo-simulasie en 'n tegniek gebaseer op die limietstaatteorie word ondersoek om die voordele wat probabilistiese analise bied te toon. Die moontlikhede wat skynintelligensietegnieke soos neurale netwerke bied, word ook bespreek. 'n Tegniek wat 'n kombinasie is tussen neurale netwerke en limietstaatteorie word voorgestel om besluitneming te ondersteun wanneer projekspesifieke inligting beskikbaar is.*

*In Hoofstuk 5 word ondersoek ingestel na die eienskappe van data en die invloed wat dit kan hê op berekeninge wat uitgevoer kan word tydens besluitneming. 'n Bestaande databasis se data word gebruik om tipiese aspekte aan te toon soos die bepaling van beduidende veranderlikes, aanpassings weens tyd en uitskieterfoute, die bepaling van gemiddeldes, standaardafwykings en korrelasie tussen veranderlikes.*

Die veranderlikheid van data wat voorkom by siviele ingenieurswese ontledings, is 'n algemene probleem. *In Hoofstuk 6 word 'n metode ontwikkel wat gebruik kan word om die invloedfaktore te identifiseer wat hierdie veranderlikheid kan veroorsaak.* Hierdie metode word verduidelik aan die hand van tekortkominge in terme van data-onderskepping wat gevind is by die eerste toepassingsvoorbeeld. In hierdie hoofstuk word daar ook gekyk na hoe die uitbreiding van die databronne die aanwending van die reeds beskikbare data kan verbeter om sodoende effektiewe benutting van al die komponente van die inligtingstelsel te bewerkstellig.

---

Die verskillende komponente van wat in die voorafgaande hoofstukke bespreek is, word in *Hoofstuk 7* saamgevat en toegepas op 'n nuwe omgewing, nl. dambou. In hierdie omgewing bestaan daar egter nie 'n sentrale databasis met inligting oor eenheidstariewe nie. *Die metodes wat in die voorafgaande hoofstukke bespreek is, word dus aangewend om te toon hoe so 'n stelsel inkrementeel ontwikkel kan word in 'n nuwe omgewing waar daar nie reeds so 'n stelsel bestaan nie.* Die verdere toepassingsmoontlikhede van so 'n stelsel word ook bespreek om sodoende effektiewe benutting van al die komponente in die stelsel te verseker.

*In Hoofstuk 8 word 'n generiese ontwikkelingmodel voorgestel, wat al die aspekte wat in die voorafgaande hoofstukke geïdentifiseer is, integreer. 'n Geïntegreerde ontwikkelingproses, wat op hierdie model bou, word bespreek om van nuuts af 'n besluitsteunstelsel te ontwikkel of 'n reeds bestaande stelsel te verbeter.* Die potensiële voordele van so 'n stelsel word ook bespreek en hoe die utiliteit van so 'n stelsel gemaksimeer kan word.

In die laaste hoofstuk, *Hoofstuk 9*, word die bevindings saamgevat wat in hierdie studie gemaak is aangaande: prosesmodelontwikkeling; databasisontwikkeling; kwantitatiewe risiko-analise tegnieke; die statistiese aard van data en die effek daarvan op onderskepping van data, asook die voordele wat 'n geïntegreerde benadering met die voorgestelde generiese ontwikkelingmodel en geïntegreerde ontwikkelingproses kan inhou vir die ontwikkeling/herontwikkeling van 'n besluitsteunstelsel.





## HOOFSTUK 2

---

### PROSESMODELONTWIKKELING

---

**P**rosesmodellering word aangewend om die verwantskappe wat tussen aktiwiteite, mense of groepe mense betrokke en die inligting wat tydens die aktiwiteite gebruik en geskep word, te beskryf. 'n Verskeidenheid van tegnieke of sogenaamde argitekture is beskikbaar om hierdie modelle mee op te stel. Vir die ontwikkeling van 'n bron van data vir risikobestuur, is dit nodig om insig te ontwikkel in die prosesse wat betrokke is, asook waar, wanneer en deur wie tersaaklike data geskep word. Prosesmodelle kan hiervoor gebruik word.

Hier volg 'n beskrywing van prosesmodellering, hoe dit aangewend kan word om die kwaliteit en kwantiteit van data te verbeter en die nut hiervan vir risikobestuur. 'n Generiese prosesmodel word ontwikkel wat die aktiwiteite, insluitend waar risikobestuurdata geskep en benodig word, modelleer. Tekortkomings in van die bestaande sagteware wat beskikbaar is om sulke prosesmodelle mee te ontwerp, is egter welbekend. 'n Tegniek is ontwikkel en geïmplementeer om te toon hoe hierdie tekortkominge aangespreek kan word om die korrektheid en volledigheid van prosesmodelle te verbeter.

---

#### 2.1 PROSESMODELLERING

---

Die voorstelling van aktiwiteite van die konstruksiebedryf met behulp van 'n prosesmodel kan aangewend word om datavloei in die bedryf te ondersoek. Weens die kompleksiteit van die konstruksiebedryf beskik modelleringsgereedskap egter nie oor voldoende funksionaliteit om die volledigheid en korrektheid van voorgestelde modelle te verseker nie.

---

---

### 2.1.1 DOEL VAN PROSESMODELLE

---

'n Model is 'n voorstelling van iets wat bestaan en vir 'n spesifieke doel gebruik word. Modelle word in alle vlakke van ons samelewing gebruik, omdat dit 'n vereenvoudigde perspektief bied op dit wat belangrik is. Die prosesse wat plaasvind in die konstruksiebedryf is kompleks, nie net is daar 'n verskeidenheid van aktiwiteite en mense betrokke nie, maar daar word ook van groot volumes data gebruik gemaak [Howard (1991)]. Om insig te kan ontwikkel in hoe hierdie drie aspekte, nl. aktiwiteite, persone en data, bymekaar aansluit is dit nodig om na die interaksie van die verskillende aspekte met mekaar te kyk.

Hierdie interaksie kan voorgestel word met prosesmodelle wat die aktiwiteite wat voorkom modelleer. 'n Proses kan beskou word as: 'n aantal aktiwiteite wat data of materiaal transformeer; die mense betrokke by die aktiwiteite en ook die data wat benodig word in die transformasie. Hierdie aktiwiteite vind nie spontaan plaas nie, maar daar is gebeurtenisse wat aanleiding gee tot die begin van een of meer aktiwiteite wat in serie of parallel kan plaasvind. 'n Voltooiende aktiwiteit lei weer tot 'n gebeurtenis, wat aanleiding kan gee tot daaropvolgende aktiwiteite, totdat die proses voltooi is. [Scheer (1992)]

Die prosesmodel vorm die raamwerk waarbinne die vloei van data plaasvind. Die datavloei-model toon die interaksie van die verskillende datastelle (historiese data – digitaal of papier, ondervinding, ontwerpkode, eksperimentele resultate, wetgewing, ens.) met mekaar, m.a.w. watter datastelle saam gebruik word, wat die oorsprong van hierdie data is en waarheen hierdie data dan weer vloei (verslae, databasisse, besluite). Hierdie interaksie en sintese van inligting vind ook nie net spontaan na mekaar plaas nie, maar vind by spesifieke aktiwiteite in die prosesmodel plaas. Die prosesmodel vorm dus 'n basis waarop die datavloei-model ontwikkel kan word.

Prosesmodellering kan op verskillende vlakke van detail en abstraksie gedoen word. Abstraksie is die groepering van detail. Aanvanklik kan op 'n hoëvlak van abstraksie met min detail 'n oorsigtelike model geskep word, maar soos die behoefte ontstaan kan elke komponent van die model op 'n laer vlak uitgebrei word deur meer detail by te voeg om die werklikheid meer realisties te modelleer. Met die verlaging in die vlak van abstraksie en toename in detail gaan die oorsigtelike beeld egter verlore. Dit staan bekend as afwaartse ontleding ("*top-down*"). By die laer vlakke van abstraksie kan die datavloei omvattend gemodelleer en ondersoek word.

---

### 2.1.2 PROSESMODELLERING IN SIVIELE INGENIEURSWESE

---

Na die Latham-verslag wat kritiek uitgespreek het oor die konstruksiebedryf se swak dienslewering, is volskaalse navorsingswerk geloods om te bepaal hoe

---

die bedryf meer kliënt-georiënteerd kan raak. Een van die metodes wat voorgestel is, was die beter optimering van konstruksieprosesse m.b.v. die modellering van die prosesse en identifisering van areas vir verbeterde optimering. [Sheath (1996)]

Hierdie prosesmodelle moet egter nie verwar word met die skeduleringsmodelle wat by die beplanning van konstruksieprojekte gebruik word nie. Daar is 'n paar definitiewe verskille, waarvan seker die belangrikste is dat daar gewoonlik nie 'n tydskaal gekoppel word aan prosesmodelle nie. Die toepassings van die twee modelle verskil ook van mekaar – die prosesmodel word oorwegend gebruik om die uitvoering van die algehele projek te beskryf, teenoor skeduleringsmodelle wat gebruik word om die hulpbronbestuur van die konstruksiefase van die projek te beskryf en te optimeer. Verder is daar ook verskille in die voorstelling van die twee modelle – prosesmodelle word gewoonlik met geometriese simbole voorgestel teenoor skeduleringsmodelle wat gewoonlik met Gantt-grafieke voorgestel word.

In die vervaardigingsindustrie is die aanwending van prosesmodellering vir optimering bestudeer ongeveer 'n dekade voor die konstruksiebedryf hierin begin ondersoek instel het. Die doel hiervan was die beter integrering van die ontwerp- en vervaardigingskomponente deur die ontwikkeling van rekenaar argitekture wat toegang verleen tot informasie en modelle (data en prosesse) tussen alle gebruikers en ook deur die her-oriëntering van organisatoriese strukture. [Sanvido (1990)]

Yates (1987) stel in 1987 voor dat beter outomatisering deur verbeterde rekenaar integrering die geleentheid vir die verbetering van produksie kan bied, wat besig was om af te neem weens fragmentasie in die konstruksiebedryf. Dit lui die era in van navorsing op die gebied van rekenaargeïntegreerde konstruksie ("*Computer Integrated Construction*", CIC). Die vordering wat gemaak is in die vervaardigingsindustrie op die gebied van nuwe produkontwikkeling is gebruik as die beginpunt vir verdere navorsing, alhoewel dit bekend was dat direkte toepassing nie moontlik sou wees nie weens groter kompleksiteit binne die konstruksiebedryf. [Sanvido (1990)]

Die navorsingswerk het gefokus op die ontwikkeling van generiese prosesmodelle wat die konstruksieprosesse kon beskryf. Die ontluiking van nuwe tegnologie soos gelyklopende ingenieurswese maak die verspreiding en gelyktydige aanwending van projekinligting moontlik. As hoeksteen tot hierdie tegnologie is gestandaardiseerde of algemene informasie modelle nodig, sodat 'n verskeidenheid van rekenaartoepassings data kan uitruil. Weens die verskeidenheid van rekenaarsistelsels wat gebruik word in die konstruksiebedryf moet ontwikkelaars van gelaagde modulêre ontwikkeling gebruik maak om probleem-

---

spesifieke oplossings te vind gebaseer op hoëvlak, meer generiese modelle wat konseptueel veralgemeen is. Die doel van hierdie hoëvlak modelle is om konsekwentheid tussen die meer gedetailleerde modelle te verseker en om direkte data-uitruiling tussen verskillende dissiplines moontlik te maak. [Froese (1995)]

In die laaste dekade was daar heelwat ontwikkelings gewees op die gebied van konseptuele prosesmodelontwikkeling, desnieteenstaande is dit duidelik dat daar nog nie 'n standaard vir prosesmodelle in siviele ingenieurswese bestaan nie. Meeste modelle word gewoonlik net uit die perspektief van die kliënt, die raadgewende ingenieur of die kontrakteur ontwikkel. Dit neem ook in sommige gevalle net 'n paar van die fases wat voorkom tydens die lewensduur van 'n fasiliteit in ag.

---

### **2.1.3 KWANTITATIEWE RISIKOBESTUUR IN PROSESMODELLERING**

---

Die aktiwiteite van besluitneming wat tydens verskillende fases gedurende die leeftyd van 'n fasiliteit uitgevoer word, vorm 'n integrale deel van die prosesse wat plaasvind. Nie net alleen is besluitneming bepalend vir watter roetes in 'n prosesmodel gevolg gaan word nie, maar die data wat benodig word vir besluitneming strek oor verskeie aktiwiteite binne en selfs buite die konstruksieprosesse.

Besluite wat uitgevoer moet word ten opsigte van, byvoorbeeld die voortsetting van 'n projek, watter ontwerp mee voortgegaan moet word en watter wysigings aangebring kan word tydens konstruksie, beïnvloed watter roetes op die prosesmodel van 'n konstruksieprojek gevolg word. Inligting wat benodig word vir hierdie besluitneming is afkomstig van verskeie bronne binne die projek. Tegnieiese inligting afkomstig van die ontwerp, data van eenheidtariewe van items en afwykings wat verwag kan word in hoeveelhede en eenheidtariewe, word byvoorbeeld benodig vir kosteberamings.

Die data wat in 'n projek voorkom kan in vier hoofgroepe verdeel word, nl. tegniese, finansiële, omgewing en regsaspekte. Tegnieiese data verwys na alle ingenieursverwante inligting om die fasiliteit te skep, bv. ontwerpstandaarde, projekspesifikasies, resultate van analyses, berekeninge wat verbandhou met die fisiese fasiliteit, tekeninge, kwaliteitskontrole, ens. Omgewingsdata verwys na die inligting wat verbandhou met die impakte van die fasiliteite op die natuurlike en menslike omgewing, bv. habitat van plant- en dierspesies, visuele impak, geraas, besoedeling, aantal mense geaffekteer, ens. Finansiële inligting verwys na die kostes en risiko's wat ontstaan tydens die projek. Die regsaspekte sluit alle wetgewing, ooreenkomste en kontrakte in.

Hierdie groepe inligting kan egter nie onafhanklik van mekaar gesien word nie, omdat die een die ander beïnvloed tydens die uitvoering van 'n projek. 'n

---

Interaksiematriks, *Figuur 2-1*, wys al die moontlike kombinasies. Tegnie se aspekte beïnvloed byvoorbeeld die impak op die omgewing, die kostes en watter wetgewing aan voldoen moet word. Die omgewing beïnvloed weer die ontwerp, die kostes en watter wetgewing van toepassing is. Netso geld dit ook vir die ander twee stelsel inligting. Dit is veral belangrik om te let op die verskillende groepe inligting wat die finansiële data kan beïnvloed.

	Tegnies	Omgewings	Finansiële	Regsaspek
Tegnies		X	X	X
Omgewing	X		X	X
Finansiële	X	X		X
Regsaspek	X	X	X	

**FIGUUR 2-1: INTERAKSIES TUSSEN VERSKILLENDE GROEPE INLIGTING**

Hierdie stelsel inligting word tydens die uitvoering van 'n projek gebruik vir projekbestuurdoelwitte. Die besluite wat egter geneem word vind in 'n omgewing van onsekerheid plaas, omdat daar nie perfekte inligting beskikbaar is om die besluite op te baseer nie. Kwantitatiewe risikobestuur, waarvan die aktiwiteit ook 'n proses vorm, is die voorkeurtegniek om besluitneming te ondersteun tydens onsekerheid. Hiervoor word in meeste gevalle egter data van historiese projekte, sowel as inligting van die huidige projek, benodig.

'n Hoof kenmerk van een van die robuuste prosesmodelle wat tot dusver ontwikkel is vir konstruksieprojekte, nl. die *Generiese Ontwerp en Konstruksie Prosesprotokol* (GOKPP) [Sheath (1996)], is inligting terugvoering. Die teruggevoerde inligting afkomstig hiervan kan tweeledig gebruik word: eerstens vir die suksesvolle bestuur van latere fases van die bestaande projek en tweedens vir die beplanning van ander soortgelyke toekomstige projekte. Dit word voorgestel dat 'n sogenaamde projekargief ontwikkel word waar inligting van verskeie projekte sentraal gestoor kan word. Hierdie inligting van historiese projekte kan aangewend word vir die vermindering in onsekerheid tydens besluitneming, omdat historiese inligting beskikbaar is. 'n Verlaging in die risiko wat kan lei tot verbeterde besluitneming en die verbetering in die prestasie van 'n projek kan sodoende teweeg gebring word. [Cooper (1998)]

Die interaksie van die verskillende groepe inligting op mekaar kan egter nie uit die oog verloor word nie. Deur slegs inligting van een groep, sê finansiële, te onderskep kan 'n skewe beeld van die historiese waardes ontstaan. Elke projek is uniek weens die verskeie faktore wat dit kan beïnvloed. Groot onsekerhede met gepaardgaande risiko's kan ontstaan indien gemiddeldes van finansiële data gebruik word sonder die aanpas van verskille tussen projekte weens tegniese, omgewings- en regsaspekte. Dit is dus nodig om te kan identifiseer watter inligting gebruik is tydens besluitneming in historiese projekte, sodat van hierdie inligting ook onderskep kan word.

Prosesmodelle kan aangewend word om te identifiseer waar kwantitatiewe risikobestuur toegepas kan word, watter inligting benodig word, waar hierdie inligting ontstaan en watter roetes gevolg word om die data by die punt van aanwending beskikbaar te stel. Verder kan die persone of groepe persone ook geïdentifiseer word wat betrokke is by die dataskepping en oordrag daarvan. Die aktiwiteite van risikobestuur vorm dus deel van die konstruksieprosesmodel en behoort derhalwe ingesluit te word tydens prosesmodellering, netso moet die datavloei proses ook beskou word tydens besluitneming om al die aspekte in ag te neem wat die projekinligting kan beïnvloed.

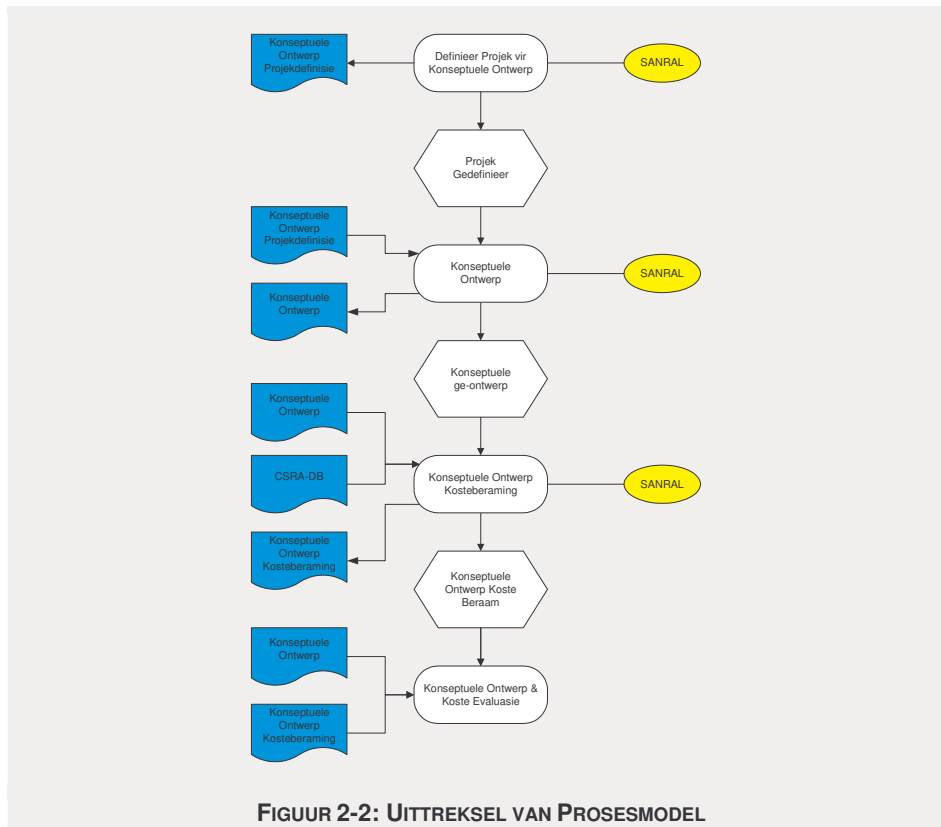
---

#### 2.1.4 ONTWIKKELING VAN 'N PROSESMODEL

---

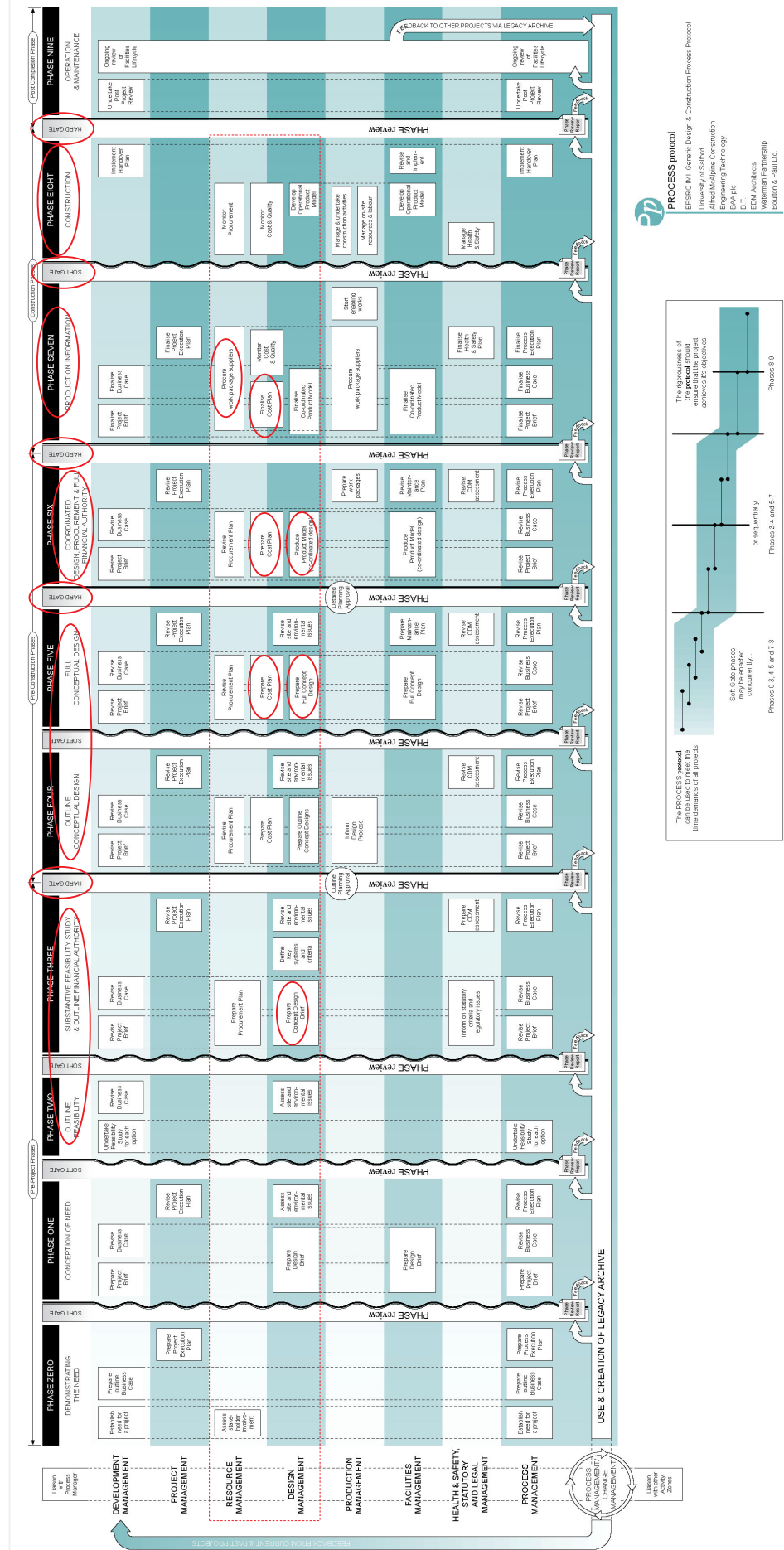
Vir hierdie studie is 'n prosesmodel ontwikkel om 'n oorsigtelike idee te vorm van waar data geskep en benodig word vir kosteberamings wat gebruik kan word in sommige van die kwantitatiewe risikobestuurstechnieke wat in *Hoofstuk 4* bespreek word. Die *Aris-toolkit* van *IDS-Scheer AG* [Aris (1999)] wat gebaseer is op 'n argitektuur bestaande uit vier komponente: funksie-blik, organisasie-blik, data-blik en beheer-blik is hiervoor gebruik, aangesien dit suksesvol in vele industrieë aangewend word om prosesmodelle te ontwikkel.

Die program bied die funksionaliteit om die konsep van abstraksie te implementeer. Vir hierdie studie is van twee vlakke van abstraksie gebruik gemaak. Eerstens 'n hoëvlak van abstraksie wat die verskillende fases toon van 'n konstruksieprojek en dus ook oorsigtelik toon hoe die hoof fases op mekaar volg. Die volgende vlak was 'n uitbreiding van die verskillende fases, met genoegsame detail, sodat die plekke waar data geskep word en waar dit benut word aangedui kan word. Verdere ontwikkeling van die model op nog laer vlakke is nie nodig geag nie, omdat die doel slegs was om die aktiwiteite te identifiseer waar inligting in die prosesmodel voorkom wat relevant is tot die risikobestuurstechnieke wat in *Hoofstuk 4* bespreek word. In *Figuur 2-2* word 'n uittreksel van die prosesmodel getoon en in *Aanhangsel A* word die volledige model wat ontwikkel is, getoon.



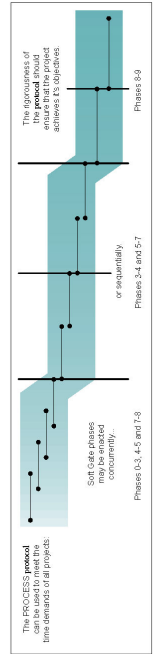
Die model wat opgestel is sluit die aktiwiteite in wat tipies sal voorkom in die sewe verskillende fases van 'n konstruksieprojek, nl. projekdefiniëring, lewensvatbaarheidstudie, konseptuele ontwerp, detail ontwerp, tender, konstruksie en afsluiting. Verder is die verskillende dokumente en generiese groepe mense getoon wat gewoonlik betrokke is by die beraming van kostes. Die doel van hierdie model is dus om oorsigtelik die besluitnemingspunte te identifiseer waar spesifiek kosteverwante data benodig word, wie hierby betrokke is en waar die verskillende data geskep word.

Dit is reeds in *Afdeling 2.1.2* genoem dat daar al menige pogings aangewend is om prosesmodelle vir konstruksieprojekte op te stel, maar dat daar gebreke by meeste is. Die *Generiese Ontwerp en Konstruksie Prosesprotokol (GOKPP)* [Sheath (1996)], *Figuur 2-3*, wat deur die *Universiteit van Selford* in samewerking met konstruksie maatskappye soos *Alfred McAlpine* en *AMC*, konstruksieklënte soos *BAA* en *BT*, raadgevende konsultante soos *Waterman Partnership Holdings Plc*, besigheidskonsultante soos *Capita*, fasiliteitsbestuurders soos *IAI* en vele meer ontwikkel is, oorkom meeste van die probleme van die tradisioneel akademiese modelle. Die model is omvattend genoeg sodat die hele lewensduurte van 'n fasiliteit, vanaf die ontstaan van die behoefte wat een of ander fasiliteit noodsaak tot en met die bedryf daarvan, beskryf kan word, maar generies genoeg sodat dit aanpasbaar is vir spesifieke toepassings.



FIGUUR 2-3: GENERIESE ONTWERP EN KONSTRUKSIE PROSESPROTOKOL

**PROCESS protocol**  
 EPSC Ltd, Generic Design & Construction Process Protocol  
 University of Salford  
 Salford  
 Engineering Technology  
 B.T.  
 EDM Accredited  
 Institution of Engineering Technicians  
 Institution of Mechanical Engineers  
 Institution of Civil Engineers





Indien die twee modelle met mekaar vergelyk word, word dit duidelik dat daar 'n opvallende verskil is in die vlak van detail ten opsigte van data en mense betrokke. Die model wat in hierdie studie ontwikkel is fokus spesifiek op die koste-inligting en wie daarby betrokke mag wees. Dit is in hierdie opsig nie so veralgemeen soos die GOKPP nie, omdat dit vir 'n spesifieke doel aangewend word, nl. om datavloei tydens kosteberekening te ondersoek.

Die onderskeie fases van die modelle stem egter sterk ooreen met mekaar, alhoewel van die fases fyner in die GOKPP afbreek word. Die prosesmodel wat in hierdie studie ontwikkel is kan feitlik netso op die GOKPP se “Resource Management” en “Design Management”-baan (aangedui met rooi stippellyn in *Figuur 2-3*) gepas word om sodoende op 'n meer gedetailleerde vlak die inligting en mense betrokke by die kosteberamings te identifiseer. Die ooreenstemmende besluitnodusse wat in beide modelle voorkom en die fases wat gedeel word, word in *Figuur 2-3* getoon. Die raakpunte wat betref die onderskeie aktiwiteite word ook aangedui, asook die invoer en afvoer van data vanaf die historiese databasis (word in diepte in *Afdeling 6.2.2* bespreek) of in die GOKPP se geval die “Legacy Archive”.

Dit blyk dus dat die model wat in hierdie studie ontwikkel is wel veralgemeen genoeg is, nie net vir padkonstruksie nie, maar ook vir ander konstruksie toepassings en dat die datavloei wat voorkom dus generies genoeg is dat dit op ander projekte toegepas kan word.

---

### 2.1.5 TEKORTKOMINGS VAN MODELLERINGSGEREEDSKAP

---

Prosesmodellering word met 'n groot mate van sukses in vele industrieë aangewend, byvoorbeeld by die motor-industrie vir die modellering, simulering en optimering van prosesse. Ondervinding met die gebruik van prosesmodelle het getoon dat dit egter nie direk toegepas kan word op die konstruksiebedryf nie, omdat daar substansiële verskille tussen die twee bestaan.

In die konstruksiebedryf is daar gewoonlik vier hoof rolspelers, nl. die kliënt, die konsultant, die kontrakteur en die verskaffer. Hierdie vier groepe is duidelik verskillend in die funksie wat elkeen uitvoer gedurende die lewensiklus van 'n fasiliteit. Elkeen van hierdie vier groepe het 'n afsonderlike organisasiestruktuur, elk met hul eie langtermyn besigheidstrategieë. Binne elkeen van hierdie groepe word hierdie gefragmenteerdheid nog verder gevoer indien daar byvoorbeeld verskillende spesialis konsultante, sub-kontrakteurs en 'n verskeidenheid van verskaffers in die onderskeie groepe voorkom. Dit lei tot 'n groot mate van gefragmenteerdheid waar koppelvlakke tussen die verskillende groepe moeilik gedefinieer kan word. [Howard (1991), Mar (1989)]

---

Konstruksieprojekte word verder meer gewoonlik net eenkeer uitgevoer en dus is die kombinasie van kliënt, konsultant, kontraktuur en verskaffer eenmalig. Tipies van die konstruksiebedryf is projekontwerpe ook nooit heeltemal finaal nie. Veranderings aan projekbeskrywings kan soms op 'n laat tydstip aangebring word, wanneer konstruksie alreeds aan die gang is. Onvoorsiene omstandighede kan ook voorkom wat noodsaak dat die projek aangepas moet word. [Sanvido (1990)]

Die tradisionele modelleringsgereedskap, alhoewel voldoende om generiese modelle soos hierbo te ontwikkel, beskik nie oor die nodige funksionaliteit om hierdie hoëvlak generiese modelle verder te ontwikkel op laer vlakke van abstraksie en toenemende detail nie. Dit kan toegeskryf word aan unieke fragmentasie van die konstruksiebedryf wat nie by die vervaardigingsbedryf voorkom nie en sodoende unieke probleme laat ontstaan. Die funksionaliteit wat benodig word om van hierdie probleme aan te spreek sluit onder andere metodes in hoe om: die effek van veranderings te bepaal; watter aktiwiteite data skep en wie verantwoordelik hiervoor is; waar data aangewend word en deur wie dit gewysig kan word; waar besluitnemingsaktiwiteite voorkom en watter data hiervoor benodig word.

By kwantitatiewe risikobestuur speel die beskikbaarheid van korrekte data 'n beduidende rol in die betroubaarheid waarmee voorspellings gemaak kan word. Funksionaliteite waarmee bepaal kan word wie, wat en wanneer data gebruik of geskep word, kan dus nuttig aangewend word by die ontleding van prosesmodelle om databronne wat reeds beskikbaar is verder uit te brei vir kwantitatiewe risikobestuur en om prosesse daar te stel wat sal verseker dat die data wel onderskep word tydens konstruksieprojekte. Prosesmodelleringsprogramme behoort dus te kan toets vir eienskappe soos volledigheid en korrektheid van prosesse.

---

## 2.2 ONTWIKKELING VAN 'N ANALISERINGSPROGRAM

---

Dit is nodig om die funksionaliteite van modelleringsprogramme uit te brei om te verseker dat prosesmodelle korrek en volledig opgestel word. Daar is 'n paar roetes wat gevolg kan word om te verseker dat dit plaasvind. Een metode is om slegs gebruik te maak van prosesse wat al voorheen gedefinieer is en wat blyk korrek te funksioneer. 'n Ander metode is om modelleringsgereedskap so op te stel dat dit verseker dat 'n model korrek is deur streng beperkings op te lê. 'n Verdere metode is om funksionaliteite in modelleringsgereedskap in te bou, sodat die prosesmodel na modellering vir korrektheid en volledigheid getoets kan word.

Laasgenoemde geval is die roete wat gevolg is vir hierdie studie, omdat dit voldoen aan 'n kritieke vereiste, naamlik veralgemening. Word daar van eers-

---

genoemde twee metodes gebruik gemaak, kan dit maklik gebeur dat die gebruiker van die program aan bande gelê word deurdat daar nie vooraf opgestelde prosesse beskikbaar is waarmee die gebruiker sy proses mee kan modelleer nie of dat beperkings wat deur die program op die gebruiker geplaas word, veroorsaak dat prosesse nie korrek gemodelleer kan word nie.

Om 'n gereedskapstuk te ontwikkel wat kan toets vir volledigheid en korrektheid is daar besluit om veralgemeende vrae oor die model te formuleer wat oor enige prosesmodel gevra kan word nadat dit opgestel is. Dit kan dan onder andere gebruik word om te bepaal: waar besluitneming plaasvind, of die korrekte data gebruik word, of die benodigde data wel onderskep word en wie hierby betrokke is, sodoende kan die aspek van volledigheid en korrektheid aangespreek word, asook die behoefte om inligting te identifiseer vir besluitsteun. Hierdie vrae het te make met die verwantskappe wat tussen die verskillende komponente van die prosesmodel voorkom. Aangesien hierdie verwantskappe tussen die komponente gesien kan word as wiskundige relasies kan die ontwikkeling van die funksionaliteit wiskundig gefundeer word.

---

## 2.2.1 WISKUNDIGE AGTERGROND

---

Prosesmodelle kan, weens die ooreenkomste wat bestaan tussen die struktuur van die prosesmodelle en grafieke, beskryf word m.b.v. basiese diskrete wiskunde van versamelings- en grafiekteorie. Vir die ontwikkeling van 'n program wat korrektheid en volledigheid moet toets, is dit handig om die bekende resultate van die diskrete wiskunde te gebruik, soos byvoorbeeld die datastrukture om die modelle in te stoor en die algoritmes om dit mee te analiseer. As agtergrond tot die ontwikkeling van gereedskap om te toets vir volledigheid en korrektheid van prosesmodelle, word basiese versamelings- en grafiekteorie derhalwe bespreek.

---

### 2.2.1.1 VERSAMELINGSTEORIE

---

Versamelingsteorie is die grondslag waarop moderne wiskunde gebaseer is om struktuur daaraan te gee. Vir die doeleindes van hierdie studie is dit nodig om 'n paar eenvoudige versamelingsteorie definisies te stel wat later gebruik sal word. As vertrekpunt word versamelings, elemente, kardinaliteite en subversamelings gedefinieer [Ince (1995), Mayberry (2000), Pahl (2000), Piff (1992)]. Voorbeelde uit die prosesmodel in *Aanhangsel A* word as verklaring tot elke definisie gebruik. (In die voorbeelde word van 'n "óf" gebruik om die voorbeelde te skei; dit is nie die logiese "of" wat gebruik word nie, maar 'n grammatiese "of".)

**Definisie 2-1:** 'n **Proposisie** is 'n stelling wat waar of vals is. •

---

Voorbeeld: Ingenieurs is 'n groep mense óf kosteberekening is 'n aktiwiteit óf data word gelees óf voorlopige studie voltooi, is 'n gebeurtenis óf (meer algemeen)  $1 + 1 = 2$  óf  $1 + 1 = 3$  (wat vals is).

**Definisie 2-2:** 'n **Predikaat** is 'n stelling  $P(x_1, \dots, x_n)$  met  $n$ -veranderlikes  $x_1, \dots, x_n$  wat 'n proposisie word as spesifieke objekte van die ruimte van bespreking in die veranderlikes  $x_1, \dots, x_n$  ingestel word. •

Voorbeeld:  $P(x) = x$  is 'n groep mense, waar  $x = \text{ingenieurs, argitekte, prokureurs, arbeiders, ens.}$  kan wees óf  $P(x) = x$  is 'n konstruksie verwante aktiwiteit, waar  $x = \text{kosteberekeninge, ontwerp, voorlopige studie, konstruksie, ens.}$  kan wees óf  $P(x) = x$  is gebeurtenisse, waar  $x = \text{kosteberekeninge voltooi, ontwerp voltooi, voorlopige studie voltooi, ens.}$  kan wees.

**Definisie 2-3:** Laat  $P(x)$  'n predikaat wees in die ruimte  $E$ . Die groep van al die objekte  $x$  in  $E$ , sodat  $P(x)$  waar is, word 'n **versameling** genoem. Dit kan geskryf word as:  $\{ x : P(x) \}$  of  $\{ x \mid P(x) \}$ . •

Voorbeeld: Gestel die ruimte  $E$  is die verskeidenheid van aktiwiteite wat voorkom, dan kan die versameling  $A = \{ x \mid x \text{ is 'n konstruksie aktiwiteit} \}$  gedefinieer word óf die ruimte  $E$  is die verskeidenheid van gebeurtenisse, dan kan die versameling  $B = \{ x \mid x \text{ is 'n gebeurtenis} \}$  gedefinieer word óf die ruimte  $E$  is die verskeidenheid van groepe mense, dan kan die versameling  $G = \{ x \mid x \text{ is 'n groep} \}$  gedefinieer word óf die ruimte  $E$  is die data wat gebruik word, dan kan die versameling  $D = \{ x \mid x \text{ is 'n data} \}$  gedefinieer word.

**Definisie 2-4:** Laat  $A = \{ x \mid P(x) \}$  en veronderstel dat  $P(a)$  waar is, dan is  $a$  'n **element** van  $A$ . Dit word geskryf as  $a \in A$ . •

Voorbeeld: Van die elemente in die versameling  $A$  hierbo is  $\{ \text{definiëring van projek; konseptuele ontwerp; ontwerp; kosteberekening} \}$  óf in die versameling  $B$  is  $\{ \text{projek gedefinieer; konseptuele ontwerp voltooi; kostes bereken} \}$  óf in die versameling  $G$  is  $\{ \text{kliënt; konsultant; kontrakteur; subkontrakteur} \}$  óf in die versameling  $D$  is  $\{ \text{kodes; projek spesifikasies; hoeveelheidslys; tenderkontrak} \}$ .

**Definisie 2-5:** Laat  $A$  'n eindige versameling wees, dan is die **kardinaliteit** die aantal elemente in die versameling. Dit kan voorgestel word as  $|A|$ . •

Voorbeeld: Die kardinaliteit van  $A$  is  $|A| = 4$ .

**Definisie 2-6:** Laat  $A$  en  $B$  twee versamelings wees. Veronderstel dat elke element  $a \in A$  voldoen aan  $a \in B$ , dan kan  $A$  'n **subversameling** van  $B$  genoem word. Dit kan geskryf word as  $A \subseteq B$ . •

**Definisie 2-7:** Laat  $A$  en  $B$  twee subversamelings wees van 'n versameling  $E$ . Die **deursnee**  $A \cap B = \{ x \in E \mid (a \in A) \text{ en } (b \in B) \}$ . •

**Definisie 2-8:** Laat  $A$  en  $B$  twee versamelings wees. Die **kartesiese produk**  $A \times B$  word dan gedefinieer as:  $A \times B = \{ (a,b) \mid a \in A, b \in B \}$ . •

Voorbeeld: Laat die aktiwiteitversameling  $A$  wees en die groepversameling  $G$  soos in *Definisie 2-4* dan is  $A \times G = \{ (\text{definiëring van projek, kliënt}); (\text{konseptuele ontwerp, kliënt}); (\text{ontwerp, kliënt}); (\text{kosteberekening, kliënt}); (\text{definiëring van projek, konsultant}); (\text{konseptuele ontwerp, konsultant}); (\text{ontwerp, konsultant}); (\text{kosteberekening, konsultant}); (\text{definiëring van projek, kontrakteur}); (\text{konseptuele ontwerp, kontrakteur}); (\text{ontwerp, kontrakteur}); (\text{kosteberekening, kontrakteur}); (\text{definiëring van projek, subkontraakteur}); (\text{konseptuele ontwerp, subkontraakteur}); (\text{ontwerp, subkontraakteur}); (\text{kosteberekening, subkontraakteur}) \}$ .

**Definisie 2-9:** Laat  $A$  en  $B$  twee versamelings wees. 'n **Relasie** tussen  $A$  en  $B$  is 'n subversameling  $R \subseteq A \times B$ . Dit kan geskryf word as  $a R b$  as  $(a,b) \in R$  of  $\{ (a,b) \mid a \in A, b \in B \mid R \}$ . •

Voorbeeld: Laat  $A \times G$  gedefinieer wees soos in *Definisie 2-8*, dan bestaan daar 'n relasie  $AG = \{ (a,g) \mid a \in A, g \in G \mid \text{uitgevoer deur} \} = \{ (\text{definiëring van projek, kliënt}); (\text{konseptuele ontwerp, konsultant}); (\text{ontwerp, konsultant}); (\text{kosteberekening, konsultant}) \}$  wat die verwantskap tussen 'n aktiwiteit wat uitgevoer word en 'n groep beskryf, net so kan die volgende relasies ook gedefinieer word wat in hierdie studie gebruik word:

- volgorde tussen aktiwiteit en gebeurtenis:  $AB = \{ (a,b) \mid a \in A, b \in B \mid \text{lei tot} \}$ ;
  - volgorde tussen gebeurtenis en aktiwiteit:  $BA = \{ (b,a) \mid b \in B, a \in A \mid \text{lei tot} \}$ ;
  - 'n aktiwiteit skep data:  $AS = \{ (a,d) \mid a \in A, d \in D \mid \text{skep} \}$ ;
  - 'n aktiwiteit skryf data:  $AK = \{ (a,d) \mid a \in A, d \in D \mid \text{skryf} \}$ ;
  - 'n aktiwiteit lees data:  $AL = \{ (a,d) \mid a \in A, d \in D \mid \text{lees} \}$ ;
  - 'n aktiwiteit verwyder data:  $AV = \{ (a,d) \mid a \in A, d \in D \mid \text{verwyder} \}$ .
-

**Definisie 2-10:** 'n **Binêre** relasie in die versameling  $A$  is 'n subversameling  $R \subseteq A \times A$ . •

Voorbeeld: Die relasie  $GG = \{ (g,g) \mid g \in G \mid \text{werk saam} \}$  beskryf die verwantskap tussen twee groepe wat saam werk.

Binêre relasies kan maklik gevorm word deur van die volgende wiskundige kunsgreep gebruik te maak. Gestel  $R \subseteq A \times B$  is 'n relasie, vorm die vereniging  $C = A \cup B$  dan is die relasie  $R' \subseteq C \times C$  met  $R = R'$ , 'n binêre relasie.

---

### 2.2.1.2 EIENSKAPPE VAN BINÊRE RELASIES

---

Die binêre relasies wat voorkom in 'n versameling kan vier basiese eienskappe hê. Vir hierdie studie was dit nodig gewees om te bepaal of enige relasies sulke eienskappe besit met die doel om dit dan aan te wend om addisionele inligting uit die prosesmodel te verkry. Die eienskappe is as volg:

**Definisie 2-11:** 'n Binêre relasie met  $a R a$  vir alle  $a \in A$ , word **refleksief** genoem. •

Voorbeeld: Geen voorbeelde hiervan in hierdie studie nie.

**Definisie 2-12:** 'n Binêre relasie  $a R b$  is **simmetries** as en slegs as  $b R a$ . •

Voorbeeld: Die binêre relasie wat ontstaan tussen die vereniging van die versamelings *Aktiwiteit* en *Groep* is simmetries.

**Definisie 2-13:** 'n Binêre relasie is **antisimmetries** as  $R \cap R^{-1} = R^0$ , waar  $R^{-1}$  die inverse relasie is en  $R^0$  die identiteitsrelasie. •

Voorbeeld: Die binêre relasie wat ontstaan met die vereniging van die aktiwiteitsversameling  $A$  en die gebeurtenisversameling  $B$  is antisimmetries. Hierdie is 'n belangrike opmerking, want dit beteken dat daar nie op dieselfde roete terug beweeg kan word nie.

**Definisie 2-14:** 'n Binêre relasie is **transitief** as  $a R b$  en  $b R c$  dan is  $a R c$ . •

Voorbeeld: Die binêre relasie wat ontstaan as die versamelings *Groep*, *Aktiwiteit* en *Data* verenig word, is transitief.

---

### 2.2.1.3 GRAFIEKTEORIE

---

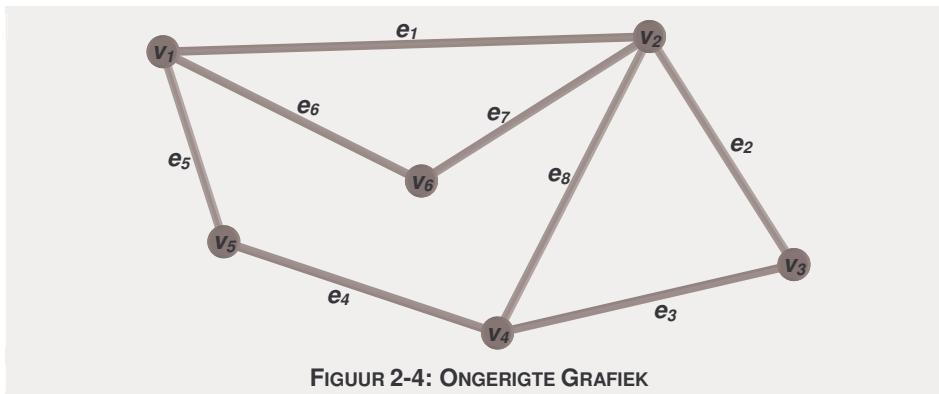
Grafiekteorie is gebaseer op versamelingsteorie, die relasie  $R$  is byvoorbeeld presies dieselfde as 'n gerigte grafiek. Aangesien daar in prosesmodelle van grafieke gebruik gemaak word om die verwantskappe tussen die verskillende entiteite voor te stel, is dit nodig om 'n paar eenvoudige grafiekteorie definisies

---

te stel wat handig gebruik kan word. [Gross (1998), Manber (1989), Swamy (1992)]

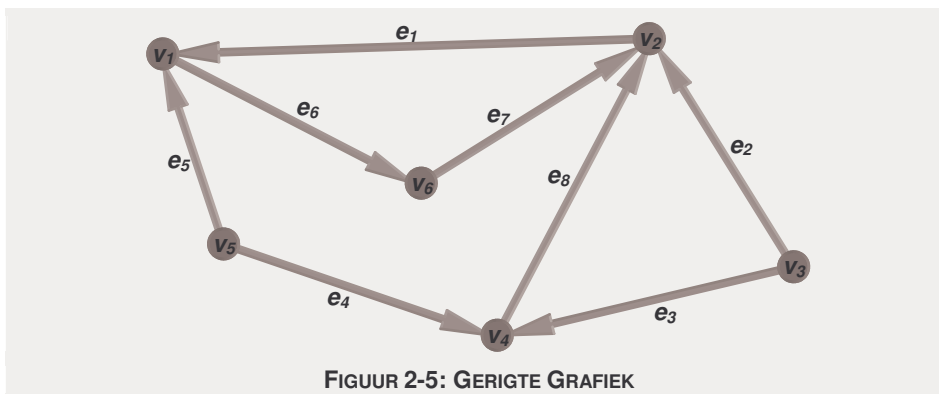
**Definisie 2-15:** Laat  $V$  'n versameling wees. Definieer dan  $V_2 = \{ A \subseteq V : |A| = 2 \}$  en noem dit die **ongeordende produk** van  $V$  met homself. •

**Definisie 2-16:** Laat  $V$  'n eindige versameling wees en laat  $E$  'n subversameling wees van  $V_2$ . Die kombinasie van  $V$  en  $E$  word genoem die **ongerigte grafiek**  $G$  met **nodus** versameling  $V$  en **lyn** versameling  $E$ . Dit kan geskryf word as die grafiek  $G = (V, E)$ . •



**Definisie 2-17:** Laat  $V$  'n versameling wees, dan is  $V^2 = V \times V$ . •

**Definisie 2-18:** Laat  $V$  'n eindige versameling wees en laat  $E$  'n subversameling van  $V^2$  wees. Die kombinasie van  $V$  en  $E$  word genoem die **gerigte grafiek**  $G$  met  $E$  die **boog** versameling. Dit kan ook geskryf word as die grafiek  $G = (V, E)$ . •

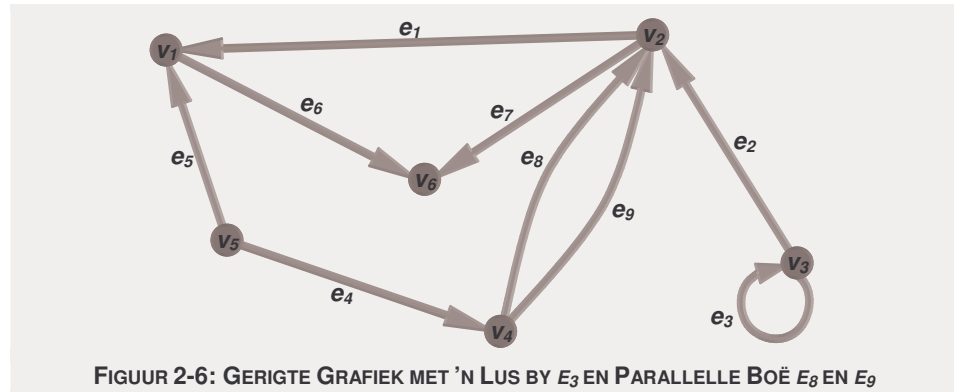


Voorbeeld: Die elemente van die vier versamelings *Aktiwiteit*, *Gebeurtenis*, *Groep* en *Data* van die prosesmodel wat in hierdie studie ontwikkel is, kan gesien word as die nodusse in die versameling  $V$ , terwyl die relasie die boë voorstel in die versameling  $E$ . Die prosesmodel kan dus met die gerigte grafiek  $G = (V, E)$  voorgestel word.

**Definisie 2-19:** 'n **Lus** is 'n lyn of boog wat 'n enkele nodus verbind. •

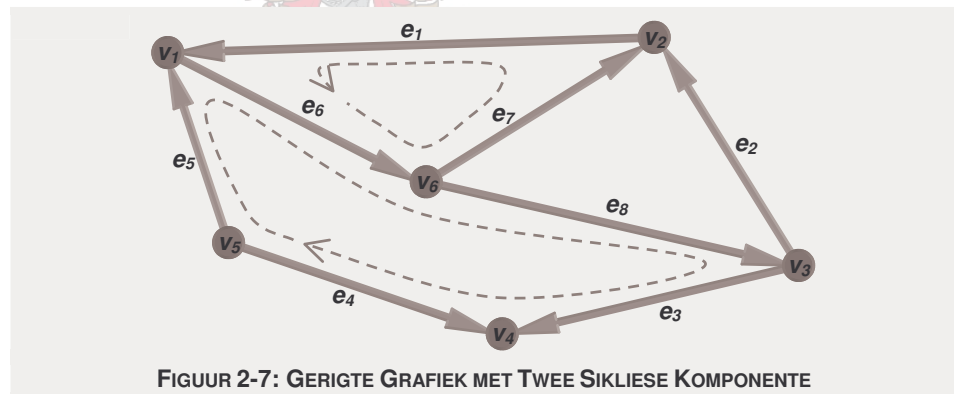
**Definisie 2-20:** Parallele lyne of boë deel dieselfde eindpunte. •

**Definisie 2-21:** 'n Eenvoudige grafiek het geen lusse. •



FIGUUR 2-6: GERIGTE GRAFIEK MET 'N LUS BY  $E_3$  EN PARALLELE BOË  $E_8$  EN  $E_9$

**Definisie 2-22:** Laat  $G = (V, E)$  'n gerigte grafiek wees en laat  $v_1, \dots, v_n \in V$  so wees dat  $v_i v_{i+1} \in E$  vir  $1 \leq i \leq n-1$ . Die reeks van  $v_1, v_1 v_2, v_2 v_3, v_3, \dots, v_{n-1}, v_{n-1} v_n, v_n$  word 'n **gerigte lynry** genoem. As die boë in 'n lynry uniek is, word dit 'n **gerigte roete** genoem en as die nodusse uniek is, behalwe miskien  $v_n = v_1$  dan word dit 'n **gerigte pad** genoem. 'n Lynry, roete of pad word **geslote** genoem as  $v_n = v_1$ , anders is dit **oop**. •

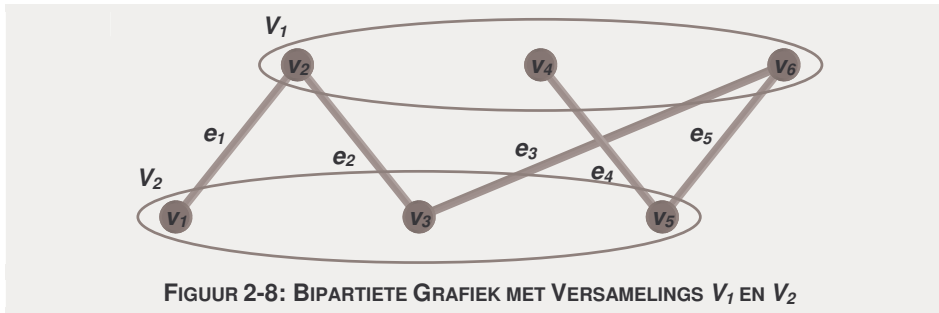


FIGUUR 2-7: GERIGTE GRAFIEK MET TWEE SIKLIESE KOMPONENTE

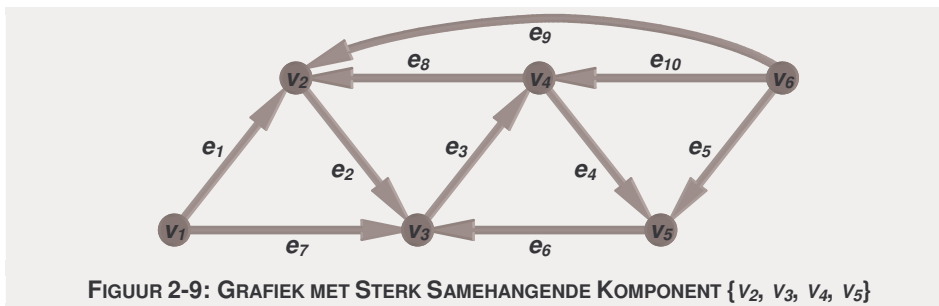
**Definisie 2-23:** 'n Geslote gerigte pad word 'n **gerigte kring** genoem. •

**Definisie 2-24:** 'n Grafiek is **bipartiet** as die versameling  $V$  verdeel kan word in disjunk, nie-leë komponente  $V_1$  en  $V_2$  met  $E = \{v_1 v_2 \mid v_1 \in V_1, v_2 \in V_2\}$ . •



FIGUUR 2-8: BIPARTIETE GRAFIEK MET VERSAMELINGS  $V_1$  EN  $V_2$ 

**Definisie 2-25:** Laat  $G$  'n gerigte grafiek wees en laat  $v_1, v_2 \in V(G)$ . Dan is  $v_1$  en  $v_2$  **sterk samehangend** as daar 'n gerigte pad van  $v_1$  na  $v_2$  is en 'n gerigte pad van  $v_2$  na  $v_1$ . 'n Maksimaal sterk samehangende subgrafiek van  $G$  word sterk samehangende komponente genoem. •

FIGUUR 2-9: GRAFIEK MET STERK SAMEHANGENDE KOMPONENT  $\{v_2, v_3, v_4, v_5\}$ 

#### 2.2.1.4 ALGORITMES

In grafiekteorie is daar 'n aantal fundamentele algoritmes wat gebruik kan word om grafieke mee te analiseer. Vir hierdie studie is daar ondersoek ingestel om te bepaal of enige van hierdie algoritmes van nut kan wees vir 'n korrektheid- en volledigheidprogram. Hieronder volg 'n bespreking van die algemeenste algoritmes [Manber (1989), Piff (1992), Swamy (1992)]. 'n Definisie van elk word gegee; die persoon aan wie die ontwikkeling van die algoritme gekoppel word en of dit aangewend kan word in die studie:

- *Transitiewe Sluiting* – die berekening van transitiewe sluiting  $x R^* y$  van 'n binêre relasie  $R$  as daar 'n reeks  $x = x_0, x_1, x_2, \dots, x_k = y$  bestaan, sodat  $x_0 R x_1, x_1 R x_2, \dots, x_{k-1} R x_k$ ; Warshall en Warren; kan gebruik word om te bepaal watter data deur watter groep gebruik word en andersom.
- *Kortste Pad* – vind die kortste pad tussen 'n spesifieke nodus  $s$  na alle ander nodusse in 'n verbindende gerigte grafiek of bereken die kortste pad tussen al die geordende nodusse in 'n verbindende gerigte grafiek; Dijkstra se algoritme bereken dit vir 'n enkele nodus, Floyd se algoritme bereken dit vir alle geordende pare nodusse; geen toepassing in hierdie studie nie, maar kan wel vir foutvoortplanting gebruik word.

- *Minimum Gewig van 'n Gespande Boom* – bereken die minimum gewig van 'n verbinde geweegde ongerigte grafiek met geen negatiewe gewigte; Kruskal of Prim se algoritmes; weens definisie geen toepassing, omdat die prosesmodel vir hierdie studie met gerigte grafieke voorgestel word.
- *Optimale vertakking* – bereken die optimum vertakking van 'n verbinde gerigte sikliese grafiek met gewigte aan die boë toegeken; Edmunds; weens definisie van geweegde boë geen toepassing.
- *Chinese posman probleem* – vind die minimum gewig van 'n geslote lynry wat elke boog ten minste eenkeer oorsteek vir 'n verbinde geweegde grafiek; Edmunds en Johnson; geen toepassing in hierdie studie nie, omdat van ongeweegde grafieke gebruik gemaak word.
- *Diepte-eerste soektog* – sistematiese metode om 'n verbinde gerigte of ongerigte grafiek sonder lusse te verken deur onderskeidelik al die lyne of boë te besoek; Tarjan; handig om saam met die volgende algoritme te gebruik.
- *Bi-samehangend en Sterk samehangend* – bepaal die maksimale samehangende sub-grafiek in 'n verbinde ongerigte grafiek, asook die maksimale sterk samehangende gerigte sub-grafiek van 'n gerigte grafiek; Tarjan; die laasgenoemde toets vir sterk samehangend is handig om geslote lynrye te vind.
- *Reduseerbaarheidstoets* – gerigte grafiek sonder parallel boë en 'n nodus is  $s$  met 'n gerigte pad na elke ander nodus in  $G$ , dit is dan reduseerbaar deur die herhaaldelik uitvoering van twee transformasies; Tarjan; geen toepassing in hierdie studie nie.
- *st-Numering van 'n Grafiek* – bereken die st-numering van  $g(v)$  van nodus  $v$ ; Even en Tarjan; word saam met die volgende algoritme gebruik.
- *Planetêre Toetsing* – toets om te sien of 'n grafiek op 'n plat vlak geteken kan word sodat boë mekaar net kruis by nodusse; Lempel, Even en Cederbaum; geen toepassing in hierdie studie nie.
- *Vloei in netwerke* – verbinde gerigte grafiek sonder lusse en wat aan die volgende voorwaardes voldoen: slegs een beginnodus, een endnodus en elke positiewe reële getal aan elke boog gekoppel; Ford en Folkerson, Dinic, Goldberg en Tarjan; geen toepassing in hierdie studie nie.

### 2.2.1.5 REKENAARMATIGE VOORSTELLING

Dit is nodig om datastrukture te hê waarbinne beskrywende inligting gestoor en gemanipuleer kan word om grafieke op 'n rekenaar voor te stel. Matrikse kan hiervoor gebruik word. Daar is vier algemene tipes matrikse wat in grafiekteorie gebruik kan word, nl. die insidensie-, siklus-, sny- en naasliggendheidmatriks. Eersgenoemde drie word in die studie van elektriese netwerke gebruik om die koëffisiënte van die Kirchhoffvergelykings voor te stel. Vir hierdie studie word daar egter net van die insidensie- en naasliggendheidmatriks gebruik gemaak [Budd (1994), Manber (1989), Swamy (1992)].

**Definisie 2-26:** Laat  $G$  'n grafiek wees met  $n$  nodusse en  $m$  lyne of boë. Die insidensiematriks  $A_c = [a_{ij}]$  van  $G$  het  $n$  rye, een vir elke nodus en  $m$  kolomme, een vir elke lyn of boog. Die element  $a_{ij}$  van  $A_c$  word dan as volg gedefinieer:

- as  $G$  'n gerigte grafiek is:

$$a_{ij} = \begin{cases} 1, & \text{as die } j\text{te boog weg gerig is van die } i\text{te nodus,} \\ -1, & \text{as die } j\text{te boog na die } i\text{te nodus gerig is,} \\ 0 & \text{as geen } j\text{te boog die } i\text{te nodus verbind nie.} \end{cases}$$

- as  $G$  'n ongerigte grafiek is:

$$a_{ij} = \begin{cases} 1, & \text{as die } j\text{te lyn die } i\text{te nodus verbind,} \\ 0, & \text{andersins.} \end{cases}$$

**Definisie 2-27:** Laat  $G = (V, E)$  'n gerigte grafiek wees met  $|V| = n$ . Die naasliggendheidmatriks  $M = [m_{ij}]$  van  $G$  is 'n  $n \times n$  matriks met  $m_{ij}$  gedefinieer as volg:

$$m_{ij} = \begin{cases} 1, & \text{as } (v_i, v_j) \in E, \\ 0, & \text{andersins.} \end{cases}$$

'n Tekortkoming van hierdie wyse van matriksvoorstellings is dat dit om twee redes nie geheue effektief is nie. Eerstens kan die matrikse baie groot raak indien daar heelwat nodusse en lyne is, omdat die matriks se grootte met  $O(|V| \cdot |E|)$  en  $O(|V|^2)$  toeneem. Tweedens kom daar ook heelwat nulle in die matriks voor, indien daar nie verbindings tussen die nodusse is nie. Die matrikse staan dan bekend as ylmatrikse. Om hierdie tekortkoming te omseil is dit moontlik om van sogenaamde insidensie- en naasliggendheidlyste gebruik te maak. Hierdie strukture se grootte behoeftes is  $O(|V|)$ . Vir hierdie studie is daar egter nie hiervan gebruik gemaak nie, omdat dit tydens implementering heelwat oorhoofse kostes verg.

---

## 2.2.2 IMPLEMENTERING

---

'n Verskeidenheid van vrae kan gevra word om te toets vir die volledigheid en korrektheid van 'n prosesmodel. Hierdie vrae hou verband met die interaksies tussen die vier versamelings wat voorkom in die prosesmodel, nl. *Aktiwiteit*, *Gebeurtenis*, *Groep* en *Data*, derhalwe kan die vrae ook volgens die versamelings gegroepeer word soos in die res van hierdie afdeling. Die vrae kan beantwoord word deur te bepaal of die relasies, soos in die voorbeeld van *Definisie 2-9*, tussen die verskillende versamelings korrek gedefinieer is. [Bester (2001)]

'n Rekenaarprogram is geïmplementeer om te toon dat dit moontlik is om hierdie konseptuele tegniek prakties te ontwikkel en toe te pas. Die rekenartaal *Java* is hiervoor gebruik. Enige ander rekenartaal kon egter ook net so goed aangewend word om die konsepte te illustreer. In *Aanhangsel B* is daar 'n kort bespreking van die program, hoe dit geïmplementeer is en die rekenarkode daarvan word ook getoon.

In die res van hierdie afdeling word die datastrukture getoon waarbinne die relasies tussen die vier versamelings gestoor kan word. Vrae word gestel wat betrekking het op die onderskeie relasies om te toets vir korrektheid en volledigheid. 'n Beskrywing word dan gegee van die algoritme wat die gevraagde antwoorde kan lewer. Die beskrywing word gevolg met die grafiese afvoer van die resultaat van die ontwikkelde rekenaarprogram wat toegepas is op 'n prototipe prosesmodel.

---

### 2.2.2.1 AKTIWITEIT EN GROEP

---

Die relasie *AG*, tussen die versameling *Aktiwiteit* en die versameling *Groep* is simmetries en kan deur die naasliggendheidsmatriks *M* beskryf word, waar:

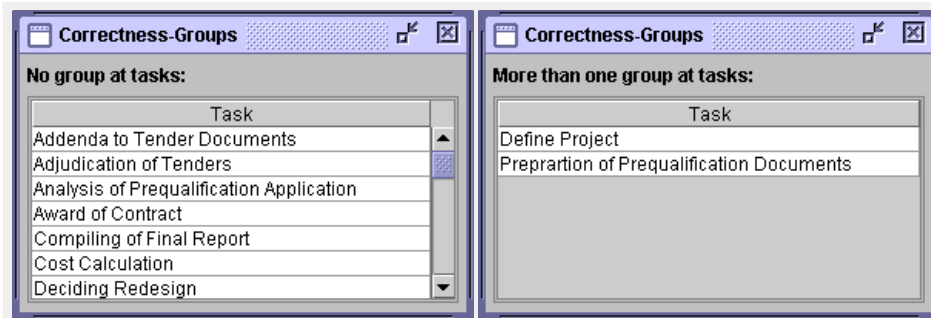
$$m_{ij} = \begin{cases} 1, & \text{as } (aktiwiteit_i, groep_j) \in AG, \\ 0, & \text{andersins.} \end{cases}$$

Die volgende vrae kan nou oor die relasie *AG* gevra word:

Vraag: Is daar enige groep betrokke by 'n aktiwiteit of is daar meer as een groep?

Algoritme: Die som van elke ry in *M* kan bereken word, dit lewer die aantal groep/e betrokke by elke aktiwiteit.

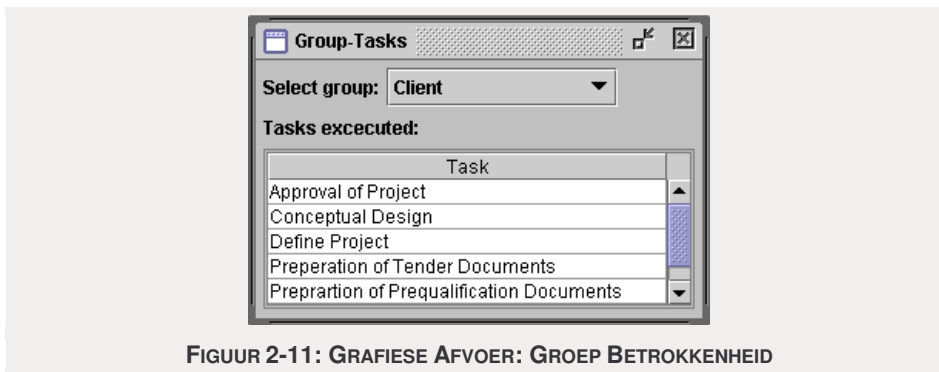
---



FIGUUR 2-10: GRAFIESE AFVOER: AANTAL GROEPE BETROKKE

Vraag: In watter aktiwiteit/e is 'n groep betrokke?

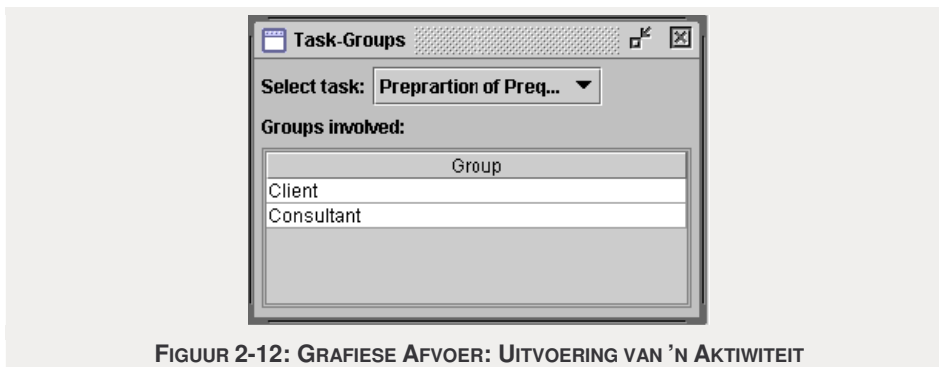
Algoritme: Soek deur die ooreenstemmende kolom van die gegewe groep in *M* en vind dan so die aktiwiteit/e in die ooreenstemmende ry/e wat aangedui word met 1'e in die kolom.



FIGUUR 2-11: GRAFIESE AFVOER: GROEP BETROKKENHEID

Vraag: Wie is betrokke by 'n spesifieke aktiwiteit?

Algoritme: Soek deur die ooreenstemmende ry van die gegewe aktiwiteit in *M* en vind so dan die groep/e in die ooreenstemmende kolom/me wat aangedui word met 1'e in die ry.



FIGUUR 2-12: GRAFIESE AFVOER: UITVOERING VAN 'N AKTIWITEIT

Met hierdie vrae kan bepaal word waar daar geen groepe gedefinieer is nie en waar daar meer as een groep gedefinieer is. Daar kan ook bepaal word of 'n groep betrokke is by die korrekte aktiwiteite en andersom, d.w.s. of die aktiwiteit se uitvoering met die regte groepe gemodelleer word. Hierdie vrae is belangrik, omdat dit aangewend kan word om te bepaal of die prosesmodel volledig en

korrek gedefinieer is t.o.v. die gebruikers wat aktiwiteite uitvoer, hetsy van tegniese, administratiewe of besluitnemingsaard.

### 2.2.2.2 AKTIWITEIT EN DATA

Tussen die versamelings *Aktiwiteit* en *Data* bestaan daar vier relasies: *skep*, *skryf*, *lees* en *uitvee*. Databasisse herken tradisioneel slegs twee tipes toegange, naamlik *lees*- en *skryf*-toegang. Vir die datamodel word *skep*, *skryf* en *uitvee* as 'n databasis *skryf*-toegang gedefinieer en *lees* as 'n *lees*-toegang. By die prosesmodel is dit egter nodig om te weet of data wel bestaan as dit gewysig, gelees of verwyder moet word. Dit is dus nodig om by die prosesmodel die toegange as volg te definieer, *skep* as 'n *skryf*-toegang, terwyl *skryf*, *lees* en *uitvee* as 'n *lees*-toegang gedefinieer kan word.

In die program wat ontwikkel is, word albei die modelle ondersteun. Die gebruiker kan kies tussen watter model ondersoek moet word deur slegs in die grafiese koppelvlak aan te dui watter model beskou moet word. (Die voorbeelde wat hier getoon word, is vir die prosesmodel datatoegang gevalle.) Dit is nodig om twee insidensiematrikse te definieer om onderskeid tussen die twee toegangsmodelle vir die rekenaarimplimentasie te kan maak, een vir die prosesmodel datatoegang, matriks *P* en een vir die standaard databasis datatoegang, matriks *D*:

$$p_{ij} = \begin{cases} 1, & \text{as aktiwiteit}_i \text{ data}_j \text{ skryf,} \\ -1, & \text{as aktiwiteit}_i \text{ data}_j \text{ lees,} \\ 0 & \text{andersins.} \end{cases}$$

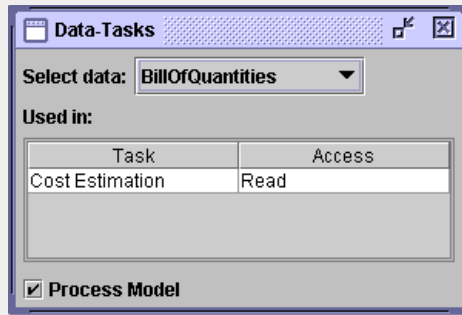
en

$$d_{ij} = \begin{cases} 1, & \text{as aktiwiteit}_i \text{ data}_j \text{ skryf,} \\ -1, & \text{as aktiwiteit}_i \text{ data}_j \text{ lees,} \\ 0 & \text{andersins.} \end{cases}$$

Met hierdie twee matrikse gedefinieer is dit moontlik om die volgende vrae te kan beantwoord:

Vraag: Watter aktiwiteit/e gebruik 'n data en wat is die toegang in die prosesmodel en datamodel?

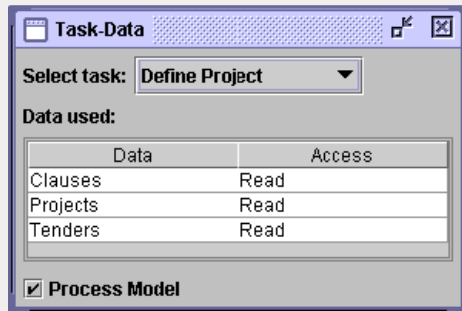
Algoritme: Soek deur die ooreenstemmende kolom van die gegewe data in matriks *P* of *D* en vind dan die aktiwiteit/e in die ooreenstemmende ry/e en ook die toegang wat sal afhang van die waarde/s van  $p_{ij}$  of  $d_{ij}$ .



FIGUUR 2-13: GRAFIESE AFVOER: WATTER AKTIWITEITE 'N DATA GEBRUIK

Vraag: Watter data word gebruik deur 'n aktiwiteit en wat is die tipe toegang in die proses- of datamodel?

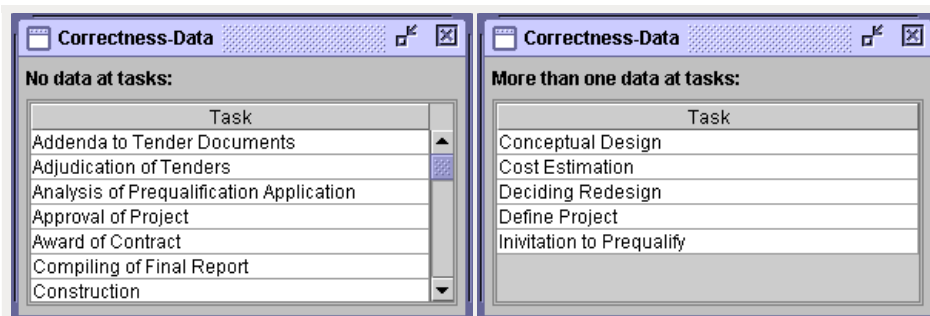
Algoritme: Soek deur die ooreenstemmende ry van die gegewe aktiwiteit in matriks  $P$  of  $D$  en vind dan die data in die ooreenstemmende kolom/me en die toegang afhangend van die waarde/s van  $p_{ij}$  of  $d_{ij}$ .



FIGUUR 2-14: GRAFIESE AFVOER: DATA DEUR 'N AKTIWITEIT GEBRUIK

Vraag: By watter aktiwiteit word daar geen of meer as een data gebruik?

Algoritme: Die som van die ooreenstemmend ry van die gegewe data in matriks  $P$  of  $D$  kan bereken word om sodoende die aantal data te vind.



FIGUUR 2-15: GRAFIESE AFVOER: AANTAL DATA BETROKKE BY AKTIWITEITE

Die volgende vrae aangaande die relasies tussen die twee versamelings kan ook gevra word:

- Is die volgorde van *skryf*- en *lees*-toegange korrek gedefinieer vir die prosesmodel en die datamodel?

- Watter aktiwiteit/e het voor hierdie aktiwiteit *lees-toegang* of *skryf-toegang* tot die data gehad?
- Watter aktiwiteit/e sal na hierdie aktiwiteit *lees-toegang* of *skryf-toegang* van hierdie data hê?
- Watter aktiwiteit/e word beïnvloed deur verkeerde data?

Dit is in die meeste gevalle nodig om 'n tydskuur te koppel aan die uitvoering van aktiwiteite om hierdie vrae te kan antwoord. Die modelle vir hierdie studie is met gerigte grafieke opgestel. Geweegde grafieke moet egter gebruik word om hierdie probleem te kan oplos. Dit is derhalwe nie verder ondersoek nie, maar behoort wel aandag te geniet tydens verdere navorsing.

Met die vrae wat wel bespreek is, kan bepaal word waar daar geen data gedefinieer is nie en waar daar meer as een data gedefinieer is by 'n aktiwiteit, asook die toegange vir beide die proses- en datamodel. Daar kan ook bepaal word of 'n data aangewend word by die korrekte aktiwiteite en andersom, d.w.s. of die aktiwiteit die regte data en toegang tot die data het. Hierdie vrae is belangrik, omdat dit aangewend kan word om te bepaal of die data- en prosesmodel volledig en korrek gedefinieer is t.o.v. die data wat aktiwiteite, hetsy tegniese, administratiewe- of besluitnemingsdata, skep of benodig.

---

### 2.2.2.3 DATA EN GROEP

---

Die verwantskap *AG*, tussen die versameling *Aktiwiteit* en *Groep* is simmetries en kan beskryf word met die naasliggendheidmatriks *M*, soos reeds in *Afdeling 2.2.2.1* gedefinieer is. Dit is weer nodig om van die twee insidensiematrikse gebruik te maak vir die prosesmodel datatoegang, matriks *P* en vir die datamodel datatoegang, matriks *D*, soos reeds gedefinieer in *Afdeling 2.2.2.2* om onderskeid te maak tussen die twee toegangsmodelle. Verder meer is dit ook nodig om die volgende twee matrikse *K* en *L* te definieer, waar *K* die relasie tussen *Groep* en *Data* vir die prosesmodel toegang voorstel en *L* vir dieselfde relasie, maar vir die datamodel toegang:

$$k_{ij} = \begin{cases} 1, & \text{as } \textit{groep}_i \textit{ data}_j \textit{ skryf}, \\ -1, & \text{as } \textit{groep}_i \textit{ data}_j \textit{ lees}, \\ 2, & \text{as } \textit{groep}_i \textit{ data}_j \textit{ lees en skryf}, \\ 0 & \text{andersins.} \end{cases}$$

en

$$l_{ij} = \begin{cases} 1, & \text{as } \textit{groep}_i \textit{ data}_j \textit{ skryf}, \\ -1, & \text{as } \textit{groep}_i \textit{ data}_j \textit{ lees}, \\ 2, & \text{as } \textit{groep}_i \textit{ data}_j \textit{ lees en skryf}, \\ 0 & \text{andersins.} \end{cases}$$


---



Bogenoemde twee matrikse kan bereken word deur van standaard algoritmes vir matriksvermenigvuldiging van die getransponeerde matriks  $M^T$  en matriks  $P$  of  $D$  gebruik te maak. Die enigste verskil is by die punt waar die som van die produkte bereken word, in die direkte algoritme word die matriksprodukt  $C$  van twee  $n \times n$  matrikse  $A$  en  $B$  gedefinieer soos volg:

$$c_{ij} = \sum_{k=1}^n a_{ik} b_{kj} \dots\dots\dots \text{Vgl. 2-1}$$

terwyl vir hierdie toepassing dit gedefinieer word as:

$$c_{ij} = \begin{cases} 1, & \text{as } \exists a_{ik} b_{kj} = 1 \text{ en } \forall a_{ik} b_{kj} \neq -1 \text{ vir } k = 1..n, \\ -1, & \text{as } \exists a_{ik} b_{kj} = -1 \text{ en } \forall a_{ik} b_{kj} \neq 1 \text{ vir } k = 1..n, \\ 2, & \text{as } \exists a_{ik} b_{kj} = -1 \text{ en } \exists a_{ik} b_{kj} = 1 \text{ vir } k = 1..n, \\ 0 & \text{andersins.} \end{cases}$$

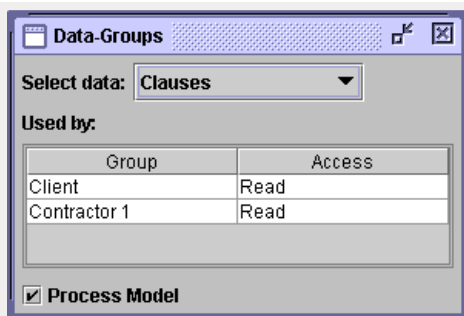
waar

- $\exists$  = eksistensiële kwantifiseerder
- = ten minste een
- $\forall$  = universele kwantifiseerder
- = vir alle

Die volgende vrae kan nou gevra word:

Vraag: Wie gebruik die data en wat is die tipe toegang in die proses- of data-model?

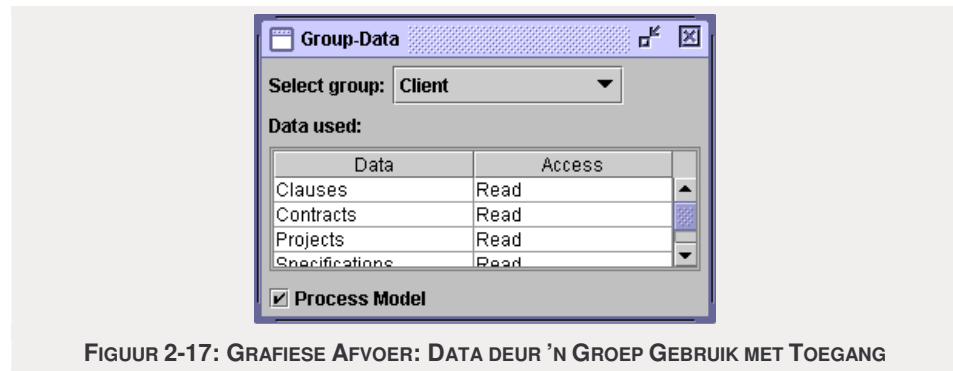
Algoritme: Soek deur die ooreenstemmende kolom van die gegewe data in matriks  $K$  of  $L$  en vind die ry/e van die ooreenstemmende groep/e, so wel as die toegang/e afhangend van die waarde van  $k_{ij}$  of  $l_{ij}$ .



FIGUUR 2-16: GRAFIESE AFVOER: WIE DATA GEBRUIK MET TOEGANG

Vraag: Watter data word deur 'n groep gebruik en wat is die tipe toegang vir die proses- of datamodel?

Algoritme: Soek deur die ooreenstemmende ry van die gegewe groep in matriks  $K$  of  $L$  en vind dan die data van die ooreenstemmende kolom/me, asook die toegang/e wat afhang van die waarde van  $k_{ij}$  of  $l_{ij}$ .



FIGUUR 2-17: GRAFIESE AFVOER: DATA DEUR 'N GROEP GEBRUIK MET TOEGANG

Die volgende vrae kan ook gevra word oor die relasies tussen die versamelings *Groep* en *Data*:

- Watter groep/e het die data voor hierdie groep gebruik en wat is die tipe toegang/e van die groep/e?
- Watter groep/e sal die data na hierdie groep gebruik en wat is hulle toegang/e?
- Watter groep/e sal beïnvloed word deur foutiewe data nadat 'n spesifieke groep toegang tot die data gehad het?

Dit is weereens nodig om van geweegde gerigte grafieke gebruik te maak om hierdie tipe vrae te beantwoord. Vir hierdie studie is dit egter nie geïmplementeer nie en derhalwe kan daar ook nie antwoorde hieroor gelewer word nie. Hierdie is ook 'n gebied wat verder ondersoek moet word.

Met die vrae wat hier gevra en beantwoord word, kan bepaal word of gebruikers toegang tot data het en ook wat die gebruiker se toegang tot die data is. Dit is belangrik om dit te kan bepaal, sodat verseker kan word dat die prosesmodel korrek gemodelleer is.

#### 2.2.2.4 AKTIWITEITE EN GEBEURE

In 'n prosesmodel kan dit gebeur dat 'n aktiwiteit of gebeurtenis onderskeidelik aanleiding gee tot twee of meer gebeurtenisse of aktiwiteite. Daar is twee redes hiervoor, nl.:

- 'n proses kan in parallel uitgevoer word en verder in die prosesmodel weer aansluit of stop terwyl 'n ander tak van die model verder gaan.
- 'n proses kan weer terugkeer na 'n vorige gebeurtenis of aktiwiteit in die model en sodoende siklusse vorm.

Albei hierdie gevalle is belangrik om uit te ken in prosesmodellering aangesien die ontwerper bewus moet wees indien sulke situasies voorkom. Dit is nodig, sodat vrae beantwoord kan word ten opsigte van of aktiwiteite wel in

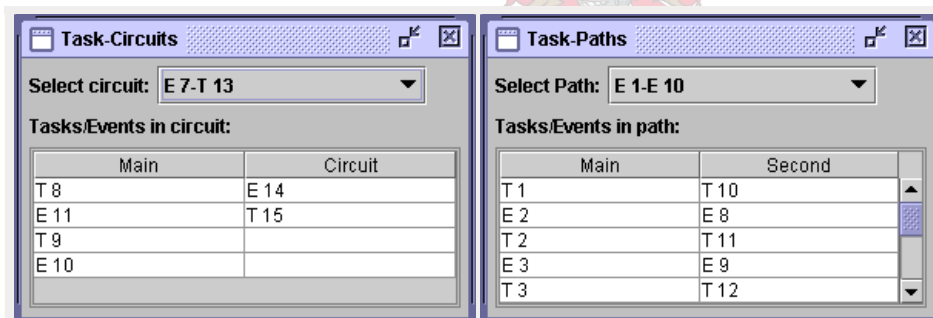
parallel uitgevoer word en of siklusse wel bestaan; wie hierby betrokke is, watter data gebruik word en wie verantwoordelik is indien daar besluit moet word of 'n siklus moet stop of herhaal. Hierdie twee vorms kan as volg bepaal word.

Skep 'n insidensie matriks  $A$  deur van die gerigte grafiek wat gebaseer is op die relasies tussen die versamelings *Aktiwiteit* en *Gebeurtenis* gebruik te maak, met die volgende eienskappe:

$$a_{ij} = \begin{cases} 1, & \text{as die } j\text{te boog weg gerig is van die } i\text{te nodus,} \\ -1, & \text{as die } j\text{te boog na die } i\text{te nodus gerig is,} \\ 0 & \text{as daar geen } j\text{te boog die } i\text{te nodus verbind nie.} \end{cases}$$

Vraag: Watter aktiwiteite en gebeurtenisse is in siklusse?

Algoritme: Loop 'n diepte-eerste soek algoritme, wat aangepas is om vir sterk samehangendheid te toets. Die takke wat gevorm word van die diepte-eerste soektog kan of paaie wees of geslote lyn rye. Die sterk samehangend komponente (takke) wat geïdentifiseer word deur die sterk samehangendheidstoets is geslote lynrye, wat siklusse in die prosesmodel is.



FIGUUR 2-18: GRAFIESE AFVOER: SIKLUSSE EN PAAIE

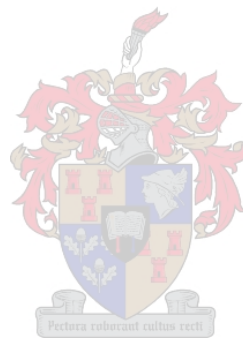
## 2.3 AANWENDING VAN ANALISERINGSPROGRAM

Die komplekse proses wat by konstruksieprojekte voorkom, asook die gefragmenteerde struktuur van die konstruksiebedryf noodsaak 'n goed geïntegreerde inligtingstelsel wat gebaseer word op 'n laevlak detail prosesmodel wat inskakel by hoëvlak generiese modelle. Hierdie gedetailleerde prosesmodelle moet ondersoek kan word vir aspekte soos volledigheid en korrektheid. Die funksionaliteit wat in hierdie studie konseptueel ontwikkel en geïmplementeer is, kan hiervoor gebruik word soos getoon in die voorafgaande afdeling.

By die vele fasette van kwantitatiewe risikobestuur kan hierdie funksionaliteit aangewend word om onder andere te bepaal: waar besluitneming plaasvind; of besluitneming wel ondersteun word met kwantitatiewe risikobestuuraktiwiteite;

die korrekte data aangewend word by 'n kwantitatiewe risikobestuuraktiwiteit tydens die onderskeie projektfases; die regte persone betrokke is by die proses; relevante kwantitatiewe risikobestuur inligting afkomstig van projekte wel onderskep word en of prosesse bestaan hoe om dit na 'n databasis vir historiese projekte te voer.

---



## HOOFSTUK 3

---

### DATABASISONTWIKKELING

---

Organisasies kan suksesvol groei slegs as korrekte inligting beskikbaar is om besluite op te baseer. Meeste organisasies skep groot volumes data wat, indien dit verwerk word, die nodige inligting kan verskaf om besluitneming te ondersteun. Die data moet egter korrek en maklik toeganklik wees. Tradisioneel was die liasseerkabinet 'n plek waar historiese data gestoor is. Met tegnologiese ontwikkelings en 'n toenemende volume data het verskillende vorms van elektroniese databasisse egter die voorkeur stooromgewing vir data geword.

Hierdie hoofstuk het te make met die strukturele ontwikkeling van 'n databasis om tenderinligting te stoor wat vir risikobestuurdoeleindes uitgebrei kan word. In die eerste gedeelte word die verskillende komponente van 'n databasis bespreek en 'n beknopte oorsig word oor die basiese teorie van databasis ontwerp gegee. Die beoordeling van die CSRA-databasis word dan gedoen en 'n voorgestelde ontwerp word bespreek. In die laaste gedeelte word daar gekyk na potensiële uitbreidings van so 'n databasis en die voordele verbonde aan die voorgestelde nuwe ontwerp.

---

#### 3.1 INLEIDING

---

Die term "databasis" is welbekend aan meeste mense, omdat dit orals in die omgewing waarin ons leef voorkom, of dit nou 'n telefoongids is, of 'n liasseerkabinet waar fakture gestoor word en of dit 'n elektroniese stoorruimte is waar inkomstebelastingdata versamel word. Tog is daar, veral by elektroniese stoor-

---

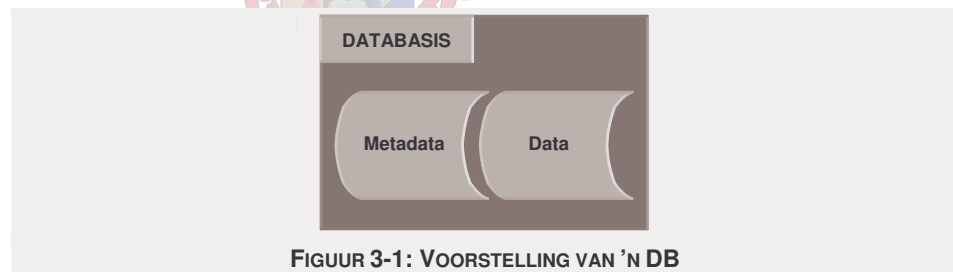
ruimtes, 'n mate van onduidelikheid t.o.v. wat al die komponente van 'n databasis is en wat die funksies daarvan is.

### 3.1.1 DATABASIS

'n Databasis (DB) stel 'n gedeelte van die wêreld waarbinne ons leef voor. Die gedeelte van die wêreld wat dit voorstel kan as die besprekingsruimte gesien word. Binne hierdie databasis of dan model van ons besprekingsruimte word 'n versameling van koherente data gestoor wat een-of-ander doel of funksie vir die mense in die besprekingsruimte het. [O'Neil (2001)]

'n Databasis is egter nie die program wat die gebruiker toegang tot die data verleen of die program wat die data beheer nie. Die databasis behoort eerder gesien te word as die inhoud van 'n elektroniese liasseerkabinet met 'n goeie struktuur wat op permanente elektroniese media gestoor word. Hierdie elektroniese liasseerkabinet word in die vorm van 'n datalêer gestoor.

Die datalêer bestaan uit twee gedeeltes, nl. metadata (data oor data) wat die struktuur van die model van die besprekingsruimte beskryf en dan die data of die rou feite wat die model bevolk. Dit kan soos in *Figuur 3-1* voorgestel word. Verskillende strukture is beskikbaar waarmee die besprekingsruimte gemodelleer kan word, nl. hiërargiese, netwerk-, relasionele en ook objekrelasionele struktuur. [Rob (1997)]



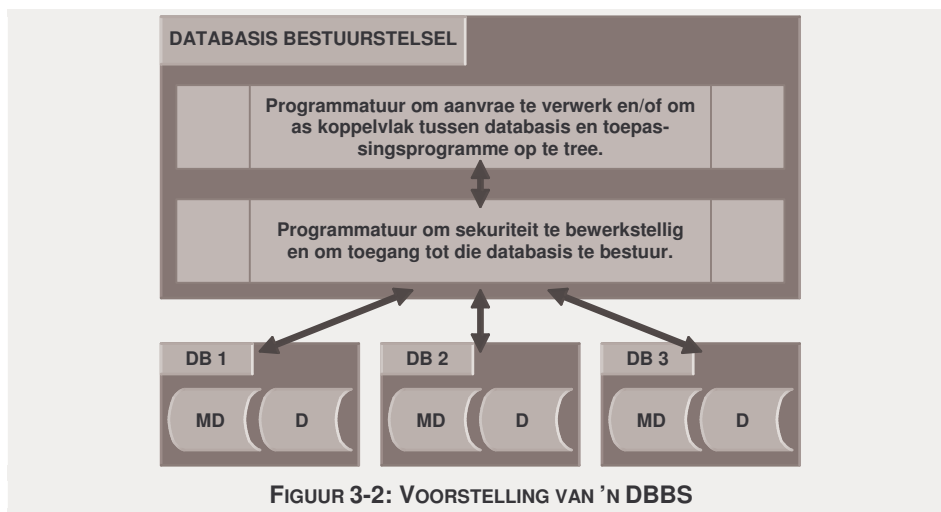
FIGUUR 3-1: VOORSTELLING VAN 'N DB

Vir hierdie studie is van die relasionele struktuur wat deur *Codd (1970)* ontwikkel is, gebruik gemaak, omdat dit tans die struktuur is wat die meeste gebruik word en die sagteware wat beskikbaar was aan die outeur om dit implementeer. Verder is dit ook 'n struktuur waarmee die besprekingsruimte van hierdie studie realisties en eenvoudig gemodelleer kan word.

### 3.1.2 DATABASIS BESTUURSTELSEL

'n Databasis bestuurstelsel (DBBS) is 'n versameling van komplekse programme om een of meer databasisse effektief en doeltreffend te beheer. Die DBBS word gebruik om databasisse te definieer, konstrueer, manipuleer en ook vir toegangsbeheer.

Tydens die definiëring van 'n databasis word datatipes, die struktuur van die data en die beperkings van die besprekingsruimte vasgelê in die metadata van die databasis, soos deur die ontwerper konseptueel en logies beplan is. Wanneer 'n databasis dus ontwerp word is dit nie die DBBS wat ontwerp word nie, maar die struktuur van die databasis. Konstruering vind plaas wanneer die DBBS die databasis op een of ander permanente elektroniese medium stoor, soos bv. 'n hardeskyf, herskryfbare kompakte skyf (KS) of starskyfie. Manipulering van die databasis vind plaas as data van die besprekingsruimte gestoor, verander, verwyder of gelees moet word. Die DBBS beheer ook wie toegang tot verskillende data in die databasis het, asook wie aan die struktuur van die databasis mag verander. In *Figuur 3-2* word die interaksie tussen 'n DBBS en drie DB'se getoon.



Die DBBS kan as 'n vlak tussen die databasis en die gebruiker gesien word. Hierdie vlak versteek die kompleksiteit wat verbonde is aan die bestuur van 'n databasis. Algemene industriële DBBS'e wat voorkom is *Oracle Server*, *DBase 2*, *MS SQL Server*. Hierdie DBBS'e is gebaseer op die relasionele model van data, alhoewel die funksionaliteit van objek-relasionele modelle ook besig is om beskikbaar te word in bogenoemde DBBS'e. Vir hierdie studie word van *MS Access* gebruik gemaak, omdat dit voldoende is om konseptueel die gedagtes oor te dra van hierdie studie.

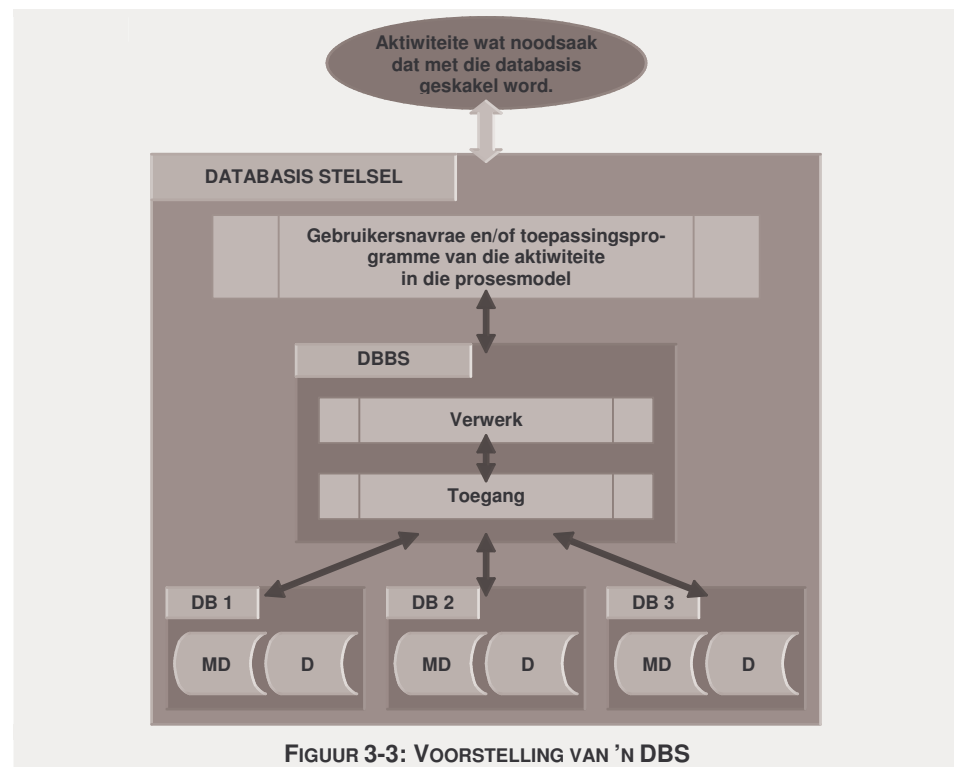
### 3.1.3 DATABASIS STELSEL

Die databasis stelsel (DBS) verwys na die organisasie van die komponente wat betrokke is by 'n databasis. Die stelsel bestaan uit drie hoof-komponente, nl. hardeware, programmatuur en data, indien die gebruikers, programmeerders en die prosesse wat die data gebruik nie hierby ingesluit word nie.

Die hardeware kan bestaan uit een of meer rekenaars, hetsy persoonlik of hoofraam, asook die randapparatuur wat daarmee saamgaan. Die programmatuur bestaan uit: die bedryfstelsel wat nodig is om al die apparatuur met mekaar

te laat skakel; die DBBS en verdere toepassingsprogramme wat met die DBBS skakel om die data te manipuleer. Die data is natuurlik die rou informasie van die besprekingsruimte soos dit in die databasis gestoor is, asook die metadata wat struktuur aan die databasis verleen.

In *Figuur 3-3* word 'n DBS getoon met die interaksie tussen die gebruikers/programmeerders. Die prosesse kan gesien word as die aktiwiteite binne die besprekingsruimte wat veroorsaak dat daar data in die databasis gestoor of daaruit onttrek word. Hierdie deel sluit aan by die datavloei en prosesmodellering van aktiwiteite wat data skep en benodig, soos in *Hoofstuk 2* bespreek is.



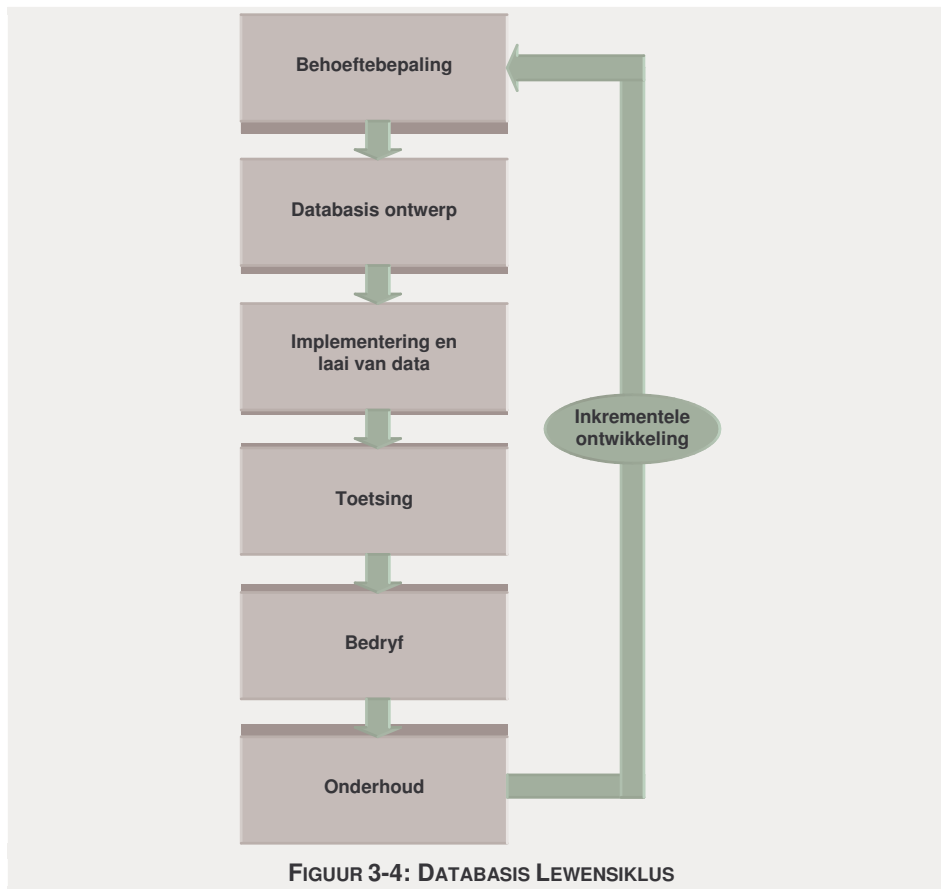
'n Verskeidenheid van databasis stelsels bestaan, afhangend van die aantal gebruikers, of dit op 'n hoofraamrekenaar of op 'n persoonlike rekenaar gehuisves word, of dit gesentraliseer of desentraliseer is en die doel waarvoor dit aangewend word, hetsy 'n transaksiedatabasis of 'n besluitsteundatabasis. Vir hierdie studie kan die databasis stelsel gesien word as 'n enkelgebruiker, gesentraliseerde databasis vir besluitsteundoelindes wat op 'n persoonlike rekenaar geïmplementeer is.

## 3.2 DATABASIS LEWENSIKLUS

Die elementêre lewensiklus van 'n databasis kan in ses fases verdeel word, nl. behoeftebepaling; ontwerp; implementering; toetsing; bedryf en onderhoud, soos in *Figuur 3-4* getoon word. Die fokus van hierdie gedeelte van die studie is die verbetering van die kwaliteit van data deur die databasis waarin dit gestoor word, inkrementeel te verbeter. Daar is dus slegs na die eerste vier fases



gekyk, omdat die wyse waarop data in 'n databasis gestoor word direk die kwaliteit van die data kan beïnvloed as die databasis nie behoorlik ontwerp is nie. Laasgenoemde twee fases het grotendeels te make met die bestuur van 'n funksionerende databasis en word dus nie beskou nie.



In die toepassingvoorbeeld, wat later gebruik word om die basiese konsepte van databasisse te verduidelik en om 'n algemene databasismodel te ontwikkel vir 'n kostedatabasis, was daar alreeds 'n bestaande databasis beskikbaar gewees. Daar word derhalwe nie veel aandag aan die formele behoeftebepalingfase geskenk nie, omdat die bestaande databasis alreeds aan die aanvanklike behoeftes voldoen het. Addisionele behoeftes wat wel geplaas word op die herontwikkelde databasis spruit voort uit die analise tegnieke vir risikobestuur wat in *Hoofstuk 4* bespreek word en waar uitbreidings aan die databasis in *Hoofstuk 6* geïdentifiseer word.

Vir gevalle waar daar nie reeds bestaande databasisse is nie, sal 'n formele behoeftebepaling gedoen moet word, soos in die toepassing van die nuwe omgewing wat in *Hoofstuk 7* bespreek word. Hier moet dus van nuuts af die tradisionele wie, waar, wat, wanneer, waarvoor en waaroms bepaal word voor die databasis ontwikkel kan word. Verder moet daar ook, soos in *Hoofstuk 6* bespreek word, ondersoek ingestel word na die analise tegnieke wat gebruik sal word om die kosteberekeninge mee uit te voer, asook die data wat hiervoor

bebenodig word, om te verseker dat daar voorsiening gemaak word in die databasis om hierdie inligting te kan stoor.

Om die veranderlikheid van konstruksiekostedata (wat lei tot onsekerheid) te verminder is dit nodig om ook beskrywende inligting oor die data te onderskep, sodat die omgewing waarbinne die data ontstaan gekarakteriseer kan word; sodoende kan moontlike risiko's in kosteberamings weens variasie in die projekomgewing of ander eksterne faktore verlaag word. Die behoeftebepaling vir die ontwikkeling van die databasis strek dus effens wyer as die “wie, waar, wat, wanneer, waarvoor en waaroms”. 'n Formele ondersoek moet dus na beskrywende inligting van kostedata uitgevoer word om die invloedrykste data te onderskep. Die tegnieke wat hiervoor gebruik kan word, word in *Hoofstuk 6* bespreek.

Die volgende stap in die lewensiklus is databasis ontwerp. Databasis ontwerp bestaan uit 'n konseptuele, logiese en fisiese ontwerp. Die konseptuele ontwerp het te make met die uitleg van tabelle en die verwantskappe daartussen. Die logiese ontwerp het te make met hoe die konseptuele ontwerp met die DBBS oorgedra word om die metadata te vorm. Tydens die fisiese ontwerp word daar gekyk na die apparatuur waarop die DBBS geïmplementeer word. Die ontwerp van die databasis word in *Afdeling 3.5* bespreek.

Tydens die implementering en laai van data word die DBBS geïnstalleer, die databasis word gedefinieer op grond van die logiese ontwerp en die data van die bestaande databasis word aangepas sodat dit gelaai word in die voorgestelde model. Hierdie gedeelte word in *Afdeling 3.6* bespreek.

Vir die toetsingsfase word vir hierdie studie gekyk of die data wat vanaf die bestaande model oorgedra is, wel korrek gedoen is. Verder moet daar ook gekontroleer word om seker te maak dat ten minste die versoeke wat op die bestaande databasis gedoen kon word ook op die voorgestelde databasismodel gedoen kan word. Dit word in *Afdeling 3.7* bespreek.

Alhoewel dit lyk asof daar 'n sogenaamde waterval volgorde, waar daar vanaf die een voltooid fase na die volgende beweeg word, gevolg word, is dit nie werklik so nie. Die lewensiklus van 'n databasis is 'n iteratiewe proses waar daar heelyd verbetering en wysigings aangebring word om aan te pas en te voldoen aan die behoeftes van die gebruikers. Hierdie inkrementele ontwikkeling kan getoon word by die CSRA-databasis.

In *Afdeling 7.1* word die ontwikkelingsvlak verskille tussen die twee voorbeeld omgewings bespreek. Hier kom dit sterk na vore dat daar alreeds 'n noemenswaardige inkrementele ontwikkeling van die CSRA-databasis plaasgevind het. Met die nuwe probabilistiese tegnieke wat toegepas kan word op die

data, ontstaan daar egter nog verdere behoeftes om die kwaliteit van die data te verbeter en die data-onderskepping uit te brei, dus is 'n volgende vlak van die ontwikkeling van die databasis nodig. Indien daar egter nie data beskikbaar was om te identifiseer watter items 'n noemenswaardige invloed het nie, sou dit nie moontlik gewees het om te kon identifiseer watter uitbreidings nodig sou wees nie en of daar enige uitbreidings nodig is nie. Hierdie inkrementele ontwikkeling word verder in *Afdeling 5.1* bespreek. Besluitsteun databasisontwikkeling moet dus as evolusionêr gesien word, soos nuwe aspekte geïdentifiseer word, moet die databasis uitgebrei kan word.

---

### 3.3 TEGNIESE AGTERGROND

---

Dit is nodig om 'n paar basiese databasisbegrippe te definieer, voordat daar gekyk kan word na die herontwerp en uitbreiding van die bestaande databasis. Verder meer is dit nodig om te kyk na hoe entiteitsverwantskapsdiagramme en normaalvorms gebruik kan word om die databasismodel te herontwerp. 'n Kort beskrywing word ook gegee van die taal wat gebruik word om met 'n DBBS te kommunikeer.

---

#### 3.3.1 BASIESE BEGRIPPE

---

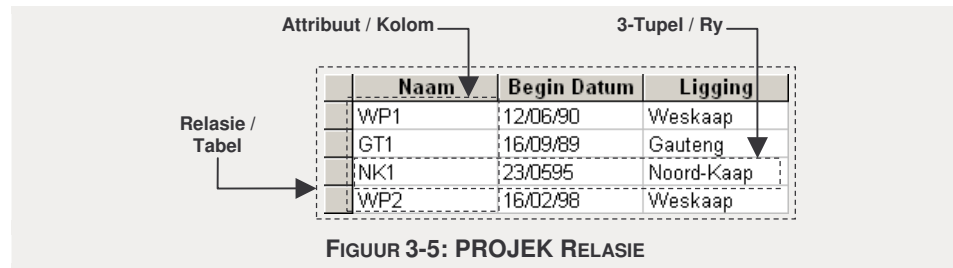
Twee vorms van benamings word gebruik vir databasisbegrippe, een is afkomstig uit die volksmond en die ander vanuit die teorie van databasisse. Vir eersgenoemde geval word daar verwys na tabelle, kolomme en rye, terwyl die ander dienooreenkomstig verwys na relasies, attribute en entiteite. Om verwarring te vermy sal van eersgenoemde benamings gebruik gemaak word, behalwe in hierdie afdeling waar die definisies gegee word vanuit die raamwerk van laasgenoemde. [Date (1995), Elmasri (1994), O'Neil (2001)]

**Definisie 3-1:** 'n **Entiteit** kan enigiets wees waarvan ons data wil versamel. Dit kan vir hierdie studie 'n projek, kontrakteur, raadgevende ingenieur, item of 'n tender wees. •

**Definisie 3-2:** Entiteite kan sekere karakteristieke eienskappe hê wat **attribute** genoem word. •

Die entiteit **PROJEK** kan 'n naam, begin datum en ligging hê. Elk van hierdie attribute het 'n domein of ongeveer die ekwivalent hiervan in databasisterme 'n datatipe met 'n beperkingsreël. 'n Projek se begindatum kan byvoorbeeld na die 1ste Januarie 1987 wees. Vir so 'n geval sal die datatipe dan "Datum" wees en die beperkingsreël sal dan bepaal dat slegs datums na die 1ste Januarie in die attribuut gestoor kan word. As 'n entiteit drie attribute het is dit 'n 3-tupel, soos in *Figuur 3-5* getoon word.

---



**Definisie 3-3:** 'n **Relasie** is die subversameling van al die voorkomste van geldige n-tupels van die entiteite onder beskouing (sien *Definisie 2-9*). •

Wanneer dus na 'n relasionele databasis verwys word, word juis na hierdie aspek van die relasie wat binne in tabelle voorkom verwys en nie verkeerdelik na die "relasies" wat tussen die verskillende tabelle voorkom nie. Hierdie "relasies" tussen tabelle is sogenaamde verwantskappe.

### 3.3.2 SLEUTELS

Die beperking, of beheerde duplisering van data om inkonsekwentheid te voorkom, is die hoofdoel waarom van relasionele databasisse gebruik gemaak word. Om bogenoemde te kan doen behoort 'n relasionele DBBS twee belangrike reëls toe te pas, nl. entiteit- en verwysingsintegriteit. Entiteitsintegriteit verwys na die uniekheid van inskrywings in 'n relasie/tabel, terwyl verwysingsintegriteit bepaal dat 'n tupel in een relasie wat verwys na 'n ander relasie moet verwys na 'n bestaande tupel in die relasie waarna verwys word. Om hierdie twee integriteitsreëls te kan toepas word van sogenaamde sleutels gebruik gemaak.

**Definisie 3-4:** 'n **Sleutel** maak van funksionele afhanklikheid gebruik om entiteitsverwantskappe te definieer, m.a.w. attribuut  $B$  is funksioneel afhanklik van attribuut  $A$  as 'n waarde van attribuut  $A$  een en slegs een waarde van attribuut  $B$  bepaal. •

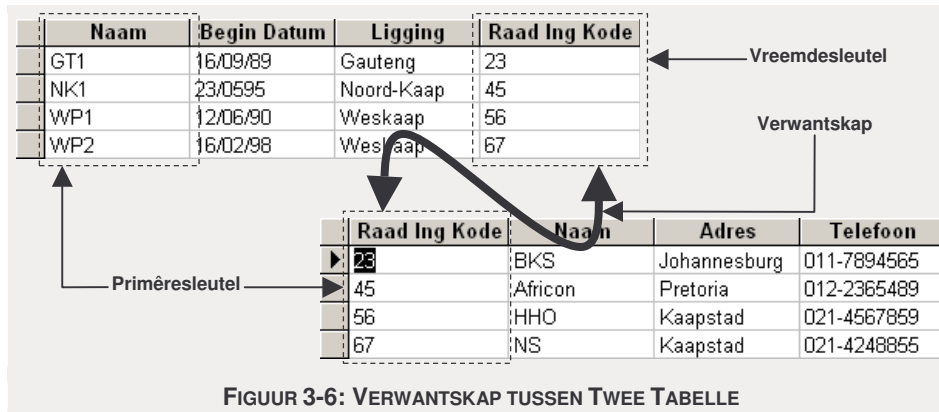
In hierdie studie sal daar slegs van drie tipes sleutels gebruik gemaak word, nl. supersleutel, primêre sleutel en die vreemde sleutel, alhoewel daar heelwat verskillende tipes sleutels voorkom in die databasis terminologie.

**Definisie 3-5:** Die **Supersleutel**, waarvan daar meer as een kan wees in 'n relasie, is 'n attribuut of versameling van attribute wat elke entiteit uniek bepaal in die relasie. •

**Definisie 3-6:** Die **Primêre sleutel**, is die gekose supersleutel wat in die relasie gebruik word. •

**Definisie 3-7:** Die **Vreemde sleutel** is 'n attribuut in een relasie waarvan die waarde moet ooreenstem met die primêre sleutel in 'n ander relasie. •

Met 'n supersleutel is dit dus moontlik om entiteitsintegriteit na te kom en met die primêre sleutel en die vreemde sleutel is dit moontlik om entiteitsverwantskappe vas te lê, soos in *Figuur 3-6*.



### 3.3.3 EV-DIAGRAMME

Die entiteitverwantskapsdiagramme (EV-diagramme) word gebruik om die besprekingsruimte konseptueel te modelleer. Hierdie wyse van modellering is in 1976 deur *Chen (1976)* aan die wêreld bekend gestel in 'n artikel "The Entity-Relationship Model: Towards a Unified View of Data". Dit is nodig om 'n paar van die konsepte wat in hierdie studie gebruik word te definieer om verwarring te voorkom, alhoewel daar groot verwantskappe tussen die relasionele model van databasis en EV-diagramme voorkom. [Benyon (1990), Halpin (2001)]

#### 3.3.3.1 ENTITEIT

'n Entiteit in 'n EV-diagram verwys na die entiteitsversameling soos dit voorkom in die relasionele model. Hier is dus 'n subtiele verskil tussen hoe die twee begrippe gebruik word in EV-diagramme en die relasionele model. Eersgenoemde verwys na 'n groep, terwyl die ander slegs na 'n enkele item in die relasie verwys.

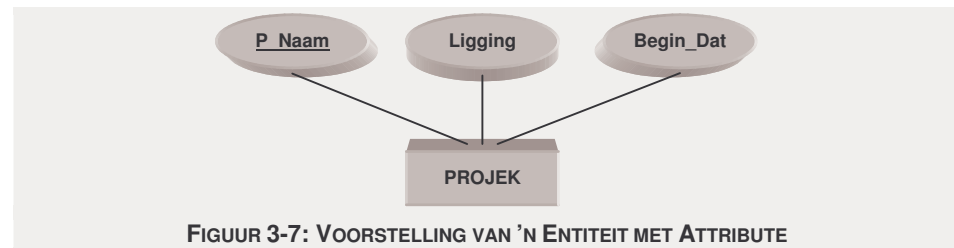
'n Entiteit in die EV-omgewing word diagrammaties voorgestel met 'n reghoek met die entiteit se naam daarin, soos in *Figuur 3-7* getoon word. Soms is dit egter nie moontlik om die volledige naam van die entiteit te gebruik nie en dan word van afkortings gebruik gemaak, bv. **Begin\_Dat**, **Hoeveelh** en **RAADGEWEND** wat die afkorting van "begin datum", "hoeveelheid" en "raadgewende ingenieur" onderskeidelik is.

Twee variasies van die gewone entiteit kom voor, nl. swak entiteite en saamgestelde entiteite. 'n Swak entiteit se bestaan is afhanklik van 'n ander

entiteit waarmee dit 'n verwantskap het, *Figuur 3-23*. 'n Saamgestelde entiteit se primêre sleutel is 'n samestelling van die primêre sleutels van die twee entiteite wat dit verbind, *Figuur 3-24*.

### 3.3.3.2 ATTRIBUUT

Attribute beskryf die eienskappe van 'n entiteit. Attribute het 'n versamelingsruimte, wat ooreenstem met 'n datatipe en beperkingsreël. Dit kan diagrammadies voorgestel word met 'n ovaal en die attribuut se naam daarin, soos in *Figuur 3-7* getoon word.

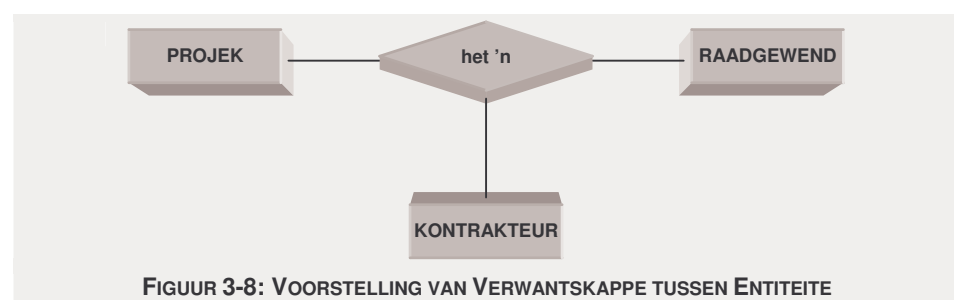


FIGUUR 3-7: VOORSTELLING VAN 'N ENTITEIT MET ATTRIBUTE

'n Verskeidenheid van tipes attribute bestaan, maar daar sal egter met hierdie algemene beskrywing volstaan word. Indien een van die attribute 'n primêre sleutel is, word dit voorgestel deur dit te onderstreep, soos in die geval van P\_Naam.

### 3.3.3.3 VERWANTSKAPPE

'n Verwantskap is 'n assosiasie tussen twee entiteite. Daar bestaan byvoorbeeld 'n verwantskap tussen die raadgewende siviele ingenieursfirma wat 'n projek bestuur en die betrokke projek. Verwantskappe word gewoonlik so benoem dat dit die verwantskap beskryf. Dit word diagrammadies voorgestel met 'n diamant, soos in *Figuur 3-8*.



FIGUUR 3-8: VOORSTELLING VAN VERWANTSKAPPE TUSSEN ENTITEITE

'n Verwantskap kan verskillende grade hê om te bepaal hoeveel entiteite betrokke is. Die mees algemene vorm is 'n binêre verwantskap waar twee entiteite betrokke is, maar daar kom ook enkel- en veelvuldige verwantskappe voor.

### 3.3.3.4 KONNEKTEERDHEID

Konnekteerdheid word gebruik om 'n verwantskap te klassifiseer. Drie tipes kom voor, nl. een-tot-een, een-tot-baie en baie-tot-baie. Aangesien eersge-

noemde en laasgenoemde nie in databasisse behoort voor te kom nie, omdat dit onderskeidelik nie van nut is nie en nie korrek is nie, word slegs die een-tot-baie konnekteerdheid bespreek.

Aan die hand van *Figuur 3-9* kan gesien word dat 'n raadgewende ingenieursfirma meer as een projek kan bestuur, daarom die *M* wat langs **PROJEK** staan. 'n Projek kan egter, argumentsonthalwe slegs een raadgewende ingenieursfirma hê wat dit bestuur. Dit word aangedui deur die *1* wat langs **RAADGEWEND** staan.



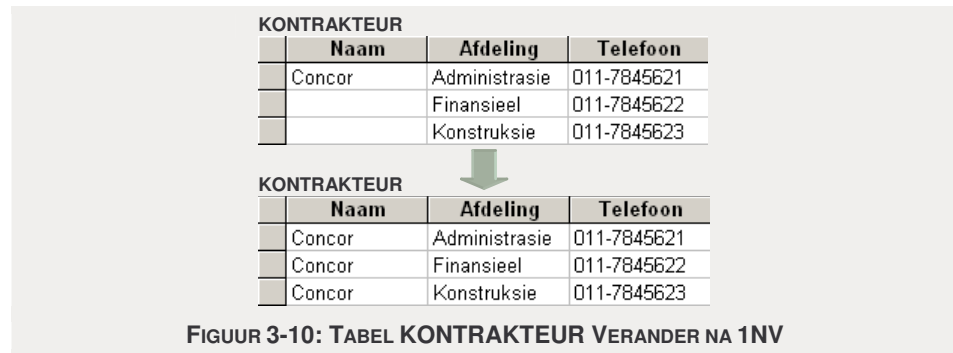
### 3.3.4 NORMAALVORMS

Alhoewel 'n konseptueel goed ontwerpte databasis m.b.v. die EV-diagramme veronderstel is om 'n goeie relasie/tabel struktuur te lewer is dit nie noodwendig so nie. Anomalieë weens onnodige dupliserings kan steeds voorkom. Die proses van normalisering, wat ook deur *Codd (1970)* ontwikkel is, kan gebruik word om te toets of die ontwerp wel voldoen aan die nodige vereiste om anomalie te voorkom.

Die proses van beoordeling bestaan uit drie toetse, nl. Eerste, Tweede en Derde Normaalvorm. 'n Vierde toets wat later ontwikkel is deur *Codd* en *Boyce* staan bekend as *Boyce-Codd* Normaalvorm (BCNV). Hoër normaalvorms is ook beskikbaar, maar vir hierdie studie word volstaan met die BCNV wat beskou word as voldoende om onnodige duplisering te beperk tot 'n minimum.

#### 3.3.4.1 EERSTE NORMAALVORM

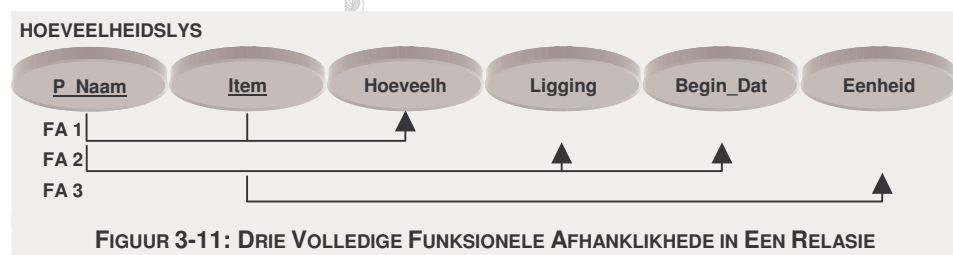
Eerste Normaalvorm (1NV) is 'n voor die hand liggende toets wat bepaal dat attribute slegs atomiese waardes mag hê om sodoende herhalende groepe te voorkom. Dit kan aan die hand van die volgende voorbeeld verduidelik word. In *Figuur 3-10* het die tabel **KONTRAKTEUR**, waar die kontrakteur se naam, afdeling en telefoonnommer in staan, 'n **Naam**, **Afdeling** en **Telefoon** attribuut. In eersgenoemde figuur is dit duidelik dat vir 'n enkele naam daar groepe met dieselfde naam voorkom wat herhaal word. Dit is dus nie in 1NV nie. Vir die tweede geval het elke ry nou 'n kontrakteur, al is dit dieselfde ene. Dit voldoen nou aan die vereiste van 1NV.



### 3.3.4.2 TWEDE NORMAALVORM

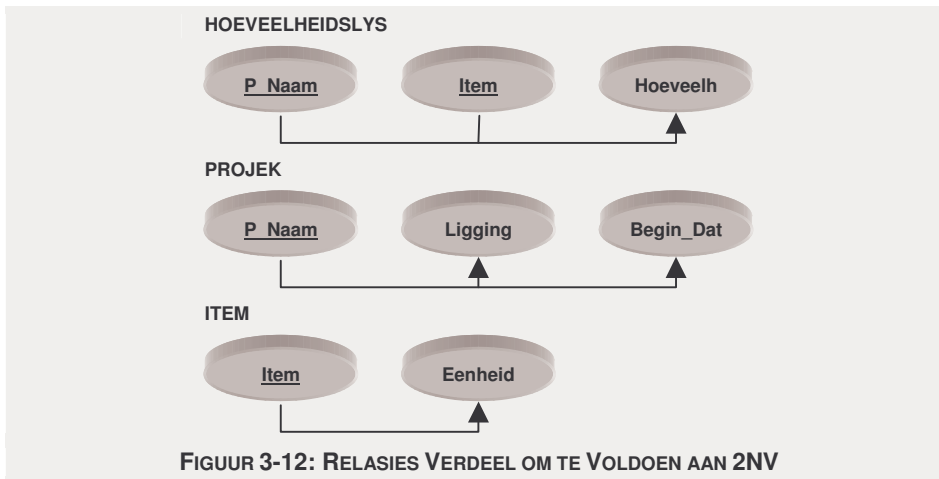
Tweede Normaalvorm (2NV) is gebaseer op die konsep van volledige funksionele afhanklikheid (FA), soos in *Definisie 3-8* verduidelik word. Indien alle attribute volledig funksioneel afhanklik is van die primêre sleutel dan is die relasie/tabel in tweede normaalvorm.

**Definisie 3-8:** 'n **Funksionele afhanklikheid**,  $X \rightarrow Y$ , is volledig funksioneel indien die verwydering van enige attribuut  $a$  van  $X$ , veroorsaak dat die afhanklikheid nie meer bestaan nie. •



In *Figuur 3-11* kan gesien word dat die relasie/tabel HOEVEELHEIDSLYS 'n saamgestelde primêre sleutel, P\_Naam en Item het. Slegs die Hoeveelh attribute word volledig funksioneel bepaal deur hierdie primêre sleutel, terwyl die ander attribute deur subversamelings van die primêre sleutel volledig funksioneel bepaal word. Hierdie relasie/tabel behoort dus verdeel te word in drie relasies soos in *Figuur 3-12* getoon word.

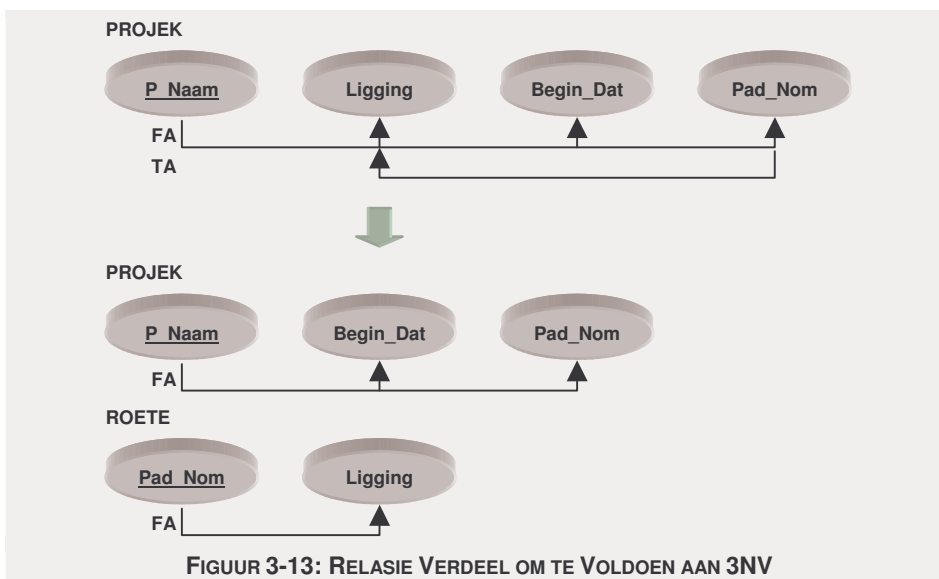




### 3.3.4.3 DERDE NORMAALVORM

Derde Normaalvorm (3NV) bepaal dat geen transitiewe afhanklikheid mag voorkom in die relasie/tabel nie tensy die primêre sleutel deur 'n nie-primêre sleutel bepaal word, waar transitiwiteit reeds bespreek is in *Definisie 2-14*.

In *Figuur 3-13* kan gesien word dat in die relasie PROJEK die primêre sleutel P\_Naam, die ander attribute bepaal. Verder kan ook gesien word dat die attribute Pad\_Nom, die attribute Ligging van die projek bepaal. Dit is dan weens hierdie afhanklikheid dat hier 'n transitiewe afhanklikheid ontstaan,  $P\_Naam \rightarrow Pad\_Nom$  en  $Pad\_Nom \rightarrow Ligging$  dus  $P\_Naam \rightarrow Ligging$ . Die relasie PROJEK is dus nie in 3NV. Onnodige duplisering kan weereens voorkom word deur die relasie in twee tabelle op te breek. Die relasies wat dan gevorm word is in 3NV.



### 3.3.4.4 BOYCE-CODD NORMAALVORM

Boyce-Codd Normaalvorm (BCNV) is slegs 'n sterk vorm van 3NV, deurdat die laaste gedeelte van die bepaling van 3NV weggelaat word. BCNV bepaal dus

slegs dat geen transitiewe afhanklikheid mag voorkom in die relasie/tabel nie, dit maak nie voorsiening dat die primêre sleutel bepaal kan word deur 'n ander attribuut nie.

---

### 3.3.5 SEQUEL

---

Die rekenaartaal wat gebruik word om opdragte aan 'n relasionele DBBS te gee staan bekend as *SQL* (“Structured Query Language”). *SQL* is 'n nie-prosedurele verklarende (“declarative”) taal wat nie stapsgewys uitgevoer word soos tradisionele rekenaartale soos *Fortran*, *C* en *Modula-2* nie [Adams (1993)]. By *SQL* word die tradisionele beheerstrukture soos “if” en “while” ook nie gevind nie. Die ooglopende verskille tussen *SQL*, wat 'n vierde generasie taal is en ander derde generasie tale is dat by *SQL* dit nie nodig is om te sê hoe 'n taak gedoen moet word nie, maar slegs wat gedoen moet word.

Die verskil kan verduidelik word aan die hand van 'n tabel wat gerangskik moet word. Indien 'n stel data van groot na klein georden moet word, is dit nodig by 'n derde generasie taal om die datastruktuur van die data te ken om sodoende 'n algoritme te ontwikkel wat dan die sortering kan uitvoer. By 'n vierde generasie taal kan egter net gesê word wat gedoen moet word. Geen kennis van die datastruktuur van die data is nodig nie en daar hoef ook nie 'n algoritme ontwikkel te word nie.

*SQL* voorsien in die drie ideale behoeftes van 'n DBBS taal, nl.:

- Datadefiniëring – skep van databasis en die struktuur van die tabelle;
- Databestuur – opdragte om data by te voeg, te verwyder, te korrigeer en te lees;
- Datanavraag – opdragte beskikbaar om data te soek volgens gegewe kriteria.

Dit is ook 'n gebruikersvriendelike taal wat gestandaardiseer is. Die mees onlangse standaard wat deur die ANSI (“American National Standards Institute”) aanvaar word, staan bekend as *SQL-92*. In *Aanhangsel C* en *Aanhangsel D* word die *SQL*-navrae wat gebruik is vir die omskakeling van die bestaande padkonstruksiekostedatabasis na die voorgestelde databasismodel getoon.

---

## 3.4 EVALUASIE

---

In enige databasis word dit verwag dat daar onakkuraathede in die data sal voorkom. Korrekte en betroubare data word egter benodig indien analyses wat sinvolle resultate moet lewer, uitgevoer word, derhalwe is dit nodig om ondersoek in te stel na die oorsprong van moontlik foute. Onakkuraathede kan ontstaan weens swak spesifisering, data wat verkeerdelik in die databasis

---

ingevoer word of weens 'n databasismodel wat nie konsekwentheid verseker nie.

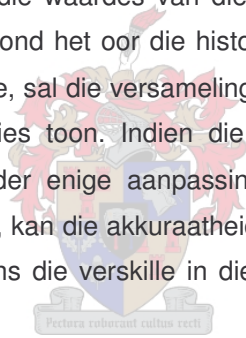
---

### 3.4.1 SPESIFISERING

---

Die padkonstruksiekostedatabasis wat hier beskou word, word hoofsaaklik vir detailkosteberekeninge gebruik. Die hoeveelheidslyste vir sulke berekening is gewoonlik afgebreek tot op die detailvlak van items wat benodig word om die projek uit te voer. Weens die spesifieke beskrywing van elke detail item en die kostes wat daarby ingesluit word, kan dit maklik gebeur dat daar subtiele variasies in die beskrywings van dieselfde item tussen projekte is, indien daar nie streng standaardspesifikasies is nie. Hierdie subtiele variasies, byvoorbeeld wanneer die gratis trekafstand vir dieselfde item tussen twee projekte met 500m verskil, het 'n invloed op die eenheidtarief van die item. Hierdie probleem strek egter verder as net verskille in die kostes wat ingesluit word. Verskille in die opmetingsmetodes wat voorgeskryf word, of soms nie voorgeskryf word nie kan ook lei tot verskille in die koste van items [Abudayyeh (1991), (Birrell (1980))].

Vir die kosteberamer wat slegs na die waardes van die eenheidtarief van verskillende projekte kyk en nie agtergrond het oor die historiese projekte van waar die eenheidtariewe afkomstig is nie, sal die versameling eenheidtariefdata vir ooreenkomstige items groot variasies toon. Indien die gemiddeldes van hierdie data netso gebruik word, sonder enige aanpassings om vir hierdie verskille tussen projekte te kompenseer, kan die akkuraatheid van die berekende koste groot onsekerheid inhou weens die verskille in die spesifisering van die items.



Die klaarblyklike oplossing is dus streng gespesifiseerde standaarditems met 'n beskrywing van die kostes wat ingesluit is by die item en ook die metode van opmeting wat gevolg moet word. Dit is egter nie altyd die oplossing nie. Indien die items te spesifiek gedefinieer word kan dit gebeur dat daar 'n groot aantal standaarditems ontstaan en dat by elke nuwe projek daar heelwat nie-standaarditems bygevoeg moet word. In die databasis moet daar dus 'n groot verskeidenheid standaarditems gestoor word met ylbevolkte eenheidtariewe vir elke item en nie-standaarditems met slegs enkel eenheidtariewe.

Die teenoorgestelde is natuurlik ook waar, indien die items te generiese gespesifiseer word kan dit gebeur dat daar verskillende aktiwiteite en dus kostes ingesluit word by dieselfde item tussen verskillende projekte. Dit sal weer dieselfde probleme veroorsaak, alhoewel minder so, as wanneer daar geen standaardspesifikasies is nie.

'n Fyn balans behoort dus gehandhaaf te word tussen te generies en te spesifieke standaardspesifikasies. In die geval van die CSRA-spesifikasies kom dit

---

voor asof die samestellers suksesvol was om die spesifikasies generies genoeg op te stel sodat die nie-standaarditems tot enkele items per projek beperk is. Hierdie spesifikasies het egter nie oornag tot stand gekom nie. Dit het deur 'n inkrementele ontwikkeling gegaan, soos in *Afdeling 7.1.1* beskryf word, netsoos die databasis self ook inkrementeel ontwikkel moet word. Hierdie proses is ook nog nie voltooi nie. Die CSRA-spesifikasies is ongeveer na 'n dekade vervang met die meer omvattende spesifikasies.

Die sukses van 'n eenheidtariefdatabasis wat gebruik kan word vir die voorspelling van eenheidtariewe in detailkosteberekening, is dus onder andere afhanklik van die kwaliteit van die items se spesifikasies wat in die databasis gestoor word. Indien spesifikasies nie sterk gedefinieer is nie kan dit tot groot onsekerhede in die kosteberekening lei of indien te streng gespesifiseer kan dit veroorsaak dat daar nie genoegsame data is om realistiese beramings uit te voer nie.

---

### 3.4.2 DATA

---

Foute in die data kan ontstaan deur foutiewe data wat korrek in die databasis ingevoer word of deur korrekte data wat foutief in die databasis ingevoer word. Eersgenoemde geval kan veroorsaak word deur datavloei wat ontoereikend is. In die geval van die databasis onder beskouing is byvoorbeeld gevind dat feitlik presies dieselfde inligting vyf keer heronderskep word. Hierdie heronderskappings kan veroorsaak dat foute insluip en netso oorgedra word na die databasis.

Laasgenoemde geval kan weer ontstaan wanneer korrekte data wat ontvang word in die databasis ingevoer word. Indien die struktuur waarin die data ontvang word nie ooreenstem met die databasismodel nie, sal dit nodig wees om die data te transformeer. Weereens kan dit aanleiding gee tot foute wat insluip as dit nie met sorg gedoen word nie. In die geval van die databasis onder beskouing is hierdie probleem heelwat meer basies. Data moes tot onlangs toe handmatig in die databasis ingevoer word!!!

Om te toets vir eersgenoemde foute is feitlik onmoontlik aangesien die oorspronklike dokumente benodig word om te kan kontroleer of oordragte korrek uitgevoer is. Die enigste manier om hierdie aspek te verbeter is om die datavloei na die databasis te verbeter. Hierdie aspek is reeds breedvoerig in *Hoofstuk 2* aangespreek. In laasgenoemde geval kan meganismes daargestel word wat verseker dat tenminste die inligting wat ontvang word korrek ingevoer word. Dit kan gedoen word deur te bepaal of die data wat ingevoer is dieselfde totale lewer as wat op die oorspronklike Tender Verslae verskyn waarvan die data afkomstig is.

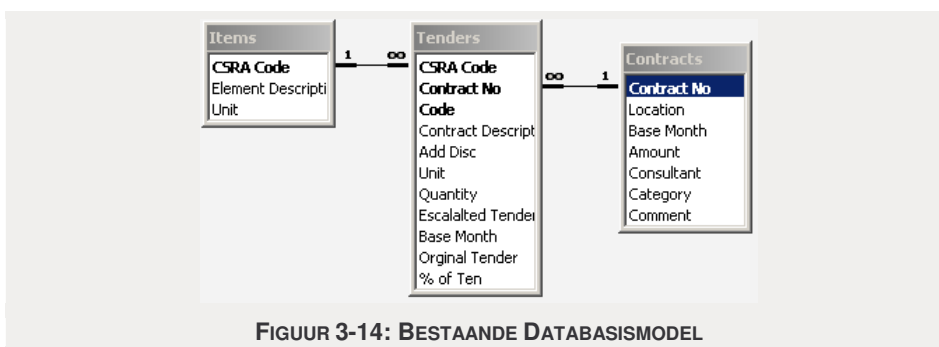
---

Op die oog af lyk dit asof hierdie toets sal verseker dat geen onakkuraathe- de kan deurglip nie, maar daar is egter in die databasis onder beskouing voorbeelde gevind waar foute wel deurgeglim het. 'n Voorbeeld hiervan is wan- neer 'n eenheidtarief van R 2.00 per ongeluk as R 20.00 ingevoer word en die hoeveelheid wat 1500 moet wees as 150 ingevoer word. In so 'n geval blyk die totale koste nog steeds korrek te wees, maar by beide die eenheidtarief en die hoeveelheid word 'n orde grootte fout gemaak. Weereens is die enigste metode om hierdie probleem op te los om die datavloei te verbeter.

### 3.4.3 DATABASISMODEL

Indien 'n databasismodel nie voldoen aan die vereistes wat gestel word deur die verskillende normaalvorms nie, kan dit gebeur dat onnodige duplisering van data in die databasis voorkom wat kan lei tot data-inkonsekwentheid. Om te bepaal of relasies korrek gedefinieer is in bestaande databasisse moet dit getoets word vir die normaalvorm waarin dit is. Hierdeur kan bepaal word of inkonsekwentheid in die data kan ontstaan.

Die bestaande CSRA-databasis bestaan uit drie hoof tabelle, *Figuur 3-14*. Dit blyk dat twee van die tabelle korrek gedefinieer is behalwe dat in sommige gevalle meer as een vorm van skryfwyse vir dieselfde attribute gebruik word. Dit is veral opmerklik by die raadgevende ingenieur attribuut, waar daar byvoor- beeld drie spelwyses van dieselfde naam voorkom, nl., SS, SScot en Steward Scot, terwyl dit eintlik "Stewart Scott" moet wees. Dieselfde situasie kom ook voor by die projekkatgorie en die eenhede. Hierdie inkonsekwentheid is egter onbelangrik aangesien dit nie die akkuraatheid van die statistieke van die data beïnvloed nie, maar wel onvolledige soekresultate kan lewer.



Die tabel waar die tenderdata in die bestaande CSRA-databasis gestoor word, toon egter enkele gebreke. Die projeknaam is 'n saamgestelde attribuut wat die projeknommer sowel as die tender bevat. Word hierdie naam egter ontbind, blyk dit dat daar heelwat funksionele afhanklikheid binne die databasis voorkom wat tot onnodige duplisering van data en inkonsekwentheid kan lei. Eenvoudige toetse wat op die data uitgevoer is, het getoon dat dit wel die geval is.

## 3.5 ONTWIKKELING VAN NUWE DATAMODEL

Aangesien die datamodel bepaal of inkonsekwentheid in die databasis kan ontstaan weens onnodige duplikasie word hier 'n generiese datamodel vir tenderinligting ontwikkel wat slegs beheerde duplikasies het. Hierdie model kan aangepas word vir ander meer spesifieke datamodelle waarin tenderdata gestoor kan word. 'n Konseptuele ontwerp is uitgevoer, asook normalisering van die relasies. Verder is die logiese ontwerp ook gedoen voor implementering van die databasis.

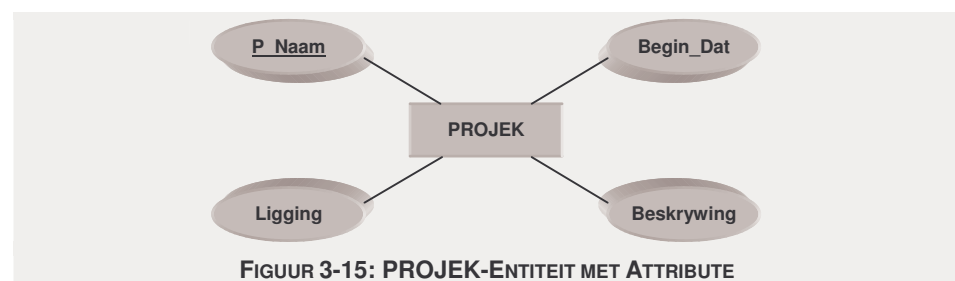
### 3.5.1 KONSEPTUELE ONTWIKKELING

Om behoorlik gevormde entiteite en die verwantskappe tussen hierdie entiteite te definieer is dit nodig om 'n konseptuele ontwikkeling te doen. Dit bestaan uit die bepaling van die entiteite, die verwantskappe tussen hulle en verdere verfyning wat nodig blyk te wees om te voldoen aan die databasisreëls. Aangesien daar nie 'n formele behoeftebepaling gedoen is nie word die bestaande CSRA-databasis as vertrekpunt gebruik om 'n voorgestelde datamodel te ontwikkel.

#### 3.5.1.1 ENTITEITE

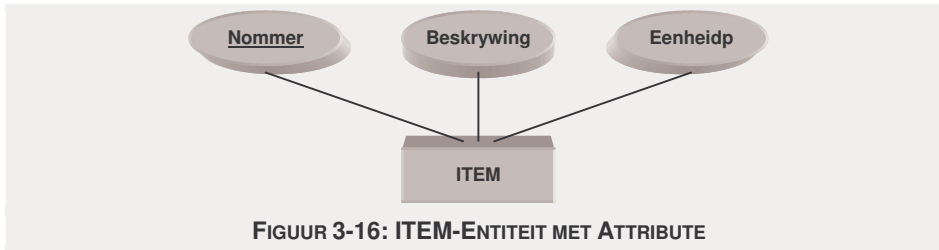
In die bestaande databasis word daar drie hoof entiteite geïdentifiseer, nl. projekte, items en tenders, *Figuur 3-14*. Hierdie entiteite word nou herontwikkel vir die voorgestelde datamodel. Verder word twee nuwe entiteite ook bygevoeg wat nie eksplisiet in die bestaande datamodel voorkom nie, maar wat die voorgestelde generiese model meer volledig sal maak.

Die entiteit **PROJEK**, waar 'n projek kan verwys na konstruksieprojekte, het vier attribute soos gesien kan word in *Figuur 3-15*. Die **P\_Naam** is die naam van die projek, **Ligging** is waar die projek geleë is, **Begin\_Dat** is wanneer konstruksie begin is en **Beskrywing** gee meer inligting oor die projek. Die primêre sleutel van hierdie entiteit is die projek se naam, omdat dit vir elke projek uniek behoort te wees.



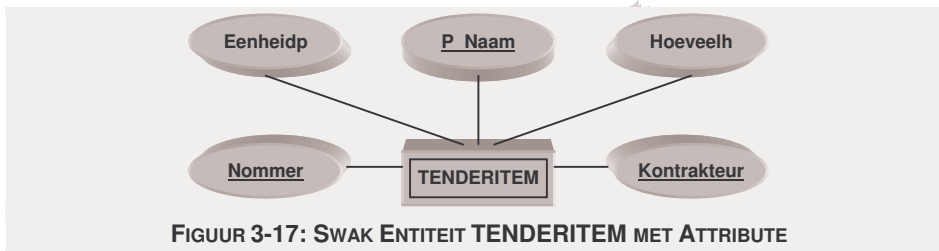
Die entiteit **ITEM**, waar 'n item kan verwys na die items soos gedefinieer in enige standaardspesifikasie, het drie attribute, *Figuur 3-16*. Die attribuut **Nommer** verwys na die standaardspesifikasie nommer wat aan 'n item toegeken is, **Beskrywing** is meer inligting oor die item en **Eenheid** verwys na die eenheid van

afmeting vir die item. Die primêre sleutel van hierdie entiteit is die nommer, omdat dit die item uniek kan bepaal.



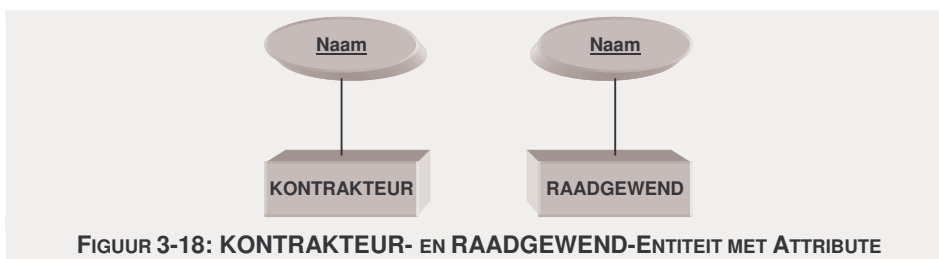
FIGUUR 3-16: ITEM-ENTITEIT MET ATTRIBUTE

Die entiteit TENDERITEM, waar hierdie entiteit verwys na die item waarvoor getender word deur 'n kontrakteur, het vyf attribute, soos getoon in *Figuur 3-17*. Die attribuut Hoeveelh verwys na die hoeveelheid van 'n item in 'n projek en Eenheidp verwys na die eenheidsprys van die item waarvoor die kontrakteur getender het. Die ander drie attribute is afkomstig van die entiteite PROJEK, ITEM en KONTRAKTEUR se attribute. Weens hierdie drie attribute is hierdie entiteit egter nie 'n gewone entiteite nie, maar word dit beskou as 'n swak entiteit, omdat dit afhanklik is van die bestaan van die drie ander entiteite.



FIGUUR 3-17: SWAK ENTITEIT TENDERITEM MET ATTRIBUTE

Verder is dit nodig om nog twee entiteite te definieer wat ook betrokke is by die tenderproses. Dit is KONTRAKTEUR en RAADGEWEND wat onderskeidelik verwys na die kontrakteur en die raadgewende ingenieur. Albei hierdie entiteite het slegs 'n enkel attribuut Naam, soos getoon word in *Figuur 3-18*.



FIGUUR 3-18: KONTRAKTEUR- EN RAADGEWEND-ENTITEIT MET ATTRIBUTE

### 3.5.1.2 VERWANTSKAPPE

Met die entiteite bekend wat betrokke is by die databasis kan daar nou gefokus word op die verwantskappe wat tussen hierdie entiteite voorkom. Die naam wat in die verwantskap staan (voorgestel met 'n diamant), beskryf gewoonlik die verwantskap, soos vroeër genoem. Weens die beperking van ruimte in die diagramme word egter soms net van afkortings gebruik gemaak.



FIGUUR 3-19: VERWANTSKAP TUSSEN PROJEK EN ITEM

Die eerste verwantskap wat gedefinieer word, is tussen PROJEK en ITEM, *Figuur 3-19*. Elke projek kan een of meer items hê en net so kan elke item ook in meer as een projek voorkom. Hierdie is dus 'n tipiese baie-tot-baie situasie wat vermy behoort te word weens die onnodige duplisering wat dit veroorsaak. As voorbeeld hiervan kan in *Figuur 3-20* gesien word dat vir slegs twee projekte, met twee items, dit tot onnodige duplikasies in ITEM lei. Hierdie probleem word aangespreek in die *Afdeling 3.5.1.3* waar die model verfyn word.

PROJEK					
	P_Naam	Nommer	Beg Dat	P_Ligging	Beskryw
	WP1	21.01	12/06/90	Weskaap	OPGR PAD
	WP1	21.02	12/06/90	Weskaap	OPGR PAD
	GT1	21.01	16/09/89	Gauteng	HERB HOOFPAD
	GT1	21.02	16/09/89	Gauteng	HERB HOOFPAD

ITEM				
	Nommer	P_Naam	Beskryw	Eenheid
	21.01	WP1	Uitgrawing vir vore	m3
	21.02	WP1	Skoonmaak en vorming van bestaande vore	m3
	21.01	GT1	Uitgrawing vir vore	m3
	21.02	GT1	Skoonmaak en vorming van bestaande vore	m3

FIGUUR 3-20: PROJEK EN ITEM VERWANTSKAP IS M:N

Tussen die entiteit PROJEK en KONTRAKTEUR kom daar ook 'n baie-tot-baie verwantskap voor, *Figuur 3-21*. Vir elke projek is daar meer as een kontrakteur wat kan tender en 'n kontrakteur kan vir meer as een projek tender. Hierdie relasie moet dus ook ontbind word om onnodige duplisering te voorkom. Dit word ook verfyn in *Afdeling 3.5.1.3*.



FIGUUR 3-21: VERWANTSKAP TUSSEN KONTRAKTEUR EN PROJEK

Tussen die entiteit RAADGEWEND en PROJEK kom daar 'n een-tot-baie verwantskap voor. Elke raadgewende ingenieursfirma kan meer as een projek bestuur, maar slegs een raadgewende ingenieursfirma staan aan die hoof van 'n projek. Hierdie verwantskap word getoon in *Figuur 3-22*.

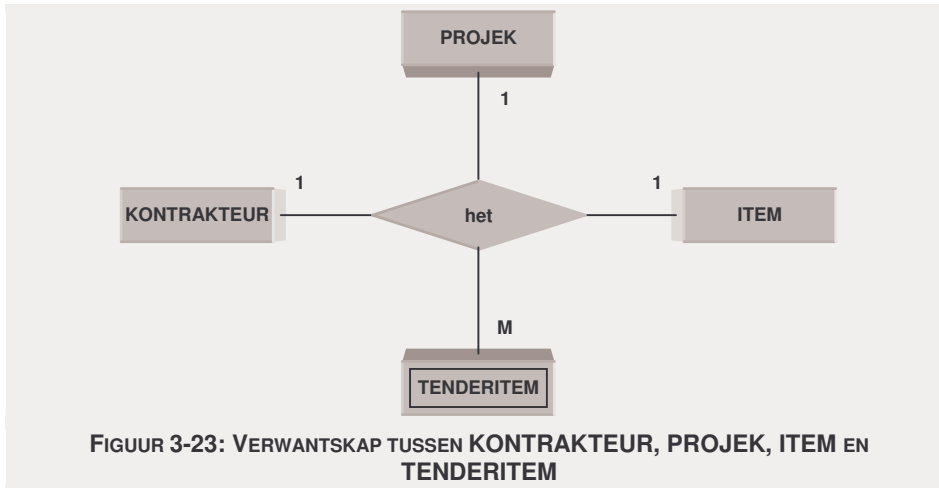


FIGUUR 3-22: VERWANTSKAP TUSSEN RAADGEWEND EN PROJEK

'n Veelvuldige verwantskap kom voor tussen die entiteite PROJEK, ITEM en KONTRAKTEUR met TENDERITEM as die sogenaamde swak entiteit. Die konekteerdheid kan as volg verklaar word – 'n kontrakteur tender vir 'n ver-



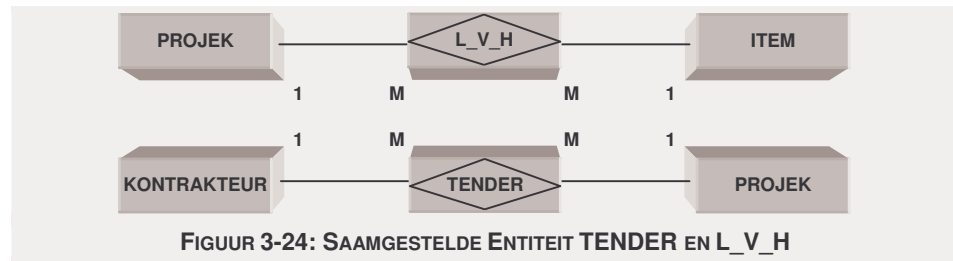
skeidenheid van items vir 'n gegewe projek (een-tot-baie), 'n projek kan 'n verskeidenheid van items hê waarvoor getender moet word (een-tot-baie) en vir 'n item kan daar ook baie tenders wees, omdat dit in 'n verskeidenheid van projekte gebruik word (een-tot-baie). Elke getenderde item, TENDERITEM, is egter uniek, omdat daar slegs een kontrakteur, projek en item vir 'n getenderde item is. Hierdie verwantskap word getoon in *Figuur 3-23*.



### 3.5.1.3 VERFYNING

Met die verwantskappe bekend is dit nou nodig om die probleme wat voorkom in die model aan te spreek. Drie onreëlmatighede kom op die oomblik in die model voor. Die eerste twee is reeds genoem, nl. die baie-tot-baie verwantskappe wat voorkom by "het" en "tender" onderskeidelik in *Figuur 3-19* en *Figuur 3-21*. Die derde fout ontstaan deurdat die model inkonsekwentheid in die data wat in die databasis gestoor word, kan laat ontstaan weens duplisering. Dit vind plaas deurdat dit moontlik is om met die verwantskap soos in *Figuur 3-23* gedefinieer, 'n inskrywing te maak in TENDERITEM, sonder dat die gepaardgaande verwantskappe tussen KONTRAKTEUR en PROJEK en ook tussen PROJEK en ITEM geskep is.

As oplossing vir eersgenoemde twee probleme is dit moontlik om die verwantskappe te herdefinieer sodat daar 'n sogenaamde saamgestelde entiteit ontstaan tussen elk van die twee pare entiteite waar die baie-tot-baie verwantskap voorkom, *Figuur 3-24*. Hierdeur word die onnodige duplisering van data tot 'n minimum beperk, soos in *Figuur 3-25* gesien kan word indien die data van die voorbeeld in *Figuur 3-20* aangepas word om op die nuwe model te pas. Indien daar ook veranderings nodig sou wees in projek- of itemdata, weens foute wat gemaak word, is dit slegs nodig om dit op een plek te verander en nie op 'n aantal plekke nie.



Die twee saamgestelde entiteite wat gevorm word se betekenis kan as volg gesien word. Tussen PROJEK en ITEM kom die L\_V\_H (Lys van Hoeveelhede) voor, *Figuur 3-25*. Elke projek in PROJEK het 'n aantal items in die lys van hoeveelhede, terwyl 'n item in die L\_V\_H slegs aan een projek behoort. Vanaf die ITEM-kant is dit ook dieselfde situasie waar 'n item op verskeie plekke in L\_V\_H voorkom, maar 'n item in die L\_V\_H slegs na een item in ITEM verwys. Netso ook tussen KONTRAKTEUR en PROJEK is die saamgestelde entiteit TENDER.

PROJEK				
	P_Naam	Beg Dat	P_Ligging	Beskryw
	WP1	12/06/90	Weskaap	OPGR PAD
	GT1	16/09/89	Gauteng	HERB HOOPPAD

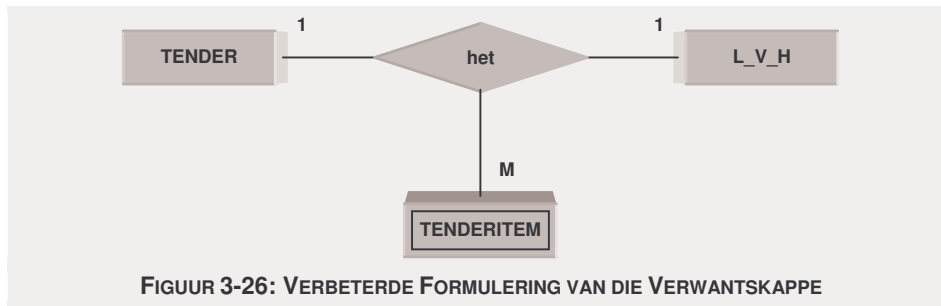
L_V_H		
	P_Naam	Nommer
	WP1	21.01
	WP1	21.02
	GT1	21.01
	GT1	21.02

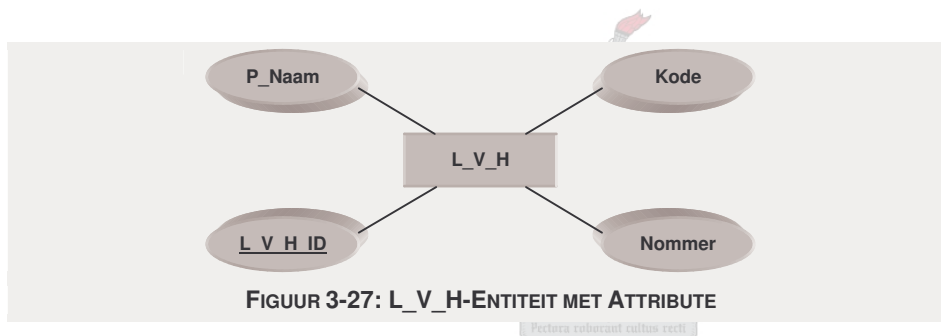
ITEM			
	Nommer	Beskryw	Eenheid
	21.01	Uitgraving vir vore	m3
	21.02	Skoonmaak en vorming van bestaande vore	m3

**FIGUUR 3-25: PROJEK EN ITEM MET SAAMGESTELDE ENTITEIT L\_V\_H**

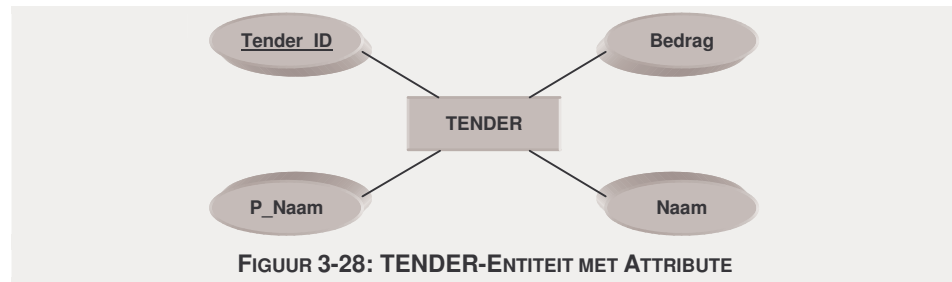
Die aanpassing wat aan hierdie twee entiteite gemaak word bied nou die geleentheid om die derde probleem van inkonsekwentheid wat by die veelvuldige verwantskap ontstaan, uit te skakel. Aangesien die saamgestelde entiteite TENDER en L\_V\_H uniek die verwantskap tussen die entiteite KONTRAKTEUR en PROJEK en tussen die entiteite PROJEK en ITEM onderskeidelik bepaal, kan hierdie saamgestelde entiteite gebruik word om die swak entiteit TENDERITEM uniek te bepaal. Dit voorkom dus 'n duplisering van data wat tot inkonsekwentheid kan lei. Die nuwe binêre verwantskap met 'n swak entiteit word getoon in *Figuur 3-26*.



Met hierdie aangepaste uitleg is dit nou moontlik om nog 'n verdere verfyning aan die model aan te bring, sodat twee spesifieke attribute wat tipies kan voorkom by enige databasis wat tenderinligting stoor, ook ingesluit word in die voorgestelde model. Die twee attribute is Kode en Bedrag. In 'n projek kan dit gebeur dat dieselfde item 'n paar keer gebruik word. Om te kan onderskei tussen dieselfde items word 'n unieke kode aan elk gekoppel. Die attribuut Kode word derhalwe ingesluit by die verwantskap wat tussen PROJEK en ITEM bestaan om sodoende meer as een van dieselfde items in 'n projek toe te laat. Dit kan dus by die samegestelde entiteit L\_V\_H wat in die verfyning ontwikkel is ingesluit word. In *Figuur 3-27* word die attribute van die entiteit L\_V\_H getoon.



Die attribuut Bedrag verwys na die totale getenderde koste vir die projek, soos deur 'n kontrakteur getender. Hierdie attribute, net soos die attribuut Kode, is gekoppel aan die verwantskap wat bestaan tussen twee entiteite, vir hierdie geval tussen die entiteite PROJEK en KONTRAKTEUR. Die samegestelde entiteit TENDER, wat die verwantskappe stoor tussen bogenoemde twee entiteite kan dus gebruik word om die attribuut Bedrag in te sluit. In *Figuur 3-28* word die attribute van die entiteit TENDER getoon.

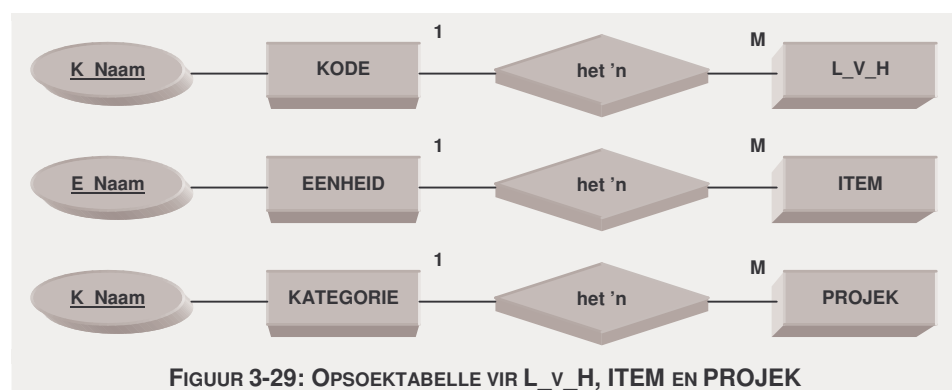


### 3.5.1.4 OPSOEKTABELLE

In *Afdeling 3.4* word daar verwys na die probleme wat voorkom weens data-inkonsekwentheid wat veroorsaak word deur die verskillende maniere waarop presies dieselfde data in die databasis gestoor word. Hierdie verskynsel, soos alreeds genoem, kom by drie van die attribute in die databasis voor, nl. raadgevende ingenieurs, kategorie en eenhede. Om hierdie probleem tot 'n minimum te beperk kan van sogenaamde opsoektabelle gebruik gemaak word.

Opsoektabelle word gebruik om waardes van attribute wat uit 'n beperkte aantal elemente bestaan te stoor, sodat indien die gebruiker 'n waarde in die attribuut moet invul, slegs 'n keuse gemaak hoef te word vanuit dit wat beskikbaar is. Hierdeur word die probleem opgelos van dieselfde data wat op verskillende wyse geskryf word, maar presies dieselfde betekenis het. Indien dit wel sou nodig wees om die versamelingsruimte uit te brei, omdat daar byvoorbeeld 'n eenheid gestoor moet word wat nog nie beskikbaar is nie, kan dit slegs by die opsoektabel gevoeg word.

Vir hierdie ontwikkeling moet die volgende entiteite gemodelleer word KODE, EENHEID en KATEGORIE. Die verwantskap van elk van die drie entiteite met die betrokke entiteit waar dit voorkom, asook 'n attribuut word getoon in *Figuur 3-29*.



### 3.5.2 NORMALISERING VAN HUIDIGE MODEL

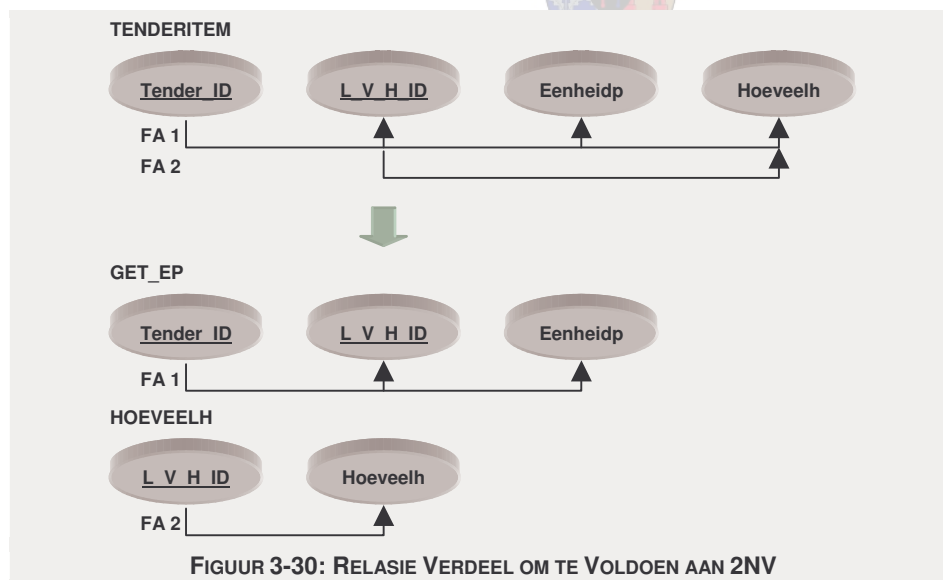
Met die konseptuele model voltooi is dit nou nodig om te kyk of al die entiteite se strukture korrek gedefinieer is, soos reeds bespreek in *Afdeling 3.3.4*. Aan-

gesien die konseptuele ontwerp goed gestruktureerde tabelle teweeggebring het, is meeste van die entiteite triviaal in die korrekte normaalvorm. Hier word dus nie formeel vir elke normaalvorm getoets nie, maar slegs die resultate word in *Tabel 3-1* getoon. Die enigste entiteit wat wel hier volledig getoets word is die TENDERITEM, omdat dit nie in die korrekte normaalvorm is nie.

**TABEL 3-1: NORMAALVORMS VAN DIE TABELLE**

Entiteit	1NV	2NV	3NV	BCNV
PROJEK	Ja	Ja	Ja	Ja
ITEM	Ja	Ja	Ja	Ja
RAADGEWEND	Ja	Ja	Ja	Ja
KONTRAKTEUR	Ja	Ja	Ja	Ja
TENDER	Ja	Ja	Ja	Ja
KODE	Ja	Ja	Ja	Ja
EENHEID	Ja	Ja	Ja	Ja
KATEGORIE	Ja	Ja	Ja	Ja
L_V_H	Ja	Ja	Ja	Ja

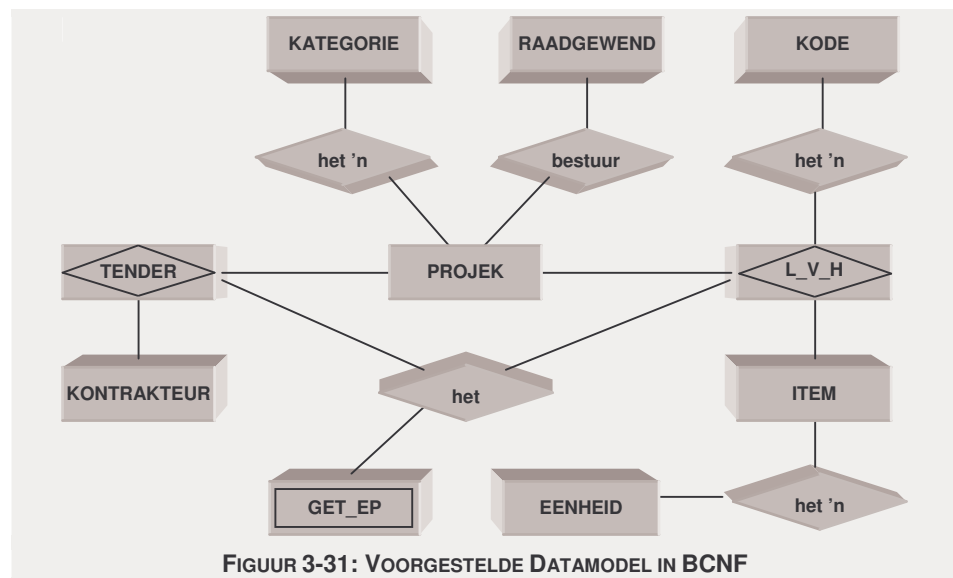
Vir die 1NV, nl. is die data atomiese waardes, slaag die relasie TENDERITEM. Vir die 2NV se vereiste van funksionele afhanklikheid slaag die relasie egter nie, omdat daar twee funksionele afhanklikhede ontstaan soos getoon word in *Figuur 3-30*. Dit kan egter ontbind word om te voldoen aan die vereiste van 2NV. Die ontbinding vorm twee relasies een wat GET\_EP genoem word, vir die getenderde eenheidspryse, en die ander wat HOEVEELH genoem word.



Dit blyk egter by nadere beskouing van die relasie HOEVEELH dat dit presies dieselfde primêre sleutel sal hê as die relasie L\_V\_H en dat die data wat hier gestoor word netso goed by die relasie L\_V\_H ingesluit word. Dit is dus nie nodig om 'n addisionele relasie te definieer nie, maar net die L\_V\_H, wat reeds aan BCNV voldoen, uit te brei sodat dit die relasie insluit. Die nuwe relasie

GET\_EP voldoen ook aan die vereistes van die derde normaalvorm se vereiste van transitiewe afhanklikheid en aan die strengere BCNV.

Daar kan dus volstaan word om te sê dat die databasismodel soos dit konseptueel uitgelê is voldoen aan BCNV. Dit beteken dus dat onnodige duplisering van data tot 'n minimum beperk word en dat data-inkonsekwentheid wat kan voorkom, minimaal is. In *Figuur 3-31* word die entiteite en verwantskappe daartussen getoon vir die voorgestelde datamodel.



### 3.5.3 LOGIESE ONTWERP

Met die konseptuele ontwerp voltooi en ook die normalisering agter die rug bestaan daar nou 'n datamodel waarvan die verwantskappe tussen die relasies vasgelê is en 'n korrekte struktuur tussen die relasies verseker is. Die volgende stap is om die logiese ontwerp uit te voer. Die logiese ontwerp van die databasis is oorwegend sagteware afhanklik en daarom moet besluit word watter tipe databasis stelsel gebruik gaan word. Vir hierdie studie is daar besluit op 'n relasionele enkelgebruiker databasis wat op 'n persoonlike rekenaar geïmplementeer word, slegs om die konsepte van die voorgestelde databasismodel te illustreer.

Soos reeds genoem, is vir hierdie studie van *MS Access* gebruik gemaak. Die logiese ontwerp wat hier voorgestel word is egter nie net beperk tot *MS Access* nie, maar kan feitlik op enige ander kragtiger databasis oorgedra word met weinig veranderings, om 'n multi-gebruiker, verspreide databasis te implementeer, sou dit nodig wees. Voordat egter voortgegaan word met die logiese ontwerp is dit nodig om 'n paar aspekte te bespreek wat betrekking het op hierdie ontwerp. Hierdie aspekte het te make met [Litwin (1997)]:

- streng definiëring van die data wat in elke attribuut toegelaat word. Hierdie aspek kan op verskillende vlakke toegepas word om te verseker dat geldige data in die databasis geplaas word. Met die funksionaliteit wat binne die DBBS beskikbaar is, kan dit in 'n mate op attribuutvlak toegepas word. Datatypes en ook die beperking van leë waardes (NULLS) kan byvoorbeeld voorgeskryf word en ander filters is ook beskikbaar;
- verbetering in die spoed van die databasis. Aangesien van die tabelle in die databasis met honderde duisende waardes en met verloop van tyd met miljoene waardes bevolk sal wees, is dit nodig om van sogenaamde indekse gebruik te maak, om soektogte in die tabelle te versnel. Indekse maak gebruik van datastrukture soos binêre bome en “hash”-tabelle om soektogte te versnel; dit word ook op attribuutvlak gedefinieer.
- definiëring van 'n addisionele attribuut by elk van die entiteite. Alhoewel elke relasie 'n primêre sleutel het, enkel of samegestel, is besluit om van 'n funksionaliteit gebruik te maak wat beskikbaar is in meeste databasisse, nl. om outomaties 'n unieke syfer te koppel aan elke inskrywing wat in 'n tabel gemaak word. Hierdie identifiserings attribuut kan gebruik word as vreemde sleutel in die tabel wat daarna verwys. Die voordeel hiervan is dat dit geheuegewys meer effektief is. Die attribuut word vir hierdie ontwerp altyd aangedui met die entiteit se naam en met die letters “ID” daar agter, byvoorbeeld vir die PROJEK-entiteit sal die attribuut se naam wees `Projek_ID`.
- voorkoming van nalatige uitwissing. Weens die verwantskappe wat tussen die tabelle gedefinieer gaan word m.b.v. die primêre sleutels en vreemde sleutels, is daar handige funksionaliteit beskikbaar in die DBBS wat sogenaamde watervalle opdaterings en verwydering van data kan veroorsaak. Die implikasie hiervan is egter dat indien 'n raadgewende ingenieursfirma se naam byvoorbeeld uit die tabel `RAADGEWEND` verwyder word, die projekte waarby die raadgewende ingenieur betrokke is ook verwyder word en so ook dan die items in die `L_V_H`, tenders in `TENDERS` en finaal dan die getenderde pryse van al die projekte waarby die raadgewer betrokke was. Dit word eksplisiet by die aanduiding van die sleutels genoem dat hierdie funksie afgeskakel word.

Die tien entiteite of tabelle wat voorkom in die voorgestelde databasis word vervolgens bespreek. Daar word weer vlugtig verwys wat die doel van elkeen is en daarna word na die attribute gekyk. By elke attribuut moet besluit word: met watter datatype dit voorgestel gaan word; of leë waardes (“Nulls”) aanvaar word; of dit 'n gewone attribuut is, 'n vreemde sleutel (VS) of 'n primêre sleutel (PS); aan watter tabel dit verwant is, indien dit 'n vreemde sleutel is en of indekse

---

gebruik word om soektogte te versnel. Die ontwerp van elk word in tabelformaat getoon. Die Engelse benaming is vir die implementasie van elke tabel gebruik.

### 3.5.3.1 PROJEKTABEL

Die PROJEK-tabel stoor inligting wat verbandhou met elke projek. Die naam van die projek, waar dit geleë is, wanneer dit begin is, wie die raadgewende ingenieursfirma is wat die projek bestuur het en watter projektype dit is, kan hier gevind word. Die logiese ontwerp vir hierdie tabel is gedefinieer soos gesien kan word in *Tabel 3–2*.

**TABEL 3–2: LOGIESE ONTWERP VIR PROJECTS**

Attribuut	Datatype	Nulls	PK/VS	Verwant	Indeks
Project_ID	AutoNumber	No	PS	-	Yes (no dupl)
Name	Char(50)	No	-	-	Yes (no dupl)
Location	Char(50)	No	-	-	-
Base_Month	Date	No	-	-	-
Consultant_ID	Long Int	No	VS (no casc)	Consultant	-
Category_ID	Long Int	No	VS (no casc)	Category	-

### 3.5.3.2 ITEMABEL

In die ITEM-tabel word inligting gestoor oor die CSRA-items of dan enige ander standaarditems. Die itemnommer, beskrywing van die item en die eenheid van die item kan hier gevind word. Die logiese ontwerp vir hierdie tabel is gedefinieer soos in *Tabel 3–3*.

**TABEL 3–3: LOGIESE ONTWERP VIR ITEMS**

Attribuut	Datatype	Nulls	PK/VS	Verwant	Indeks
Item_ID	AutoNumber	No	PS	-	Yes (no dupl)
Item_Num	Char(50)	No	-	-	Yes (no dupl)
Description	Char(50)	No	-	-	-
Unit_ID	Long Int	No	VS (no casc)	Unit	-

### 3.5.3.3 L\_V\_H-TABEL

Die L\_V\_H-tabel stoor die items wat voorkom in elke projek. Die projek, die itemnommer, die kode om onderskeid te maak as dieselfde item meer as eenkeer voorkom, die hoeveelheid en 'n addisionele beskrywing as daar iets uitsonderlik omtrent die item is kan hier gevind word. Die logiese ontwerp van hierdie tabel is gedefinieer soos in *Tabel 3–4*.



**TABEL 3-4: LOGIESE ONTWERP VIR BILLOFQUANTITIES**

Attribuut	Datatype	Nulls	PK/VS	Verwant	Indeks
Bill_ID	AutoNumber	No	PS	-	Yes (no dupl)
Project_ID	Long Int	No	VS (no casc)	Projects	Yes (no dupl)
Item_ID	Long Int	No	VS (no casc)	Items	Yes (no dupl)
Code_ID	Long Int	No	VS (no casc)	Codes	Yes (no dupl)
Quantity	Real	No	-	-	-
Add_Describ	Char(50)	No	-	-	-

### 3.5.3.4 TENDERTABEL

In die TENDER-tabel word data gestoor oor wie vir watter projekte tender en ook die totale koste daarvan. Die projek, kontrakteur en totale tenderbedrag kan hier gevind word. Die logiese ontwerp vir hierdie tabel kan gesien word in *Tabel 3-5*.

**TABEL 3-5: LOGIESE ONTWERP VIR TENDER**

Attribuut	Datatype	Nulls	PK/VS	Verwant	Indeks
Tender_ID	AutoNumber	No	PS	-	Yes (no dupl)
Project_ID	Long Int	No	VS (no casc)	Projects	Yes (no dupl)
Contractor_ID	Long Int	No	VS (no casc)	Contractors	Yes (no dupl)
Total_Amount	Real	No	-	-	-

### 3.5.3.5 GETENDERDE PRYSE-TABEL

In die GET\_EP-tabel word die prys gestoor van elke tender van 'n kontrakteur vir 'n item in die Lys van Hoeveelhede van 'n projek. Die tender, die hoeveelheidslys item en eenheidsprys kan hier gevind word soos getoon in *Tabel 3-6*.

**TABEL 3-6: LOGIESE ONTWERP VIR FIXED\_PRICES**

Attribuut	Datatype	Nulls	PK/VS	Verwant	Indeks
Tender_ID	Long Int	No	VS (no casc)	Tenders	Yes (no dupl)
Bill_ID	Long Int	No	VS (no casc)	BillsOfQuant.	Yes (no dupl)
Unit_Price	Real	No	-	-	-

### 3.5.3.6 OORBLYWENDE TABELLE

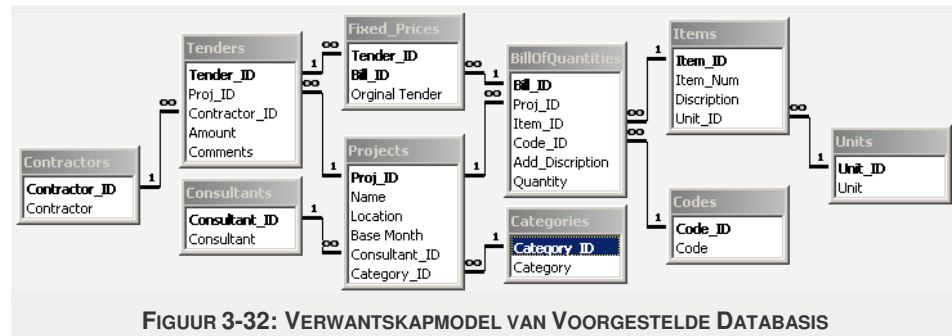
Die oorblywende tabelle, RAADGEWEND, KONTRAKTEUR, KODE, EENHEID en KATEGORIE, is die sogenaamde opsoektabelle wat gebruik word om te voorkom dat dieselfde benamings verkeerdelik geskryf word. Die struktuur van die vyf tabelle is presies dieselfde met slegs die benamings van die twee attribute wat elk bevolk wat verskil. Hier word dus net die logiese ontwerp vir KONTRAKTEUR getoon in *Tabel 3-7*.

**TABEL 3-7: LOGIESE ONTWERP VIR CONTRACTORS**

Attribuut	Datatype	Nulls	PK/VS	Verwant	Indeks
Tender_ID	Long Int	No	VS (no casc)	Tenders	Yes (no dupl)
Bill_ID	Long Int	No	VS (no casc)	BillsOfQuant.	Yes (no dupl)

### 3.5.3.7 VERWANTSKAPSMODEL

As laaste stap van die logiese ontwerp word die volledige verwantskapsmodel getoon vir die voorgestelde databasismodel soos dit dan in *MS Access* geïmplementeer is. In *Figuur 3-32* word die tabelle getoon met elkeen se attribute. Die primêre sleutel soos in die logiese ontwerp bepaal is word in vet letters getoon. Die verwantskappe wat tussen die tabelle bestaan word getoon met die swart lyne wat gemerk is met 'n 1 en ∞ om die een-tot-baie verhouding voor te stel.



FIGUUR 3-32: VERWANTSKAPMODEL VAN VOORGESTELDE DATABASIS

## 3.6 OORDRAG VAN DATA NA NUWE MODEL

Met die ontwerp van die databasis nou voltooi en ook geïmplementeer in *MS Access* is dit nodig om die data oor te dra vanaf die oorspronklike databasis na die voorgestelde databasis. Hierdie oordrag kan nie direk gedoen word nie aangesien die data in die oorspronklike databasis oor drie tabelle versprei is, teenoor die voorgestelde databasis waar die data oor 10 tabelle versprei is. Met die verdeling van die data is dit belangrik om die relasies wat in elk van die drie tabelle voorkom sodanig oor te dra na die voorgestelde databasis dat dit in die vorm van die verwantskappe tussen die nuwe tabelle behoue bly.

Die korrekte oordrag van data was egter nie die enigste aspek wat in gedagte gehou moes word nie. Die voorgestelde datamodel pas aansienlik sterker data-integriteit toe as die bestaande model. Inkonsekwente data wat in die bestaande model voorkom sal dus nie in die voorgestelde model geplaas kan word nie, omdat die DBBS dit sal voorkom. Tydens die oordrag van die data is dit dus ook nodig om die inkonsekwente data te verwyder, naas die behoud van die relasies in die bestaande databasis. Dit is belangrik om wel daarop te let dat slegs inkonsekwentheid verwyder kan word, maar foutiewe data wat weens vingerfoute deurgelip het nie deur hierdie proses opgespoor kan word nie.

Verskillende metodes en “gereedskap” kan gebruik word om die oordrag te doen. Vir die doeleindes van hierdie studie was *MS Excel* en *MS Visual Basic*-makroprogramme geskryf en gebruik om deur die proses te gaan om inkonse-

kwenthede op te spoor en die data oor te dra. Hier word volstaan deur net een voorbeeld van 'n inkonsekwentheid te noem wat voorgekom het.

In die bestaande databasis is die attribuut `Contract_No`, wat verwys na die projeknommer, 'n sogenaamde samegestelde attribuut. Die naam bestaan uit die werklike projeknaam en die tendernommer om onderskeid te maak tussen die 1ste, 2de en 3de tenders. In die voorgestelde databasis word dit van mekaar geskei. Die eerste gedeelte van die attribuut `Contract_No` word geplaas in die `P_Naam`-attribuut van `PROJEK`, terwyl die laaste gedeelte van die naam, of dan die tender, geplaas word in die attribuut `Naam` van `KONTRAKTEUR`.

Om die verdeling te kon doen na die vorm soos getoon in *Figuur 3-33* moes dit geskei word met behulp van 'n makro wat in *MS Excel* geskryf is. Dit het egter geblyk om nie so eenvoudig te wees nie weens die inkonsekwentheid wat voorgekom het in hoe die verskillende tenders aangedui is. Vir tender 1 was daar verskillende skryfwyses gewees, byvoorbeeld: "1", "Tend 1", "Tend1" en "TEND1". Hierdie tendens is ook gevind by tender 2 en tender 3. Die soekfunksie wat die skeiding maak moes al hierdie verskillende vorms in ag neem. In *Aanhangsel C* word die volledige metode van omskakeling bespreek.



### 3.7 TOETSING VAN OORGEDRAGTE DATA

Met die voltooiing van die oordrag van die data was dit nodig gewees om twee aspekte te toets van die voorgestelde databasis, nl. of dieselfde informasie-aansigte soos in die ou databasis verkry kan word in die nuwe databasismodel en of die data korrek oorgedra is. In eersgenoemde geval is dit nodig om te bepaal of die data sodanig saamgestel kan word sodat dit dieselfde struktuur het as die oorspronklike data, m.a.w. dat die data saamgevoeg kan word om die oorspronklike drie hoof tabelle te vorm. Dit moet gedoen word sodat toepasingsprogramme wat vir die bestaande DB geskryf is wel die voorgestelde DB kan benut. Die tweede aspek is om te bevestig dat die oordragproses korrek plaasgevind het en dat geen van die relasies verlore is nie.

Hierdie twee toetse is in een stap uitgevoer deur eerstens data in 'n soektog op die voorgestelde databasis dieselfde struktuur te gee as die bestaande databasis. Hierdie data is dan vergelyk met die data afkomstig uit die bestaande databasis se tabelle. Addisionele toetse is ook uitgevoer om te bevestig dat berekeninge wat gebaseer is op data afkomstig van die twee databasisse dieselfde resultate lewer. Dit is gedoen deur die som van eenheidspryse en

hoeveelhede te bereken en dan ook totale projekkostes te bereken deur die som van die produk van die hoeveelhede en eenheidtariewe te vind. Hierdie toets is uitgevoer deur die data van die twee databasisse in *MS Excel* te laai en dan 'n vergelykende toets met *MS Excel* se ingeboude funksies op die waardes in die sigblad te doen. Weens die grootte van die tabelle in die databasisse moes die data oor twee sigblaaie verdeel word om die toets in *MS Excel* te kon uitvoer. Om 'n idee te vorm van die grootte van toetse kan na die grootte van die twee sigblaaie gekyk word. Dit is elk in die orde van 13 megagreep groot. 'n Volledige bespreking kan in *Aanhangsel D* gesien word.

Uit die toets het dit eerstens geblyk dat die struktuur van die bestaande databasis wel herskep kon word. Die toetse om te bepaal of die oordragproses korrek plaasgevind het, het egter getoon dat in al drie vergelykings met die hoof tabel van die bestaande databasis verskille voorgekom het. By nadere ondersoek het dit egter geblyk dat hierdie foute eerder verskille is wat voortspruit uit die korreksies wat aangebring is weens die strenger integriteit van die voorgestelde databasis. Hierdie bevindings is bevestig toe die finale toets gedoen is om te bepaal of die totale klop. In 100% van die ongeveer 120 000 items is die eenheidtariewe korrek oorgedra, terwyl by ongeveer 1% van die 512 projektenders verskille in die hoeveelhede voorgekom het. Dit is veroorsaak deurdat in die bestaande databasis daar verskille voorgekom het in die hoeveelheidslyste van dieselfde projek. Met die samevoeging is daar egter slegs van die laagste tender se hoeveelheidslys gebruik gemaak, wat dan die inkonsekwentehede uitgewys het. Aangesien daar in hierdie studie oorwegend gefokus word op die eenheidtariewe van items is daar nie verdere aandag geskenk om die foute te probeer verwyder nie. Die toets het egter getoon dat die oordrag van die data na die verskillende tabelle in die voorgestelde model suksesvol was en dat daar geen inligting verlore geraak het nie.

---

### 3.8 UITBREIDING VAN DIE NUWE MODEL

---

Alhoewel die databasis aanvanklik slegs ontwikkel was vir die doel om die getenderde eenheidtariewe van items in projekte se hoeveelheidslyste te stoor is dit in hierdie studie geïdentifiseer dat addisionele inligting in die databasis gestoor moet word. Dit spruit voort uit addisionele projekdata-insameling en die storing daarvan vir die nuwe analise tegnieke wat in hierdie studie ontwikkel is. Hierdie addisionele inligting van projekte word meer breedvoerig in *Hoofstuk 6* bespreek waar dit geïdentifiseer word en ook getoon word hoe maklik dit in die voorgestelde databasismodel ingepas kan word.

Daar is egter drie ander stelle inligting beskikbaar wat formeel betrokke is by hetsy die tenderproses of by die afsluiting van die projek, maar nie in die databasis gestoor word nie. In volgorde van nuttigheid is dit die finale hoeveelhede

---

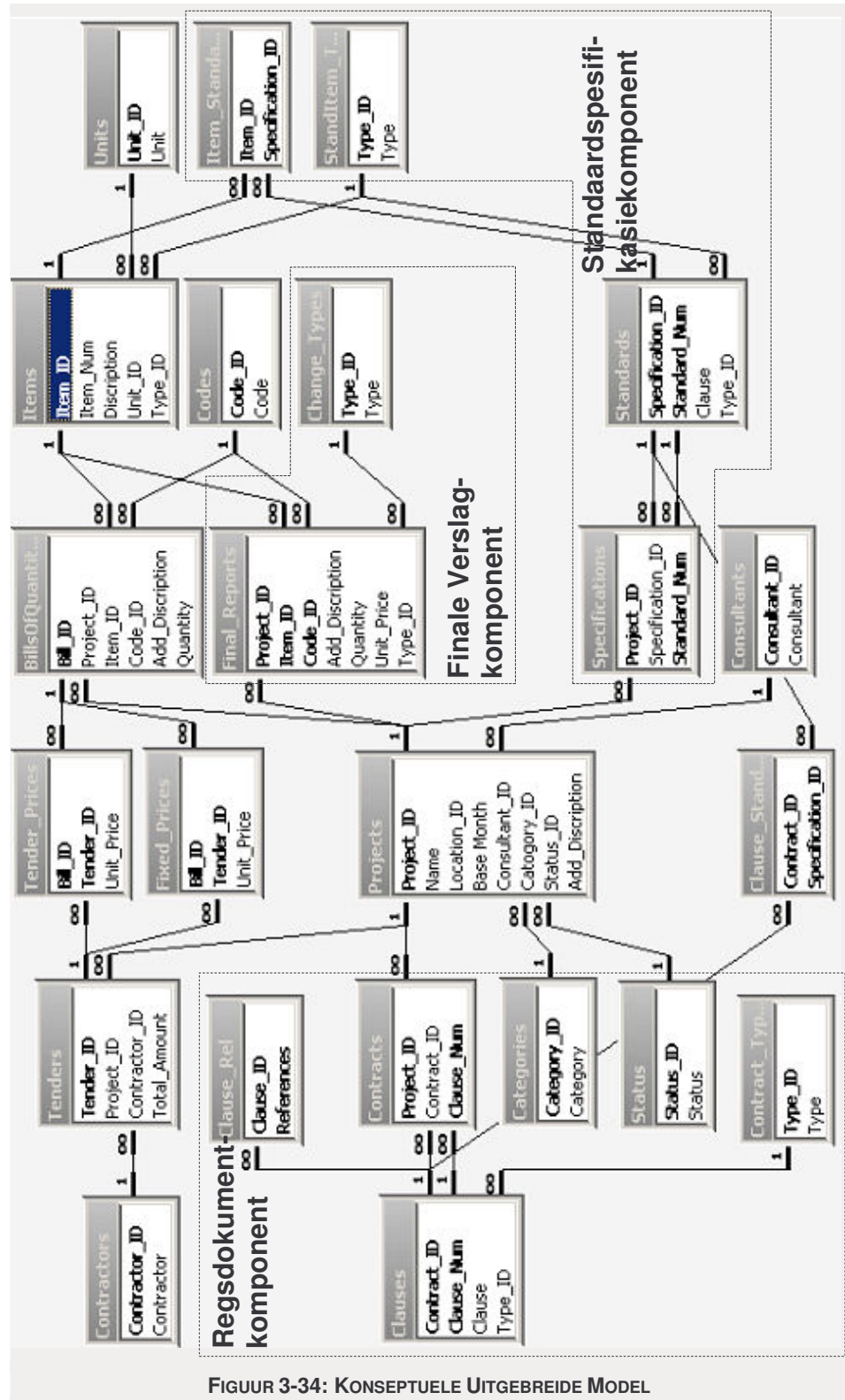
per item in 'n projek, die standaardspesifikasies en die regskontrakte. Die nut daarvan om die finale hoeveelhede van items in 'n databasis te stoor, sodat dit vergelyk kan word met beraamde hoeveelhede spreek vanself en word ook later nog meer in diepte in *Hoofstuk 6* bespreek.

Indien die spesifikasies vir elke projek onveranderd gebly het sou die enigste nut miskien gewees het om dit vir dokumentasie doeleindes in die databasis te stoor, sodat die verwantskappe tussen item en spesifikasie elektronies gedefinieer is. Die realiteit is egter dat die standaardspesifikasie nie so "standaard" is nie en dat daar in feitlik elke tenderdokument 'n paar veranderings aan sommige klousules aangebring word. Hierdie wysigings, wat miskien kan lig werp op die groot veranderlikheid van die eenheidtariewe, is egter nie beskikbaar nie.

Die rede hoekom dit heelwaarskynlik nie onderskep en gestoor word nie, is, omdat dit skynbaar 'n groot taak sal wees om vir elke projek al die spesifikasies in te voer. Indien die databasis egter korrek ontwerp word, is dit slegs een keer nodig om al die CSRA-spesifikasies in die databasis in te voer. Indien daar dan wel projekspesifieke veranderings in die klousules aangebring moet word, is dit slegs nodig dat hierdie veranderings gestoor moet word. Langs dieselfde weg kan aangevoer word dat die regsdokument ook in die databasis gestoor behoort te word, hoewel die voordele nie so voor die handliggend is nie.

In *Figuur 3-34* word slegs 'n konsepuele uitbreiding getoon van al drie die komponente wat by die voorgestelde model ingepas kan word. Hierdie datamodel kan in konsep al die data stoor van 'n aantal van die formele dokumente wat geskep word tydens die verskillende prosesse in die tenderfase asook tydens voltooiing van die projek wanneer volledige data beskikbaar is oor finale hoeveelhede. Sodoende word die uitbreikbaarheid van die databasis verseker vir latere behoeftes. [Severn (1995)]

---



FIGUUR 3-34: KONSEPTUELE UITGEBREIDE MODEL

### 3.9 VOORDELE VAN NUWE ONTWERP

Die voordele verbonde aan die uitbreiding van die databasis kan gesien word in die verbetering van die benutting van ruimte in die databasis, verbetering in integriteit en vergemakliking van uitbreiding.

### **3.9.1 OPTIMALE RUIMTE BENUTTING**

---

Die bestaande databasis is ongeveer 18 megagreep groot. 'n Groot hoeveelheid van hierdie data is slegs duplisering van presies dieselfde data wat alreeds êrens anders in die model gestoor is. Nie net skep dit die moontlikheid van inkonsekwentheid wat kan ontstaan nie, maar dit lei ook tot onnodige tyd wat gespandeer word aan die invoer van data. In die voorgestelde databasis kom slegs noodsaaklike beheerde duplisering voor. Dit is bevestig nadat die bestaande databasis se data oorgedra is. Dit het geblyk dat die voorgestelde databasis se grootte ongeveer met die helfte gekrimp het na 10 Megagreep, terwyl dit presies dieselfde informasie bevat as in die bestaande databasis.

### **3.9.2 INTEGRITEIT**

---

Weens die strenger definiëring van die voorgestelde datamodel en ook omdat dit in die korrekte normaalvorm is, help dit dat inkonsekwentheid wat in die bestaande model voorkom, nie meer toegelaat word in die voorgestelde model nie. Hierdie strenger definisies sal ook help om die kwaliteit van die data te verbeter, alhoewel vingerfoute wat gemaak word nooit uitgeskakel kan word nie. Dit kan wel minimeer word deur die datavloei m.b.v. prosesmodelle te ontleed en te verbeter.

### **3.9.3 UITBREIDING VERGEMAKLIK**

---

Die voorgestelde datamodel bied die moontlikheid om die databasis verder uit te brei, nie net vir die stoor van tender eenheidspryse nie, maar ook om heelwat meer van die data wat voorkom in die tenderdokumente te stoor. Alhoewel dit ook sou moontlik wees in die bestaande databasis maak die nuwe model dit moontlik om hierdie data op 'n effektiewe en konsekwente wyse te stoor sonder onnodige duplisering.

---

---

## HOOFSTUK 4

---

### KWANTITATIEWE RISIKO-ANALISE

---

**T**radisionele deterministiese tegnieke neem nie risiko wat ontstaan weens onsekerheid tydens die verskillende kosteberamingfases van konstruksieprojekte in ag nie. Kwantitatiewe risiko-analise wat van probabilistiese tegnieke gebruik maak om risiko te identifiseer en te integreer is egter beskikbaar om besluitneming tydens onsekerheid te ondersteun.

In hierdie hoofstuk word onderskeid gemaak tussen die risiko en onsekerheid wat bestaan tydens kosteberamings. Verskillende aspekte van kosteberamings word bespreek, asook die verskillende vorms van die kostevergelyking. Tegnieke wat gebruik kan word om risiko's wat tydens kosteberekening voorkom te integreer word getoon en 'n metode word bespreek hoe onsekerheid van data verminder kan word en saam met probabilistiese tegnieke gebruik word.

---

#### 4.1 RISIKO EN ONSEKERHEID

---

Dit word gereeld gesê dat die toekoms onsekerheid en risiko inhou, omdat dit nie voorspelbaar is nie. Hierdie algemene stelling is veral van toepassing op die konstruksiebedryf waar kapitaalintensiewe projekte oor etlike maande en selfs jare strek. Kritieke besluite moet geneem word voor die aanvang van projekte, gebaseer op aannames van wat toekomstige gebeure kan inhou vir koste, tyd en kwaliteit van die projek. Die vraag kan egter gevra word of daar 'n wesentliche verskil tussen die twee begrippe onsekerheid en risiko is, wat hierdie verskil is en of daar nie eerder 'n verband tussen die twee bestaan nie.

---



In *Raftery (1994)* word aanvanklik die stelling gemaak dat risiko tradisioneel gesien is as kwantifiseerbaar en onsekerheid as onkwantifiseerbaar. Dit word ondersteun deur te sê dat risiko beskryf kan word as:

*Risiko = waarskynlikheid van gebeurtenis  $\times$  impak van gebeurtenis ..... Vgl. 4-1*

Hierteenoor kan daar by onsekerheid nie 'n waarskynlikheid vir die gebeurtenis bepaal word nie en dus kan daar nie 'n waarde aan onsekerheid gekoppel word nie. Die onderskeid tussen die twee begrippe word egter vervaag deur 'n argument te ontwikkel waar aangevoer word dat dit nie 'n substansiële verskil is nie, maar dat onsekerheid wel kwantifiseerbaar is, maar minder so as risiko.

*Ang (1975)* stel hierteenoor dat risiko voortspruit uit onsekerheid. Onsekerheid word gekwalifiseer as die verspreiding van uitkomste wat 'n gebeurtenis kan hê. Hierdie verspreiding kan gekwantifiseer word met die statistiese begrip variansie, wat die verspreiding van waarskynlike uitkomste meet. ('n Dimensioeneel meer gerieflike vorm is die standaardafwyking,  $\sigma$ , wat die vierkantswortel van die variansie is.) Risiko word dan beskryf as die ongunstige gevolg van onsekerheid, wat gemeet kan word in terme van tyd, koste en kwaliteit [*Ker-Fox (1998)*]. In *Figuur 4-1* word vir 'n item,  $i$ , getoon hoe onsekerheid weens die veranderlikheid in dié item se koste 'n risiko te weeg kan bring, omdat dit 'n uitgawe meer as die verwagte waarde,  $\mu_i$ , veroorsaak vir 'n gegewe vlak van vertroue,  $\beta$ .



Hierdie siening integreer die risiko's verbonde aan besluitneming, met en sonder inligting, op 'n elegante wyse. Die onsekerheid verbonde aan besluitneming sonder inligting is groot, omdat daar niks konkreet is waarop die besluite gebaseer kan word nie. Die risiko verbonde aan hierdie besluit in terme van tyd, koste of kwaliteit is dus hoog. Indien inligting egter bekom word waarop die besluit gebaseer kan word, is dit nie te sê dat onsekerheid verdwyn het nie, dit het net afgeneem. Die inligting mag 'n groot of klein veranderlikheid toon wat nog steeds gesien word as onsekerheid. Hierdie onsekerheid kan dus 'n groot of klein risiko veroorsaak. Slegs vir die ideale geval wanneer besluitneming met perfekte inligting gedoen word (sonder enige veranderlikheid) is daar geen risiko aan verbonde nie.

Vir hierdie studie word *Ang (1975)* se definisie gevolg, omdat dit aansluit by hoe onsekerheid verminder kan word met meer en beter inligting en daardeur die risiko verbonde aan besluitneming kan verminder.

---

## 4.2 RISIKOBESTUUR

---

In die breë sin kan risikobestuur gedefinieer word as die proses van neem en implementering van besluite met die doel om die nadelige effekte van onverwagte verliese te demp [Porro (2001)]. Hierdie proses kan verdeel word in drie fases: risiko-identifisering; risiko-analise en hantering van risiko [Rafferty (1994)].

Risiko-identifisering bestaan uit die identifisering van die risiko's wat voorkom, asook die kwalitatiewe moontlikheid van voorkoms en die effek wat dit kan hê. Prioritiserings van hierdie risiko's verseker dat aandag geskenk word aan die klaarblyklik meer kritieke risiko's as die wat 'n kleiner impak sal hê.

Risiko-analise is 'n iteratiewe proses wat te make het met die ontwikkeling van 'n model wat die probleem beskryf, die versameling van data en die ontwikkeling van 'n oplossing vir die model. Verskillende tipes tegnieke, soos deterministiese en probabilistiese, is beskikbaar om probleme mee te modelleer. Deterministiese tegnieke en sensitiviteitsanalise, alhoewel voldoende vir voorlopige analises, is onvoldoende om onsekerheid en die verwantskappe tussen die parameters van sulke modelle te modelleer. Probabilistiese tegnieke, soos Monte Carlo-simulasie en betroubaarheidanalise, *Afdeling 4.5 & 4.4*, spreek hierdie tekortkominge aan.

Verdere tegnieke wat aangewend kan word, indien beskrywende inligting beskikbaar is om voorspellings mee te maak, is liniêre regressie, nie-liniêre regressie of skynintelligensietegnieke soos neurale netwerke, *Afdeling 4.6*. Diskrete tegnieke, waar daar tussen verskillende strategieë gekies moet word in die lig van inherente risiko verbonde aan elk, kan ook gebruik word. Hierdie modelle, wat gewoonlik die vorm van besluitnemingsbome aanneem, se parameterwaardes is egter afkomstig van bogenoemde deterministiese, probabilistiese of skynintelligensietegnieke.

Hantering van die risiko het te make met hoe die bestuur die informasie afkomstig van die analise aanwend om die risiko/s te hanteer. Vier generiese strategieë is beskikbaar hiervoor [Thompson (1992)]:

- aanvaar die risiko sonder om enigiets te doen in gevalle waar dit nie geregverdig is om 'n klein risiko te beheer nie;
  - beheer die risiko deur die waarskynlikheid van die gebeurtenis te verminder of om die koste van die gebeurtenis te verlaag;
-

- oordrag van risiko na 'n ander party wat die risiko beter kan bestuur of in die vorm van versekering wat uitgeneem word;
- verwyder die risiko deur 'n ander strategie te volg waar die risiko nie voorkom nie.

In konstruksieprojekte, waar die onsekerheid en gevolglike risiko groot is, kan risikobestuur op verskeie areas toegepas word soos byvoorbeeld: skedule-ring, kwaliteit, tegniese prestasie, veiligheid, die omgewing, reputasie, ens. Een hiervan is kosteberamings wat tydens verskillende fases van 'n projek uitgevoer word. Onsekerheid weens projekbeskrywende inligting, fisiese omgewing, ekonomiese klimaat, politieke toestande, ens., veroorsaak groot risiko's wat voorkom in die berekende waardes wat gebruik word vir besluitneming. Tegniese wat egter vir kosteberamings gebruik word neem nie altyd hierdie onsekerheid in ag nie.

---

### 4.3 KOSTEBERAMING

---

Die doel van 'n kosteberaming is om te voorspel wat die koste van 'n projek gaan wees gegewe 'n beskrywing van die projek [Peurifoy (1989)]. Die beskrywing kan konseptueel of gedetailleer wees, afhangend van die tydperk in die projek se leeftyd waarop die kosteberekening uitgevoer word. Hierdie beraamde koste-inligting kan aangewend word vir besluitneming en monitering [Yeo (1990)].

Kosteberamings kan in drie hoofklasse verdeel word, nl. langtermyn, konseptueel en detail. Langtermyn kosteberamings is essensieel wanneer begrotings opgestel moet word vir kapitaalintensiewe projekte wat vir 'n aantal jare in die toekoms beplan word [Herbsman (1983)]. Konseptuele kosteberamings word uitgevoer om die lewensvatbaarheid van 'n projek te bepaal en om alternatiewe konseptuele ontwerpe te vergelyk [Oberlender (2001)]. Detailkosteberekening word tydens die ontwerpfasie uitgevoer om verskillende ontwerpe met mekaar te vergelyk [Birrell (1980)]. Gepaardgaande hiermee is dan ook die doel waarvoor dit aangewend word, nl. beplanning, besluitneming en monitering [Birrell (1980)].

Alhoewel konstruksieprojekte uniek is, is daar in meeste gevalle drie hoof partye betrokke, nl. die eienaar (kliënt), ontwerper (raadgewende ingenieur) en die kontrakteur, wat tydens die verskillende fases kosteberamings uitvoer [Peurifoy (1989)]. Die kliënt sal vir begrotings- en lewensvatbaarheidsdoeleindes kosteberamings uitvoer wat vir beplanning gebruik word, terwyl die raadgewende ingenieur dit sal doen vir besluitneming tydens konseptuele en detail ontwerp. Die kontrakteur voer weer tydens die tenderfasie kosteberamings

---

uit om die tenderbedrag te bepaal. Hierdie waardes word dan gebruik om werklike uitgawes tydens die konstruksiefase te monitor.

'n Ander komponente wat ook belangrik is by kosteberaming is die vlak van detail en volledigheid [Carr (1989), Teicholz (1993)]. As 'n bepaalde vlak van onsekerheid (of dan variasie in die data) voldoende is vir die doel waarvoor die inligting aangewend gaan word, dan is die vlak van detail wat hierdie vlak van onsekerheid lewer, voldoende. In Yeo (1990) word vyf klasse van onsekerheid voorgestel wat gekoppel is aan die verskillende fases, *Tabel 4-1*. Hierdie klasse is vergelykbaar met die kosteberaming klassifikasiesistelsel van die AACE [AACE (1997), AACE (1998)] waar die verband tussen projekdefinisievlak, doel, metodologie, verwagte akkuraathede en voorbereidingspoging getoon word. Volledigheid bepaal verder dat alhoewel daar verskillende vlakke van detail mag wees, elke vlak wel al die kostes moet insluit.

**TABEL 4-1: KLASSE VAN ONSEKERHEID**

Klas	Beraming	Waarskynlike fout	Doel
V	Orde grootte	± 25 tot 40%	Lewensvatbaarheid
IV	Faktor beraming	± 15 tot 25%	Vroeë beoordeling
III	Begroting beraming	± 10 tot 15%	Voorlopige begroting
II	Definitiewe beraming	± 5 tot 10%	Finale begroting
I	Finale beraming	± 0 tot 5%	Finale koste

Met bogenoemde in ag genome kan die volgende twee deterministiese vorme van die kostefunksie voorgestel word wat tydens die verskillende fases vir benaderde kosteberamings en detailkosteberekeninge gebruik word. Eersgenoemde funksie vir benaderde kostes lyk soos volg:

$$k_{\text{Totaal}} = \sum_{i=1}^n k_i$$

waar

$$k_{\text{Totaal}} = \text{totale projek koste; } \dots\dots\dots \text{Vgl. 4-2}$$

$$k_i = \text{koste van komponent } i;$$

$$n = \text{aantal komponente.}$$

Die kostekomponente,  $k_i$ , kan as 'n versameling van 'n aantal items (som van die produk van eenheidtariewe en hoeveelhede) gesien word.

Op gedetailleerde vlak word hierdie funksie gewoonlik by hoeveelhedslys-kontrakte gebruik waar hoeveelhede vasgestel word en die kontrakteur vir die eenheidtariewe moet tender. Die detail koste funksie kan soos volg voorgestel word:

$$k_{\text{Totaal}} = \sum_1^n h_i e_i$$

waar

$$\begin{aligned} k_{\text{Totaal}} &= \text{totale projek koste; } \dots\dots\dots \text{Vgl. 4-3} \\ h_i &= \text{hoeveelheid van item } i; \\ e_i &= \text{eenheidskoste van item } i; \\ n &= \text{aantal items.} \end{aligned}$$

Bostaande voorstellings van die kostefunksie is egter nie omvattend genoeg om die onsekerheid wat by van die veranderlikes mag voorkom in ag te neem nie. Dit kan byvoorbeeld gebeur dat 'n komponent in *Vgl. 4-2* vanuit die deterministiese beskouing nie 'n groot bydrae tot die totale koste van 'n projek lewer nie. Indien die onsekerheid wat bestaan t.o.v. die werklike koste van die komponent egter in ag geneem word, kan dit 'n groot risiko inhou en sodoende die totale koste noemenswaardig beïnvloed.

'n Meer realistiese benadering is dus om van die komponente wat die grootste bydrae kan lewer tot onsekerheid van die totale koste van die projek as ewekansige (stochastiese) veranderlikes te beskou. Die kostefunksie kan dan as volg uitgebrei word:

$$k_{\text{Totaal}} = \sum_1^k K_i + \sum_{k+1}^n k_i$$

waar

$$\begin{aligned} k_{\text{Totaal}} &= \text{totale projek koste; } \dots\dots\dots \text{Vgl. 4-4} \\ k_i &= \text{deterministiese kostekomponent } i; \\ K_i &= \text{stochastiese kostekomponent } i; \\ n &= \text{aantal kostekomponente;} \\ k &= \text{aantal stochastiese komponente.} \end{aligned}$$



Vir die detail kostefunksie is daar onsekerheid omtrent hoeveel die werklike (getenderde) eenheidtariewe gaan wees en ook wat die werklike (finale) hoeveelhede gaan wees. Die funksie kan dus uitgebrei word tot die volgende vorm om die onsekerheid in ag te neem:

$$k_{\text{Totaal}} = \sum_1^k H_i E_i + \sum_{k+1}^n h_i e_i$$

waar

$$\begin{aligned} k_{\text{Totaal}} &= \text{totale projek koste;} \\ h_i &= \text{hoeveelheid van item } i; \\ e_i &= \text{eenheidskoste van item } i; \dots\dots\dots \text{Vgl. 4-5} \\ H_i &= \text{stochastiese hoeveelheid } i; \\ E_i &= \text{stochastiese eenheidskoste } i; \\ n &= \text{aantal items;} \\ k &= \text{aantal stochastiese items.} \end{aligned}$$

Hiermee kan die totale koste van verskillende projekte beter met mekaar vergelyk word in die lig van onsekerheid wat bestaan omtrent die werklike waardes van die eenheidstarief en hoeveelheid van items.

'n Verdere toepassing hiervan is tydens tenderbeoordeling. Dit is bekend dat kontrakteurs hulle tenders struktureer deur die eenheidstarief van een item te laai met die van 'n ander, omdat beraamde hoeveelhede van die raadgevende ingenieurs nie altyd akkuraat is nie. Pryse wat op hierdie manier gestruktureer is, hou egter vir die kliënt 'n risiko in deurdat 'n tender wat op 'n deterministiese vergelykende metode gekies word, omdat dit 'n laer prys as 'n ander tender het, in werklikheid 'n groter risiko inhou weens onsekerheid wat bestaan in die hoeveelhede van die gelaaide items.

Om die effek wat hierdie onsekerheid t.o.v. eenheidstariewe en hoeveelhede van die items op die kostefunksie het, in ag te neem, kan van probabilistiese tegnieke gebruik gemaak word om die risiko te bepaal. Twee van hierdie tegnieke, nl. Monte Carlo-simulasie en Limietstaat Kostefunksie, word vervolgens bespreek.

---

#### 4.4 MONTE CARLO-SIMULASIE TEGNIEK

---

Wanneer historiese data beskikbaar is of skattings van die bestek van parameters se waardes deur spesialiste gemaak word, kan die veranderlikes in die model beskou word as ewekansige getalle in plaas van enkele vaste waardes. Die model van die probleem lewer dan nie net 'n enkel vaste waarde nie, maar 'n waarskynlikheidsverdeling met 'n bereik van waardes, teenoor die enkel vaste punt parameter model wat deterministiese resultate lewer. Hierdie "globale" waarskynlikheidsverdeling wat die model beskryf bied die geleentheid om nie net 'n afvoerwaarde te bereken nie, maar ook die vertroue (betroubaarheid) waarmee hierdie waarde voorspel word te bepaal en sodoende insig te ontwikkel in die risiko wat gekoppel word aan 'n spesifieke waarde [Boles (2000)].

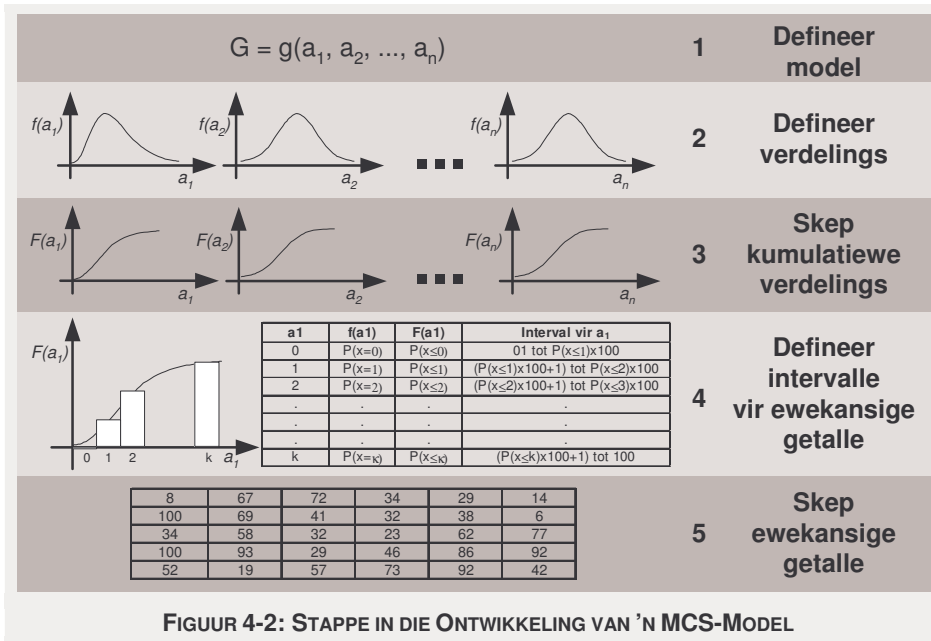
Hierdie tegniek wat ontwikkel is om 'n model van 'n probleem te bou gebaseer op die waarskynlikheidsverdelings van die parameters van die model staan bekend as Monte Carlo-simulasie (MCS). Die tegniek is ontwikkel uit militêre en operasionele spele teorie deur die bekende wiskundige *von Neuman* wat navorsing gedoen het op die ewekansigheid van die neuron. Hierdie ewekansigheid het die aanwending van die roulettewiel gesuggereer om die waarskynlikheid te hanteer. Weens die spele geaardheid het *von Neuman* dit Monte Carlo-simulasie genoem om die wette van kans te bestudeer. Met die ontwikkeling van die rekenaar het die aanwending van hierdie tegniek aansienlik veld gewen en word dit vandag in feitlik elke moontlik studierigting aangewend. Dit kan ook in vele van die fases van konstruksieprojekte waar kosteberekening gedoen moet word aangewend word [Render (2000)].

---

Vir kosteberekeninge kan hierdie tegniek met sukses aangewend word deur die eenheidtariewe en ook die hoeveelhede as ewekansige getalle te beskou. Die kostefunksie word dan as 'n globale waarskynlikheidsverdeling ontwikkeling wat die model se uitkoms beskryf. 'n Enkele koste word dus nie gevind nie, maar 'n verskeidenheid van kostes, elk met 'n vertroue van uitkoms.

#### 4.4.1 ONTWIKKELING VAN 'N MONTE CARLO-SIMULASIEMODEL

Die ontwikkeling van 'n Monte Carlo-simulasiemodel word aan die hand van ses stappe bespreek waar vyf van die stappe in *Figuur 4-2* getoon word.



- *Stap 1* – Definiëring van die model: 'n wiskundige model moet opgestel word met die nodige veranderlikes wat die probleem beskryf en daar moet besluit word watter van hierdie veranderlikes stochasties of deterministies modelleer gaan word.
- *Stap 2* – Definiëring van waarskynlikheidsverdelings: om 'n model se veranderlikes as stochastiese veranderlikes te modelleer, is dit nodig om die veranderlikes te beskryf in terme van diskrete of kontinue waarskynlikheidsverdelings. Om dit te kan doen moet die tipe verdeling gekies word en die parameters van die verdeling moet bepaal word. Twee bronne van inligting, nl. historiese data en ingenieurservaring, is beskikbaar om die verdelings en parameters te bepaal.
- *Stap 3* – Vorm kumulatiewe verdelings: die kumulatiewe waarskynlikheidsverdelings,  $F(x)$ , word ontwikkel vanaf die waarskynlikheidsmassafunksie,  $f(t)$ , van die diskrete verdelings:

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t) \dots\dots\dots \text{Vgl. 4-6}$$

of vir die waarskynlikheidsdigtheidsfunksie,  $f(t)$ , van die kontinue verdelings:

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt \dots\dots\dots \text{Vgl. 4-7}$$

- *Stap 4* – Stel van ewekansige getal intervalle: vir diskrete verdelings is dit nodig om sogenaamde vakkies van waarskynlikheid te definieer wat gekoppel is aan 'n uitkomstswaarde, bv. 'n uitkoms van sê 5 sal 'n waarskynlike vakkie hê wat tussen  $P(x \leq 4) = 45\%$  en  $P(x \leq 5) = 71\%$  is.
- *Stap 5* – Skep van ewekansige getalle: ewekansige getalle tussen 0 en 1 word gebruik om die waarskynlikhede van gebeurtenisse te simuleer. Hierdie waarskynlikhede word dan gebruik om die uitkomste te bepaal vir elk van die onafhanklike veranderlikes se verdeling. Die ewekansige getalle kan verkry word van verskeie bronne. Een van die bekendste bronne is die *Rand Korporasie* se publikasie “*A Million Random Digits with 100000 Normal Deviates*” [Rand (1955)], indien daar egter groot hoeveelhede ewekansige getalle nodig is kan dit ook maklik gegenereer word met 'n ewekansige getal generator. 'n Betroubare generator word deur *Reiser (1992)* aanbeveel:

$$z_{n+1} = a \times z_n \text{ mod } m$$

waar

$$a = 16807;$$

$$m = 2^{31} - 1.$$

..... Vgl. 4-8

Vir die doeleindes van 'n simulاسie kan die ewekansige getal dan gevind word met:

$$r = \frac{z_n}{m}$$

waar

$r =$  ewekansige getal. .... Vgl. 4-9

- *Stap 6* – Simulasie: laastens word 'n aantal simulاسies uitgevoer. 'n Simulasie bestaan uit die trek van 'n aantal ewekansige getalle gelykstaande aan die aantal onafhanklike veranderlikes. Van elk van die onafhanklike veranderlikes se kumulatiewe verdelings word dan 'n uitkoms verkry gegewe 'n ewekansige getal wat getrek is. Hierdie uitkomste word vervolgens in die model soos gedefinieer in *Stap 1* ingestel en 'n afvoer resultaat word verkry. Die proses word 'n aantal kere herhaal. 'n Histogram word opgestel van die resultate wat verkry is en sodoende kan 'n globale waarskynlikheidsmassafunksie hierop ontwikkel word.



Hierdie tegniek is toegepas op 'n denkbeeldige projek, wat in *Hoofstuk 5* bespreek word, om die werking van die MCS-tegniek te toon en ook om die resultate wat dit lewer te vergelyk met die LSKF-tegniek wat in *Afdeling 4.5* bespreek word. Vir die analise is van *MS Excel* gebruik gemaak met 'n uitbreiding van die sigbladprogram se funksionaliteit met *Crystal Ball*, 'n program wat Monte Carlo-simulasie modellering binne *MS Excel* vergemaklik. Die berekening kan in *Aanhangsel F* gevind word.

---

## 4.4.2 KOMMENTAAR OOR MONTE CARLO-SIMULASIE

---

Die Monte Carlo-simulasie tegniek bied die geleentheid om beter insig te ontwikkel in die onsekerheid wat heers by kosteberamings. In plaas daarvan om slegs 'n enkele koste van 'n projek te lewer, is daar 'n waarskynlikheidsverdeling beskikbaar waarmee 'n koste en 'n gekoppelde vertroue van voorspelling bepaal kan word.

Die akkuraatheid van die “globale”-verdeling is afhanklik van twee aspekte, nl. die wiskundige model wat die werklikheid modelleer, *Stap 1*, en die data wat gebruik word om die model te dryf, *Stap 2*. Vir kosteberekeninge is die funksie wat in die model gebruik word presies dieselfde vereenvoudigde funksie (nie al die veranderlikes word probabilisties beskou nie, *Vgl. 4-4* en *Vgl. 4-5*) as die werklike funksie wat gebruik word en derhalwe word dit verwag dat daar geen onrealisme kan ontstaan nie. Die enigste bron van onakkuraatheid lê dus in die data wat vir die model gebruik word. Drie belangrike aspekte uit die literatuur is geïdentifiseer wat hierdie onakkuraatheid kan veroorsaak, nl. die tipe verdeling; die veranderlikheid van die data en korrelasie tussen die data. Alhoewel hierdie drie aspekte voorts toegelig word met voorbeelde afkomstig van Monte Carlo-simulasies is dit ook van toepassing by ander probabilistiese tegnieke.

---

### 4.4.2.1 VERDELINGS

---

Aangesien daar verskillende tipes verdelings is om die veranderlikes van die model mee te modelleer moet daar 'n keuse gemaak word van watter verdeling die geskikste is. Alhoewel *Bradley (1990)* voel dat dit nie soseer die verdeling is wat belangrik is nie, maar eerder die verspreiding van die data, is dit tog volgens *Boles (2000)* en *Touran (1992)* een van die belangrikste besluite wat geneem moet word.

*Boles (2000)* gaan voort en stel die vyf eienskappe voor wat so 'n verdeling vir eenheidtariewe van items behoort te hê:

- enige eenheidtariefberamings het 'n bo en onder grens waarbuite enige kosteberamer seker is dat die eenheidtarief nie sal wees nie, dus moet die verdeling geslote wees;
-

- 'n kosteberaming kan enige koste aanneem tussen die grense, dus behoort die verdeling kontinuu te wees;
- dit word aangeneem dat die waarskynlikheid van voorkoms afneem na die bo en onder grense, dus moet die verdeling 'n konvekse vorm hê;
- dit word verwag dat daar slegs een koste sal wees met 'n mees waarskynlik voorkoms, dus moet die verdeling unimodaal wees;
- omdat die werklike koste groter vryheid gaan hê om hoër te wees as laer behoort daar in die kosteberekening voorsiening gemaak te word vir skeefheid.

In *Tabel 4–2* word verskillende kontinue verdelings met mekaar vergelyk in terme van die bogenoemde eienskappe. Die verdelings word in *Aanhangsel E* getoon.

**TABEL 4–2: VERSKILLENDE Tipes KONTINUE VERDELINGS**

Verdeling	Begrens	Unimodaal	Skeefheid
Normaal	Nee	Ja	Nee
Log-normaal	Onder	Ja	Ja
Uniform	Ja	Nee	Nee
Gamma	Onder	Ja	Ja
Driehoek	Ja	Ja	Ja
Beta	Ja	Somtyds	Ja
Gumbell	Nee	Ja	Ja

Indien daar geen data beskikbaar is om 'n verdeling op te pas nie en die enigste bron van inligting is 'n deskundige se ervaring om eenheidtariewe te skat, is die driehoeksverdeling die mees geskikte. Om die verdeling te definieer hoef die deskundige slegs 'n verwagte waarde en 'n bo- en ondergrens te skat.

Vir die denkbeeldige projek wat ontwikkel is vir die MCS-tegniek was data afkomstig uit die CSRA-databasis. Die denkbeeldige projek word in *Afdeling 5.7* bespreek. Hierdie data is tweeledig gebruik: eerstens om te bepaal of die data werklik van die eienskappe vertoon wat genoem is en verder om te bepaal wat die effek van verskillende verdelings op die totale koste is.

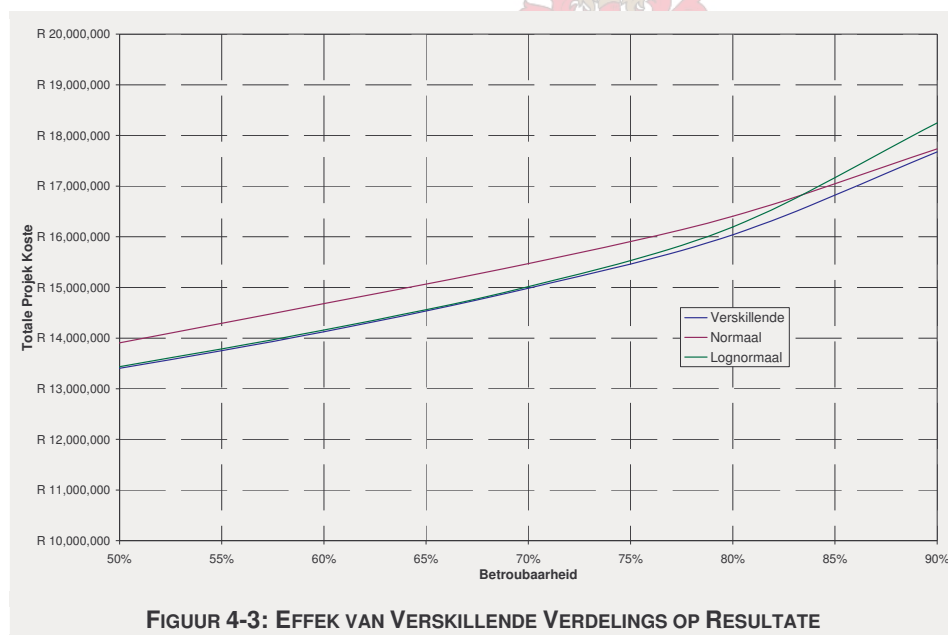
Om te bepaal watter verdelings die meeste voorkom is daar gebruik gemaak van 'n funksionaliteit in *Crystal Ball* genaamd *Batch-Fit* om die passings uit te voer. Daar is van Chi-kwadraat toetsing gebruik gemaak om te toets watter kontinue verdelings die beste pas. Die resultate daarvan word in *Tabel 4–3* getoon in volgorde van hoogste voorkoms. Dit blyk hieruit dat die eerste drie verdelings al drie die eienskap van skeefheid besit, maar dat slegs die derde verdeling die eienskap van begrensdeheid het. Dit wil dus voorkom dat geslotenheid by modellering van die databasis se data nie so kritiek is nie, maar dat skeefheid wel belangrik is. Dit is egter wel realisties om geslotenheid by die

kosteberaming van 'n gegewe projek te aanvaar aangesien 'n beramer met sekerheid kan sê dat 'n eenheidtarief binne sekere grense sal wees.

**TABEL 4-3: VOORKOMS VAN VERDELINGS**

Verdeling	Voorkoms
Gumbell	29%
Lognormaal	27%
Driehoek	12%
Normaal	10%
Gamma	7%
Beta	2%
Ander	12%

Om die invloed van die tipe verdeling wat vir modellering gebruik word te bepaal, is drie simulaties uitgevoer, elk met verskillende verdelings. Vir die eerste simulatie is die verdelings gebruik net soos dit bepaal is met *Batch-Fit*, m.a.w. die verdelings wat die beste op die data gepas het. Vir die tweede simulatie is al die stochastiese veranderlikes met normaalverdelings gemodelleer en by die laaste simulatie met lognormaalverdelings. Die resultate hiervan word getoon in *Figuur 4-3*. Hieruit kan gesien word dat die verdelings wat gebruik word wel 'n invloed het op die uitkoms van die model.



**FIGUUR 4-3: EFFEK VAN VERSKILLEDE VERDELINGS OP RESULTATE**

Indien die eerste simulatie, waar die bes passende verdelings gebruik is, beskou word as die verwysingsmodel, blyk dit dat die model waar normaalverdelings aanvaar is, afwyk van die verwysingsmodel. Die simulatie waar net lognormaalverdelings aanvaar is stem egter weer sterk ooreen met die verwysingsmodel tot by vlakke van betroubaarheid van ongeveer 75%. Dit strook met wat *Touran (1992)* voorstel, naamlik dat statistiese verdelings wel belangrik is en dat eenheidtariewe van items eerder met lognormaalverdelings behoort gemodelleer te word as enige ander verdeling.

---

#### 4.4.2.2 VERANDERLIKHEID VAN DATA

---

Die verwantskap wat bestaan tussen risiko en onsekerheid is vroeër reeds bespreek. Risiko ontstaan weens onsekerheid of dan die veranderlikheid in die data wat die modelle dryf. Dit kom voor by hoeveelhede en eenheidtariewe van items afhangend van wanneer die kosteberekeninge uitgevoer word. Die effek hiervan tree veral sterk na vore by die finale kosteberekening wanneer die eenheidtariewe en ook die werklike hoeveelhede nie bekend is nie.

Die hoeveelhede van items wat tydens die kosteberekening bepaal is, verskil van die werklike hoeveelhede wat tydens die uitvoering van die projek benut is. Verskeie redes kan hiervoor aangevoer word: vlak van detail van ontwerp, terrein toestande, onvoorsiene gebeure, wysigings weens kliënt se behoeftes, ens. [Oberlender (2001)]. Geen inligting is egter beskikbaar om te bepaal hoe groot die verwagte verskille tussen berekende en werklike hoeveelhede is nie en derhalwe is daar ook geen inligting oor die veranderlikheid daarvan nie, slegs deskundiges se beramings hieroor is beskikbaar.

Hierteenoor is historiese getenderde tariewe in meeste gevalle beskikbaar. Vir padkonstruksie kan van die CSRA-databasis gebruik gemaak word om eenheidtariewe van items te bepaal. Vir hierdie komponente van die kosteberekening word dit dus verwag dat onsekerheid verminder kan word, omdat historiese data beskikbaar is. Dit blyk egter dat daar groot variasies, soos in *Hoofstuk 5* bepaal is, in die eenheidtarief van sommige items voorkom.

Hierdie variasie bestaan uit twee komponente, nl. 'n verklaarbare en 'n onverklaarbare gedeelte [Amekudzi (2001)]. Die verklaarbare gedeelte word veroorsaak deur verskillende faktore wat by konstruksieprojekte voorkom, soos: wisselende terrein toestande, klimaat, produktiwiteit van werkers, ens. Die verklaarbare komponente is dus 'n funksie van die vlak van inligting wat beskikbaar is vir 'n gegewe projek, hetsy die huidige projek waarvoor die eenheidtariefinligting bepaal moet word of die historiese projekte waarop dit gebaseer moet word. Inligting hieroor en die effek wat dit vir elke item kan inhou is egter nie beskikbaar vir die historiese projekte in 'n databron soos die CSRA-databasis nie. Die onverklaarbare gedeelte ontstaan weens 'n inherente ewekansigheid wat voorkom.

Hierdie groot veranderlikheid wat by eenheidtariewe en hoeveelhede voorkom veroorsaak dat tydens die kosteberamings die “globale”-verdeling ook 'n groot veranderlikheid toon. Indien 'n hoëvlak van betroubaarheid vir kosteberaming deur 'n kliënt verwag word, blyk dit dat projekte onbekostigbaar duur word, omdat daar so 'n hoë risiko verbonde aan die werklike koste is weens die onsekerheid in die kostedata (naas ander onsekerhede). Dit kan veroorsaak dat lewensvatbare projekte nooit aangepak word nie. Indien dit moontlik is om die

---

variasie in hierdie data te verklein, sal dit moontlik wees om meer realistiese kosteberamings uit te voer. Hierdie aspek word verder in *Hoofstuk 6* bespreek.

---

#### 4.4.2.3 KORRELASIE

---

Die metode wat in *Afdeling 4.4.1* bespreek is om 'n MCS-model te ontwikkel, neem nie die effek van korrelasies tussen die veranderlikes in ag nie. Beide *Touran (1993)* en *Bradley (1990)* beweer egter dat dit kritiek is dat korrelasie, indien dit voorkom, ingesluit word. *Touran (1997)* ondersoek hierdie kwessie in diepte om te bepaal of korrelasie tussen eenheidtariewe korrek in MCS-modelle ingesluit kan word. Hy bevind dat rekenaarpakkette, soos *Crystal Ball*, wat van sogenaamde “ge-ordende”-korrelasie gebruik maak, dit suksesvol implementeer.

By die verskillende vorms van die kostefunksie waar stochastiese veranderlikes gebruik word, kan korrelasie tussen hierdie veranderlikes bestaan. Tydens die finale kosteberekening, wanneer beide die eenheidstarief sowel as die hoeveelheid van die beduidende items as stochastiese veranderlikes beskou word, kan korrelasie tussen die verskillende eenheidspryse, die verskillende hoeveelhede en ook tussen eenheidsprys en hoeveelheid voorkom. Die verskillende meganismes hoe korrelasie tussen die verskillende komponente van die matriks ontstaan word in *Afdeling 5.5* bespreek.

Met die data wat beskikbaar was in die CSRA-databasis, was dit moontlik om korrelasiekoëffisiënte vir die eenheidtariewe van die items in die denkbeeldige projek te bepaal, soos in *Hoofstuk 5* bespreek word. In *Figuur 4-4* kan gesien word wat die verskil in die projek koste is as dit met en sonder korrelasie gemodelleer word. Hieruit blyk dit dat *Touran (1997)* se stelling dat inagneming van korrelasie lei tot verhoogde projek kostes (of dan verhoogde kostes om betroubaarheid te verbeter weens 'n verhoging in die totale variansie as korrelasie in ag geneem word) wel korrek is. By lae vlakke van betroubaarheid lei dit weer tot 'n verlaging in die geassosieerde koste.

---



FIGUUR 4-4: EFFEK VAN KORRELASIE OP RESULTATE

## 4.5 LIMIETSTAAT KOSTEFUNKSIE

Die Limietstaat Kostefunksie-tegniek (LSKF-tegniek) is net soos die MCS-tegniek ook 'n probabilistiese analisemetode wat gebruik kan word, indien inligting beskikbaar is, om die globale betroubaarheid van 'n sisteem te bereken. Die tegniek verskil egter fundamenteel van die MCS-tegniek waar van simulاسie gebruik gemaak word om betroubaarheid te bereken teenoor die LSKF-tegniek waar dit direk bereken kan word. Hierdie tegniek is gebaseer op limietstaatteorie wat onder andere gebruik word om ontwerpsfaktore in struktuurkodes te bepaal vir 'n gegewe betroubaarheid.

### 4.5.1 LIMIETSTAATTEORIE

By ingenieursprobleme kan betroubaarheid beskryf word as die bepaling van 'n voorsieningskapasiteit om aan 'n vereiste kapasiteit te voldoen, byvoorbeeld die invloei (voorsiening) na 'n dam moet groter of gelyk wees aan die onttrekking (vereiste) van water uit die dam anders kan die watersuiweringsaanleg nie funksioneer nie. Dit kan met 'n sogenaamde prestasiefunksie voorgestel word:

$$m(x,y) = x - y$$

waar

- $x$  = deterministiese voorsieningskapasiteit; ..... Vgl. 4-10
- $y$  = deterministiese vereiste kapasiteit.

In meeste gevalle kan daar egter nie op deterministiese waardes berus word nie, omdat daar onsekerheid heers oor wat die werklike waardes van die veranderlikes is. Dit is derhalwe meer realities om die veranderlikes as ewekansige

getalle (aangedui met hoofletters) te beskou. Die prestasiefunksie verander dan na die volgende vorm:

$$M(X, Y) = X - Y \dots\dots\dots \text{Vgl. 4-11}$$

Vir ingenieurstoepassings is dit nodig om sover as moontlik te verseker dat die sisteem nie faal nie en dus dat  $Y$  oorwegend kleiner as  $X$  sal wees. Die waarskynlikheid dat dit nie so is nie en dat die vereiste kapasiteit die voorsieningkapasiteit wel oorskry is:

$$\begin{aligned} P_f &= P(X < Y) \\ &= P(M < 0) \dots\dots\dots \text{Vgl. 4-12} \end{aligned}$$

Indien  $X$  en  $Y$  statisties onafhanklik is, m.a.w. daar bestaan geen korrelasie tussen hulle nie, kan die verwagte waarde van  $M$  beskryf word deur:

$$\begin{aligned} \mu_M &= \mu_X - \mu_Y \\ \text{waar} & \dots\dots\dots \text{Vgl. 4-13} \\ \mu_X &= \text{verwagte waarde van } X; \\ \mu_Y &= \text{verwagte waarde van } Y. \end{aligned}$$

en die standaardafwyking van  $M$  kan beskryf word deur:

$$\begin{aligned} \sigma_M &= \sqrt{\sigma_X^2 + \sigma_Y^2} \\ \text{waar} & \dots\dots\dots \text{Vgl. 4-14} \\ \sigma_X &= \text{standaardafwyking van } X; \\ \sigma_Y &= \text{standaardafwyking van } Y. \end{aligned}$$



As verder aangeneem word dat beide  $X$  en  $Y$  normaalverdeel is, met hulle onderskeie parameters van verwagte waardes en standaardafwykings, dan kan die waarskynlikheid van falings,  $P_f$ , met die standaard kumulatiewe normaal waarskynlikheidstabel bereken word as:

$$P_f = \Phi\left(\frac{0 - \mu_M}{\sigma_M}\right) = 1 - \Phi\left(\frac{\mu_M}{\sigma_M}\right) \dots\dots\dots \text{Vgl. 4-15}$$

Die ooreenstemmende waarskynlikheid van sukses,  $P_s$ , is dan:

$$P_s = 1 - P_f = \Phi\left(\frac{\mu_M}{\sigma_M}\right) \dots\dots\dots \text{Vgl. 4-16}$$

Die betroubaarheid is dus 'n funksie van  $\mu_M / \sigma_M$  wat die betroubaarheidsindeks,  $\beta$ , genoem word. Hiermee kan die betroubaarheid van 'n sisteem wat uit twee parameters bestaan beskryf word. In werklikheid kan dit gebeur dat

die sisteem egter uit meer as twee veranderlikes bestaan. 'n Veralgemeende vorm van die prestasiefunksie lyk dan soos volg:

$$g(X) = g(X_1, X_2, X_3, \dots, X_n) \dots\dots\dots \text{Vgl. 4-17}$$

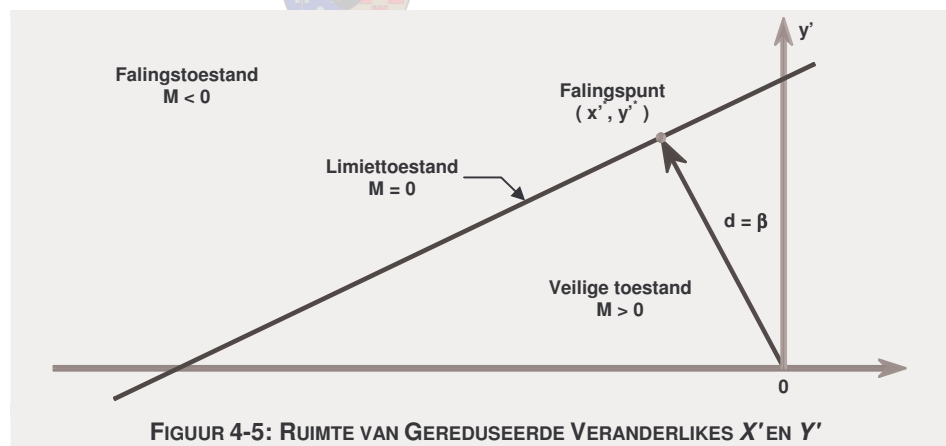
Hierdie funksie kan voorgestel word as 'n *n*-dimensionele geometriese oppervlakte. Drie toestande kan voorkom by hierdie funksie:

- $g(X) > 0$  – dan is die sisteem in 'n veilige toestand;
- $g(X) < 0$  – dan is die sisteem in 'n onveilige toestand;
- $g(X) = 0$  – dan is die sisteem by die limiet toestand.

Die limiettoestande kan gesien word as die skeidingsvlak tussen die veilige en onveilige toestand as die gereduseerde veranderlike,  $X_i'$ , gevind word van elke veranderlike,  $X_i$ , *Figuur 4-5*. Die gereduseerde veranderlike,  $X_i'$ , kan as volg bepaal word:

$$X_i' = \frac{X_i - \mu_{X_i}}{\sigma_{X_i}} \dots\dots\dots \text{Vgl. 4-18}$$

Daar kan bewys word dat die naaste punt op hierdie vlak aan die oorsprong van die assestelsel, die falingspunt,  $x^{'*}$ , met komponente  $(x_1^{'*}, x_2^{'*}, x_3^{'*}, \dots, x_n^{'*})$  is en dat die afstand *d* die betroubaarheidsindeks,  $\beta$ , is [Ang (1984)].



Die volledige ontwikkeling van hoe die falingspunt bepaal kan word, word in *Ang (1984)* bespreek. Dit is gevind dat die falingspunt,  $x^{'*}$ , se komponente,  $x_i^{'*}$ , bereken kan word deur die sogenaamde rigtingkosinus,  $\alpha_i$ , vir elke variant te vind:

$$\alpha_i = \frac{\left( \frac{\partial g}{\partial X_i} \sigma_{X_i} \right)_*}{\sqrt{\sum_i^n \left( \frac{\partial g}{\partial X_i} \sigma_{X_i} \right)_*^2}} \dots\dots\dots \text{Vgl. 4-19}$$



en die betroubaarheidsindeks,  $\beta$ , wat as volg gedefinieer word:

$$\beta = \frac{-\sum_i^n X_i \left( \frac{\partial g}{\partial X_i} \sigma_{X_i} \right)^*}{\sqrt{\sum_i^n \left( \frac{\partial g}{\partial X_i} \sigma_{X_i} \right)^*}} = \frac{\mu_g}{\sigma_g} \dots\dots\dots \text{Vgl. 4-20}$$

Die komponente,  $x_i^*$ , is dan:

$$x_i^* = \mu_{X_i} - \alpha_i \beta \sigma_{X_i} \dots\dots\dots \text{Vgl. 4-21}$$

Dit is egter nodig om die rigtingkosinusse van elke komponent van die falingspunt by die falingspunt te bepaal, derhalwe word 'n iteratiewe proses in Ang (1984) voorgestel totdat die betroubaarheidsindeks,  $\beta$ , konvergeer.

### 4.5.2 VORMS VAN LIMIETSTAAT KOSTEFUNKSIE

Die LSKF-tegniek, wat deur Ker-Fox (1998) vanuit die limietstaatteorie ontwikkel is, kan gebruik word om betroubaarheidanalises op die kostefunksie uit te voer. Weens die verskillende vorms van die kostefunksie, is daar twee verskillende vorms van die LSKF-tegniek, nl. 'n liniêre en 'n nie-liniêre vorm.

#### 4.5.2.1 LINIÊRE PRESTASIEFUNKSIE

Die liniêre vorm van die LSKF-tegniek ontstaan wanneer slegs die hoeveelhede of eenheidtariewe van items as stochastiese veranderlikes beskou word in die kostefunksie. Die koste prestasiefunksie, kan dan as volg voorgestel word:

$$g(X) = m - \sum_i^n a_i X_i$$

waar

- $m$  = beskikbare fondse; ..... Vgl. 4-22
- $a_i$  = deterministiese veranderlike  $i$  (hoeveelheid of eenheidskoste);
- $X_i$  = stochastiese veranderlike  $i$  (eenheidskoste of hoeveelheid);
- $n$  = aantal stochastiese items.

Weens die liniêre koste prestasiefunksie wat so ontstaan kan die komponente,  $x_i'^*$ , van die falingspunt,  $x'^*$ , direk bereken word. Die rigtingkosinus,  $\alpha_i$ , kan dan geskryf word as:

$$\alpha_i = \frac{-a_i \sigma_{X_i}}{\sqrt{\sum_i^n (a_i \sigma_{X_i})^2}} \dots\dots\dots \text{Vgl. 4-23}$$

en die betroubaarheidsindeks,  $\beta$ :

$$\beta = \frac{\mu_g}{\sigma_g} = \frac{m - \sum_i^n a_i \mu_{X_i}}{\sqrt{\sum_i^n (a_i \sigma_{X_i})^2}} \dots\dots\dots \text{Vgl. 4-24}$$

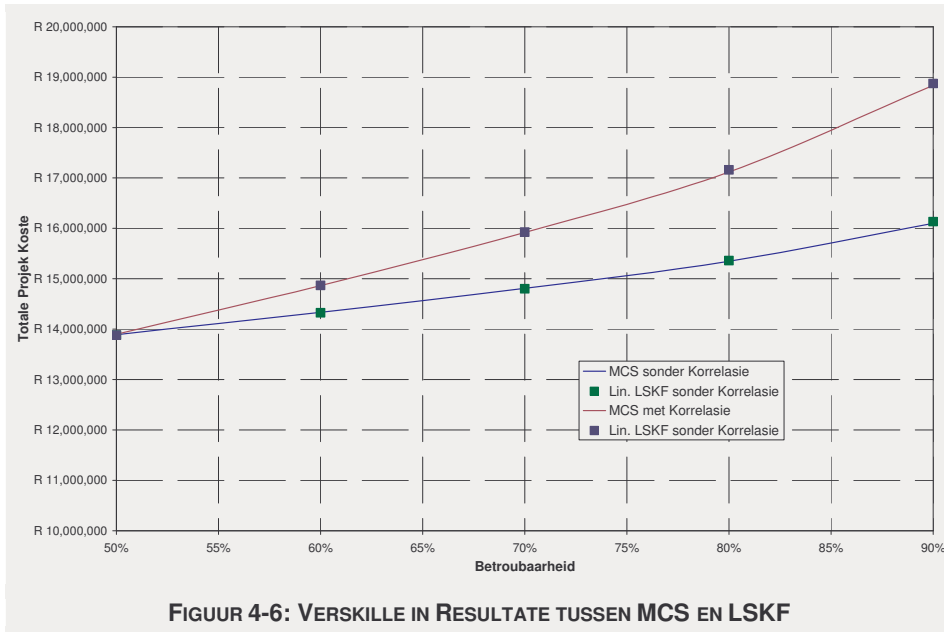
Verder stel *Ker-Fox (1998)* 'n metode voor om die komponente van die falingspunte direk te bereken vir 'n liniêre prestasiefunksie met korrelasie tussen die veranderlikes. Die rigtingkosinus,  $\alpha_i$ , kan dan as volg gevind word:

$$\alpha_i = \frac{\sum_j^n a_j \rho_{ij} \sigma_{X_j}}{\sqrt{\sum_i^n \sum_j^n a_i a_j \rho_{ij} \sigma_{X_i} \sigma_{X_j}}} \dots\dots\dots \text{Vgl. 4-25}$$

en die betroubaarheidsindeks,  $\beta$ :

$$\beta = \frac{\mu_g}{\sqrt{\sum_i^n \sum_j^n a_i a_j \rho_{ij} \sigma_{X_i} \sigma_{X_j}}} \dots\dots\dots \text{Vgl. 4-26}$$

Beide hierdie twee vorms van die prestasiefunksie kan getoets word teen die resultate wat die MCS-tegniek gelewer het. Aangesien die LSKF-tegniek egter ontwikkel is vanuit die veronderstelling dat die data normaalverdeel is, is van die resultate gebruik gemaak waar vir die MCS-tegniek aanvaar was dat die data normaalverdeel is. (Die metodes om nie-normaalverdelings in die LSKF-tegniek te akkommodeer is egter geredelik beskikbaar.) Die berekeninge van die LSKF-tegniek word in *Aanhangsel G* getoon en die resultate in *Figuur 4-6*. Hieruit blyk dit dat met en sonder korrelasie albei tegnieke feitlik dieselfde resultate lewer, derhalwe die moeilike onderskeid wat gemaak kan word op die grafiek tussen die twee pare resultate van die MCS- en LSKF-tegniek.



**4.5.2.2 NIE-LINIËRE PRESTASIEFUNKSIE MET KORRELASIE**

Die nie-liniëre vorm van die koste prestasiefunksie ontstaan wanneer beide die hoeveelhede en die eenheidstaries van die items as stochastiese veranderlikes beskou word. Die koste prestasiefunksie vir die nie-liniëre geval kan as volg geskryf word:

$$g(X) = m - \sum_{i=1}^n X_{1i} X_{2i}$$

waar

- $m$  = beskikbare fondse;
- $X_{1i}$  = stochastiese eenheidskoste veranderlike van item  $i$ ;
- $X_{2i}$  = stochastiese hoeveelheid veranderlike van item  $i$ ;
- $n$  = aantal stochastiese items.



Vgl. 4-27

Om die falingspunt,  $x'^*$ , vir 'n nie-liniëre prestasiefunksie met korrelasie tussen die veranderlikes te bepaal, moet die gekorreleerde veranderlikes omgeskakel word na 'n stel ongekorreleerde veranderlikes. Die omskakelingsproses is egter 'n omslagtige prosedure waartydens 'n ortogonale transformasie van die gereduseerde standaard normaal veranderlikes gedoen moet word. Ker-Fox (1998) het 'n direkte metode voorgestel wat gebruik kan word as die basiese veranderlikes normaal verdeel is. Hiermee kan die rigtingkosinusse,  $\alpha_i$ , bepaal word:

$$\alpha_i = \frac{\sum_j^n \left( \frac{\partial g}{\partial X_j} \right)_* \rho_{ij} \sigma_{X_j}}{\sqrt{\sum_i^n \sum_j^n \left( \frac{\partial g}{\partial X_i} \right)_* \left( \frac{\partial g}{\partial X_j} \right)_* \rho_{ij} \sigma_{X_i} \sigma_{X_j}}} \dots \dots \dots \text{Vgl. 4-28}$$

Hierdie vorm van die LSKF-tegniek is ontwikkel en kon gedeeltelik getoets word teen die resultate van die LSKF-tegniek waar die liniêre geval van die koste funksie met en sonder korrelasie beskou is. Dit kon egter nie getoets word teen die resultate van die MCS-tegniek nie, omdat daar 'n perk van maksimaal 50 ten volle gekorreleerde waardes in *Crystal Ball* is.

Om die hoeveelhede as stochastiese veranderlikes te kon modelleer is aanvaar dat:

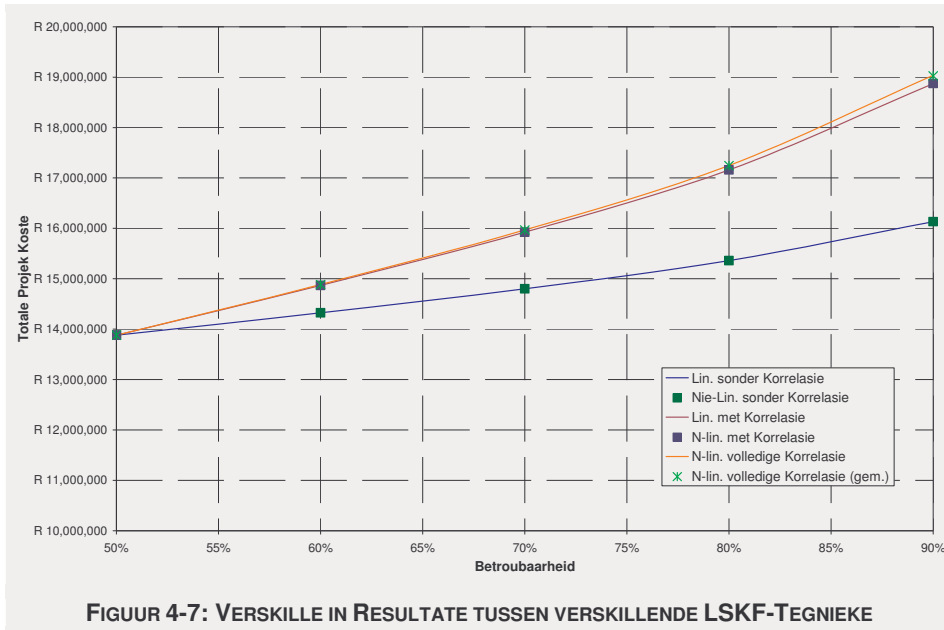
- die korrelasie wat bestaan tussen die berekende hoeveelhede dieselfde is vir die werklike hoeveelhede;
- al die hoeveelhede 'n koëffisiënt van variasie van 10% het (die verwagte akkuraatheid van hoeveelhede tydens die finale kosteberekening).

In *Figuur 4-7* word die resultate van die berekeninge getoon, asook die resultate soos gevind is vir die liniêre vorm van die LSKF. Die berekeninge word in *Aanhangsel G* getoon. Drie gevalle van die LSKF-tegniek is vir die nie-liniêre vorm bereken, nl. by die falingspunte sonder korrelasie, by die falingspunte met korrelasie en by die gemiddeldes met korrelasie.

Dit is interessant om te sien dat vir die geval waar korrelasie nie aanvaar is nie die nie-liniêre vorm feitlik presies ooreenstem met die liniêre geval waar korrelasie ook nie in ag geneem is nie. Dit word eintlik verwag dat daar 'n toename in die totale koste sal wees, weens die onsekerheid van die hoeveelhede wat ook nou in berekening gebring word.

Die rede vir die ooreenstemming kan egter gevind word as die rigtingkosinusse beskou word. Dit blyk dat weens die klein koëffisiënt van variasie van slegs 10%, wat aanvaar is vir die hoeveelhede, relatief tot die eenheidstariewe se koëffisiënt van variasie, wat tot 50% en meer is, dat die rigtingkosinusse van die hoeveelhede feitlik nul is en derhalwe nie die kostes noemenswaardig beïnvloed nie. Word korrelasie egter in ag geneem blyk dit 'n invloed te hê deur die kostes te verhoog.

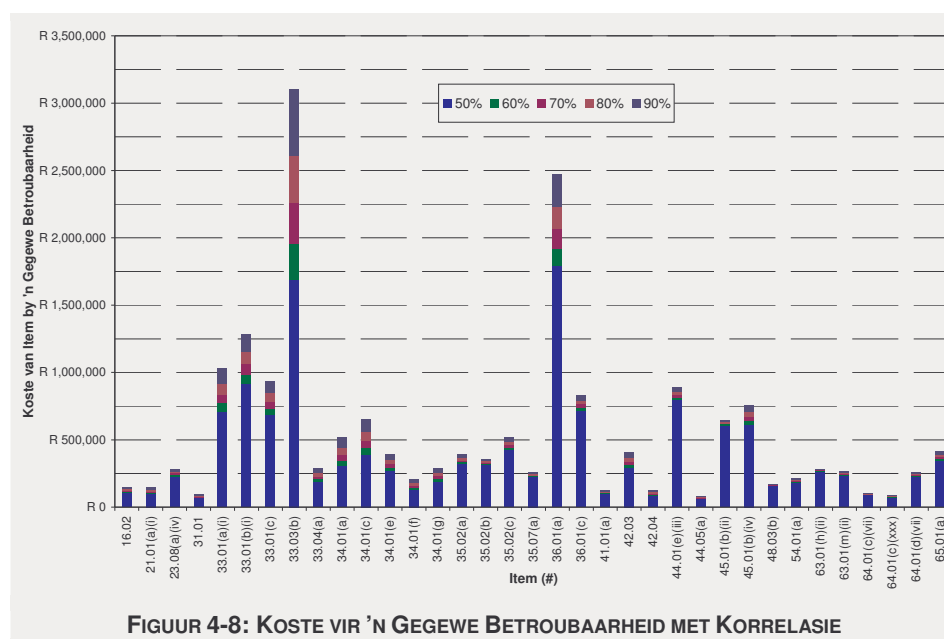
Laastens is 'n benaderde metode gevolg soos deur *Ker-Fox (1998)* aanbeveel om nie die rigtingkosinusse by die falingspunte,  $x_i^{**}$ , te beskou nie, maar by die gemiddelde waardes,  $\mu_{x_i}$ . Hierdie direkte metode noodsaak nie iteratiewe berekeninge, soos by die eerste twee gevalle om die falingspunte te vind nie. Die verskil in resultate van hierdie benaderde metode en die bogenoemde metode waar 'n iteratiewe tegniek gevolg is, is minder as 0.1%.



### 4.5.3 KOMMENTAAR OOR LSKF-TEGNIK

Alhoewel daar ooreenkomste tussen die MCS-tegniek en die LSKF-tegniek is, is daar een groot voordeel wat die LSKF-tegniek bied. Dit bied die moontlikheid om presies te bepaal hoe elke veranderlike bydra tot die totale veranderlikheid van die sisteem. Vir kosteberekening beteken dit dat vir elke item bepaal kan word hoe groot die risiko is wat dit bydra tot die totale projek se kostes.

Netsoos die werklike kostefunksie by die MCS-tegniek gebruik word om die kosteberamings probabilisties te modelleer word dit ook by die LSKF-tegniek gedoen. Geen onakkuraathede word dus verwag t.o.v. die LSKF-model nie. Wat egter meer kritiek by die LSKF-tegniek as die MCS-tegniek is, is die statistieke wat gebruik word om die stochastiese veranderlikes mee te modelleer. Die LSKF-tegniek, soos genoem, bied parametries insig in hoe risiko versprei word tussen die verskillende items. Hierdie eienskap kan aan die hand van *Figuur 4-8* gesien word wat die resultate toon van die berekening wat in *Aanhangsel G* uitgevoer is vir die denkbeeldige projek, waar korrelasie tussen eenheidtariewe en 'n item se hoeveelheid vir die nie-liniêre geval bereken is.



**FIGUUR 4-8: KOSTE VIR 'N GEGEVE BETROUBAARHEID MET KORRELASIE**

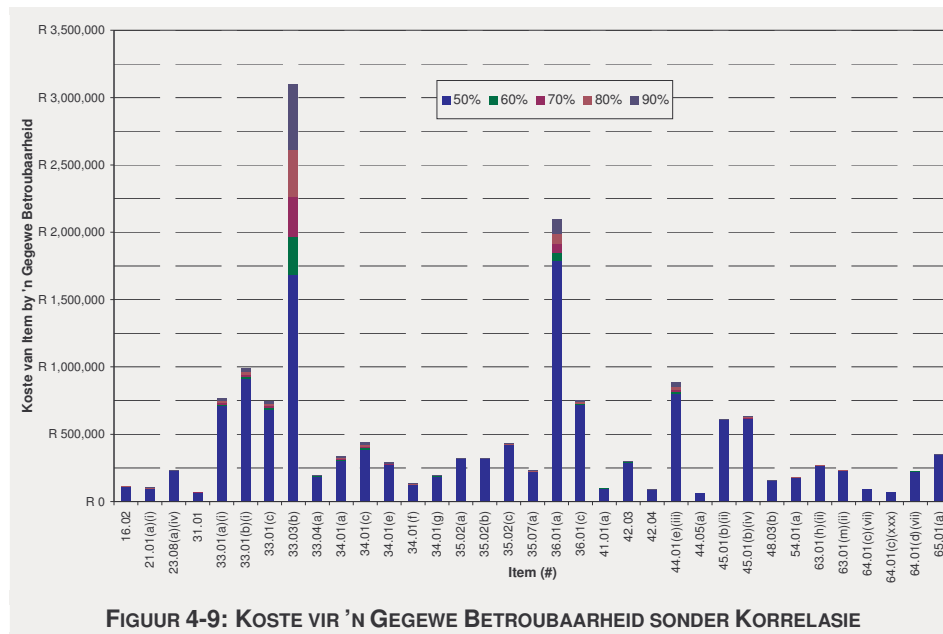
Sommige items soos 33.01(a)(i) (“Verdigting van materiaal per m<sup>3</sup>”), 34.01(a) (“Verdigting van gruiskeurlaag tot 90% gewysigde AASHTO per m<sup>3</sup>”) en 34.01(c) (“Verdigting van gruiskeurlaag tot 95% gewysigde AASHTO per m<sup>3</sup>”), blyk relatief nie sulke beduidende items te wees as deterministiese resultate van die koste van items met mekaar vergelyk word nie (stem vir hierdie berekening ooreen met 'n betroubaarheid van 50%). In elkeen se geval is daar ander items wat duurder blyk te wees en derhalwe meer aandag behoort te geniet tydens die kosteberekening. Word die variansie (onsekerheid) van die eenheidtariewe in ag geneem, blyk dit dat hierdie items by betroubarhede hoër as ongeveer 70% egter relatief duur raak weens die hoë risiko wat by elkeen voorkom. Hierdie items behoort dus in werklikheid meer aandag te geniet as van die ander items wat meer kritiek was toe onsekerheid nie in ag geneem is nie.

Die akkuraatheid waarmee hierdie risiko's voorspel word, is afhanklik van die akkuraatheid waarmee die variansies van die items voorspel word. Word die variansies verkeerdelik te groot of te klein bereken, veroorsaak dit dat items wat dienooreenkomstig nie kritiek sou gewees het nie wel as kritiek beskou word of items wat in werklikheid kritiek is weens die onsekerheid wat dit inhou, nie as belangrik geag word nie. Die hoë prioriteit van beduidende items kan dus verskuif word na minder belangrike items as die data nie korrek is nie.

Die verkeerdelike bepaling van 'n item se risiko weens die variansie wat onakkuraat is strek ook verder as die enkel item se foutief berekende variansie. Tydens die bepaling van die komponente,  $x_i^*$ , van die falingspunt,  $x^*$ , word die bydrae wat elke komponent lewer bereken relatief tot die totale bydrae deur die rigtingkosinus,  $\alpha_i$ . Indien die variansie van 'n item dus verkeerdelik te hoog of te

laag bepaal is, veroorsaak dit dat die ander items se individuele risiko's die-nooreenkomstige te laag of te hoog bereken word.

Die ander aspek wat 'n belangrike rol speel is die van korrelasie. In *Afdeling 4.4.2.3* is reeds getoon wat die effek op die berekende "globale"-verdeling is as korrelasie nie in ag geneem word nie. Die LSKF-tegniek se eienskap om parametrisie insig te verleen aan hoe risiko versprei is oor die verskillende items, noodsaak nog soveel te meer dat akkurate korrelasie-informasie beskikbaar is. In *Figuur 4-9* kan die effek gesien word wanneer korrelasie nie in ag geneem word nie, teenoor *Figuur 4-8* waar dit wel in ag geneem is. Hieruit kan gesien word dat met inagnome van korrelasie die risiko van ander items aansienlik meer krities blyk te wees as wanneer korrelasie nie in ag geneem word nie. Dit toon dus dat korrelasie inligting essensieel tot die berekeninge is.



FIGUUR 4-9: KOSTE VIR 'N GEGEWE BETROUBAARHEID SONDER KORRELASIE

## 4.6 SKYNTINTELLIGENSIE

Die veld van skynintelligensie word met groot vrug aangewend op 'n verskeidenheid van gebiede in siviele ingenieurswese. Voorbeelde hiervan in die veld van hidroulika en hidrologie is waar dit aangewend word vir die ontwikkeling van histerese sensitiewe modelle vir vloeïe-diepte kurwes [Tawfik (1997)] en droogtebeplanningbestuur [Palmer (1987)]. In vervoeringenieurswese word dit aangewend vir plaveisel onderhoud [Alsugair (1998)] en in geotegniese ingenieurswese in klassifisering van grondtipes [Goh (1995)]. In struktureeringenieurswese vind dit toepassing tydens die optimalisering van staalprofiel wat gebruik word in die konstruksie van geboue. In die veld van kosteberamings kan dit ook met groot sukses aangewend word, onder andere by die skatting van produktiwiteit [Chao (1994)] en beraming van projekkoste [Faghri (2000), Adeli (1998)].

Die twee skynintelligensietegnieke wat hiervoor gebruik kan word is besluitnemingsbome en neurale netwerke. Eersgenoemde tegnieke word oorwegend by ekspertstelsels gevind waar heuristieke (“duimreëls”) wat in die vorm van “gevalle” gestoor word, aangewend word om besluite te neem. Neurale netwerk tegnieke word weer oorwegend gebruik vir klassifisering en funksiebenaderings wat meer tolerant is vir onakkuraathede teenoor ekspertstelsel waar “harde” reëls voorkom. Neurale netwerke word dus oorwegend gebruik by kosteberaming waar die faktore wat ’n invloed het nie eksak is nie, maar heelwat data beskikbaar is. [Giarratano (1994)]

Verskillende tipes neurale netwerke is beskikbaar om die berekeninge uit te voer. Die neurale netwerk wat die meeste toepassing in siviele ingenieurswese vind is toesighoudende voorwaartsvoerende multigelaagde netwerke (“supervised feed forward multilayer”) wat gebruik maak van die algemene terugwaartsvoortplantende (“back propagation”) formulering. Toesighoudendheid verwys daarna dat daar data beskikbaar is waarteen die resultate getoets kan word en met voorwaartsvoerendheid word bedoel dat aan die eenkant data ingevoer word en daar is dan ’n afvoer. Multigelaagdheid verwys na die een of meer versteekte vlakke binne die netwerk. Die begrip “terugwaartsvoortplantende” verwys na die tegnieke wat gebruik word om die gewigte van die verbindings tussen die verskillende vlakke in die netwerk aan te pas om die netwerk te leer met die data. Hierdie tegniek stem ooreen met nie-liniêre regressie wat in statistiek voorkom. [Russel (1995)]

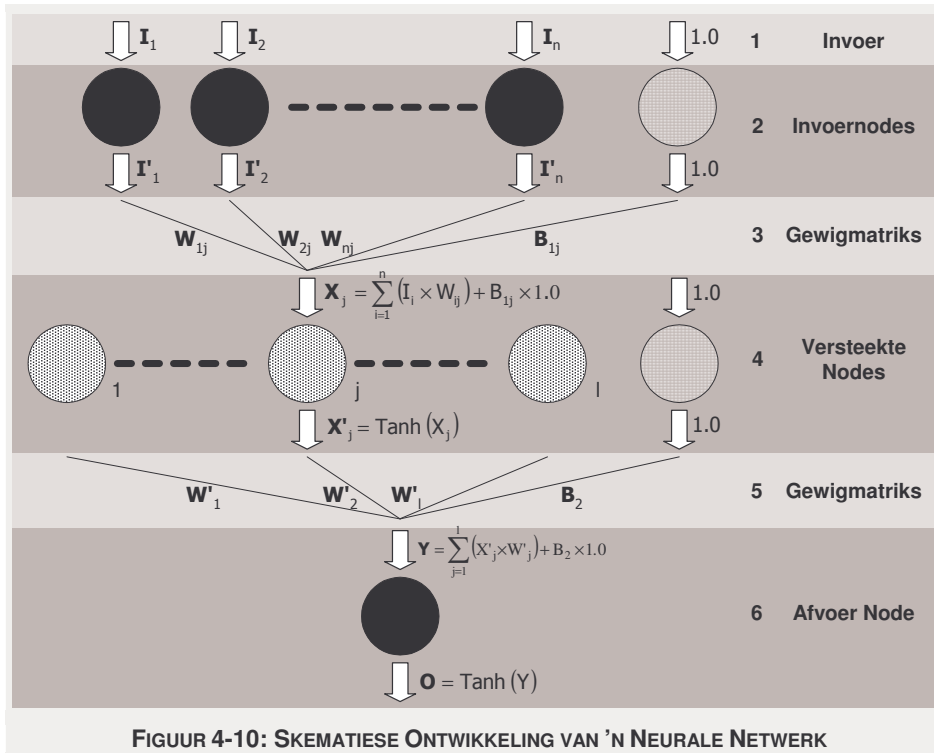
---

#### 4.6.1 ONTWIKKELING VAN ’N NEURALE NETWERK

---

Ses van die sewe stappe wat gevolg word tydens die voorwaartsvoerende berekening van ’n multigelaagde netwerk word in *Figuur 4-10* getoon en word vervolgens aan die hand van ’n referaat van *Hegazy (1998)* bespreek.





- *Stap 1 – Data-organisering:* voordat 'n neurale netwerk ontwikkel kan word is dit nodig om die data deeglik te analiseer, sodat die  $N$  onafhanklike veranderlikes en  $O$  afhanklike veranderlikes bepaal kan word. Vir  $P$  datastelle kan die data dan in 'n  $(N+O) \times P$  – matriks gestoor word.
- *Stap 2 – Dataskaling:* in hierdie stap word die invoerdata geskaal om slegs waardes vanaf  $[-1$  tot  $1]$  aan te neem. Dit word gedoen deur die invoermatriks liniêr te transformeer met die volgende funksie:

$$\text{Geskaalde waarde} = \frac{2 \times (\text{Ongeskaalde waarde} - \text{Kolom Min})}{(\text{Kolom Maks} - \text{Kolom Min})} - 1 \dots\dots \text{Vgl. 4-29}$$

Die matriks wat gevorm word is 'n  $N \times P$  matriks plus 'n kolom vir die verswaring ("bias") wat gevul word met slegs 1'e.

- *Stap 3 – Gewigsmatriks:* hierdie matriks stel die gewigte voor wat gekoppel word aan die verbindings tussen die invoernodusse en die versteekte nodusse. Indien daar  $L$  versteekte nodusse is sal daar 'n matriks gevorm word van  $N \times L$  plus 'n kolom vir die gewigte van die verswaring.
- *Stap 4 – Versteekte Nodusse:* ervaring bepaal dat versteekte nodusse ongeveer die helfte moet wees van die aantal invoernodusse ( $L = N/2$ ). Die invoer tot elk van die nodusse word bereken deur die somprodukt van elke geskaalde invoer en die gewig daaraan gekoppel. Hierdie invoer staan bekend as die sogenaamde aktivering  $X_j$ , wat gedefinieer word as:

$$X_j = \sum_{i=1}^n (I_i \times W_{ij}) + B_{1j} \times 1.0 \dots\dots\dots \text{Vgl. 4-30}$$

Die afvoer van die nodus word bereken met 'n aktiveringsfunksie, waarvan daar 'n aantal beskikbaar is om van te kies. Dit is gevind dat die *hiperboliese tan*-funksie bevredigende antwoorde lewer. Die afvoer  $X'_j$  kan dus as volg gedefinieer word:

$$X'_j = \text{Tanh}(X_j) \dots\dots\dots \text{Vgl. 4-31}$$

- *Stap 5* – Gewigsmatriks: netsoos in *Stap 3* word 'n gewigsmatriks geskep van  $1 \times L$  plus 'n verswaringnode.
- *Stap 6* – Finale Afvoer: hierdie stap stem ooreen met *Stap 4* waar die som van die produk tussen die afvoer van die versteekte nodusse en die gewigsmatriks bereken as die aktivering,  $Y_j$ , wat as volg gedefinieer is:

$$Y_j = \sum_{j=1}^l (X'_j \times W'_j) + B_2 \times 1.0 \dots\dots\dots \text{Vgl. 4-32}$$

Die afvoer,  $O$ , word ook weer met 'n *hiperboliese tan* aktiveringsfunksie verkry wat as volg gedefinieer word:

$$O = \text{Tanh}(Y) \dots\dots\dots \text{Vgl. 4-33}$$

- *Stap 7* – Terugskaling en foutberekening: die finale afvoer word nou terug geskaal deur die omgekeerde van die funksie in *Stap 2* gebruik te maak:

$$\text{Afvoer terug geskaal} = \frac{(\text{Afvoer} + 1)(\text{Maks.afvoer} - \text{Min.afvoer})}{2} + \text{Min.afvoer} \dots\dots\dots \text{Vgl. 4-34}$$

Om te toets vir korrektheid word die fout wat gemaak word bereken tussen die verskil van die neurale netwerk se afvoer en die werklike afvoer:

$$\% \text{ fout} = \frac{(\text{Neurale netwerk afvoer} - \text{Werklike afvoer})}{\text{Werklike afvoer}} \times 100 \dots\dots\dots \text{Vgl. 4-35}$$

Tydens die analisering van die data word die data in twee stappe verdeel. Die een stel data word gebruik om die neurale netwerk te leer terwyl die ander datastel gebruik word om te toets of die neurale netwerk korrek geleer is. Hierdie twee stappe data staan bekend as die leerversameling en toetsingversameling. Ervaring bepaal dat die toetsingversameling ongeveer 'n derde van die grootte van al die beskikbare data moet wees.

Hierdie sewe stappe beskryf die metode om die model op te stel. Die laaste fase is om die model se verskillende gewigte sodanig te balanseer dat dit bevredigende resultate lewer in terme van die foute wat gemaak word met die leerversameling en toetsingversameling data. Verskillende tegnieke is beskikbaar hiervoor waarvan die algemeenste die terugwaartse voorplantende metode is. Hierdie metode maak gebruik van 'n dalende helling benadering deur die fout wat na elke aanpassing van die gewigsmatrikse gemaak word, te gebruik om 'n nuwe gewigsmatriks te skep.

Ander metodes is die *simplex*-metode wat aanvanklik 'n moontlike oplossing vind en dan van die punt af die veranderlikes aanpas om beter oplossings te vind. Hierdie kragtige tegniek is beskikbaar in *MS Excel* in die vorm van die "Solver"-funksionaliteit. 'n Fundamentele ander metode as die simplekse metode is die genetiese algoritme. Hierdie tegnieke berus op die aanname dat populasies van sommige spesies na verskeie evolusies beter aanpas by hulle omgewing.

In die gids "Aanhangsel H" op die meegaande KS word 'n implementering van hierdie sewe stappe getoon. In die voorbeeld word die totale konstruksiekoste van 'n pad met sekere karakteristieke eienskappe bepaal.

---

## 4.6.2 KOMMENTAAR OOR NEURALE NETWERK

---

Die beginsel hoe 'n neurale netwerke kostebepaling doen verskil van die ander twee tegnieke deurdat daar nie van die kostefunksie gebruik gemaak word om 'n model op te stel nie. Die neurale netwerk bepaal self wat die model is wat die verwantskappe tussen die verskillende onafhanklike veranderlikes en die afhanklike veranderlike (in die geval die koste) beskryf. Neurale netwerke het ook verder die inherente eienskap om aspekte soos die effek van inflasie en korrelasie in ag te neem tydens die balansering van die model en hoef dus nie vooraf by die analise van die brondata in ag geneem te word nie.

Wat wel 'n vereiste is by die ontwikkeling van 'n neurale netwerk is klassifiseerders (m.a.w. waar die projek geleë is, tipe geologie, die jaartal, ens.) wat die gedrag van die afhanklike veranderlikes beskryf. Dit is egter hier waar historiese konstruksieprojekdata 'n leemte toon. Kriteria wat benodig word om die konstruksieprojekte te kan klassifiseer is nie beskikbaar nie. In *Hoofstuk 6* word hierdie aspek breedvoerig bespreek hoe klassifiseerders van konstruksieprojekte uitgebrei kan word.

Die tipe afvoer wat die neurale netwerk lewer verskil ook van eersgenoemde twee tegnieke. Slegs 'n enkele afvoerwaarde word bepaal per stel van invoerdata, daar is dus nie, soos in die eerste twee gevalle 'n versameling van antwoorde met gekoppelde betroubaarhede nie. Die enigste toets om te bepaal

---

hoe akkuraat die neurale netwerk voorspellings maak, is om te bepaal hoe akkuraat die netwerk teen die toetsversameling presteer.

Die grootste nadeel van 'n neurale netwerk is dat dit nie insig toon in hoe klassifikasie plaasvind nie. Dit is dus nie moontlik om te bepaal watter van die klassifiseerders 'n bepalende bydrae gelewer het nie. Neurale netwerke is egter in meeste gevalle die beste metode om klassifikasie te doen as daar voldoende inligting oor die klassifiseerder bestaan teenoor ander tegnieke soos nie-liniêre regressie. [Smith (1996)]

---

## 4.7 VERSKILLE EN INTEGRERING

---

In plaas daarvan om die twee vorms van voorspelling teenoor mekaar af te weeg sou dit meer vrugbaar wees om die unieke eienskappe wat elkeen van hierdie twee tegnieke bied aan te wend om verbeterde kosteberamings te maak. Een van die hoofredes waarom kosteberamings nooit perfek is nie is die groot veranderlikheid in die eenheidtariefbrondata en die geskatte hoeveelhede. Hierdie veranderlikheid kan verdeel word in verklaarbare variasie en onverklaarbare variasie. Dit stem in 'n mate ooreen met risikobestuur waar daar verwagte risiko's voorkom, maar ook die onverwagte risiko's.

Verklaarbare variasie by data ontstaan as gevolg van verskillende faktore wat die data beïnvloed. Hoe meer van die faktore bekend is hoe beter kan die data voorspel word. In *Hoofstuk 6* word hierdie aspek van invloedfaktore wat eenheidtariefdata beïnvloed, bespreek. Weens verskeie redes is dit egter nie moontlik om al die invloedfaktore te bepaal nie en dus kan die effek daarvan op die variasie van elke item nie voldoende gekwantifiseer word nie. Daar bestaan dus altyd 'n gedeelte onverklaarbare variasie weens 'n tekort aan voldoende inligting wat in die kosteberekening geïntegreer moet word om 'n globale onsekerheid te bepaal vir die totale koste van die projek.

Die eienskappe van die voorafgaande tegnieke kan nuttig aangewend word om albei hierdie aspekte aan te spreek. Neurale netwerke kan aangewend word om beter voorspellings te maak van die verwagte eenheidtarief met die invloedfaktore wat bekend is. Hierdie voorspellings is egter nie perfek nie, omdat daar aanvaar kan word dat nie al die invloedfaktore bekend is nie. 'n Mate van onsekerheid bestaan dus oor die resultate wat verkry word van die voorspellings. Hierdie onsekerheid is bekend, omdat dit as een van die afvoere van die neurale netwerk, in die vorm van die akkuraatheid van die voorspelling per item, gelewer word.

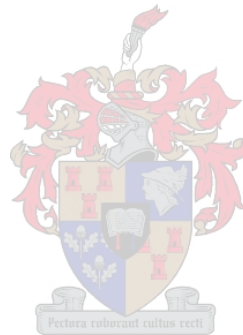
Om hierdie onsekerhede te integreer kan enige van die twee probabilistiese tegnieke wat bespreek is, aangewend word om die globale statistiese verdeling van die totale projek koste onsekerheid te bepaal met die waardes wat van die

---

neurale netwerk verkry is vir die eenheidtariewe en hoeveelhede. Hierdie gekombineerde tegniek het die voordeel dat die probabilistiese invoerdata die projek eienskappe ingesluit het weens die neurale netwerk se klassifisering eienskap en die onsekerheid waarmee die neurale netwerk die voorspellings van elke veranderlike gemaak het. Die resultaat wat vir die totale projekkoste verkry word is dus nie net projekspesifiek nie, maar daar kan ook 'n betroubaarheid aan gekoppel word.

Die tegniek het egter die nadeel dat die korrelasie tussen die veranderlikes nie in ag geneem word nie en dat die onsekerheid wat gepaardgaan met die akkuraatheid waarmee die raadgevende ingenieur die invloedfaktore gekwantifiseer het ook nie in berekeninge gebring word nie. Hierdie aspekte word verder in *Afdeling 6.8.5* bespreek.

---



---

## HOOFSTUK 5

---

### DATA-ONTLEDING

---

**D**ie geaardheid van data, of dan statistieke daarvan, bepaal hoe data vir besluitondersteuningsdoeleindes tydens analise gemanipuleer behoort te word. Die geaardheid van data sluit onder meer die eerste en tweede momente, die verdelings wat dit beskryf en korrelasies in. Data word egter ook beïnvloed deur die omgewing (ruimtelike of temporaal) waar binne dit gebruik word. In sommige gevalle is dit dus nie voldoende om slegs data se statistieke te beskou nie, maar die aspekte binne die omgewing wat dit mag beïnvloed behoort ook in ag geneem te word.

In hierdie hoofstuk word die CSRA-databasis se data as voorbeeld gebruik om die eienskappe wat die eenheidtariewe van padkonstruksie besit, te toon. Indiepte analyses van al die data in die databasis is egter nie nodig nie en daarom word daar gefokus op items wat gewoonlik 'n noemenswaardige bydrae tot die koste van projekte het. As eerste stap tot die ontleding van die data word die identifisering van hierdie beduidende items ondersoek. Die data word ook gebruik vir die ontwikkeling van 'n denkbeeldige projek wat in die voorafgaande hoofstuk gebruik is.

---

#### 5.1 IDENTIFISERING VAN BEDUIDENDE ITEMS

---

In die ideale stelsel waar alle data onderskep word vir besluitondersteuningsdoeleindes is dit nie nodig om met steekproewe uit versamelings van data te werk nie. Die praktiese realiteite van koste (geld en tyd) is egter 'n struikelblok om altyd alle data beskikbaar te hê. In 'n werklike omgewing waar daar 'n

---

bepierking op finansiële hulpbronne en tyd is, is 'n meer gefokusde benadering nodig.

Die uitbreiding van data-onderskepping met die doel om die variasie van data in kosteberekening te verminder is reeds genoem en word verder in *Hoofstuk 6* bespreek. Om onderskepping uit te brei is dit nodig om te bepaal watter veranderlikes (hoeveelhede en eenheidtariewe van items) 'n substansiële invloed het op die kosteberekening wat uitgevoer word en wat die geaardheid van hierdie data is, sodat hierdie beduidende veranderlikes korrek gemodelleer word in die kosteberekening.

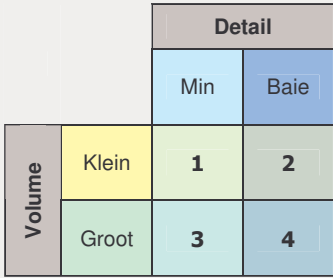
In terme van die omvang van data-onderskepping is daar twee aspekte wat in ag geneem moet word, nl. die volume van data wat onderskep word en die vlak van detail wat oor die data onderskep word. In eersgenoemde geval, nl. die volume data, kan die volgende twee ekstreme situasies beskou word:

- Groot volume: alle data word op 'n gegewe vlak van detail onderskep;
- Klein volume: slegs geselekteerde data word op 'n gegewe vlak van detail onderskep.

In die geval van laasgenoemde, nl. die vlak van detail van onderskepping, kan daar ook twee gevalle beskou word:

- Baie detail: inligting oor verskillende vlakke van detail word vir 'n gegewe data onderskep;
- Min detail: slegs die inligting op 'n gegewe vlak word onderskep.

Verskillende kombinasies van hierdie twee aspekte van data-onderskepping kan deur die ontwikkelaar van 'n besluitsteunstelsel gekies word. Die verskillende moontlike kombinasie kan voorgestel word met 'n onderskeppingstrategie matriks soos getoon in *Figuur 5-1*.



		Detail	
		Min	Baie
Volume	Klein	1	2
	Groot	3	4

**FIGUUR 5-1: ONDERSKEPPINGSTRATEGIE MATRIKS**

Kwalitatief beskou in terme van koste (tyd en geld) is benadering 1 vir die onderskepping van die data laag, maar so is ook die moontlike inligting wat daaruit geput kan word waarskynlik heeltewel te beperk vir besluitondersteu-

ningsdoeleindes. Groot onsekerheid kan voorkom weens groot variasies in die data wat ingesamel is as gevolg van te min detail oor die omgewing waarbinne dit ontstaan het. Te min data kan ook beskikbaar wees om modelle volledig te modelleer. Hierteenoor is benadering 4 vir besluitondersteuningsdoeleindes die ideale situasie waar groot volumes data met baie detail onderskep word. Die koste kan egter uiters hoog wees en veroorsaak dat hierdie opsie nie oorweeg word nie. Die twee meer aantreklike opsies blyk dus 2 of 3 te wees.

Die verskynsel van klein hoeveelheid met 'n groot impak (80-20-reël) of wat in die volksmond as "*Pareto se Beginsel*" bekend staan (alhoewel verkeerdelik so volgens *Juran (2003)*) kan die voorkeur in die guns van benadering 2 draai. Om egter die keuse te kan maak van watter veranderlikes kritiek is en wat nie is nie, is dit nodig om data aanvanklik op 'n breë front te onderskep. In die ontwikkeling van enige besluitsteunstelsel behoort daar dus by benadering 3 begin te word en met verloop van tyd oor te skuif na benadering 2. Hierdie moet egter nie gesien word as 'n enkele onderskeppingsbenadering verskuiwing nie, maar eerder 'n ontwikkelingsproses waar daar meer male vanaf benadering 3 na 2 beweeg moet word soos insig ontwikkel word in wat essensieel is op verskillende vlakke van detail van onderskepping.

In die geval van die padkonstruksie voorbeeld is daar alreeds 'n bestaande databasis beskikbaar waar feitlik al die eenheidtariewe van items wat in padkonstruksie gebruik word, naastenby 3000-items, onderskep word. Feitlik geen inligting word egter oor die omgewing waarbinne die data ontstaan, onderskep nie. Die benadering wat dus hier gevolg word kan gesien word as benadering 3. Om die stelsel te verbeter behoort daar dus vanaf een onderskeppingstrategie op 'n gewewe vlak na 'n volgende beweeg te word deur beduidende items te identifiseer en dan te bepaal watter detail onderskep moet word. Die CSRA-databasis word vervolgens gebruik om hierdie items te identifiseer.

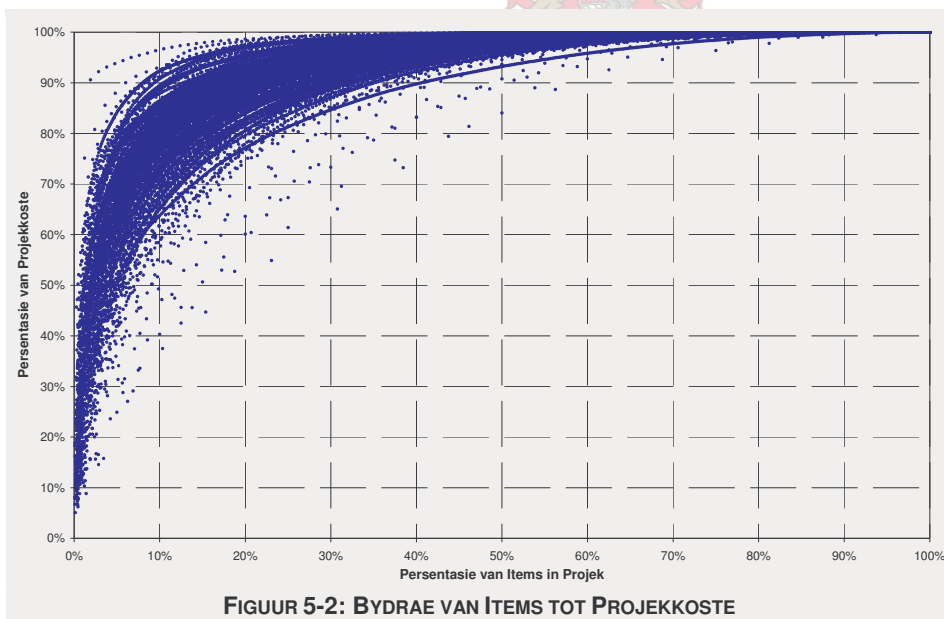
Die items van die databasis kan in nege afdelings verdeel word, wat bekend staan as Series. In Tabel 5–1 word die verskillende Series getoon met 'n beskrywing van elk. Die voorkoms en bydrae van hierdie items tot die totale koste van projekte varieer drasties. Sommige items kom feitlik in elke projek voor, terwyl ander items slegs in 'n paar projekte gebruik word. Die bydrae van items tot die totale koste wissel ook. Sommige items lewer 'n aansienlik groter bydrae tot die totale koste van 'n projek as ander items. Die items wat gereeld sulke noemenswaardige bydraes maak is gevind deur twee stelsel berekeninge uit te voer.



TABEL 5-1: LYS VAN DIE CSRA-SERIES

Serie	Beskrywing
1000	Algemeen
2000	Dreinerings
3000	Grondwerke en plaveisel van gruis materiaal of gebreekte klip
4000	Asfaltplaveisel
5000	Diverse Padwerke
6000	Strukture
7000	Diverse Strukture
8000	Diverse
9000	Aanvullende Bepalings

In die eerste geval is bepaal wat die verwantskap is tussen die persentasie items van 'n projek en die kumulatiewe bydrae wat hierdie items tot die totale koste per projek maak. Die koste per item in die CSRA-databasis is verwerk om hiervoor te toets. Die items in elke projek is gerangskik van grootste tot kleinste koste en die kumulatiewe bydrae tot die projek koste is bepaal. Die berekening word in *Aanhangsel I* getoon. Dit is gevind dat tussen 10% en 20% van die items 'n bydrae van gemiddeld tussen 80% en 90% tot die totale projek koste lewer, *Figuur 5-2*. Dit bevestig dus dat vir kwantitatiewe analise doeleindes deur slegs 'n beperkte aantal items van 'n projek te analiseer 'n groot gedeelte van die projek koste in werklikheid onder beskouing geneem word.



Tweedens is ondersoek ingestel na watter items gereeld sulke noemenswaardige bydraes lewer. Die identifisering van die beduidende items binne elke *Serie* is in drie stappe gedoen:

- die items se data in die databasis is verdeel volgens die verskillende *Series*;

- die items binne elke *Serie* is gevind wat 'n bydrae groter as 1% maak tot die totale koste van 'n projek;
- die items wat in die vorige stap gevind is, is geselekteer indien dit in meer as 5% van die projekte voorkom.

Die items wat aan hierdie kriteria voldoen is vanuit die CSRA-databasis verkry deur van twee SQL soektogte gebruik te maak. Die uitvoering hiervan word in *Aanhangsel J* getoon. Aangesien die eenheidtariewe van eerste en tweede tenders in die databasis gestoor is, is die ondersoek vir beide tenders gedoen om te bepaal of daar enigsins 'n noemenswaardige verskil tussen die voorkoms van die beduidende items tussen die twee tenders is. Dit kan in die laaste kolom van *Tabel 5–2* gesien word dat die items feitlik presies dieselfde is vir albei tenders. Daar is in totaal 44-items gevind wat aan die kriteria voldoen het. Die resultate van die aantal items wat per *Serie* gevind is en watter items dit is word onderskeidelik in *Tabel 5–2* en *Tabel 5–3* getoon.

**TABEL 5–2: VOORKOMS VAN ITEMS MET VERSKILLENDE KRITERIA**

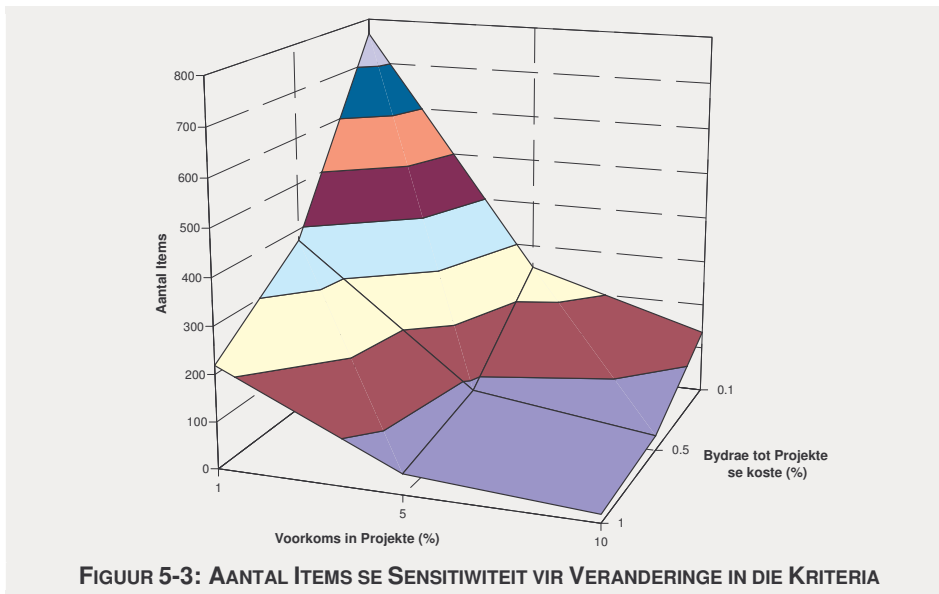
Serie	Projekitems		Bydrae > 1%		Voorkoms > 5%		Tender 1 & 2
	Tender 1	Tender 2	Tender 1	Tender 2	Tender 1	Tender 2	
1000	22372	21868	1660	1569	9	9	8
2000	13915	13556	505	497	2	2	2
3000	10433	10119	1384	1355	17	18	17
4000	6389	6323	923	909	9	9	9
5000	18575	18016	494	463	3	1	1
6000	14675	13700	877	831	6	7	6
7000	354	357	30	29	0	0	0
8000	642	595	111	92	1	1	1
9000	2623	2612	237	221	0	0	0

Dit blyk hieruit dat die *Serie 1000*, *5000* en *6000* die meeste projekitems het, terwyl die *Serie 7000* en *8000* die minste het. As hierdie resultate gefilter word om te bepaal hoeveel projekitems per *Serie* 'n noemenswaardige bydrae tot die projek koste lewer blyk dit dat die *Serie 1000* saam met die *Serie 3000* na vore tree. As hierdie resultate nog verder gefilter word om te bepaal presies watter items by meer as 5% van die projekte 'n noemenswaardige bydrae lewer is dit die *Serie 1000*, *3000* en *4000* wat die meeste items het. Dit stem goed ooreen met 'n studie van *Nel (1987)* waarin hy gevind het dat 36% van alle padkonstruksie uitgewes afkomstig is van die *Serie 3000* se items.

In die CSRA-databasis is daar 416-projekte en 'n totaal van 91323-projekitems (waarvan daar 'n eerste, tweede en soms 'n derde tender is). Die gemiddelde aantal items wat in 'n projek voorkom is dus 219. As konserwatief aangeneem word dat slegs 50% van die beduidende items wat geïdentifiseer is in 'n gemiddelde projek voorkom, m.a.w. 22-items, maak dit 10% uit van die items vir 'n gemiddelde grootte projek. In *Figuur 5-2* kan gesien word dat 10%

van die aantal items ongeveer 'n 80% bydrae tot die totale koste van 'n projek sal hê.

Die aantal items verteenwoordig dus vir die gemiddelde projek 'n substansiële bydrae van die totale koste. Deur egter die kriteria te verander is dit moontlik om die aantal items wat as bepalend beskou word te verminder of te vermeerder. In *Figuur 5-3* kan die effek gesien word indien die twee kriteria gevarieer word. Vir hierdie studie is daar egter met die 44-items volstaan.



Met die beduidende items nou bekend kan daar gekyk word na die statistiese geaardheid van hierdie items, sodat bepaal kan word watter items verder ondersoek behoort te word. Dit is egter nodig om eers twee ander aspekte te ondersoek, naamlik die foute in data en die tydveranderlikheid van die data, voordat die statistiese ontledings gedoen word.

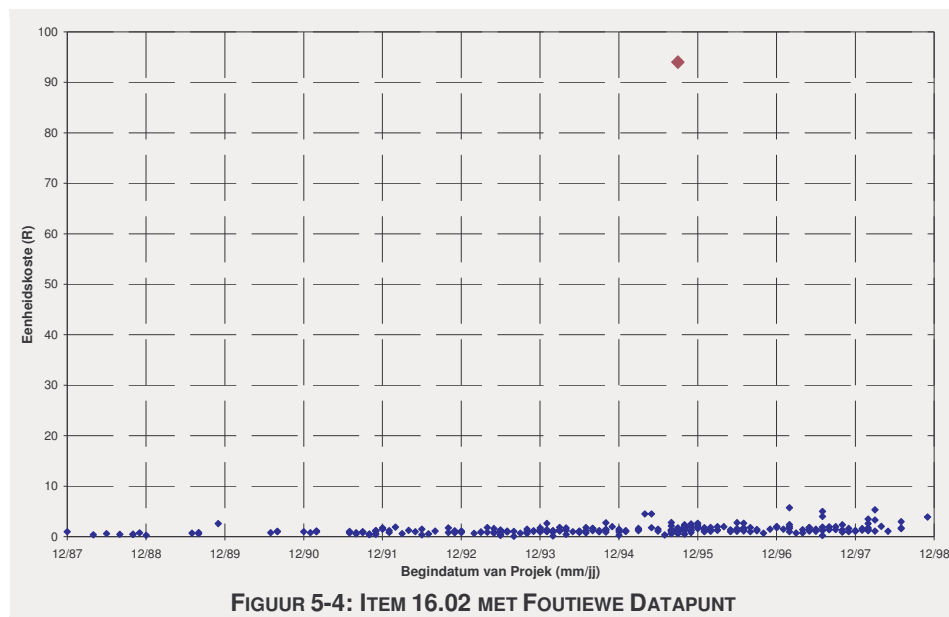
TABEL 5–3: BEDUIDENDE ITEMS MET BESKRYWINGS

Item	Beskrywing	Eenheid
13.01(a)	Aannemer se vaste verpligtinge	GB
13.01(b)	Aannemer se waarde verwante verpligtinge	GB
13.01(c)	Aannemer se tydverwante verpligtinge	maand
14.03(b)(ii)	Hanteringskoste en wins ten opsigte van telefoonkoste	%
14.07(a)	Voorlopige bedrag vir die voorsiening van gehuurde behuising	VB
14.07(b)	Hanteringskoste en wins ten opsigte van gehuurde behuising	%
15.01	Neem van verkeer en instandhouding van verleggings	km
16.02	Bykomende vervoer vir materiaal verder as 1,0 km	m <sup>3</sup> -km
21.01(a)(i)	Uitgrawing van sagte materiaal 0 m tot 1,5 m diep vir dreinerings	m <sup>3</sup>
23.08(a)(iv)	Betonvoering vir vore van Klas 20/19-beton vir ander vore	m <sup>3</sup>
31.01	Verwydering van oormaatbelaag oor groefmateriaal	m <sup>3</sup>
33.01(a)(i)	Gruismateriaal in verdigte laagdiktes van 200 mm en minder - verdig tot 90% van gewysigde AASHTO-digtheid	m <sup>3</sup>
33.01(b)(i)	Gruismateriaal in verdigte laagdiktes van 200 mm tot 500 mm - verdig tot 90% van gewysigde AASHTO-digtheid	m <sup>3</sup>
33.01(c)	Rotsopvullings uit deurgrawings en groewe insluitende gratisvervoer tot 1km	m <sup>3</sup>
33.03(b)	Ekstra bo en behalwe item 33.01 vir die uitgrawing en opbreek vir harde uitgrawing	m <sup>3</sup>
33.04(a)	Weggooi van deurgrawingsmateriaal insluitende gratisvervoer vir sagte uitgrawing	m <sup>3</sup>
34.01(a)	Gruiskeurlaag verdig tot 90% van gewysigde AASHTO-digtheid gebou met gruis verkry uit groewe of deurgrawing	m <sup>3</sup>
34.01(c)	Gruisstutlaag verdig tot 95% van gewysigde AASHTO-digtheid gebou met gruis verkry uit groewe of deurgrawing	m <sup>3</sup>
34.01(e)	Gruiskroonlaag verdig tot 97% van gewysigde AASHTO-digtheid gebou met gruis verkry uit groewe of deurgrawing	m <sup>3</sup>
34.01(f)	Gruisskouers verdig tot 93% van gewysigde AASHTO-digtheid gebou met gruis verkry uit groewe of deurgrawing	m <sup>3</sup>
34.01(g)	Gruisslytlaag verdig tot 93% van gewysigde AASHTO-digtheid gebou met gruis verkry uit groewe of deurgrawing	m <sup>3</sup>
35.02(a)	Chemiese stabiliseermiddel - portlandsement	t
35.02(b)	Chemiese stabiliseermiddel - portlandhoogoond sement	t
35.02(c)	Chemiese stabiliseermiddel - gebluste padkalk, kalsium en/of magnesium tipe	t
35.07(a)	Chemiese stabiliseermiddel - anioniese bitumenemulsie (60% netto bitumen)	l
36.01(a)	Gebreekte klipkroonlaag gebou met tipe-G1 materiaal wat uit handelsbronne verkry word	m <sup>3</sup>
36.01(c)	Gebreekte klipkroonlaag gebou met tipe-G2 materiaal wat uit handelsbronne verkry word	m <sup>3</sup>
41.01(a)	Drenklaag: RTH-3/12P-teer drenkmiddel of RTL-3/12P-teer drenkmiddel	l
42.02(a)(xiii)	30 mm dik asfaltdeklaag deur gebruikmaking van 60/70 penetrasiegraad bitumen	m <sup>2</sup>
42.03	Ingerolde klipskerwe in deklaag, 19,0 mm nominale grootte	t
42.04	Kleeflaag van 30%-stabiele graad emulsie	l
44.01(e)(iii)	Enkelseels met 13,2-mm-aggregaat	m <sup>2</sup>
44.05(a)	Voorafbedekking van aggregaat met Sacrosote 70 of 'n goedgekeurde ekwivalent	m <sup>3</sup>
45.01(b)(ii)	Dubbelseëls met 19,0-mm- en 9,5-mm-aggregaat en 80/100 penetrasiegraad bitumen	m <sup>2</sup>
45.01(b)(iv)	Dubbelseëls met 19,0-mm- en 9,5-mm-aggregaat en ander middels	m <sup>2</sup>
48.03(b)	Tipe 3 flodder wat met die hand aangewend word	m <sup>3</sup>
54.01(a)	Skram relings en houtpale se versinking	m
63.01(h)(ii)	Hoëvloei spanningstaalstawe, warmgewalste staalwapening vir dek	t
63.01(m)(ii)	Hoëvloei spanningstaalstawe, warmgewalste staalwapening vir hele struktuur	t
64.01(c)(vii)	Klas 30/19 in situ-gegiete beton vir die dek	m <sup>3</sup>
64.01(c)(xxx)	Klas 30/19 in situ-gegiete beton vir ander dele	m <sup>3</sup>
64.01(d)(vii)	Klas 40/19 in situ-gegiete beton vir die dek	m <sup>3</sup>
65.01(a)	Spankabels in die langsgigting	MN-m
81.02(b)	Vordering op inkoopprys	%

## 5.2 FOUTE IN DATA

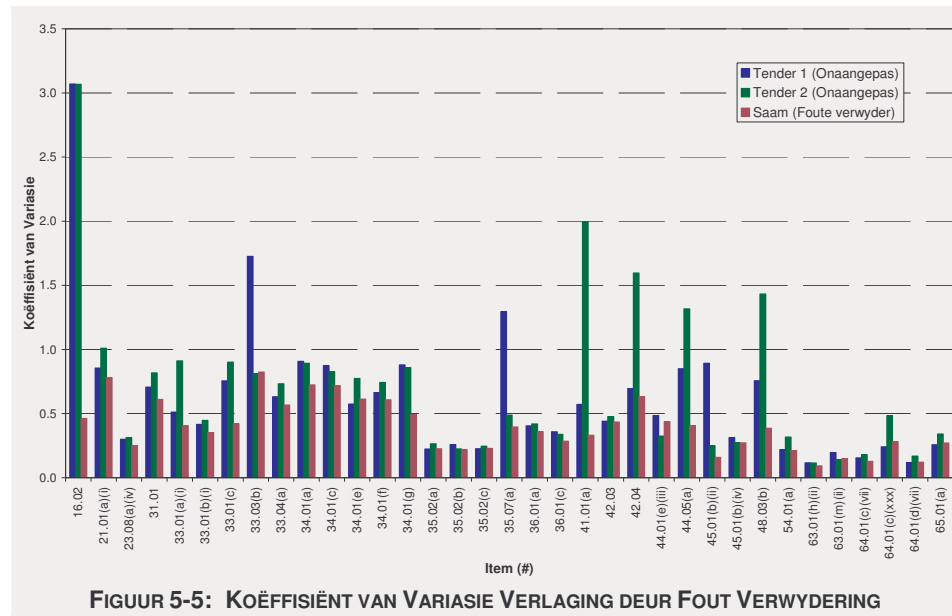
Die oorsake van foute in data is alreeds bespreek in *Afdeling 3.4*. Dit is genoem dat dit feitlik onmoontlik is om foute na die tyd te identifiseer en te korrigeer, omdat die oorspronklike dokumente benodig word om dit te kan doen. In die geval van die CSRA-databasis waar daar ongeveer 180 000-getenderde eenheidtariefinskrywings is, is dit ook nie haalbaar nie, omdat van die dokumente nie meer bestaan nie en dit 'n geweldige arbeidsintensiewe oefening is. Dit is egter nie te sê dat nuwe inligting wat ingevoer word nie behoort gekontroleer te word nie.

Indien die data van die CSRA-databasis beskou word, is dit egter maklik om van die foute in die datastel van elke item te identifiseer. Hierdie foute kan uitgeken word deurdat die foutiewe datapunte in meeste gevalle 'n orde groter is as die res van die datapunte in die reeks. In *Figuur 5-4* word 'n werklike geval getoon vir een van die beduidende items. Die datapunt moes R0.93 gewees het, maar dit is foutiewelik as R93.00 ingevoer – twee ordes te groot!



FIGUUR 5-4: ITEM 16.02 MET FOUTIEWE DATAPUNT

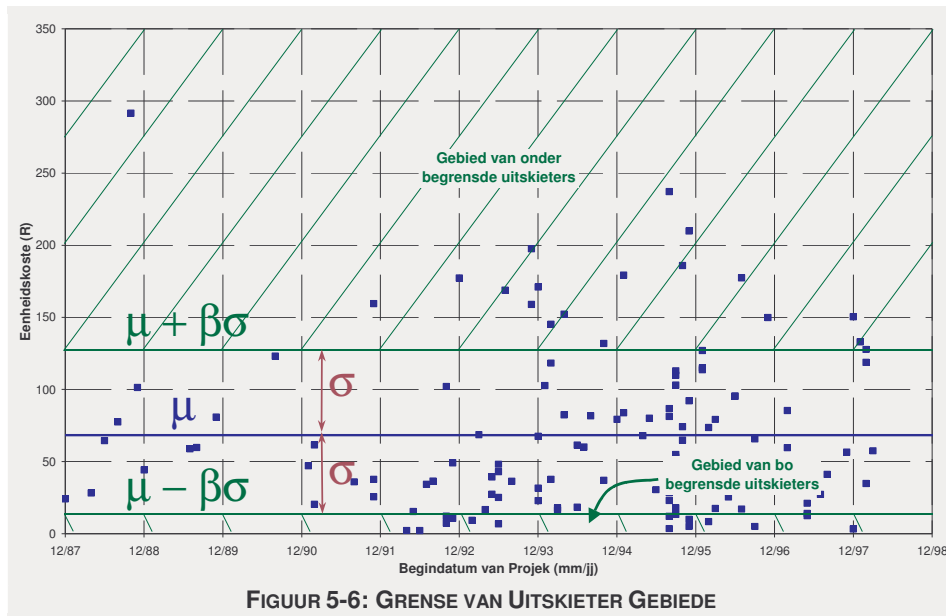
Hierdie foute kan 'n noemenswaardige impak op die variansie van 'n item se eenheidtariewe hê. In die geval van *Item 16.02* ("Bykomende vervoer verder as 1km per m<sup>3</sup>-km") is daar 275-datapunte in die reeks, tog beïnvloed hierdie enkele foutiewe datapunt die variansie dramaties. In *Figuur 5-5* kan gesien word dat die koëffisiënt van variasie (word in *Afdeling 5.4* gedefinieer) vanaf ongeveer 3.1 verlaag het na minder as 0.5 vir hierdie item, indien die foutiewe punt verwyder word. In die figuur kan ook gesien word dat heelwat van die ander items se koëffisiënte van variasie ook dramaties verlaag is deur slegs hierdie "orde grootte"-uitskieters te verwyder.



Die data kan nog verder gefilter word vir verdere minder opsigtelike foute deur visuele inspeksie of 'n filter kan in die databasis opgestel word om uitskieters te identifiseer. Om die data te filter kan gebruik gemaak word van die gemiddelde en die standaardafwyking van die populasie om bo- en ondergrense te bepaal, nadat bogenoemde “orde-grootte”-uitskieters en tydafhanklikheid (*Afdeling 5.3*) verwyder is.

Die voordeel van die gebruik van die gemiddeld en die standaardafwyking om grense te bepaal teenoor visuele inspeksie, is dat 'n sekere vlak van betroubaarheid, m.b.v. die normaalverdeling, gekoppel kan word aan die omhullingskurwe wat gevorm word deur die bo- en ondergrense. Dit word egter aangeneem dat die eenheidtariefdata van elke item normaalverdeel is, wat nie heeltemal waar is nie, soos alreeds genoem in *Afdeling 4.4.2.1*. Hierdie metode is egter prakties om benaderde vlakke van betroubaarheid te koppel aan die filter. Dit word vervolgens aan die hand van 'n voorbeeld verduidelik.

Dit is bepaal in *Afdeling 5.4* dat *Item 33.03 (b)* (“Harde uitgrawing per m<sup>3</sup>”) se gemiddelde waarde,  $\mu$ , en standaardafwyking,  $\sigma$ , onderskeidelik R68.33 en R56.48 is, nadat die “orde-grootte”-uitskieters en tydafhanklikheid verwyder is. Indien die bo- en ondergrense met een standaardafwyking vanaf die gemiddelde gekies word, *Figuur 5-6*, m.a.w. R124.81 en R11.85 onderskeidelik, is die betroubaarheidsindeks,  $\beta$ , gelyk aan een. Dit stem ooreen met 'n betroubaarheid van ongeveer 84% vanaf die gestandaardiseerde normaalverdeling. Daar is dus 'n 16% waarskynlikheid dat 'n punt in die gebied van onder of bo begrensde uitskieters nie 'n werklik uitskieter is nie of dan 84% dat dit wel ene is.



In hierdie studie is die data wel verder gefilter; nie vir die doel om verdere foute te verwyder nie, maar om projekte te identifiseer waarin daar 'n groot voorkoms van hierdie uitskieters is. Die doel hiervan was om in te fokus op hierdie projekte en dan te bepaal of daar in die brondokumente (die tenderverslae) enige moontlike verklaring gevind kon word vir die noemenswaardige afwykings. Hier word kortweg verduidelik wat gedoen is, maar 'n volledige beskrywing van die metode wat gebruik is word in *Aanhangsel K* gegee.

Aanvanklik is die “orde grootte”-foute verwyder uit al die beduidende items wat in *Afdeling 5.1* geïdentifiseer is. Dit is m.b.v. 'n visuele inspeksie van elke datastel gedoen, gevolg deur 'n bepaling van die koëffisiënt van variasie van elke datastel om te bepaal of enige van die “orde grootte”-uitskieters nie verwyder is nie. Items met 'n koëffisiënt van variasie van groter as 'n sekere drumpelwaarde is herondersoek. (Die drumpel waarde is bepaal vir elke item gebaseer op die aantal datapunte in die versameling en 'n gemiddelde koëffisiënt van variasie van 0.2 vir al die punte, behalwe een punt wat 'n orde groter is as die gemiddelde ander punte.)

Hierna is die data aangepas om voorsiening te maak vir die tydafhanklikheid wat voorkom. Die gefilterde data is verder gefilter om die items te bepaal wat buite die grense van een standaardafwyking om die gemiddelde van die datastel onder beskouing val. Projekte met vyf of meer sulke items is dan geïdentifiseer. In *Tabel 5-4* word die resultate van die voorkomste van uitskieters getoon en in *Tabel 5-5* die projekte.

Dit kan in die resultate van hierdie analise gesien word dat daar dubbel soveel van hierdie onderbegrensde items is as bo-begrensde items. Dit is in ooreenstemming met wat in *Afdeling 4.4.2.1* oor die vorm van verdelings gesê word, nl. dat eenheidtariefdata skeef na regs verdeel is, omdat die werklike

koste groter vryheid het om hoër as laer te wees. Dit is verder ook bevestig toe die data ontleed is vir die voorkoms van die tipe verdeling, *Tabel 4–2* en *Tabel 4–3* (met in ag name van hoe onderbegrensheid en bo-begrensheid in hierdie afdeling gedefinieer word teenoor hoe dit vir 'n statistiese verdeling gedefinieer word).

**TABEL 5–4: VOORKOMS VAN UITSKIETER ITEMS EN PROJEKTE**

Begrensing	Aantal Items	Aantal Projekte	Aantal Projekte wat Bo- en Onderbegrensde Uitskieters Deel
Onder	539	18	0
Bo	275	6	

Met die projekte bekend waarin daar 'n noemenswaardige aantal van die uitskieters in voorkom, is 'n ondersoek geloods om die tenderdokumente op te spoor. Die doel van hierdie ondersoek was tweeledig van aard. Eerstens moes bepaal word of uitskieters wel uitskieters is en of dit nie foute is wat gemaak is tydens die invoer van die data in die databasis nie. Tweedens is ondersoek ingestel na 'n moontlik verklaring vir hoekom die uitskieters soveel afwyk vanaf die gemiddelde van die item.



**TABEL 5–5: UITSKIETER PROJEKTE**

Bo-begrens		Onderbegrens	
Naam	Aantal Items	Naam	Aantal Items
SAPR 2220051	6	08/1997	5
SAPR 2501002	5	09/1997	6
SAPR N01167051	5	10/1997	10
SAPR N1701018/1	5	15/1995	6
TT 10701-88	5	17/1994	6
TT 10733-90	6	19/1997	5
-	-	35/1992	6
-	-	C401.1	6
-	-	C410.3	7
-	-	NC-01/1997	7
-	-	NC-05/1997	6
-	-	NT 41360R	7
-	-	RT 8659KB	16
-	-	RT(T) 10760KB	6
-	-	RT(T) 10762KB	12
-	-	SAPR 0606023	9
-	-	SAPR N0209024/1	5
-	-	ZNT 88-T	8

Ongeveer 'n derde van die oorspronklike brondokumente is geïdentifiseer in die argief van die bestuurders van die databasis. Hierdie dokumente is ondersoek om bogenoemde doel te bereik. Dit het geblyk uit hierdie steekproef van die dokumente dat feitlik al die uitskieters wel korrek in die databasis ingevoer is vanaf die brondokumente, slegs een fout kon opgespoor word. Geen verklaring kon egter gevind word vir die groot afwykings vanaf die verwagte eenheidtarief



nie, omdat daar feitlik geen ander inligting saam met eenheidtariefdata was wat die projek beskryf het nie.

Hierdie bevinding, alhoewel teleurstellend in terme van die inligting wat opgespoor kon word, is kardinaal tot hierdie studie, nl. historiese inligting, alhoewel met min detail, maar oor 'n breë front onderskep, is nodig om die databasis inkrementeel te kan ontwikkel en te verfyn soos die behoeftes van die gebruikers verander en meer gesofistikeerde ontledings gebruik begin word. Dit illustreer dus die twee hoof dimensies van databasisontwikkeling wat geoptimeer moet word, nl. die doel waarvoor die databasis aangewend gaan word of dan die voordeel wat verkry kan word en die koste verbonde aan die onderskepping.

---

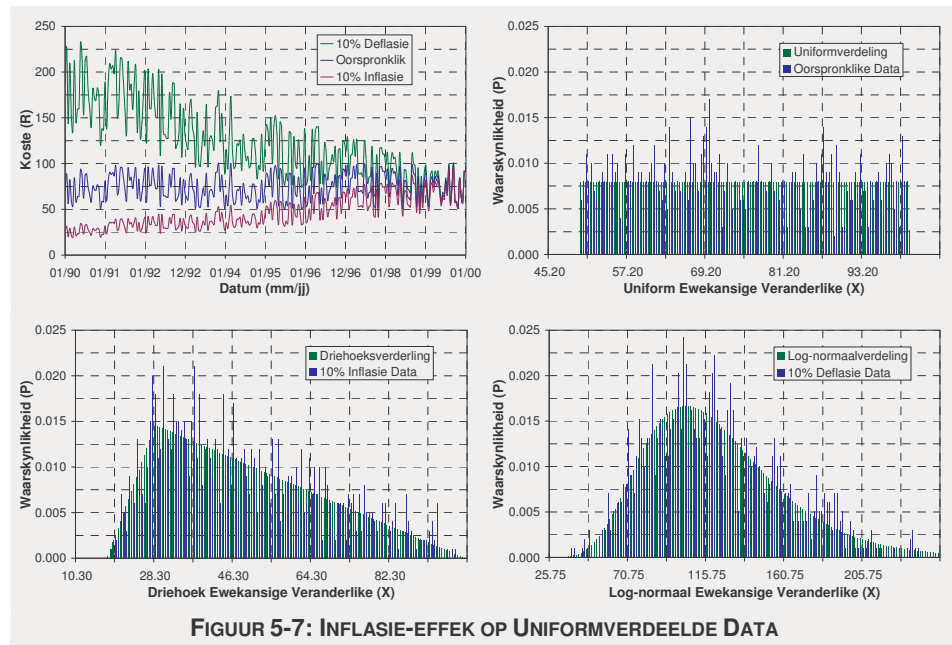
### 5.3 TYDVERANDERLIKHEID VAN DATA

---

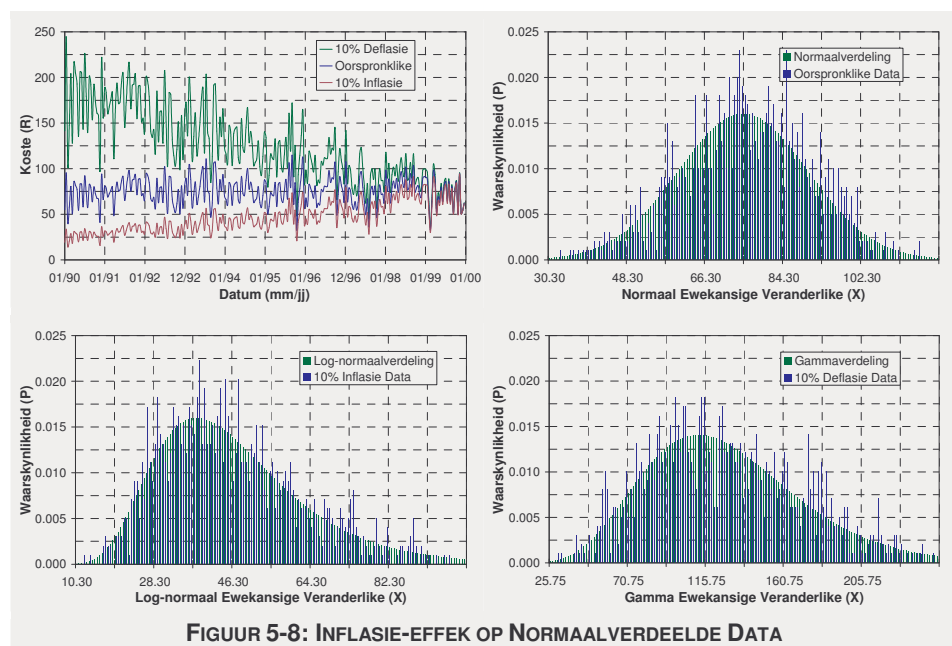
Alhoewel tydveranderlikheid van data as een van die faktore wat data mag beïnvloed beskou kan word en derhalwe eerder in *Hoofstuk 6* behoort bespreek te word, word dit egter nou reeds aangespreek. In meeste gevalle is tydsverwante inligting oor die ontstaan van data maklik en geredelik beskikbaar. Die hoofrede hoekom dit egter essensieel is om te bepaal of data tydsveranderlik is, is om te verseker dat enige basiese statistiese analise wat uitgevoer word hierdie verskynsel in ag neem, of die effek daarvan uitgeskakel word indien dit nodig is.

Seker die algemeenste voorbeelde van tydveranderlikheid van data, naas seisoenale veranderinge, is inflasie of deflasie van die eenheidtarief van items. Eenheidtariewe toon jaarliks 'n toename (of afname). Om data van vorige jare op dieselfde basis in terme van tyd te kan vergelyk is dit noodwendig nodig om die historiese data aan te pas om te kompenseer vir die effek van inflasie. Indien hierdie aanpassings nie gedoen word nie kan die resultate van benaderde statistiese beskrywings (soos die sentrale moment en die mate van verspreiding) verkeerdelik bereken word. Die effek van inflasie (of dan die ignorering daarvan) beïnvloed ook die vorm en parameters van statistiese verdelings. Hierdie effek kan maklik getoets word en word vervolgens bespreek aan die hand van die berekening wat in *Aanhangsel L* uitgevoer is.

---



Indien ewekansige getalle met 'n gegewe verdeling gegenereer word en 'n datum aan elke getal gekoppel word, kan hierdie getalle met inflasie (of deflasie) aangepas word. Die oorspronklike data se statistiese eienskappe kan dan as maatstaf gebruik word om te vergelyk hoe die aangepaste data se eienskappe verander het. Twee gevalle word hier getoon, nl. die van 'n uniform- en normaalverdeling. Albei verdelings se oorspronklike data vir 'n periode van 10-jaar is met 'n inflasie- en deflasiesyfer van 10% per jaar aangepas. In *Figuur 5-7* en *Figuur 5-8* word die oorspronklike, inflasie en deflasie aangepaste data, asook die benaderde vorm van die statistiese verdelings wat die aangepaste data aanneem vir die uniformverdeelde en normaalverdeelde data getoon. Die benaderde vorms van die verdelings is met *Crystal Ball* in *MS Excel* bepaal.



Die digtheidverdelings van die twee oorspronklike stelle data (regs bo in elke figuur) toon die verwagte vorms te hê van statistiese verdelings waarop dit gebaseer is. Indien die aangepaste data egter beskou word, blyk dit dat die vorms van die verdelings verander het. In die geval van die uniformverdeelde data het dit verander na naastenby driehoekverdeel en lognormaalverdeel vir 10% inflasie en 10% deflasie onderskeidelik. Dit is ook duidelik van *Tabel 5–6* dat die benaderde statistiese beskrywings verander het.

**TABEL 5–6: AANGEPASTE STATISTIEKE VAN UNIFORMVERDEELDE DATA**

	Oorspronklik	10% Inflasie	10% Deflasie
$\mu$	75	47	124
$\sigma$	15	17	41
$\delta$	0.2	0.36	0.33
Verdeling	Uniform	Driehoek	Log-normaal

Hierdie veranderinge is ook duidelik by die normaalverdeelde data wat verander het na 'n lognormaalverdeel en gammaverdeel vir 'n 10% inflasie en 10% deflasie onderskeidelik. Weereens het die benaderde statistiese beskrywings ook verander, maar soos verwag kan word is hierdie veranderinge feitlik dieselfde as vir die oorspronklik uniformverdeelde data, *Tabel 5–7*.

**TABEL 5–7: AANGEPASTE STATISTIEKE VAN NORMAALVERDEELDE DATA**

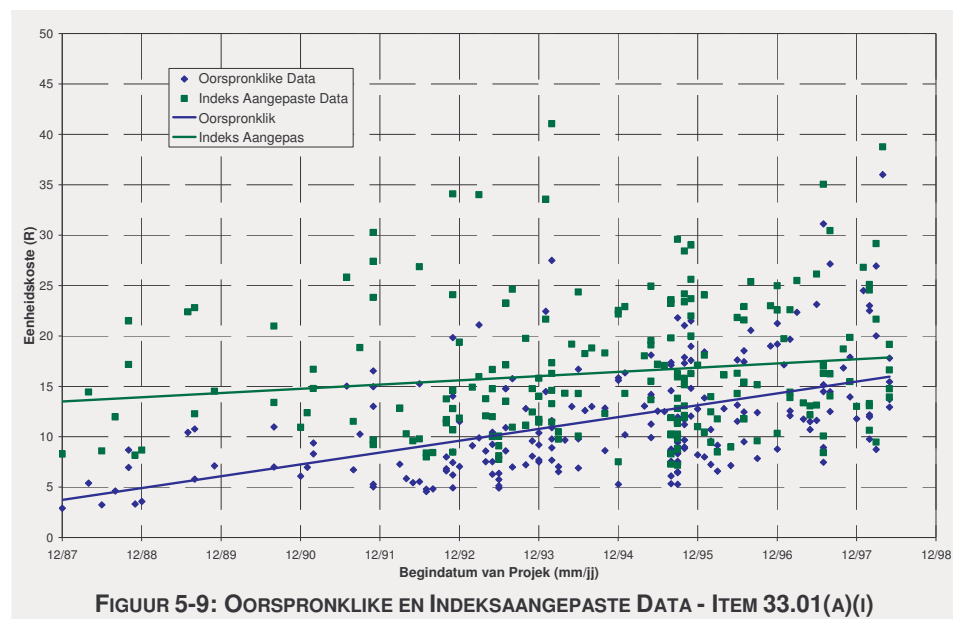
	Oorspronklik	10% Inflasie	10% Deflasie
$\mu$	75	47	125
$\sigma$	15	17	42
$\delta$	0.2	0.36	0.34
Verdeling	Normaal	Log-normaal	Gamma

*Pectus roburant cultus recti*

In die geval van die CSRA-databasis is dit kritiek dat hierdie veranderinge weens tyd tydens die bepaling van die statistieke van die data wel in ag geneem word. Indien gemiddelde kostes argumentsonthalwe bereken word op data wat nie aangepas is vir tydveranderlikheid nie, kan dit gebeur dat kostes weens inflasie onderskat of oorskat word. Dit is ook duidelik van bogenoemde voorbeelde dat die koëffisiënt van variasie in die inflasie en deflasie aangepaste data toegeneem het. Die data in die CSRA-databasis moet dus ook aangepas word om te verseker dat die variansie nie oorskat word weens die effek van inflasie of deflasie nie.

Die aanpassings kan gedoen word met inligting wat beskikbaar is oor die inflasiekoers van vorige jare. Aangesien verskillende artikels, hetsy arbeid, kos, motors, elektroniese toerusting, ens. nie teen dieselfde koers toeneem nie, is daar verskillende koste indekse wat gebruik kan word om die tendens te beskryf. Hierdie inligting word op 'n maandelikse basis deur die Sentrale Statistiekdiens beskikbaar gestel vir verskillende streke in die land.

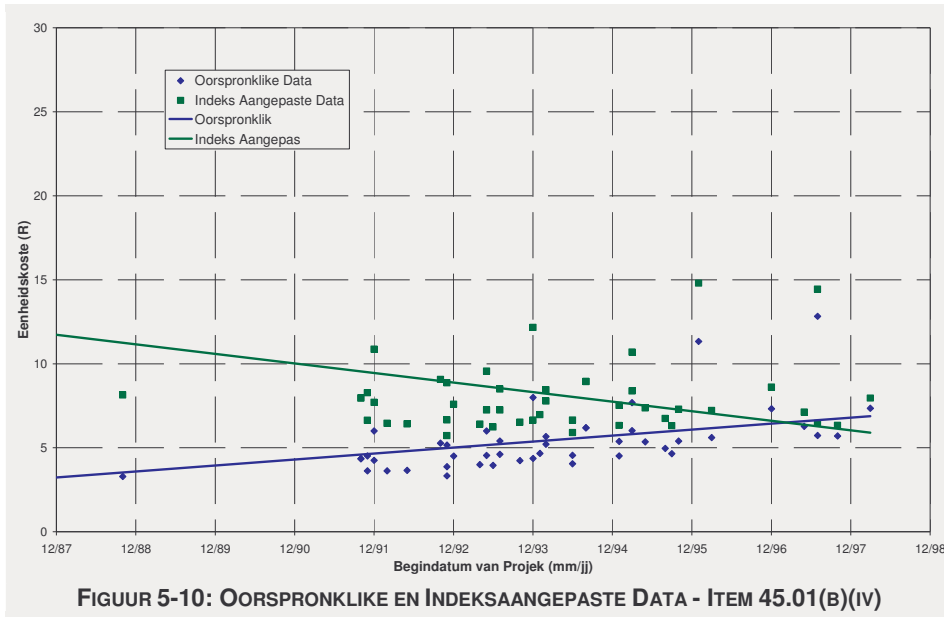
In padkonstruksieprojekte is daar van vier indekse gebruik gemaak om kostes aan te pas, nl. arbeid, toerusting, materiaal en brandstof. Vir 'n gegewe projek is hierdie indekse asook die bydrae van elkeen gewoonlik in die kontrakdokument voorgeskryf, sodat die eenheidtariewe aangepas kan word soos die projek vorder. (Die Sentrale Statistiekdiens het egter op 'n stadium opgehou om hierdie individuele indekse te publiseer en het slegs 'n veralgemeende indeks bekendgemaak wat gebruik moes word vir inflasie-aanpassings. Dit blyk egter dat daar weer terug beweeg is na die vorige stelsel.) In die CSRA-databasis is daar ook sulke aanpassings aan die oorspronklike eenheidtariewe gemaak. Hierdie aangepaste waardes is naas die oorspronklike eenheidtariewe beskikbaar vir analise doeleindes.



FIGUUR 5-9: OORSPRONKLIKE EN INDEKSAANGEPASTE DATA - ITEM 33.01(A)(I)

Indien die oorspronklike historiese data teenoor tyd geplot word, word dit verwag dat daar 'n dalende of stygende tendens in die data sal wees, of in die rare geval geen tendens nie. By die aangepaste waardes word dit egter verwag dat hierdie dalende of stygende tendens tot 'n groot mate uitgeskakel sal wees, weens die inflasie-aanpassings. Vir die beduidende items in die CSRA-databasis blyk dit egter nie so te wees nie.

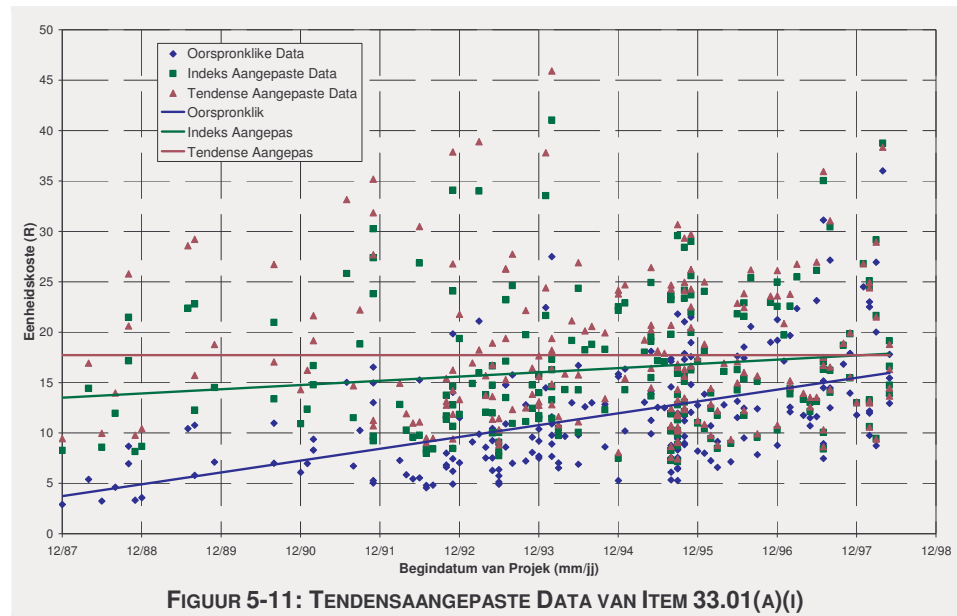
'n Regressielyn is deur die aangepaste eenheidtariewe van elke item gepas. Die berekening wat uitgevoer is word getoon in *Aanhangsel M* en die grafieke wat verkry is, kan gevind word op meegaande KS, in die gids "Aanhangsel M". Indien die aangepaste datastelle geen tendens getoon het nie, soos dit verwag word, moes hierdie lyne horisontaal gewees het. Dit blyk egter dat sommige items, bv. *Item 33.01(a)(i)* ("Verdigting van materiaal per  $m^3$ "), se eenheidtariewe nog steeds 'n sterk stygende neiging toon, *Figuur 5-9*, terwyl ander items, bv. *Item 45.01(b)(iv)* ("Dubbelseëls per  $m^2$ "), se eenheidtarief weer ooraangepas is en 'n dalende neiging toon, *Figuur 5-10*.



'n Eenvoudige tegniek is vervolgens ontwikkel om die data aan te pas totdat die stygende of dalende tendens verwyder is en om ook sodoende die tendens-gebaseerde inflasie te bepaal. Dit is aanvaar dat die inflasiekoers oor die ongeveer tien jaar periode, waarvan daar projekte is, konstant gebly het vir die individuele items. Die getenderde eenheidtariefdata van elke item is aangepas met 'n verdiskonteringsfunksie vir samegestelde rente wat gebruik word vir die berekening van huidige waarde gegewe 'n historiese waarde:

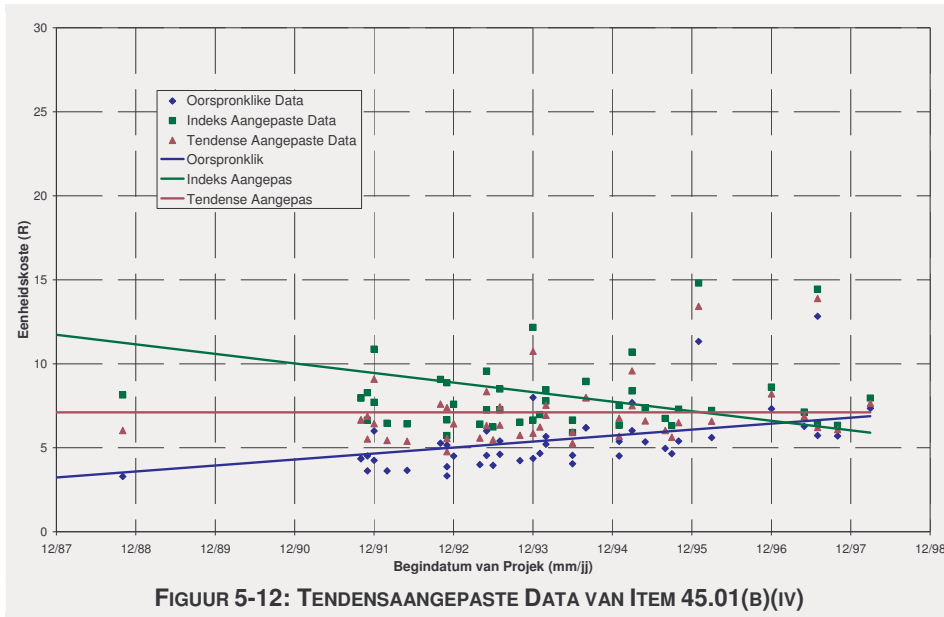
$$\begin{aligned}
 \text{Huidige waarde} &= \text{Historiese waarde} \times (1+i)^p \\
 \text{waar} & \\
 i &= \text{jaarlikse inflasiekoers;} \dots\dots\dots \text{Vgl. 5-1} \\
 p &= \text{aantal jare.}
 \end{aligned}$$

Die huidige getenderde eenheidtariewe van die historiese data van elke item kon sodoende bepaal word. 'n Regressielyn is deur hierdie aangepaste data van elke item gepas om sodoende die helling te bepaal van die tendens wat die data toon. Indien die helling nie presies nul was nie, m.a.w. waar die data wel 'n tendens getoon het, is die inflasiekoers, *i*, aangepas totdat die helling van die regressielyn nul geword het. In *Figuur 5-11* en *Figuur 5-12* word die tendens-aangepaste resultate getoon vir die twee items wat as voorbeelde gebruik is. Die berekening word breedvoerig bespreek in *Aanhangsel M* en die resultate kan ook gevind word op meegaande KS, in die gids "Aanhangsel M".

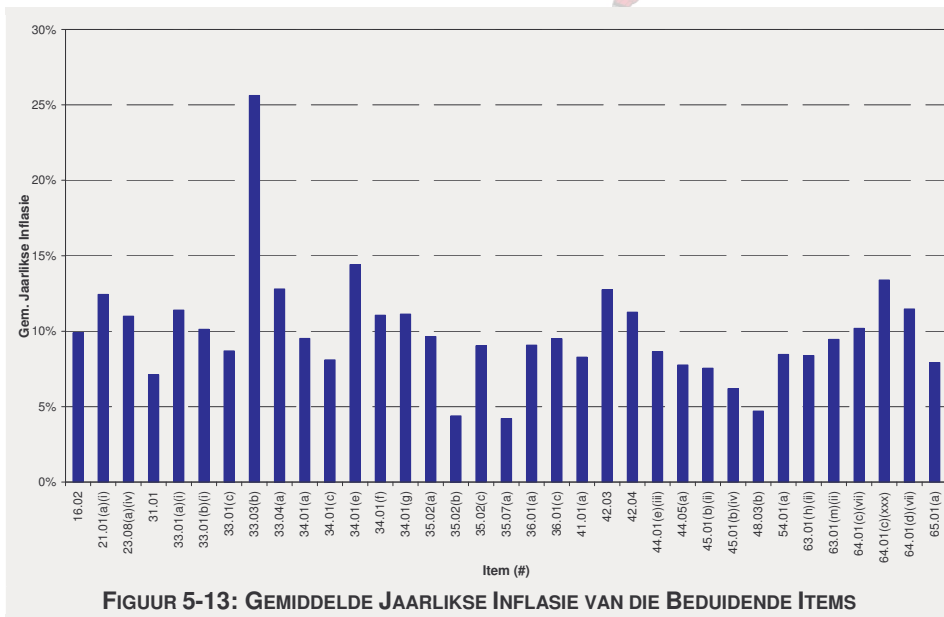


In *Figuur 5-13* word die resultate getoon van die berekende jaarlikse inflasie. Hierin kan gesien word dat die berekende inflasiesyfer tussen ongeveer 7.5% en 12.5% is, slegs enkele items se inflasie is onder 7.5% of bo 12.5%. Dit stem goed ooreen met die jaarlikse inflasie in die laat jare tagtig tot laat jare negentig. *Item 33.03(b)* ("Harde uitgraving per m<sup>3</sup>") se berekende inflasietendens blyk egter heelwat uit te wyk. Hierdie items se inflasie is bo die 25%, tien persentasie punt bo die naaste ander inflasiekoers. Dit het geblyk, nadat die item se brondata weer beskou is, dat daar geen rede is om te vermoed dat daar enige fout is met die data nie, omdat daar veral so 'n groot aantal datapunte is. Die herberekende inflasiekoers is dus as korrek aanvaar.

*Item 35.07 (a)* ("Anioniese Bitumenemulsie per l") se berekende tendens inflasie blyk weer heeltemal na die anderkant uit te wyk. Dit is ongeveer 4%. Hierdie item se data is ook weer beskou. Daar kon ook geen rede gevind word hoekom dit so noemenswaardig afwyk nie, veral nie weens die groot versameling van datapunte wat ook hier voorkom nie.



Met die “orde grootte”-foute en die tydafhanklikheid uit die data van die items verwyder kan met die volgende aspekte van die data-ontleding voortgegaan word, nl. die benaderde statistieke.



## 5.4 BENADERDE STATISTIESE BESKRYWINGS

Gegewe die vorm van 'n statistiese verdeling en ook die parameters daarvan, is dit moontlik om die probabilistiese eienskappe van die stochastiese veranderlike volledig te beskryf. Wanneer 'n persoon egter die eerste keer met 'n groot hoeveelheid data gekonfronteer word en die vorm van die statistiese verdeling nie bekend is nie, kan die benaderde beskrywings gebruik word vir die stochastiese veranderlike. Twee van die mees algemene beskrywings is die sentrale moment en die mate van verspreiding van die stochastiese veranderlike [Ang (1975)].

Die sentrale moment (verwagte of gemiddelde waarde) kan vir 'n kontinue (eenheidstariewe van items is nie diskreet nie) stochastiese veranderlike,  $X$ , met 'n waarskynlikheidsdigtheidsfunksie,  $f_X(x)$ , bepaal word deur:

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx \dots\dots\dots \text{Vgl. 5-2}$$

en vir die mate van verspreiding (variansie) met die volgende:

$$\text{Var}(X) = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx \dots\dots\dots \text{Vgl. 5-3}$$

'n Dimensioneel meer gerieflike mate van verspreiding is die vierkantswortel van die variansie, nl. die standaardafwyking,  $\sigma$ , wat as volg voorgestel kan word:

$$\sigma_X = \sqrt{\text{Var}(X)} \dots\dots\dots \text{Vgl. 5-4}$$

Direkte vergelyking tussen verskillende standaardafwykings is egter nie moontlik nie, omdat dit nie skaalvry is nie. Die skaalvrye vorm van die mate van verspreiding wat wel beskikbaar is staan bekend as die koëffisiënt van variasie en word as volg voorgestel:

$$\delta_X = \frac{\sigma_X}{\mu_X} \dots\dots\dots \text{Vgl. 5-5}$$

Laasgenoemde beskrywing lewer 'n waardevolle bydrae tot die besluit of 'n stochastiese veranderlike tydens besluitondersteuning deterministies of stochasties gemodelleer moet word. Indien dit gevind word dat die koëffisiënt van variasie klein is relatief tot van die ander veranderlikes kan die veranderlike deterministies gemodelleer word, teenoor 'n veranderlike wat 'n relatiewe groot koëffisiënt van variasie het wat stochasties gemodelleer behoort te word. Juis om hierdie rede is dit belangrik om "order grootte"-uitskieters, *Afdeling 5.2* en tydafhanklikheid, *Afdeling 5.3*, te verwyder.

In die DBBS, *MS Access*, wat gebruik is om die statistieke van die beduidende items te bereken, is die steekproefgemiddelde en variansie bereken. Dit stem presies ooreen met bogenoemde waar die geweegde gemiddeldes van  $X$  en  $(X-\mu)^2$  bepaal word. Die steekproef gemiddelde kan as volg bereken word:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \dots\dots\dots \text{Vgl. 5-6}$$

en die steekproef variansie as:



$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \dots\dots\dots \text{Vgl. 5-7}$$

Bogenoemde vorm van die variansie is die partydige vorm (“biased”). Die partydigheid kan verwyder word deur die som te deel deur (n-1). Die onpartydige variansie word in MS Access as volg bereken:

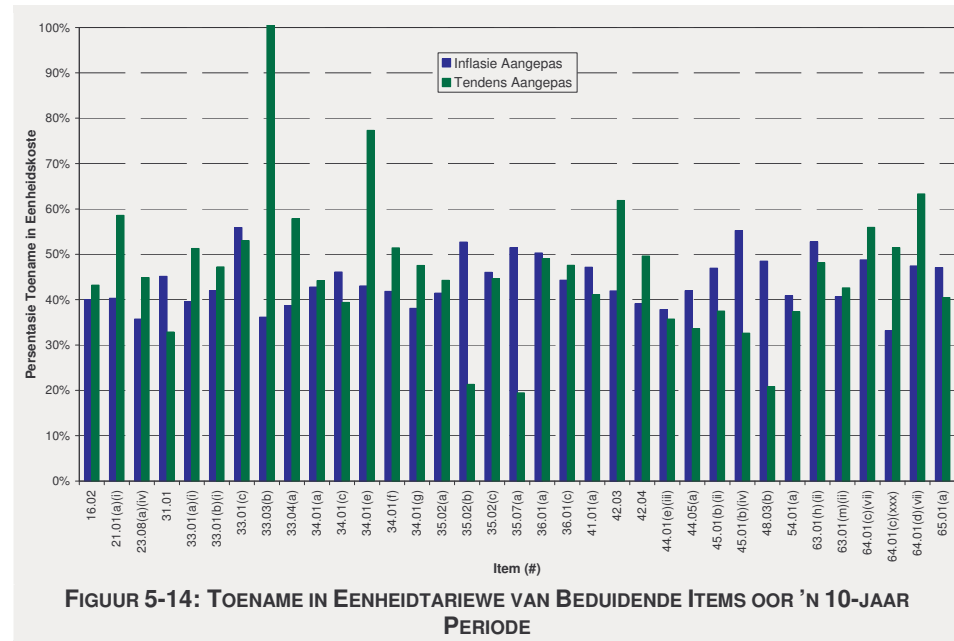
$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \dots\dots\dots \text{Vgl. 5-8}$$

Vir groot steekproewe, soos in hierdie geval waar daar soms tot 100 eenheidariewe is maak dit egter nie veel van ’n verskil nie.

**TABEL 5–8: EERSTE EN TWEDE MOMENTE VAN BEDUIDENDE ITEMS**

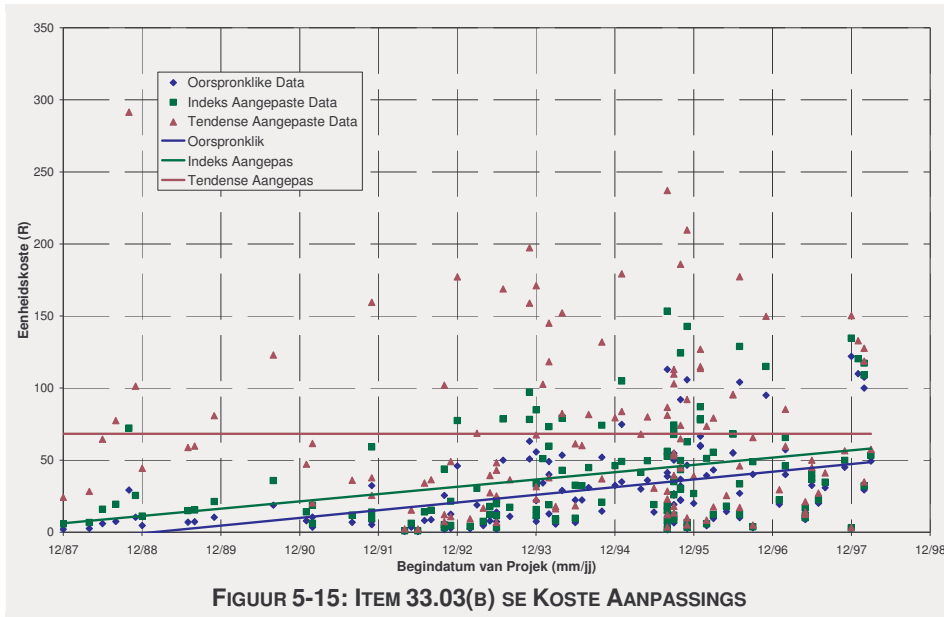
Item	Oorspronklike Data		Inflasie-aangepaste Data		Tendensaangepaste Data	
	μ	σ	μ	σ	μ	σ
16.02	R 1.36	R 0.71	R 1.90	R 0.92	R 1.94	R 0.90
21.01(a)(i)	R 15.42	R 13.00	R 21.63	R 16.93	R 24.44	R 19.14
23.08(a)(iv)	R 367.04	R 135.02	R 498.11	R 129.34	R 531.61	R 133.39
31.01	R 4.62	R 3.05	R 6.70	R 4.17	R 6.14	R 3.75
33.01(a)(i)	R 11.72	R 5.59	R 16.35	R 6.66	R 17.72	R 7.21
33.01(b)(i)	R 8.86	R 3.73	R 12.58	R 4.49	R 13.04	R 4.63
33.01(c)	R 9.80	R 4.67	R 15.28	R 6.54	R 14.99	R 6.38
33.03(b)	R 28.99	R 28.17	R 39.46	R 35.06	R 68.33	R 56.48
33.04(a)	R 8.95	R 5.64	R 12.40	R 7.10	R 14.12	R 8.01
34.01(a)	R 17.05	R 13.20	R 24.34	R 17.88	R 24.58	R 17.84
34.01(c)	R 21.31	R 16.09	R 31.13	R 22.56	R 29.69	R 21.37
34.01(e)	R 22.48	R 18.41	R 32.16	R 21.73	R 39.87	R 24.56
34.01(f)	R 19.23	R 14.62	R 27.28	R 17.47	R 29.12	R 17.76
34.01(g)	R 15.84	R 9.93	R 21.86	R 10.99	R 23.37	R 11.52
35.02(a)	R 347.09	R 102.78	R 490.83	R 110.36	R 500.66	R 113.46
35.02(b)	R 323.36	R 78.39	R 493.66	R 214.60	R 392.23	R 86.44
35.02(c)	R 388.46	R 120.10	R 567.16	R 164.85	R 561.86	R 128.74
35.07(a)	R 1.12	R 0.47	R 1.70	R 1.02	R 1.34	R 0.53
36.01(a)	R 110.67	R 48.76	R 166.30	R 64.37	R 164.95	R 59.54
36.01(c)	R 139.86	R 49.71	R 201.79	R 57.76	R 206.40	R 59.28
41.01(a)	R 1.45	R 0.54	R 2.13	R 0.81	R 2.04	R 0.68
42.03	R 203.45	R 126.51	R 288.70	R 139.21	R 329.28	R 144.01
42.04	R 1.10	R 0.81	R 1.54	R 1.00	R 1.65	R 1.05
44.01(e)(iii)	R 4.70	R 2.23	R 6.48	R 2.85	R 6.38	R 2.80
44.05(a)	R 22.08	R 9.65	R 31.34	R 14.01	R 29.49	R 12.03
45.01(b)(ii)	R 4.38	R 0.93	R 6.43	R 1.01	R 6.01	R 0.97
45.01(b)(iv)	R 5.36	R 1.82	R 8.32	R 3.17	R 7.11	R 1.96
48.03(b)	R 577.63	R 237.00	R 857.73	R 482.09	R 697.87	R 271.69
54.01(a)	R 70.46	R 19.98	R 99.25	R 21.00	R 96.77	R 20.42
63.01(h)(ii)	R 2,414.08	R 504.24	R 3,688.03	R 328.24	R 3,575.83	R 332.72
63.01(m)(ii)	R 2,738.14	R 705.98	R 3,850.14	R 585.71	R 3,904.16	R 583.51
64.01(c)(vii)	R 314.90	R 83.73	R 468.41	R 62.45	R 491.00	R 63.25
64.01(c)(xxx)	R 389.42	R 150.14	R 518.53	R 159.31	R 589.75	R 167.10
64.01(d)(vii)	R 336.74	R 82.25	R 496.36	R 69.01	R 549.83	R 67.84
65.01(a)	R 45.98	R 14.44	R 67.61	R 17.87	R 64.56	R 17.75

Die gemiddelde waarde en die standaardafwyking van die kritiese items word in *Tabel 5–8* getoon vir die oorspronklike data, inflasie-aangepaste data in die databasis en tendensaangepaste data, soos dit met *MS Access* bepaal is. Die berekening word in *Aanhangsel N* getoon. Die toename in koste van elke komponent word in *Figuur 5-14* getoon soos dit verander het met die aanpassing (inflasie) in die databasis en die heraanpassing (tendens) wat vir hierdie studie gedoen is.

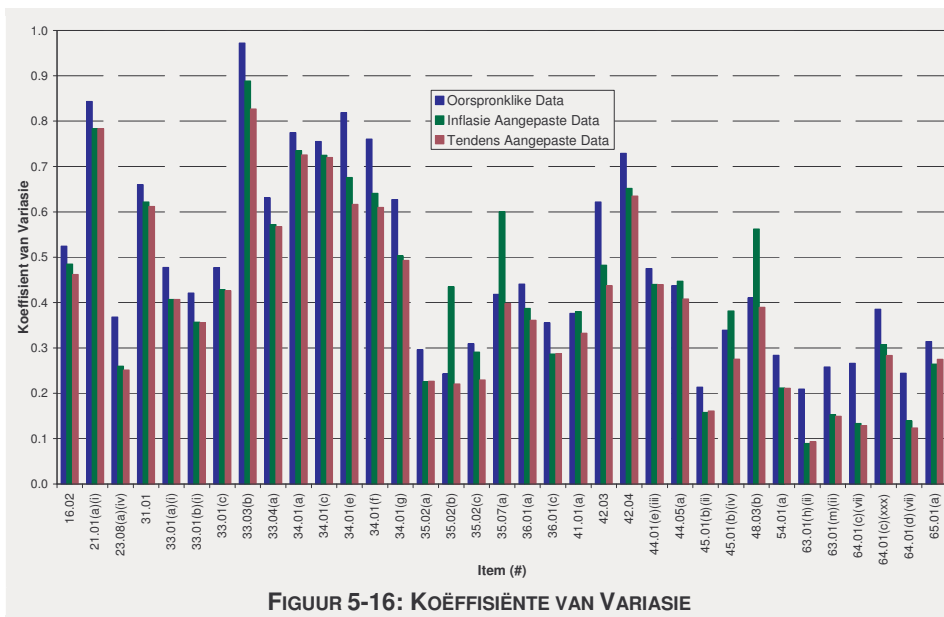


Dit blyk dat meeste items se eskalاسie onderskat is en dat feitlik al die items duurder behoort te wees as wat deur die inflasie-aangepaste waardes voorspel word, veral by *Item 33.03(b)* (“Harde uitgraving per m<sup>3</sup>”) se aangepaste waardes blyk die effek wat inflasie op die item se kostes gehad het vër onderskat te wees. In *Figuur 5-15* word ’n dataplot getoon van die oorspronklike data, inflasie-aangepaste data en tendensaangepaste data van hierdie item. Dit kan gesien word dat beide die oorspronklike data en die inflasie-aangepaste data ’n sterk stygende tendens toon. Dit is slegs na die heraanpassing, gebaseer op die tendens van die data, dat die stygende neiging verwyder is.

Enkele items soos *35.02(b)* (“Portlandhoogoond sement per t”) en *45.01(b)(iv)* (“Dubbelseëls per m<sup>2</sup>”), *Figuur 5-12*, se gemiddelde tendensaangepaste eenheidtarief is laer as die inflasie-aangepaste eenheidtarief. Hierdie afnames kan toegeskryf word aan die ooraanpassing wat plaasgevind het met die inflasie-indeks en sodoende hoër gemiddelde waardes veroorsaak het as wat dit behoort te gewees het. Die tendensaanpassing het die dalende neiging weens ooraanpassing verwyder met ’n gepaardgaande daling in die gemiddelde eenheidtarief, alhoewel die koste nog steeds van jaar-tot-jaar styg.



Dit is ook interessant om die koëffisiënt van variasie wat bereken is, nadat foute verwyder is, vir die oorspronklik data, inflasie-aangepaste data en tendensaangepaste data van elke items met mekaar te vergelyk, *Figuur 5-16*. By meeste items is daar tussen die oorspronklike en inflasie-aangepaste data 'n afname in die koëffisiënt, soos verwag kan word, maar weens ooraanpassings in sommige gevalle neem dit egter toe, soos by *Item 35.07(a)* ("Anioniese Bitumen-emulsie per l") en *45.01(b)(iv)* ("Dubbelseëls per m<sup>2</sup>"). Dit is ook opmerklik dat die tendensaangepaste data se koëffisiënt van variasie die laagste is. Veral by items soos *33.03(b)* ("Harde uitgraving per m<sup>3</sup>") en *34.01(e)* ("Gruiskroonlaag per m<sup>3</sup>") is dit opmerklik dat daar 'n noemenswaardige afname was.



Dit blyk ook uit *Figuur 5-16* dat die koëffisiënt van variasie, alhoewel daar 'n afname was, nog steeds besonder hoog is vir feitlik al die items. Slegs een item se koëffisiënt van variasie is kleiner as 0.1, terwyl meeste ander tussen 0.2 en

0.6 geleë en een selfs groter as 0.8 is. Dit kan verwag word dat hierdie groot variansies 'n noemenswaardige invloed op die onsekerheid van 'n projek se totale koste sal meebring. Dit is egter nie moontlik om spesifiek te bepaal watter een van die items die grootste risiko inhou nie, omdat die grootte van die risiko gekoppel is aan die hoeveelheid van elke item en dus projekspesifiek is. 'n Ondersoek behoort dus geloods te word om te bepaal hoe die variansie van die beduidende items verklein kan word en hoe dit per projek die onsekerheid van die kosteberaming beïnvloed. Hierdie aspekte word verder in *Hoofstuk 6* bespreek.

---

## 5.5 KORRELASIE

---

Wanneer twee veranderlikes gedeeltelik of geheel van mekaar afhang bestaan daar 'n korrelasie (of kovariansie) tussen die veranderlikes, bv. as die petrolprys styg sal die eenheidtariewe van items waar materiaal vervoer word ook styg. Die korrelasie meet die sterkte van hierdie assosiasie. Indien die assosiasie sterk is tussen twee veranderlikes word 'n groot korrelasie verwag, terwyl 'n klein positiewe korrelasie verwag word indien die assosiasie swak is.

Dit kan ook gebeur dat negatiewe korrelasie bestaan. Indien water byvoorbeeld vrylik beskikbaar is, word dit verwag dat items waar groot hoeveelhede water gebruik word se eenheidtariewe sal daal. Weereens sal 'n groot negatiewe korrelasie toon dat daar 'n sterk omgekeerde verband tussen die twee veranderlikes bestaan. Hierdie negatiewe en positiewe assosiasies is nie noodwendig weens kousale effekte nie, maar eksterne gedeelde faktore kan dit veroorsaak [Ang (1975)]. Hierdie aspek word in *Hoofstuk 6* verder bespreek.

Korrelasie tussen twee kontinue stochastiese veranderlikes,  $X$  en  $Y$ , met 'n waarskynlikheidsdigtheidsfunksie  $f_{XY}(x,y)$  kan as volg bereken word:

$$\sigma_{XY} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) f_{XY}(x, y) dx dy \dots\dots\dots \text{Vgl. 5-9}$$

In *MS Excel* word korrelasie tussen twee steekproefversamelings bereken met die volgende funksie:

$$c_{XY} = \frac{1}{n} \sum_{i=1}^n (x_i - \mu_X)(y_i - \mu_Y) \dots\dots\dots \text{Vgl. 5-10}$$

Die dimensielose vorm van korrelasie, wat waardes tussen  $-1$  en  $1$  kan aanneem, staan bekend as die korrelasiekoëffisiënt,  $\rho$ :

$$\rho_{X,Y} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{c_{XY}}{\sigma_X \sigma_Y} \dots\dots\dots \text{Vgl. 5-11}$$

Dit kan ook bepaal word met die Pearson-produkmoment korrelasiekoëffisiënt  $r$ .

---

$$r = \frac{n(\sum XY) - (\sum X)(\sum Y)}{\sqrt{[n(\sum X^2) - (\sum X)^2][n(\sum Y^2) - (\sum Y)^2]}} \dots\dots\dots \text{Vgl. 5-12}$$

Hierdie vorm van korrelasie reflekteer die sterkte van die liniêre verwantskap wat bestaan tussen die veranderlikes  $X$  en  $Y$ . Indien die verwantskap egter nieliniêr is, kan dit gebeur dat alhoewel daar 'n sterk assosiasie bestaan dit nie getoon word nie. Die Spearman korrelasiekoëffisiënt, wat gebaseer is op die rangorde van die data van die stochastiese veranderlikes, kan gebruik word om dit te oorkom [Touran (1997)]. Dit kan as volg beskryf word:

$$R = \frac{\sum_{i=1}^n (r_i - \mu_R)(s_i - \mu_S)}{\sqrt{\sum_{i=1}^n (r_i - \mu_R)^2 \sum_{i=1}^n (s_i - \mu_S)^2}}$$

waar ..... Vgl. 5-13

- $r_i$  = rang van veranderlike 1  
 $\mu_R$  = gem. van rang vir veranderlike 1  
 $s_i$  = rang van vernaderlike 2  
 $\mu_S$  = gem. van rang vir veranderlike 2

Hierdie vorm van korrelasieberekening het die voordeel dat die stochastiese veranderlikes waartussen die korrelasie bepaal word nie dieselfde verdelings hoef te hê nie. Wanneer dit vir simulatie doeleindes gebruik word, soos in *Hoofstuk 4* bespreek word, word die data wat gegenereer word vanaf die gedefinieerde verdelings ook nie gewysig nie, maar net herrangskik om die benodigde korrelasie te vorm. Die oorspronklike verdeling word dus netso behou, sonder dat daar enige transformasie uitgevoer hoef te word. [Touran (1992)]

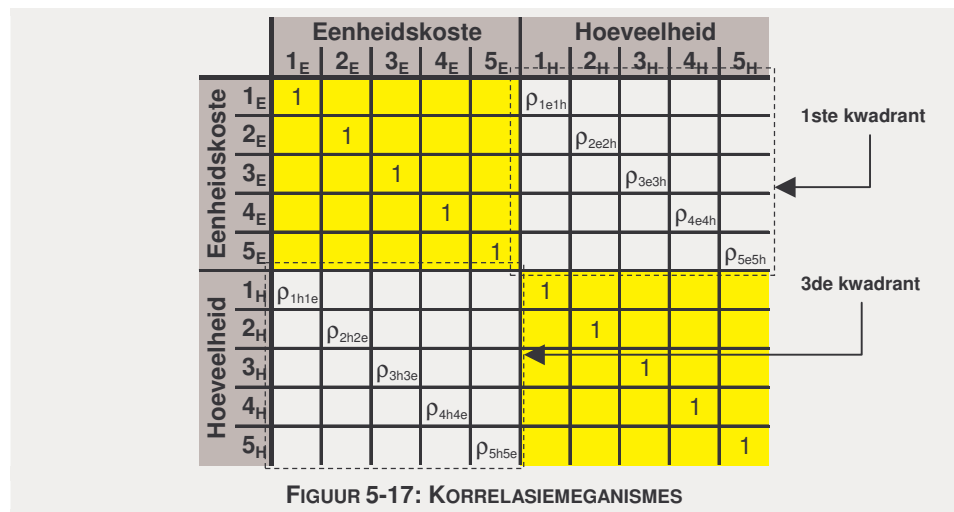
Tussen die data van eenheidtariewe en hoeveelhede kan daar drie vorms van korrelasie voorkom. Hierdie korrelasievorme kan weens die volgende korrelasie-meganismes ontstaan:

- assosiasie tussen items se eenheidtariewe: sommige items se eenheidtarief is hoër as ander items se eenheidtarief hoër is of die omgekeerde dat sommige items se eenheidtariewe sal daal as ander daal;
- assosiasie tussen items se eenheidtariewe en hoeveelhede: 'n toename in die hoeveelheid van die item lei tot 'n gepaardgaande afname in die eenheidtarief;
- assosiasie tussen die hoeveelhede van items: indien daar meer van 'n item benodig word is dit noodwendig dat daar meer van 'n ander item benodig word – die omgekeerde is egter ook waar, indien een item in die

plek van 'n ander benodig word, sal die een item toeneem terwyl die ander se hoeveelheid sal afneem.

Die korrelasies wat ontstaan weens hierdie assosiasies kan voorgestel word in 'n matriks. Gestel daar is 5-items, elkeen met 'n eenheidstarief en hoeveelheid waartussen daar bogenoemde assosiasies bestaan, dan kan die korrelasies met 'n 10x10-matriks voorgestel word soos in *Figuur 5-17*. Hierdie korrelasie matriks sal twee kenmerkende eienskappe hê. Eerstens sal die matriks 'n identiteitsdiagonaal hê, want dieselfde item se eenheidstarief is presies met homself gekorreleer en netso ook die hoeveelheid van 'n item.

Tweedens sal die matriks simmetries om die diagonaal wees, want korrelasie weens 'n assosiasie word nie deur rigting bepaal nie. Verder kan dit ook verwag word dat in die 1ste en 3de kwadrant die enigste noemenswaardige korrelasies langs die diagonale van hierdie twee kwadrante behoort voor te kom, want sterk korrelasie tussen een item se eenheidstarief en 'n ander item se hoeveelheid word nie verwag nie.



FIGUUR 5-17: KORRELASIEGANISMES

Die berekeninge wat uitgevoer is op die data van die beduidende items afkomstig van die CSRA-databasis om die Pearson-korrelasies te bepaal word in *Aanhangsel O* getoon en die volledige matriks kan op die KS onder die gids "Aanhangsel O" gevind word vir die oorspronklike data, inflasie-aangepaste data en die tendensaangepaste data. Die Spearman-korrelasies is outomaties deur *Crystal Ball* bepaal en word in die sigblaaie in die gids "Aanhangsel F" getoon.

Vervolgens word slegs uittreksels van elke kwadrant van die Pearson-korrelasiematriks in *Figuur 5-18*, *Figuur 5-19* en *Figuur 5-20* vir die oorspronklike data getoon, omdat die matriks te groot is om volledig in te pas. Die figure is onderskeidelik uittreksels van die tweede, vierde en eerste kwadrant van die matriks wat in *Figuur 5-17* getoon word. Die 17-items wat getoon word, is afkomstig uit die *Serie 3000* wat as *Serie* die grootste bydrae tot die totale koste

van projekte lewer soos vroeër reeds genoem. In die figure word slegs die korrelasies getoon wat tussen -0.1 en -1.0 is of tussen 0.1 en 1.0 voorkom. Negatiewe korrelasies word in rooi getoon.

	31_01	33_01(a)(i)	33_01(b)(i)	33_01(c)	33_03(b)	33_04(a)	34_01(a)	34_01(c)	34_01(e)	34_01(f)	34_01(g)	35_02(a)	35_02(b)	35_02(c)	35_07(a)	36_01(a)	36_01(c)
31_01	1.0	0.4	0.2	0.3	0.4	0.4	0.3	0.2	0.1	0.4	0.3	0.4	0.2	0.4	0.3	0.2	
33_01(a)(i)	0.4	1.0	0.8	0.8	0.6	0.7	0.8	0.5	0.8	0.7	0.5	0.5	0.3	0.5	0.5	0.6	0.4
33_01(b)(i)	0.2	0.8	1.0	0.8	0.7	0.6	0.6	0.5		0.6	0.6	0.5	0.1	0.4		0.5	
33_01(c)	0.3	0.8	0.8	1.0	0.5	0.7	0.6	0.6	0.4	0.4	0.5	0.2	0.3	0.4		0.3	
33_03(b)	0.4	0.6	0.7	0.5	1.0	0.5	0.5	0.6	0.3	0.5	0.5	0.4				0.4	
33_04(a)	0.4	0.7	0.6	0.7	0.5	1.0	0.6	0.5	0.4	0.6	0.5	0.4	0.4	0.5		0.5	0.4
34_01(a)	0.3	0.8	0.6	0.6	0.5	0.6	1.0	0.9	0.7	0.9	0.9	0.3	0.5	0.4	0.1	0.5	0.4
34_01(c)	0.2	0.5	0.5	0.6	0.6	0.5	0.9	1.0	0.6	0.7	0.8	0.3	0.3	0.5		0.5	0.2
34_01(e)	0.1	0.8		0.4	0.3	0.4	0.7	0.6	1.0	1.0	0.9	0.4	0.3	0.4	0.2	0.4	0.3
34_01(f)	0.4	0.7	0.6	0.4	0.5	0.6	0.9	0.7	1.0	1.0	0.9	0.4	0.3	0.5	0.2	0.5	0.3
34_01(g)	0.3	0.5	0.6	0.5	0.5	0.5	0.9	0.8	0.9	0.9	1.0	0.6	0.3	0.6	0.5	0.6	0.5
35_02(a)	0.4	0.5	0.5	0.2	0.4	0.4	0.3	0.3	0.4	0.4	0.6	1.0	0.7	0.8	0.5	0.7	0.7
35_02(b)	0.2	0.3	0.1	0.3		0.4	0.5	0.3	0.3	0.3	0.7	1.0	0.8	0.6	0.6	0.5	
35_02(c)	0.4	0.5	0.4	0.4	0.4	0.5	0.4	0.5	0.4	0.5	0.6	0.8	0.8	1.0	0.6	0.7	0.6
35_07(a)							0.1		0.2	0.2	0.5	0.6	0.6	1.0	0.1	0.3	
36_01(a)	0.3	0.6	0.5	0.3	0.4	0.5	0.5	0.5	0.4	0.5	0.6	0.7	0.5	0.7	0.1	1.0	0.9
36_01(c)	0.2	0.4				0.4	0.4	0.2	0.3	0.3	0.5	0.7	0.6	0.3	0.9	1.0	

FIGUUR 5-18: KORRELASIAMATRIKS VAN EENHEIDTARIEF

Uit bostaande matrics blyk dit dat daar sterk korrelasie bestaan tussen die eenheidtariewe van items in die *Serie 3000*. Hier kom geen negatiewe korrelasie voor nie. Dit kan dus verwag word dat hierdie sterk positiewe korrelasie die totale kostes van 'n projek sal verhoog, so getoon is in *Afdeling 4.4.2.3*. Weens hierdie sterk positiewe korrelasie kan ook aangeneem word dat dieselfde eksterne faktore hierdie items se kostes beïnvloed, omdat daar nie 'n kousale verband tussen hierdie eenheidtariewe bestaan nie. Hierdie aspek van gedeelde invloedfaktore wat die bron van die korrelasies is, word verder ontwikkel in *Hoofstuk 6*. Die simmetriese vorm van die matrics, soos reeds vroeër in hierdie afdeling genoem, is duidelik sigbaar.

	31_01	33_01(a)(i)	33_01(b)(i)	33_01(c)	33_03(b)	33_04(a)	34_01(a)	34_01(c)	34_01(e)	34_01(f)	34_01(g)	35_02(a)	35_02(b)	35_02(c)	35_07(a)	36_01(a)	36_01(c)
31_01	1.0	0.2	0.4	0.7	0.4	0.4	0.4	0.5	0.6	0.4		0.4	0.7	0.3	0.2	0.5	0.4
33_01(a)(i)	0.2	1.0	0.5	0.3	0.4	0.6	0.6	0.3	0.1			0.2	0.3	0.4	0.2	0.1	0.4
33_01(b)(i)	0.4	0.5	1.0	0.5	0.7	0.3	0.3	0.1		0.6	0.1					0.4	
33_01(c)	0.7	0.3	0.5	1.0	0.9	0.5	0.3	0.2	-0.1	0.3	0.1	0.3	0.2				0.3
33_03(b)	0.4	0.4	0.7	0.9	1.0	0.6	0.3	0.4	-0.1	0.2		0.4	0.2	0.2		0.4	0.5
33_04(a)	0.4	0.6	0.3	0.5	0.6	1.0	0.3	0.2	0.2	0.1		0.1	0.3	0.3		0.2	0.2
34_01(a)	0.4	0.6	0.3	0.3	0.3	0.3	1.0	0.6	0.3	0.4		0.5	0.4	0.6		0.5	0.5
34_01(c)	0.5	0.3	0.1	0.2	0.4	0.2	0.6	1.0	0.5	0.2		0.6	0.2	0.6		0.8	0.4
34_01(e)	0.6	0.1		-0.1	-0.1	0.2	0.3	0.5	1.0	0.1	0.4	0.4	0.6	0.4		0.3	-0.4
34_01(f)	0.4		0.6	0.3	0.2	0.1	0.4	0.2	0.1	1.0		0.3	0.4	0.2	0.3	0.4	0.6
34_01(g)			0.1	0.1					0.4		1.0		0.2	0.4	-0.2	0.2	0.3
35_02(a)	0.4	0.2		0.3	0.4	0.1	0.5	0.6	0.4	0.3		1.0				0.4	0.6
35_02(b)	0.7	0.3	0.4	0.2	0.2	0.3	0.4	0.2	0.6	0.4	0.2		1.0	0.1	0.5	0.4	
35_02(c)	0.3	0.4			0.2	0.3	0.6	0.6	0.4	0.2	0.4	0.4	0.1	1.0		0.5	-0.1
35_07(a)	0.2	0.2							0.3	-0.2			0.5		1.0		
36_01(a)	0.5	0.1	0.2	0.3	0.4	0.2	0.5	0.8	0.3	0.4	0.2	0.6	0.4	0.5		1.0	
36_01(c)	0.4	0.4			0.5	0.2	0.5	0.4	-0.4	0.6	0.3	0.6	0.4	-0.1	0.1	0.9	1.0

FIGUUR 5-19: KORRELASIAMATRIKS VAN HOEWELHEID

Uit bostaande matrics blyk dit dat daar ook sterk positiewe korrelasie bestaan tussen die hoeveelhede van die items. Dit kan dus verwag word dat indien afwykings tussen die berekende en die werklike hoeveelhede voorkom dan sal 'n aantal items se hoeveelhede gesamentlik styg of daal. Enkele negatiewe korrelasies kom ook voor, wat bevestig dat sommige items se hoeveelhede afneem indien 'n ander se hoeveelheid styg. Dit is ook weereens

duidelik dat die matrigs simmetries is. Die eenvoudige meganisme van die bron van hierdie vorm van korrelasie, nl. projekomvang, maak hierdie korrelasiemeganisme triviaal en word dus nie verder ondersoek nie.

	31_01	33_01(a)(i)	33_01(b)(i)	33_01(c)	33_03(b)	33_04(a)	34_01(a)	34_01(c)	34_01(e)	34_01(f)	34_01(g)	35_02(a)	35_02(b)	35_02(c)	35_07(a)	36_01(a)	36_01(c)
31_01	-0.3	-0.1	0.4	-0.2	-0.1	-0.1	-0.2	-0.2	-0.2	-0.1	-0.1	-0.2	-0.3	-0.1	-0.2	-0.1	
33_01(a)(i)	-0.3	-0.4	-0.4	-0.3	-0.3	-0.3	-0.3	-0.3	-0.2			-0.3	-0.1	-0.3		-0.3	-0.2
33_01(b)(i)	-0.3	-0.4	-0.4	-0.3	-0.4	-0.2	-0.1	-0.2		-0.3	-0.4		-0.1	-0.2			-0.3
33_01(c)	-0.2	-0.3	-0.3	-0.3	-0.3	-0.2	-0.2						-0.1				
33_03(b)	-0.2	-0.3	-0.3	-0.4	-0.3	-0.2	-0.3	0.2	-0.3			-0.3		-0.2	-0.2	-0.2	-0.3
33_04(a)	-0.3	-0.3	-0.2	-0.2	-0.2	-0.2	-0.3	-0.2	-0.3	-0.2	-0.2	-0.2	-0.2	-0.3		-0.3	-0.2
34_01(a)	-0.2	-0.2	-0.2	-0.2	-0.2	-0.1	-0.3	-0.2	-0.3	-0.2	-0.2	-0.1	-0.1	-0.2		-0.3	-0.3
34_01(c)	-0.2	-0.2	-0.2	-0.1	-0.2	-0.2	-0.3	-0.3	-0.3	-0.1	-0.2			-0.1	-0.3		-0.2
34_01(e)		-0.2		0.2			-0.2	-0.2	-0.3		-0.2	-0.2	-0.1	-0.2		-0.2	0.1
34_01(f)	-0.2	-0.2	-0.1	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.1	-0.2	-0.1	-0.1	-0.3			-0.2
34_01(g)	-0.1	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.3	-0.2	-0.2	-0.2	-0.2	-0.1	-0.3		-0.2	-0.1
35_02(a)		-0.2		-0.2	-0.3	-0.2		-0.1				-0.3	0.2	-0.2		-0.1	-0.1
35_02(b)			0.2	0.4	0.3	-0.1	0.1	0.2	-0.1	0.3	-0.2	0.2	0.1				0.3
35_02(c)		-0.2	0.2		-0.1	-0.2			0.2	0.1	-0.2	-0.1	0.2	-0.3			0.3
35_07(a)			0.0	0.0	0.2	0.2					-0.3				-0.2		0.2
36_01(a)	-0.3	-0.2				-0.2	-0.2	-0.1	-0.2		-0.2	-0.2		-0.1		-0.3	-0.2
36_01(c)	-0.2	-0.1			-0.3	-0.3	-0.3	-0.3	-0.7	0.1	-0.2	-0.4		-0.4	0.1	-0.4	-0.1

FIGUUR 5-20: KORRELASIEMATRIKS VAN HOEVEELHEID EN EENHEIDTARIEF

Uit bostaande matrigs blyk dit dat die korrelasies wat voorkom oorwegend negatief en relatief swak is. By hierdie matrigs behoort egter slegs die diagonaal beskou te word soos reeds bespreek. Hier blyk dit dan dat daar orals ’n geringe omgekeerde verband bestaan tussen die eenheidtarief en ooreenstemmende volume materiaal vir ’n item. Dit kan toegeskryf word aan “skaal voordeel”. Die matrigs is onsimmetries soos verwag is.

Om meer insig in die korrelasies te kry is dit in vier groepe, nl. sterk, medium, swak en geen, gegroepeer vir elk van die drie korrelasiemeganismes wat voorkom. Die grense van elk groep is as volg gedefinieer:

- Sterk* :  $0.7 < \rho \leq 1.0$  of  $-0.7 > \rho \geq -1.0$
- Medium* :  $0.4 < \rho \leq 0.7$  of  $-0.4 > \rho \geq -0.7$
- Swak* :  $0.1 < \rho \leq 0.4$  of  $-0.1 > \rho \geq -0.4$
- Geen* :  $-0.1 \leq \rho \leq 0.1$

Indien die voorkomste van pare datapunte waartussen die korrelasies moes bereken word minder as 10 was, is dit as “Onbepaalbaar” geklassifiseer om te voorkom dat toevallige korrelasie nie in ag geneem word by klein versamelingruimtes nie. In *Tabel 5–9* en *Tabel 5–10* word die resultate van onderskeidelik die oorspronklike data en tendensaangepaste data getoon vir elk van die korrelasiemeganisme en verdeel tussen positiewe of negatiewe korrelasies.



**TABEL 5–9: VOORKOMS VAN KORRELASIES TUSSEN BEDUIDENDE ITEMS (OORSPRONKLIK)**

Korrelasie	Tussen Eenheidtariewe <sup>#</sup>		Tussen Hoeveelhede <sup>#</sup>		Tussen Eenh. en Hoeveelheid	
	Negatief	Positief	Negatief	Positief	Negatief	Positief
Sterk	0 (0.0%)	171* (13.0%)	0 (0.0%)	69* (3.2%)	0 (0.0%)	4 (0.4%)
Medium	2 (0.2%)	450 (42.9%)	4 (0.4%)	214 (20.4%)	45 (4.2%)	25 (2.3%)
Swak	22 (2.1%)	386 (36.8%)	150 (14.3%)	402 (38.4%)	551 (50.9%)	149 (13.8%)
Geen	52 (5.0%)		244 (23.3%)		309 (28.5%)	
Onbepaalbaar	142 (11.8%) <sup>§</sup>		142 (11.8%) <sup>§</sup>		142 (11.6%) <sup>§</sup>	

\*Die identiteitsdiagonale van die tweede en vierde kwadrante moet nie ingereken word by die sterk positiewe korrelasie nie.

<sup>#</sup>Weens simmetrie moet al die waardes van die tweede en vierde kwadrante deur twee gedeel word, nadat die identiteitsdiagonaal afgetrek is.

<sup>§</sup>Persentasie van totale maontlike voorkomste.

Indien die twee tabelle se resultate met mekaar vergelyk word, is dit duidelik dat die grootste verskil voorkom by die korrelasies tussen eenheidtariewe. 'n Noemenswaardige verlaging van korrelasiesterkte vind plaas indien die temporale tendense verwyder word. Dit dui daarop dat 'n groot aantal van die items gedeelde eksterne faktore het wat deur temporale veranderinge beïnvloed word. 'n Substansiële aantal van die items toon egter nog steeds 'n noemenswaardige sterkte van verband tussen mekaar, nadat temporale tendense verwyder is. Die belangrike resultaat bevestig dat daar eksterne faktore moet bestaan wat die eenheidtariewe van hierdie items in dieselfde projek laat daal of styg relatief tot die verwagte eenheidtariewe van die onderskeie items, *Hoofstuk 6*. Indien die faktore bepaal kan word wat hierdie korrelasies veroorsaak, asook die meganisme wat dit voortbring, kan dit gebruik word om die veranderlikheid te verlaag.

**TABEL 5–10: VOORKOMS VAN KORRELASIES TUSSEN BEDUIDENDE ITEMS (TENDENSA-ANGEPAS)**

Korrelasie	Tussen Eenheidtariewe <sup>#</sup>		Tussen Hoeveelhede <sup>#</sup>		Tussen Eenh. en Hoeveelheid	
	Negatief	Positief	Negatief	Positief	Negatief	Positief
Sterk	0 (0.0%)	79* (4.2%)	0 (0.0%)	69* (3.2%)	0 (0.0%)	5 (0.5%)
Medium	6 (0.6%)	194 (18.5%)	4 (0.4%)	214 (20.4%)	46 (4.3%)	16 (1.5%)
Swak	96 (9.2%)	466 (44.5%)	150 (14.3%)	402 (38.4%)	561 (51.8%)	135 (12.4%)
Geen	242 (23.0%)		244 (23.3%)		320 (29.5%)	
Onbepaalbaar	142 (11.8%) <sup>§</sup>		142 (11.8%) <sup>§</sup>		142 (11.6%) <sup>§</sup>	

\*Die identiteitsdiagonale van die tweede en vierde kwadrante moet nie ingereken word by die sterk positiewe korrelasie nie.

<sup>#</sup>Weens simmetrie moet al die waardes van die tweede en vierde kwadrante deur twee gedeel word, nadat die identiteitsdiagonaal afgetrek is.

<sup>§</sup>Persentasie van totale maontlike voorkomste.

Die voorkoms van korrelasies tussen hoeveelhede is presies dieselfde tussen die twee tabelle, soos verwag kan word. Alhoewel dit ongeveer dieselfde sterkte van verband toon as die van die eenheidtariewe nadat temporale veranderlikheid verwyder is, is dit egter 'n aspek wat nie verder in hierdie studie aangespreek is nie.

Dit blyk uit bogenoemde twee tabelle dat die eenheidtarief-hoeveelheid korrelasies oorwegend onsensitief is vir temporale tendense. Die korrelasie tussen

eenheidtarief en hoeveelheid van 'n item is oorwegend swak of geen. Dit is tog wel interessant dat daar so 'n groot hoeveelheid negatief gekorreleerde voor-komste is teenoor slegs 'n klein aantal positiewe korrelasies. Die verklaring hiervoor kan gevind word in die sterk positiewe korrelasie wat voorkom tussen hoeveelhede.

Indien 'n item se hoeveelheid styg in 'n projek sal daar waarskynlik ook 'n styging in die ander items se hoeveelhede wees, derhalwe die positiewe korrelasie wat gevind word tussen hoeveelhede, *Figuur 5-19*. Weens die negatiewe korrelasie wat bestaan tussen 'n item se hoeveelheid en eenheidtarief, alhoewel swak, sal die item se koste daal soos die hoeveelheid styg, maar omdat die ander items se hoeveelhede ook sal toeneem sal hul eenheidtarief ook daal. Dus indien 'n item se hoeveelheid toeneem is daar 'n moontlikheid dat sy eenheidtarief asook die ander items se eenheidtarief sal daal; van daar die groot hoeveelheid swak negatiewe korrelasie. Die omgekeerde is natuurlik ook waar.

---

## 5.6 EIENSKAPPE VAN DATA

---

Die drie kenmerkendste eienskappe van die data wat in die studie geïdentifiseer is was die noemenswaardige bydrae van 'n klein aantal items, die groot variasie van die data en die sterk korrelasies tussen eenheidtariewe van die beduidende items.

---

### 5.6.1 NOEMENSWAARDIGE BYDRAE

---

Om detail inligting van alle items in 'n databasis te onderskep is nie moontlik nie, omdat dit tydsgewys en finansiële nie lonend is nie. Die meeste versameling van data toon egter die eienskap van 'n klein aantal met 'n noemenswaardige impak of die sogenaamde 80-20-reël, waar 80% van dit wat ondersoek word deur slegs 20% bepaal word.

Hierdie eienskap is ook geïdentifiseer by die items in die CSRA-databasis, waar 'n klein aantal items 'n noemenswaardige bydrae gelewer het tot die totale koste van projekte. Dit het uit 'n aanvanklike analise geblyk dat die verband tussen die persentasie van items in 'n projek en die bydrae van hierdie items tot die totale koste selfs effens sterker as die 80-20-reël is. Dit is gevind dat gemiddeld 10% van die aantal items in 'n projek 80% van die totale kostes bydrae en dat gemiddeld 20% van die items 90% van die totale projek-kostes bydrae. Verdere ondersoeke het getoon dat hierdie noemenswaardige bydrae in elke projek afkomstig is van feitlik dieselfde items.

Die voordeel verbonde aan hierdie eienskap wat die eenheidtariefdata toon is dat daar 'n meer gefokusde benadering tot die ontleding van die data gevolg kan word. Verdere ontledings, bevindings en aanbevelings wat op 'n klein steekproef van die totale aantal items in die databasis gedoen word lewer dus

---

'n verteenwoordigende en betekenisvolle bydrae tot die eienskappe van 'n tipiese projek se totale kostes.

---

## 5.6.2 VERANDERLIKHEID

---

Dit is reeds genoem dat veranderlikheid die verspreiding van data om die gemiddelde van 'n datastel is. Die voorafgaande berekeninge het getoon dat hierdie verspreiding besonder groot kan wees vir van die beduidende items. Die drie bronne van veranderlikheid wat gevind is, is: tydveranderlikheid, foutiewe data en epistemiese veranderlikheid (alhoewel eersgenoemde twee ook hieronder geklassifiseer kan word, word dit afsonderlik bespreek).

---

### 5.6.2.1 TYDVERANDERLIKHEID

---

Tydafhanklikheid ontstaan wanneer daar 'n verband tussen die tendense wat data toon en tyd is. Eenheidtariewe van konstruksieprojekte is sensitief vir inflasie weens die impak van die ekonomiese omgewing waarbinne projekte uitgevoer word. Dit veroorsaak dat daar 'n stygende of dalende tendens in die eenheidtariefdata van items is.

Indien hierdie temporale tendense nie uit die data verwyder word nie, kan die basiese statistieke asook die vorms van die verdelings wat dit beskryf beïnvloed word. Dit lei tot 'n oor- of onderskatting van gemiddeldes en veroorsaak altyd 'n oorskatting van die veranderlikheid van die data. Hierdie temporale tendense moet dus verwyder word.

Om die tydafhanklikheid uit die data te verwyder kan daar van inflasie-indekse gebruik gemaak word om die data aan te pas. Hierdie indekse is slegs veralgemeende weerspieëlings van die inflasie. Spesifieke items se kostes neem egter nie noodwendig met 'n gegewe indeks of samestelling van 'n aantal hiervan toe nie. Dit kan dus gebeur dat indeksaangepaste data nog steeds tydafhanklikheid toon.

Dit is in hierdie studie gevind dat meeste van die eenheidtariewe na inflasie-indeksaanpassing nog steeds tydafhanklike tendense getoon het, weens onder- of ooraanpassings. Veranderlikheid weens tydafhanklikheid het dus nog steeds voorgekom. Slegs nadat die data aangepas is om enige temporale tendense te verwyder is die veranderlikheid ook geminimeer en is meer realistiese gemiddeldes bepaal.

---

### 5.6.2.2 FOUTIEWE DATA

---

Foutiewe data is data wat verskil van die data in die oorspronklike brondokument. Hierdie verskille kan veroorsaak word deur: verkeerde data wat korrek in die databasis gevoer word, korrekte data wat verkeerdelik in die databasis gestoor word en opdaters in die databasis wat foute kan veroorsaak.

---

Die oorsake van eersgenoemde twee kan toegeskryf word aan swak data-vloei. Dit is reeds genoem dat die heronderskepping van data in die CSRA-databasis tenminste vyf keer plaasgevind het voordat dit in die databasis gestoor is. Foute kon maklik gedurende hierdie proses voorkom, soos gesien kan word in die korreksies wat in die tenderdokumente van die kontrakteurs moes aangebring word. Data kon ook nie direk digitaal in die databasis ingevoer word nie, maar moes handmatig ingevoer word.

'n Aantal van hierdie “vinger”-foute kon geïdentifiseer word, omdat dit 'n orde grootte en selfs meer verskil het van die res van data vir 'n gegewe item. Die impak wat sulke foute op die statistiese beskrywers lewer, is noemenswaardig. Verwydering van hierdie ooglopende foute het 'n merkbare verlaging in die veranderlikheid van die data te weeg gebring.

---

### 5.6.2.3 EPISTEMIESE VERANDERLIKHEID

---

Onsekerheid kan in twee kategorie verdeel word, nl. aleatories en epistemies. Aleatoriese onsekerheid ontstaan weens die veronderstelling dat sekere natuurlike gebeurtenisse 'n inherente ewekansigheid het teenoor epistemies onsekerheid wat ontstaan weens onvolledige inligting (bv. modelonsekerheid, data-onsekerheid, parameteronsekerheid).

Nadat tydveranderlikheid en “orde grootte”-foute verwyder is, toon die data nog steeds 'n groot veranderlikheid. Hierdie veranderlikheid is oorwegend epistemies van aard, omdat dit ontstaan weens foute in die data wat nie geïdentifiseer kon word nie en/of 'n tekort aan eksterne faktore wat onder andere die ruimtelike veranderlikheid van die data beskryf. In die gevalle waar die veranderlikheid van die data voorheen by kosteberekeninge in ag geneem is, soos by die berekening in *Hoofstuk 4*, is hierdie epistemiese onsekerheid as aleatoriese onsekerheid benader, omdat daar nie 'n metode was om die epistemiese veranderlikheid te beskryf nie.

In hierdie studie word gepoog om die epistemiese veranderlikheid te verlaag deur: die datavloei te verbeter (verlaging van data-onsekerheid) en die data-onderskepping te verbeter sodat die ruimtelike veranderlikheid verminder kan word (verlaging van parameteronsekerheid). Hierdie aspek word verder in *Hoofstuk 6* bespreek.

---

### 5.6.3 KORRELASIE

---

Korrelasie is alreeds gedefinieer as 'n aanduiding van die sterkte van die assosiasie tussen twee ewekansige veranderlikes. In die geval van kosteberekeninge kan daar drie korrelasiemeganismes ontstaan, nl. tussen eenheidtariewe en eenheidtariewe, tussen hoeveelhede en hoeveelhede en tussen hoeveelhede en eenheidtariewe. Dit is gevind dat die korrelasies wat

---

tussen die eenheidtariewe van items bestaan sterk positief is en dat tussen hoeveelhede dit minder sterk positief is. Swak negatiewe korrelasie is tussen hoeveelhede en eenheidtariewe gevind.

Die sterk korrelasie tussen die eenheidtariewe dui daarop dat daar eksterne faktore moet wees wat die kostes beïnvloed, omdat daar geen kousale effek tussen die eenheidtariewe bestaan nie. Hierdie faktore is die oorsaak vir die groot veranderlikheid in van die eenheidtariewe se data. Die korrelasie tussen hoeveelhede is meganisties van aard, omdat 'n toename in die een volume gepaard sal gaan met 'n styging in die van 'n ander. Hierdie aspek word dus nie verder in die studie ondersoek nie, maar dit raak wel belangrik wanneer afwykings van die beraamde kostes voorkom. Die klein negatiewe korrelasie tussen hoeveelhede en eenheidtariewe bevestig dat daar nie noemenswaardige skaal effekte betrokke is by die bepaling van eenheidtariewe is nie.

---

## 5.7 VOORBEELD

---

Die tegnieke wat in *Hoofstuk 4* bespreek word, besit die eienskappe dat dit onsekerhede in data en ook korrelasies wat tussen veranderlikes voorkom in berekening kan bring. Die data wat in hierdie hoofstuk ondersoek is vertoon groot veranderlikheid en ook sterk korrelasies tussen veranderlikes. Aangesien die items wat ontleed is ook 'n noemenswaardige bydrae lewer tot die totale koste van die gemiddelde projek in die databasis is hierdie data as verteenwoordigend beskou van wat in 'n tipiese projek voorkom.

Die voorbeeld is derhalwe uit die items wat in hierdie hoofstuk ontleed is, saamgestel om die statistiese eienskappe van 'n tipiese projek te reflekteer en die invloed hiervan op die kostes van 'n projek te toon in dien dit probabilisties gemodelleer word. Vir die eenheidtariewe van die items is die gemiddelde tendensaanangepaste waardes gebruik soos in *Tabel 5-8* getoon is en vir die hoeveelhede is die ooreenkomstige gemiddelde hoeveelhede van die items vanuit die databasis bepaal. Om die veranderlikheid van die eenheidtariewe te reflekteer is van die standaardafwykings wat ook in *Tabel 5-8* getoon word gebruik. Die korrelasies wat bepaal is in *Afdeling 5.5* tussen die eenheidtariewe is as invoer tot die kosteberekeninge gebruik waar korrelasie in ag geneem is. Hierdie resultate is in die kostefunksies wat in *Vgl. 4-5* en *4-27* getoon word ingestel en die totale projek koste is sodoende by verskillende vlakke van betroubaarheid bepaal.

---

---

## HOOFSTUK 6

---

### UITBREIDING VAN DATA-ONDERSKEPPING

---

**G**egewe die veranderlikhede en korrelasies wat eenheidtariefdata toon behoort daar eksterne faktore te wees wat hierdie statistiese verskynsels veroorsaak. Indien hierdie faktore geïdentifiseer en onderskep word, kan dit moontlik aangewend word om die epistemiese veranderlikheid te verlaag en insig te ontwikkel in die korrelasiemeganismes wat voorkom. Hierdie faktore kan ook vir kosteberekeninge wat vroeër in die leeftyd van konstruksieprojekte uitgevoer word, gebruik word om sodoende die waarde van die inligting te verhoog.

In hierdie hoofstuk word ondersoek ingestel na watter faktore hierdie veranderlikheid en korrelasie kan veroorsaak en watter van hierdie faktore gedurende die projek leeftyd onderskep behoort te word om kosteberekening te verbeter. Simplistiese invloeddiagrame word gebruik as gereedskapstuk om die interaksie tussen faktore wat die veranderlikheid van kostes en korrelasies veroorsaak te modelleer. Verder word ook geïdentifiseer hoe die bestaande onderskepte data se aanwending verder uitgebrei kan word om tydens kosteberekening in die vroeë begrotingsfase en konseptuele fase gebruik te word.

---

#### 6.1 PROJEKFAKTORE EN ITEMSPESIFIKASIES

---

Om sinvolle en betroubare kosteberekeninge gebaseer op historiese data te kan doen, is dit nodig om beskrywende inligting oor die herkoms van die data te hê. Dit moet beskikbaar wees om te bepaal hoe die historiese data gekonsilieer moet word, sodat dit vir die kosteberekening van 'n spesifieke projek aangewend kan word [Birrell (1980)]. Indien daar nie sulke data beskikbaar is nie en

---

die kosteberamer gebruik meganisties historiese data van vorige projekte sonder die wete of daar enige ooreenkomste bestaan tussen die projekte, kan dit lei tot groot onakkuraathede [Carr (1989)].

Hierdie probleem van swak spesifisering kom nie net op langtermyn en konseptuele vlak van kosteberamings voor nie, maar ook op die gedetailleerde vlak van kosteberamings, wanneer ooreenkomstig benoemde items van verskillende projekte nie ooreenkomstige kostes insluit nie. Dit word vererger deurdat die metode van opmeting van volumes en oppervlaktes vir items nie eers behoorlik gespesifiseer word nie [Birrell (1980)]. Dit is dus te wagte dat daar groot variasies in die eenheidtariewe van ooreenkomstig benoemde items sal voorkom.

In die Suid-Afrikaanse padkonstruksie omgewing bestaan daar 'n unieke situasie waar daar gestandaardiseerde items is wat in konstruksieprojekte gebruik word. Laasgenoemde probleem word grotendeels hierdeur aangespreek deur die duidelike en omvattende beskrywings wat in standaardspesifikasies, soos die CSRA-spesifikasies van padkonstruksie items, voorgeskryf word. Nie net word daar uitgespel presies watter koste by elke item ingesluit moet word nie, maar daar word ook duidelik voorgeskryf presies hoe opmetings van volumes en oppervlaktes gedoen moet word en watter aktiwiteite hierby betrokke is. Klein variasies in eenheidtariewe word dus verwag indien ooreenkomstige benoemde items se data van verskillende historiese projekte beskou word.

Die statistiese analise van die eenheidtariewe in *Hoofstuk 5* toon egter dat dit nie die geval is nie. Sommige items, nadat daar vir die effek van inflasie gekompenseer is en "orde grootte"-foute verwyder is, se standaardafwykings is feitlik dieselfde as die gemiddelde koste van die item. Hierdie tendens is egter nie uniek tot die databasis wat beskou is nie. In 'n artikel deur *Bell (1987)* om die nut van 'n databasis wat tenderinligting stoor, te illustreer, blyk dit dat meeste van die beduidende items wat in hierdie artikel geïdentifiseer is se eenheidtariewe 'n koëffisiënt van variasie tussen 0.5 en 1.0 het – dieselfde orde wat by die CSRA-databasis gevind is.

Die ooglopende probleem is dat daar nie voldoende projekspesifieke data beskikbaar is wat die omgewing, waarbinne die konstruksieprojekte plaas gevind het, beskryf nie. Feitlik geen geografiese, klimatologiese, politiese of ekonomiese inligting wat die kostes van die projekte beïnvloed het, is beskikbaar nie. Aan die eenkant is daar dus detail beskrywings van koste-items, maar slegs oppervlakkige beskrywings van die projekomgewing. Om sinvolle kosteberamings uit te voer is dit nodig om deurlopend dieselfde vlak van detail te gebruik [Carr (1989), Birrell (1980)].

---

Hierdie deurlopendheid wat betref projek- en itemspesifikasie ontbreek in die meeste databasisse wat historiese projekinligting stoor. Dit is derhalwe nodig om insig te ontwikkel in waar sulke projekfaktore benodig word in die verskillende tipes kosteberekeninge, *Afdeling 6.2*, en wat hierdie invloedfaktore moontlik kan wees, *Afdeling 6.3*. Die databasis kan sodoende uitgebrei word, *Afdeling 6.8*, om inligting te stoor om 'n meer eenvormige vlak van projekspesifikasie en itemspesifikasie daar te stel vir die verskillende kosteberekeninge wat tydens die leeftyd van 'n projek uitgevoer word.

In hierdie hoofstuk is die fokus dus om die epistemiese onsekerheid in die eenheidtariewe van detailkosteberekeninge te verminder deur faktore te identifiseer wat moontlik die veranderlikheid van eenheidtariewe tussen verskillende projekte veroorsaak. 'n Omvattende benadering tot die verbetering in dataonderskepping word egter hier gevolg deur nie net te fokus op detailkosteberekeninge nie, maar om ook te bepaal of van die inligting wat in die detail berekening geïdentifiseer word ook aangewend kan word om die ander kostebereamings te verbeter.

---

## 6.2 MODELLE VAN DATA-AANWENDINGSG

---

Tydens die verskillende kosteberekeninge, soos langtermyn, konseptueel en detail, word verskillende tipes kostedata benodig. Hierdie data kan wissel van gedetailleerde kostes (eenheidtariewe en hoeveelhede afgebreek tot op item vlak) tot ge-agregeerde kostes (koste van verskillende groepe items of die totale koste van verskillende projekte), asook die data wat benodig word om projekte te beskryf op verskillende vlakke van detail. Twee modelle, nl. die bestaande en die ideale model, van hoe hierdie data onderskeidelik aangewend word en aangewend kan word, word bespreek.

---

### 6.2.1 BESTAANDE MODEL

---

Die bestaande model, alhoewel triviaal in eenvoud, word volledig bespreek, sodat dit as voorbeeld kan dien om die uitleg van die diagram wat die model voorstel te verduidelik en om vir vergelykende doeleindes met die ideale model gebruik te kan word.

---

#### 6.2.1.1 VORM VAN BESTAANDE MODEL

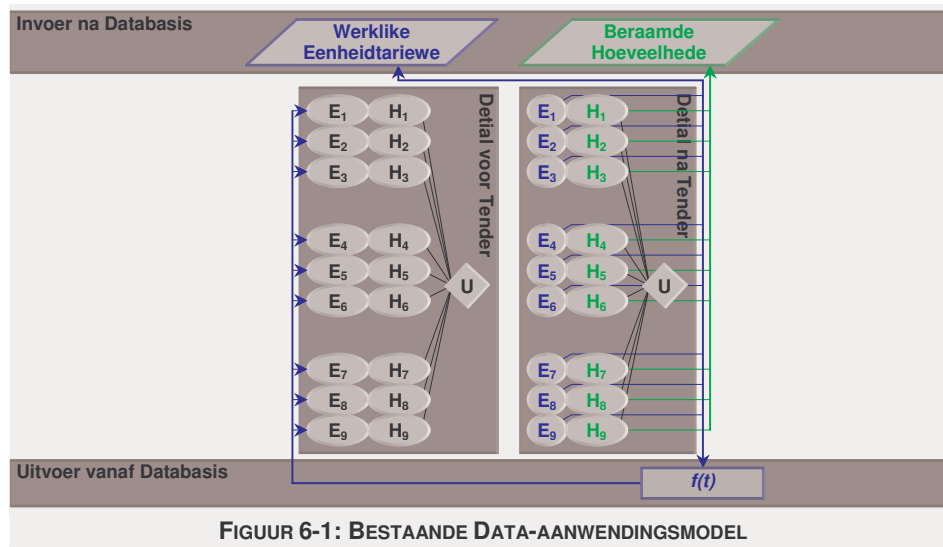
---

In die bestaande model word die data in die tenderverslae van nuwe projekte onderskep. Hierdie dokumente lewer twee stappe inligting, nl. die beraamde hoeveelhede en die getenderde eenheidtariewe van die items wat deur die raadgewende ingenieur voorsien is en in die projek gebruik gaan word. Van hierdie twee stappe inligting word slegs die historiese eenheidtariewe gebruik in die detailkosteberekening om eenheidtariewe te beraam voor die tenderfase

---



van nuwe projekte. Die model wat hierdie data-aanwending beskryf kan met die diagram in *Figuur 6-1* voorgestel.



Die model bestaan uit drie gedeeltes, nl. die invoer van inligting na die databasis, die verskillende tipes berekening wat uitgevoer word en die uitvoer van inligting vanaf die databasis. In die bestaande model is die detailkosteberekening voor en na die tenderfase die enigste berekening waarby inligting van die databasis onderskeidelik benut of geskep word. Inligting oor die getenderde (werklike) eenheidariewe,  $E_i$ , en die beraamde hoeveelhede,  $H_i$ , van elke item,  $i$ , per projek word na die databasis afgevoer, afkomstig uit die kosteberekening na die tenderfase. In die detailkosteberekening voor die tenderfase word slegs die eenheidariefinligting, met die nodige transformasies vir temporale veranderlikheid,  $f(t)$ , van die databasis gebruik.

Tydens die kosteberekening wat voor die tenderfase uitgevoer word, is die werklike eenheidariewe en hoeveelhede van die items onbekend (hierdie waardes word deur die raadgevende ingenieur beraam met 'n gepaardgaande onsekerheid), derhalwe die ovaalvorm waarmee dit in die diagram voorgestel word. Die struktuur van die detailkosteberekening na die tenderfase is soortgelyk as die voorafgaande detailkosteberekening, met slegs een verskil – die eenheidariewe is bekend. Die eenheidariewe se voorstelling in die diagram verander dus vanaf 'n ovaal na 'n sirkel. Dit is ook tydens hierdie fase dat die data onderskep word, dus word die afvoere wat na die databasis ontstaan aangetoon. Die groepering van drie items in drie versamelings, stel die groepering van items in drie denkbeeldige groepe, of soos in *Afdeling 6.4.2* bespreek word *Series*, in die figuur voor.

### 6.2.1.2 INVOER NA DATABASIS

In hierdie model is die enigste invoer afkomstig vanaf die detailkosteberekening na die tenderfase. Die eenheidariewe wat gestoor word is werklike waardes

wat afkomstige is vanaf die kontrakteurs se tenders. Die hoeveelhede wat gestoor word is egter beraamde hoeveelhede soos bereken deur die raadgevende ingenieur.

Geen inligting word na die databasis afgevoer ten opsigte van die werklike hoeveelhede van die items nadat die projek voltooi is nie. Die onsekerheid wat bestaan ten opsigte van die beraamde hoeveelhede kan dus nie bepaal word nie. Geen koste-inligting word ook onderskep van die onvoorsiene items wat tydens konstruksie gebruik is nie. Inligting wat gebruik kan word om die projekte te beskryf word ook nie onderskep nie. Daar is dus geen manier om onderskeid te maak tussen verskillende projekomgewings wat die ruimtelike veranderlikheid in eenheidtariewe kan veroorsaak nie.

---

### 6.2.1.3 KOSTEBEREKENINGE

---

Die twee kosteberekening wat getoon word is die detail berekeninge voor en na die tenderfase. Tydens hierdie berekening word daar slegs aandag geskenk aan die (beraamde en getenderde) eenheidtariewe en die beraamde hoeveelhede. Die effek van moontlike afwykings in die beraamde hoeveelhede word nie in ag geneem nie en so ook nie die addisionele koste wat onvoorsiene items kan veroorsaak nie. In die berekening word wel deterministies voorsiening gemaak vir onvoorsiene kostes deur gewoonlik 10% van projek koste by die totale projek koste te voeg.

Ander kosteberekening, bv. langtermyn en konseptueel, gebaseer op die historiese data in die databasis is ook moontlik. Dit is egter nie realisties om net gemiddeldes van historiese waardes met oppervlakte of afstande te skaal vir nuwe projekte nie. Ruimtelike veranderlikheid veroorsaak variasie in kostes tussen projekte; inligting om hierdie veranderlikhede te verminder bestaan egter nie.

---

### 6.2.1.4 UITVOER VANAF DATABASIS

---

Die enigste uitvoer van die databasis is eenheidtariewe van historiese projekte se items wat gebruik word in die detail kosteberekeninge voor die tenderfase van nuwe projekte. Hierdie data word slegs vir die effek van temporale veranderlikheid aangepas, omdat daar nie inligting beskikbaar is om die data sodanig te transformeer dat dit ruimtelike veranderlikheid in ag neem nie.

Met die inligting wat oor die hoeveelhede van historiese items in die bestaande model onderskep word kan daar, naas korrelasie ontledings, niks mee gedoen word nie, omdat daar geen inligting beskikbaar is oor die werklike hoeveelhede nie. Onsekerhede wat gepaardgaan met die beraamde hoeveelhede kan dus nie bepaal word en as uitvoer vanaf die databasis in detail kosteberekeninge gebruik word nie.

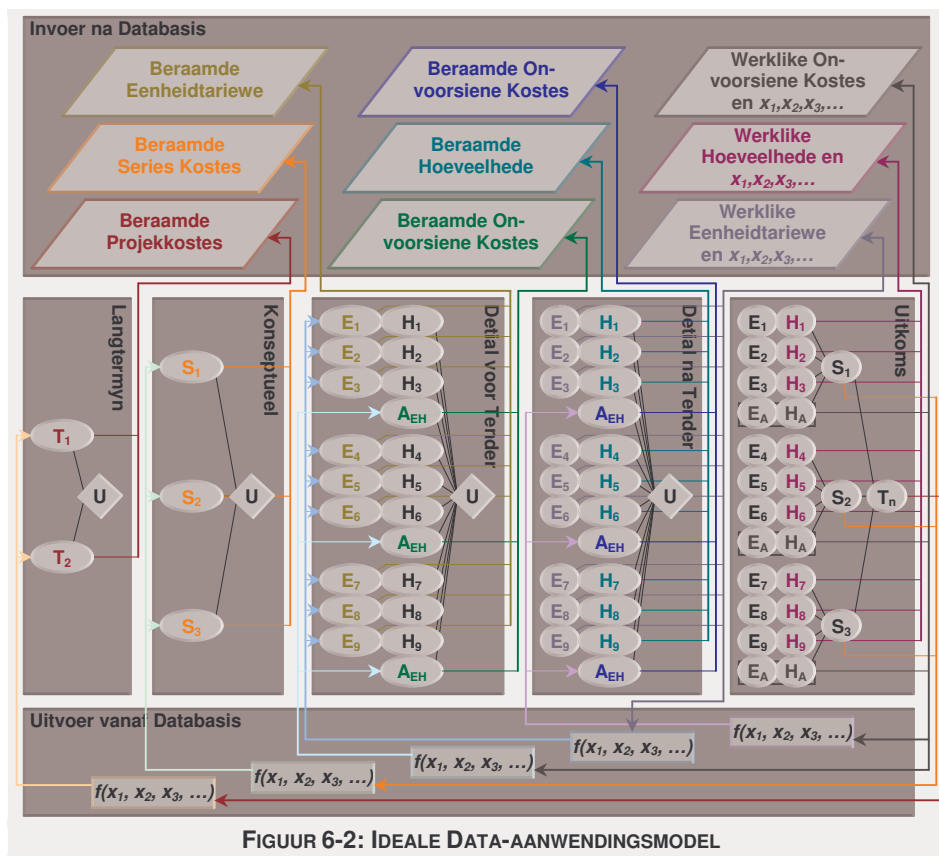
---

## 6.2.2 IDEALE MODEL

Die ideale model se onderskepping en aanwending van data is aansienlik meer omvangryk as die bestaande model. Feitlik alle inligting wat geskep word tydens die verskillende kosteberekeninge word onderskep en weer aangewend vir die beraming van nuwe projekkosies en die onsekerhede wat daarmee gepaardgaan.

### 6.2.2.1 VORM VAN IDEALE MODEL

In die ideale model word die tender- en die projekvoltooiingsverslag onderskep, asook addisionele beskrywende data oor die projek. Verder word beraamde eenheidtariewe en beraamde kosies tydens die ander fases ook onderskep. Hierdie data word aangewend nie net om eenheidtariewe te beraam nie, maar ook om kosies tydens die konseptuele fase en vir langtermyn doeleindes te bepaal saam met die gepaardgaande onsekerhede. Die model word in *Figuur 6-2* getoon.



In hierdie model word nege stelde data gestoor wat in twee groepe verdeel kan word. Die eerste groep bestaan uit ses stelde beraamde waardes van historiese projekte, nl. die kosies van projektipies, kosies van *Series*, eenheidtariewe, onvoorsiene kosies per *Serie* voor die tenderfase, hoeveelhede en onvoorsiene kosies per *Serie* na die tenderfase. Die tweede groep bestaan uit drie stelde werklike waardes van historiese projekte, nl. die eenheidtariewe,

hoeveelhede en onvoorsiene kostes per *Serie*. Verder meer kan die invloedfaktore wat hierdie drie stelle data beskryf ook na die databasis afgevoer. Presies watter data word in *Afdeling 6.4* bespreek.

Vyf tipiese kosteberekeninge kan in die leeftyd van 'n projek voorkom. Die eerste twee, nl. langtermyn en konseptueel, bestaan uit ge-agregeerde eenheidtariefdata wat onderskeidelik volgens projektipies, *T*, en *Series*, *S*, gegroepeer is, teenoor die daaropvolgende drie kosteberekeninge wat op item detailvlak (hoeveelhede, *H*, en eenheidtariewe, *E*) gedefinieer word. Die inligting wat benodig word vir die eerste vier berekeninge is afkomstige van die databasis met die nodige transformasies om voorsiening te maak vir temporale en ruimtelike veranderlikheid.

---

### 6.2.2.2 INVOERE NA DATABASIS

---

Invoere na die databasis vanaf die verskillende kosteberekeninge tydens die onderskeie fases kom gedurende die hele leeftyd van 'n projek voor. By elk van die daaropvolgende fases is daar egter 'n uitbreiding in die vlak van detail van inligting wat na die databasis gevoer word. Die totale koste van die spesifieke projektype by die langtermyn berekeninge word afgevoer teenoor die konseptuele berekeninge waar die kostes per *Serie* in 'n projek afgevoer word en by die detailkosteberekeninge waar die afvoer tot op itemvlak afgebreek word.

Hierdie aanvanklike afvoere is slegs die beraamde waardes van nuwe projekte gebaseer op vorige historiese projekte se werklike kostes, nadat dit geskaal is vir die nuwe projekte. Wanneer daar egter werklike inligting vanaf 'n projek beskikbaar begin word, soos byvoorbeeld: die werklike getenderde eenheidtariewe; werklike hoeveelhede en werklike kostes van onvoorsiene items, is dit nodig om beskrywende inligting te onderskep oor die projekomgewing. Hierdie beskrywende inligting moet omvattend genoeg wees, sodat dit voldoende inligting bevat om die omgewing van die projek te karakteriseer, vir die onderskeie kosteberekeninge wat tydens die leeftyd van 'n nuwe projek uitgevoer kan word. Die identifisering van hierdie inligting word verder in *Afdeling 6.8* bespreek.

Alhoewel dit voorkom asof daar 'n dramatiese toename in data-onderskepping moet plaasvind is dit nie die geval nie. Meeste van die inligting wat na die databasis ingevoer word, naas projekbeskrywende inligting, word alreeds in harde kopie dokumente vasgevang. Dit is slegs nodig om die data-vloei van hierdie dokumente na die databasis te verbeter. Sewe van die groepe inligting wat getoon word is afkomstig, met geringe aanpassings, uit die twee tradisionele verslae, nl. die tender- en die projekvoltooiingsverslag.

---

- *Tenderverslag* – Die inligting wat gewoonlik in die tenderverslag verskyn is die beraamde hoeveelhede en beraamde eenheidtariewe van die raadgewende ingenieur, asook die getenderde (werklike) eenheidtariewe van die verskillende kontrakteurs. Die ander twee stelle inligting wat maklik in hierdie dokument bygevoeg kan word, is die beraamde onvoorsiene koste per *Serie* voor die tenderfase (gebaseer op die raadgewende ingenieur se kosteberekening) en die beraamde onvoorsiene kostes na die tenderfase vir elk van die tenders wat deur die verskillende kontrakteurs ingedien word.

Die doel van die onderskepping van die beraamde eenheidtariewe is om dit te kan vergelyk met die getenderde eenheidtariewe, sodat die onsekerheid bepaal kan word waarmee hierdie beramings gepaardgaan. Die onderskepping van die beraamde onvoorsiene kostes per *Serie* is ook vir die bepaling van die onsekerheid van hierdie beramings ten opsigte van die werklike onvoorsiene kostes per *Serie*.

Die rede hoekom die onvoorsiene kostes per *Serie* twee keer onderskep moet word (voor en na tenderfase) is te wyte aan 'n verbetering in die inligting waarop die kosteberaming gedoen word. Die onsekerheid weens variasie in die raadgewende ingenieur se beraamde eenheidtariewe en hoeveelhede voor die tenderfase, veroorsaak dat onsekerheid oor die onvoorsiene kostes per *Serie* afkomstig is van twee bronne. Na die tenderfase voltooi is en die getenderde eenheidtariewe beskikbaar is kan verbeterde kosteberekening van die onvoorsiene koste vir elk van die *Series* gedoen word, omdat slegs die werklike hoeveelhede nou onbekend is. Uit die tenderverslag word die beraamde onvoorsiene koste per *Serie* voor die tenderfase (gebaseer op die raadgewende ingenieurs se kosteberekening) en na die tenderfase (gebaseer op die verskillende kontrakteurs se kosteberekening) na die databasis ingevoer.

Die enigste stel inligting wat werklik addisionele insette van die raadgewende ingenieur vereis, is die kwantifisering van die faktore wat die projekomgewing beskryf om sodoende die epistemiese onsekerheid te verlaag. In *Afdelings 6.3* tot *6.6.2* word van hierdie moontlike faktore verder bespreek en in *Afdeling 6.8* word getoon hoe die invloedrykste faktore bepaal en in die databasis gestoor kan word om sodoende die voordeel van invloedfaktoronderskepping te maksimeer.

- *Projekvoltooiingsverslag* – Die inligting afkomstig uit die projekvoltooiingsverslag wat na die databasis afgevoer moet word, is die werklike hoeveelhede en die werklike onvoorsiene kostes per *Serie*. Dit geskied nadat die konstruksie voltooi is en die verskille oor hoeveelhede wat vir
-

betaling ge-eis word opgelos is en dispute geskik is. In die diagram is die oorsprong van hierdie inligting en die faktore wat die onder/oorskatting van hoeveelhede en onvoorsiene kostes veroorsaak het, die uitkomstkos-teberekeninge. Hierdie inligting kan saam met die beraamde kostes per *Serie* en itemhoeveelhede gebruik word om die onsekerhede te bepaal wat tussen beraamde en werklike waardes voorkom.

Die oorblywende twee groepe inligting is die beraamde koste in die langtermyn en konseptuele kosteberekeninge. Dit kan weereens aangewend word om die onsekerheid van hierdie beramings te bepaal indien dit met die werklike kostes van projekte vergelyk word. Die datavloei van hierdie beraamde waardes is egter nie so duidelik vasgelê soos van die voorafgaande twee bron dokumente nie. Daar kan ook verder geargumenteer word dat die insameling van hierdie waardes nutteloos is, omdat projekte sodanig kan verander tydens die verskillende fases dat daar nie meer 'n verband tussen die oorspronklike projekdefinisie waarop die kosteberekeninge gedoen is en die werklike projekte te bespeur is nie.

Indien die prosesmodelle in *Afdeling 2.1.4* egter beskou word is dit duidelik dat hierdie twee kosteberekeninge wel gedefinieer is in die prosesmodel van 'n tipiese projek; verder meer word daar ook duidelike afvoere in die vorm van dokumente geskep. Hierdie inligting is dus beskikbaar en kan onderskep word; dit is net 'n geval dat dit bestuur moet word. Die antwoord op die vraag of hierdie historiese kosteberekeninge nutteloos is weens projekveranderinge en projekkruip lê in die vraag. Weens projekveranderinge en projekkruip is dit juis nodig om te kan bepaal watter onsekerheid geassosieer is met hierdie vroeë vlakke van kosteberamings. Dit kan slegs gedoen word mits hierdie inligting onderskep word.

---

### 6.2.2.3 KOSTEBEREKENINGE

---

Die verskillende strukture van die kosteberekeninge gedurende die leeftyd van 'n projek wat in *Figuur 6-2* getoon word, is gebaseer op die bespreking van die kosteberekeningsmodelle in *Afdeling 6.4*. Hier word verduidelik hoe kostes ge-agreëer word tydens die verskillende kosteberekeninge.

Die langtermyn kosteberekening, hoëvlak koste samevoeging, se koste is afkomstig van individuele projektipes of saamgestel uit die verskillende projektipes, nl. instandhouding, rehabilitasie, herkonstruksie en nuwe padkonstruksie. Dit kan in die uitkomstmodel van *Figuur 6-2* gesien word dat die projekkkoste gebaseer is op getenderde (werklike) eenheidtariewe, die werklike hoeveelhede van die items en die werklike onvoorsiene kostes per *Serie*. Die koste is dus die totale werklike projekkkoste vir individuele projektipes of 'n samestelling van 'n aantal projektipes se totale werklike projekkkostes.

---

Die kostes van 'n konseptuele projek is afgebreek tot op 'n laer vlak van koste samevoeging en word dus saamgestel uit die kostes van verskillende itemgroepe, of soos in *Afdeling 6.4.2* bespreek word, die verskillende CSRA-Series. Die verskillende *Series* is 'n versameling van 'n aantal items wat verbandhou met mekaar. Die koste van 'n *Serie* is soos in die geval van die langtermyn kosteberekening saamgestel uit die werklike eenheidtariewe en hoeveelhede per item, asook die werklike onvoorsiene koste per itemgroep of dan *Serie*.

In die detailkosteberekening, laagste vlak van samevoeging, word die projekte se koste saamgestel uit die items wat deur die raadgewende ingenieur voorsien is gebruik gaan word se beraamde eenheidtariewe en hoeveelhede, asook onvoorsiene kostes per *Serie*. Die eenheidtariewe is gebaseer op die historiese eenheidtariewe wat aangepas word om dit projekspesifiek te maak. Die hoeveelhede wat in hierdie berekening gebruik word, is gebaseer op die hoeveelhede wat deur die raadgewende ingenieurs bereken word tydens die ontwerp. In enige projek kan dit egter verwag word dat daar items in die konstruksieproses gebruik gaan word wat nie aanvanklik voorsien is nie. Om presies te identifiseer wat hierdie items in 'n nuwe projek gaan wees, hoeveel daarvan benodig gaan word en wat die eenheidtariewe hiervan gaan wees, is egter feitlik onmoontlik. Daar word dus eerder 'n enkel bedrag per *Serie* vir onvoorsiene kostes beraam, gebaseer op die totale voorsiene koste van die ooreenstemmende *Serie*.

By die detailkosteberekening na die tenderfase verander die kosteberekening. Die beraamde hoeveelhede bly dieselfde (soms tydens die tenderfase word dit wel aangepas). Die werklike eenheidtariewe is egter nou bekend, dus kan daar 'n verbeterde beraming gemaak word van die totale voorsiene koste per *Serie*. Hierdie verbeterde beraming kan dan gebruik word om 'n verbeterde beraming van die onvoorsiene koste per *Serie* te maak.

In die uitkomsteberekening, nadat die projek voltooi is, is die eenheidtariewe, die werklike hoeveelhede en die koste van die onvoorsiene items per *Serie* bekend. Die totale werklike koste per *Serie* en ook die totale werklike koste van die gegewe projekte kan nou bereken word. Die werklike koste per *Serie* en die totale koste van die projekte word egter nie na die databasis afgevoer nie, omdat dit deur die samevoeging van die itemkoste en onvoorsiene kostes per *Serie* bereken kan word.

---

#### 6.2.2.4 AFVOER VANAF DATABASIS

---

In die ideale model word die data van historiese projekte getransformeer om die temporale en ruimtelike verskille tussen die historiese en die nuwe projek waar

---

die data aangewend word in ag te neem. Vyf verskillende tipe transformasies kan voorkom.

Die eerste groep transformasies stel die aanpassings voor wat per projektepe gedoen moet word vir langtermyn kosteberekening. Afhangende van die langtermyn kosteberekening kan dit gebeur dat net een projektepe se langtermyn kostes bereken moet word of dat meer as een bereken moet word. Die afvoere na hierdie transformasies is die totale werklike projektkostes per tipe historiese projek wat aangepas word met die invloedfaktore soos bespreek in *Afdeling 6.5*.

Die tweede groep transformasies stel die aanpassing voor wat per *Serie* gedoen word vir die konseptuele kosteberekening. Die afvoer vanaf die databasis is die werklike historiese koste per *Serie*, na die ooreenstemmende transformasie. Weereens kan die konseptuele berekening net bestaan uit een *Serie* (wat natuurlik effens onrealisties is) of soveel as al nege *Series*. Hierdie kostes word omskep om dit spesifiek te maak vir die projek onder beskouing met die invloedfaktore soos bespreek in *Afdeling 6.6*.

Tydens die detailkosteberekening voor die tenderfase word inligting benodig vir die eenheidtariewe en die beraamde onvoorsiene kostes. Twee groepe transformasies word hiervoor gebruik. Die eenheidtariewe van die nuwe projek kan saamgestel word vanaf die werklike getenderde historiese eenheidtariewe. Die invloedfaktore wat dit mag beïnvloed word in *Afdeling 6.6.2* bespreek. Die beraamde onvoorsiene koste is gebaseer op die werklike onvoorsiene koste per *Serie* van vorige projekte. Dit word aangepas met die nuwe projek se koste per *Serie*, wat gebaseer is op die beraamde eenheidtariewe en hoeveelhede.

Die laaste groep transformasies wat voorkom is weereens die van die onvoorsiene kostes wat by die detailkosteberekening na die tenderfase gebruik word. Hierdie transformasie is nodig, omdat daar nuwe inligting bekend geraak het betreffende die eenheidtariewe van die voorsiene items. 'n Verbeterde beraming kan nou gemaak word van die koste van elke *Serie* en sodoende ook van die verwagte onvoorsiene kostes per *Serie*.

---

### 6.3 INVLOEDSDIAGRAMME

---

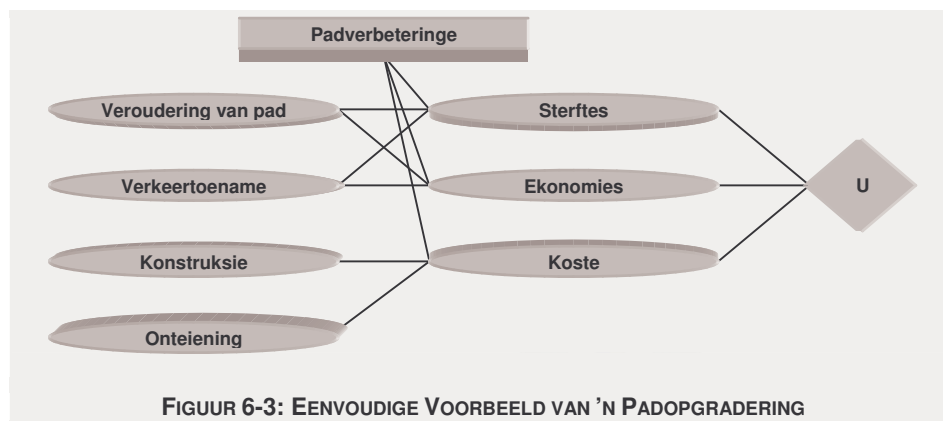
In die eenvoudigste vorm kan 'n invloeddiagram (besluitnemingsnetwerk) beskryf word as 'n gereedskapstuk om die topologie (rangskikking van boë en nodusse) van 'n stelsel of sisteem te beskryf. Die nodusse stel die huidige toestand, moontlike aksies, toekomstige toestande weens aksies en die utiliteit van hierdie aksies voor, terwyl die boë die verwantskappe hiertussen voorstel [Russel (1995)].

---



In *Figuur 6-3* word 'n voorbeeld getoon van 'n paar aspekte wat die keuse van die opgradering van 'n pad kan beïnvloed. Dit illustreer die drie tipes nodusse wat kan voorkom:

- **Kansnodus (ovaal):** stel die ewekansige veranderlikes voor. Die veroudering van 'n pad, verkeertoename, konstruksie en onteiening is onsekerhede asook die sterftes, ekonomiese impak en koste wat afhanklik is van die verskillende padverbeteringsmoontlikhede.
- **Besluitnodus (reghoek):** stel die punt voor waar 'n besluit geneem word; Die tipe rehabilitasie kan 'n impak hê op die sterftes, die ekonomiese voordele en die koste daarvan.
- **Utiliteitnodus (diamant):** stel die utiliteitsfunksie voor en het die kansnodusse wat direk die utiliteit beïnvloed as kinders. In hierdie geval kan die utiliteitnodus 'n kostefunksie wees, met monetêr kwantifiseerbare impakte, wat geminimeer moet word.



Bogenoemde voorbeeld van 'n invloeddiagram dui 'n tipiese struktuur aan van inligting wat benodig word by die keuse van die tipe opgradering van 'n pad. Dit is duidelik dat heelwat ander inligting ook by so 'n besluit betrokke is, naas die konstruksiekoste. Elk van hierdie kansnodusse kan in meer diepte ontwikkel word om die aspekte wat dit beïnvloed meer omvattend te modelleer. In hierdie studie, soos reeds genoem, word daar net gefokus op die konstruksiekoste en derhalwe is dit die enigste komponent wat verder ontwikkel word in hierdie hoofstuk.

## 6.4 KOSTEBEREKENINGMODELLE

Die inligting wat benodig word om verbeterde kosteberekeninge mee te doen kan met generiese invloedmodelle voorgestel word. Verskillende kosteberekeninge soos langtermyn, konseptueel en detail word egter op verskillende fases in die leeftyd van 'n projek gedoen, met gepaardgaande verskille in die vlak van detail, soos reeds bespreek in *Afdeling 4.3*. Weens hierdie logiese verskille in

vlakke van detail a.g.v. doel en tydstip van berekening kan elke tipe kosteberekening met unieke invloeddiagramme voorgestel word.

In hierdie afdeling word die vyf verskillende modelle, nl. die langtermynmodel, konseptuele model, die detailberekeningmodel voor en na die tenderfase en die finale uitkomstmodel nadat die projek voltooi is, oorsigtelik bespreek. Slegs die “eerste-vlak” invloedfaktore van hierdie modelle word bespreek; in die daaropvolgende afdelings word die volgende vlakke van invloedfaktore van elke model meer in diepte ondersoek.

---

#### 6.4.1 LANGTERMYNKOSTEMODEL

---

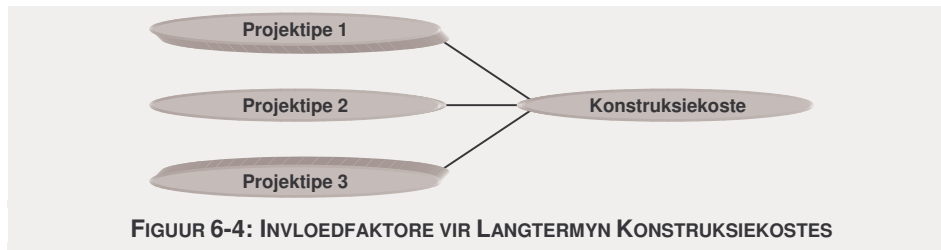
Langtermyn kosteberamings word gedoen vir die beplanning en begroting van kapitaal uitgawes. Dit kan projekspesifiek wees of dit kan oor 'n aantal projekte en projektipos span vir 'n sekere streek, provinsie of vir die hele land. Die vooruitskattingshorison vir langtermyn kosteberekeninge is ongeveer 5-jaar en meer. Akkuraathede van beter as tussen –25 en +40% word nie verwag nie [Yeo (1990)].

Weens die langtermyn aard en hoë vlak van abstraksie is geen detail van enige spesifieke projekte bekend nie, behalwe die grootte daarvan, waar dit miskien geleë mag wees en die tipe projek. Die invloedfaktore wat vir hierdie tipe kosteberaming gebruik word kan dus nie projekspesifiek wees nie, maar moet eerder veralgemeen genoeg wees om tendense te beskryf wat 'n impak kan hê op alle projekte.

Verskillende tipes padprojekte het verskillende kostes verbonde daaraan, soos byvoorbeeld: roetine instandhoudingswerk; verbeteringswerk; die herkonstruksie van 'n bestaande pad en die konstruksie van 'n algehele nuwe pad. Hierdie verskille ontstaan weens die aantal en tipe aktiwiteite wat voorkom by 'n gegewe projektype. Dit is tydens gesprekke met die CSRA-databasisbestuurders bevestig dat die onderskeie projektipos elk verskillende hoeveelhede en kombinasies van hulpbronne benut.

Indien daar dus kostes vir langtermyn doeleindes beraam moet word, is dit nodig om onderskeid te maak tussen die kostes wat die verskillende projektipos tot die totale koste bydra. Hierdie “eerste vlak” invloedfaktore van langtermyn kosteberamings is dus op 'n hoë vlak van abstraksie, wat bestaan uit die koste-bydrae van die verskillende tipes padprojekte (soos reeds hierbo genoem). Die vereenvoudigde model word in *Figuur 6-4* getoon en word in *Afdeling 6.5* na die volgende vlak uitgebrei.

---



FIGUUR 6-4: INVLOEDFAKTORE VIR LANGTERMYN KONSTRUKSIEKOSTES

## 6.4.2 KONSEPTUELE KOSTEMODEL

Die doel van die konseptuele kosteberekening is om die lewensvatbaarheid van 'n projek te bepaal of om alternatiewe konseptuele ontwerpe relatief tot mekaar te vergelyk. Die vooruitskattingshorison is ongeveer 2-jaar en meer. Die beste akkuraatheid wat met hierdie tipe kosteberekening geassosieer word is in die orde van  $-15$  tot  $+25\%$  [Yeo (1990)]. Weereens is daar 'n aantal invloed-faktore naas die konstruksiekoste wat die utiliteitsfunksie beïnvloed. Daar word egter in hierdie studie, soos reeds genoem, slegs gefokus op die konstruksie-kostes en die gepaardgaande invloedfaktore.

Tydens die konseptualisasie van 'n projek is meer projekspesifieke inligting beskikbaar as in die geval van die langtermyn kosteberekening, *Afdeling 6.4.1*, waar die berekening oor 'n verskeidenheid van projekte kan span en min detail van individuele projekte beskikbaar is. Basiese inligting soos ligging en die definisie van die projekomvang behoort al in hierdie fase van die projek beskikbaar te wees. Detail ontwerp-inligting wat gebruik kan word om volumeberekening mee uit te voer waarop detailkosteberekening gebaseer is, is egter nie tydens hierdie fase bekend nie en so ook nie die eenheidtariewe van items nie. Die faktore wat die kostes beïnvloed kan derhalwe meer konkreet gemaak word om die meer omvattende inligting wat beskikbaar is oor die projek in ag te neem, maar dit kan nie tot op itemspesifieke vlak ontleed word nie.

*Herbsman (1984)* verwys spesifiek na die tipiese kostepunte van 'n projek, nl. arbeid, materiaal, toerusting en oorhoofse koste, wat tydens voorlopige kosteberekening in gedagte gehou moet word. Meeste van die ander literatuur wat bestudeer is handel oor die onderwerp van empiriese metodes om die **akkuraatheid** van die kosteberekening te voorspel gegewe die omgewing waarbinne die kosteberekening gedoen is. [Oberlender (2001), Knight (2002), Attalla (2001), Trost (2003)]. Faktore soos die ervaring van die ontwerpspan, detail van ontwerp, tyd en historiese koste is gevind as bepalende faktore wat die akkuraatheid van kostes kan beïnvloed. Hierdie faktore beskryf egter slegs van die oorsake wat die afwykings tussen werklike en beraamde kostes kan beïnvloed, maar nie die faktore wat die kostes bepaal nie.

Die CSRA-spesifikasies bied 'n goeie vertrekpunt om sulke hoofkostekomponente, waaruit die konstruksiekostes saamgestel is, te identifiseer. Die

historiese detailkosteberekeninge se data is bekend, dus is dit slegs nodig om versamelings van gemeenskaplike items te vorm. Die groepering van ooreenkomstige items is egter alreeds in die verskillende *Series* van die CSRA-spesifikasies, *Tabel 5–1*, gedoen. Die eerste vlak van faktore waaruit die konseptuele kostemodel saamgestel is, kan dus gebaseer word op die reeds bestaande groeperings in hierdie spesifikasie.

Die kostes wat verbandhou met grondwerke, plaveisels van gruis of gebreekte klip val onder die koste-item *Grondwerke*, bv. maak van leengroewe, breek, plasing en kompaktering van materiaal, verskuiwing van grond, ens. Die *Asfalt en Seëls* kostekomponent sluit al die kostes in wat te doen het met die verkryging, plasing en kompaktering van asfaltplaveisel en –seëls. Koste wat verbandhou met erosiebeskerming, plasing van skanskorwe, padtekens en omgewingsverwante werk groepeer saam onder *Padwerke*. *Allerlei* omskryf aspekte wat verbandhou met die toets van materiale, kwaliteitskontrolle en verwerk. Die koste van duikers, dreine, randstene of asfaltberms val onder *Dreinerings*. *Algemeen* sluit die oprigting van die konstruksiekamp en die kostes van padverleggings in. Die *Strukture* kostekomponent sluit al die kostes van beton bouwerke in, bv. steierwerk en bekisting, staal en beton, betonafronding, ens.

In *Figuur 6-5* word die verskillende invloedfaktore getoon. Elke kostekomponent sluit al die kostes vir materiaal, brandstof, toerusting en arbeid in, asook die oorhoofse kostes en wins. Daar word dus voldoen aan die vereiste wat deur *Carr (1989)* gestel word dat die gepaste vlak van detail gebruik word en dat alle kostes volledig ingesluit moet word. Die volgende vlak van invloedfaktore word in *Afdeling 6.6* bespreek.

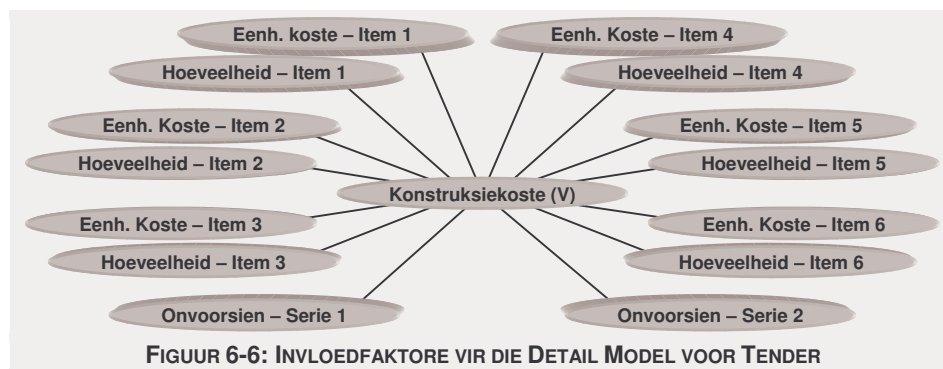


### 6.4.3 DETAILKOSTEMODEL

Detailkosteberekening word uitgevoer om verskillende ontwerpe met mekaar te vergelyk en om die finale kostes van 'n projek te bepaal. Die akkuraatheid wat verwag word vir die doel om projekte te vergelyk is tussen –10 en +15% terwyl finale kosteberekening se akkuraatheid tussen -5 en +10% is [Yeo (1990)]. Die vooruitskattingsperiode is gewoonlik minder as 'n jaar.

Tydens die detailkosteberekening behoort die ontwerp feitlik voltooi te wees en al die nodige inligting beskikbaar te wees om die items wat gebruik gaan word en hul hoeveelhede te bepaal. Die konstruksiekoste kan dus op 'n meer gedetailleerde vlak tydens hierdie fase bereken word as in die vorige kosteberekening. Die werklike eenheidtariewe en so ook die werklike hoeveelhede is egter nie bekend nie weens 'n aantal faktore wat afwykings tussen die beraamde en die werklike kostes en hoeveelhede van die items kan veroorsaak.

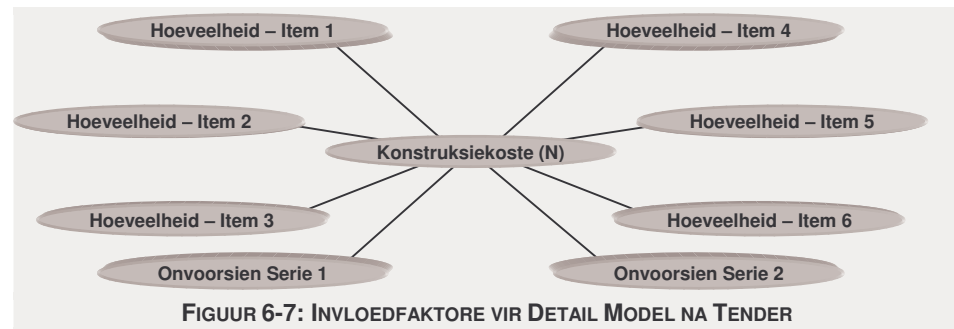
Die items waaruit die projek saamgestel is, is die standaarditems. Vir groot projekte kan daar tot 600 items wees, maar dit moet in gedagte gehou word dat slegs 'n klein gedeelte van die items – die met die grootste impak – probabilisties gemodelleer hoef te word, terwyl die res deterministies gemodelleer kan word. Die probabilistiese items se eenheidtariewe en hoeveelhede word voorgestel met kansnodusse, *Figuur 6-6*.



Tydens die detailkosteberekening word die volumes met die beskikbare ontwerp-inligting so ver as moontlik akkuraat bepaal. Dit kan egter weens onvoorsiene omstandighede gedurende die konstruksiefase gebeur dat meer materiaal uitgegrawe, verskuif of gebruik moet word, of dat vertragings voorkom. Addisionele werk of wysigings wat deur die kliënt aangevra word kan ook veroorsaak dat addisionele volumes materiaal of tyd benodig word. Naas die afwyking wat dit in die beraamde hoeveelhede van die items meebring, kan dit gebeur dat items wat nie voorsien is nie benodig word. Aangesien daar tydens hierdie kosteberekening geen manier is om te bepaal wat hierdie items is nie en nog minder die hoeveelhede en eenheidtariewe hiervan, word dit slegs as 'n enkele kansnodus, Onvoorsiene koste per *Serie*, voorgestel.

Na die tenderfase voltooi is en die projek aan die kontrakteur toegeken is, lyk die invloedsmodel van die kosteberekening effens anders. Die onsekerheid ten opsigte van die item se eenheidtariewe is uit die weg geruim en derhalwe bly slegs die onsekerheid van die hoeveelhede oor van die items wat tydens die detailkosteberekening geïdentifiseer is, *Figuur 6-7*. Die hoeveelheid kansnodusse word dus in hierdie model getoon as ook die kansnodus Onvoorsiene

koste per *Serie* wat nog vir dieselfde redes as vroeër ingesluit moet word, maar wat nou effens beter beraam kan word.



#### 6.4.4 UITKOMSKOSTEMODEL

Wanneer die konstruksie gedeelte van 'n padprojek voltooi is, al die nodige sertifikate van voltooiing deur die kontrakteur ingedien is en alle dispute en verskille afgehandel is kan die finale kosteberekening van die projek gedoen word. Die totaal kan gesien word as die werklike uitkoms (koste) van die projek. Hierdie bedrag kan gebruik word om te bepaal hoe goed die prestasie van konsultante was om die konstruksiekoste te bereken en/of te bestuur het. Aangesien daar geen onsekerheid meer verbonde behoort te wees aan die hoeveelhede, eenheidariewe en die onvoorsiene kostes nie, kan die hele model reduseer tot slegs die utiliteitfunksie.

Die realiteit is ongelukkig anders. Verskille wat ontstaan tussen hoeveelhede wat ge-eis word asook onvoorsiene items kan veroorsaak dat dispute lank sloer. 'n Mate van onsekerheid bly dus oor wanneer die projekvoltooiingsverslag beskikbaar gemaak word.

#### 6.5 LANGTERMYNKOSTEMODEL INVLOEDFAKTORE

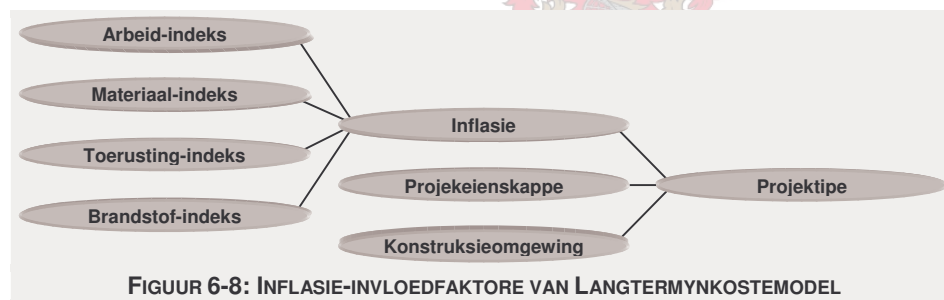
In die literatuur wat handel oor kosteberekeninge is daar nie veel inligting beskikbaar oor langtermyn kosteberekeninge in konstruksieprojekte nie. Meeste navorsing wentel slegs om konseptuele en detailkosteberekeninge. Die bevindings van hierdie navorsings is gevolglik meer projekspesifiek en die faktore is nie veralgemeen genoeg om in 'n langtermynkostemodel gebruik te word nie.

Slegs twee artikels, *Herbsman (1984)* en *Wilmont (2003)*, kon geïdentifiseer word wat hierdie onderwerp aanspreek. *Herbsman (1984)* identifiseer twee aspekte wat die koste van projekte kan beïnvloed, nl. werklike insetkoste en sogenaamde "influence factors". *Wilmont (2003)* is ook eens dat die werklike insetkoste 'n bepalende rol speel, veral die eskalasiestkoste van bepalende koste-items word spesifiek uitgelig, asook twee ander invloedfaktore, nl. projek-eienskappe en kontrakteursomgewing. Alhoewel *Wilmont (2003)* se projekteienskappe baie spesifiek is kan dit saam met *Herbsman (1984)* se

faktore gevoeg word om drie algemene invloedfaktore voor te stel, nl. inflasie, projekeienskappe en konstruksie-omgewing.

Die eersgenoemde invloedfaktor, inflasie, se impak op individuele projekiteme is reeds getoon in *Afdeling 5.3*. Die groot impak wat inflasie op die langtermyn kostes van projekte kan hê, selfs al bly die reële besteding presies dieselfde, is nie te betwyfel nie. Dit is heelwaarskynlik een van die belangrikste invloedfaktore wat in ag geneem moet word by langtermynkostemodelle. Vir konstruksiekostes word inflasie beïnvloed deur die prysverandering van die vier kostekomponente waaruit konstruksie aktiwiteite saamgestel is, nl. arbeid, materiaal, toerusting en brandstof.

Elke projektype het 'n unieke sensitiwiteit ten opsigte van die inflasie van hierdie komponente weens unieke kombinasies van aktiwiteite. Die prysstyging van hierdie komponente word weerspieël in die styging van die inflasie-indeks van elk. Elk van hierdie indekse kan waarskynlik nog verder afgebreek word tot by die faktore wat dit beïnvloed. Dis egter feitlik onmoontlik om al die faktore te bepaal wat prysstyging kan veroorsaak en nog minder moontlik om langtermynvoorspellings hiervan te maak. Daar word derhalwe volstaan met die vier komponente as die invloedfaktore tot die inflasiefaktor, *Figuur 6-8*.



Die projekeienskappe invloedfaktor, 'n veralgemening van *Wilmont (2003)* se projekeienskappe word gebruik om die gemiddelde koste van 'n spesifieke projektype te bepaal. Die twee mees veralgemeende projekeienskap invloedfaktore wat beskikbaar is tydens beplanningsdoeleindes en begrotingberamings is die omvang van die projek/te en die ligging/s daarvan. Die omvang invloedfaktor, alhoewel triviaal gegee die aard daarvan, word gebruik om twee belangrike langtermyn tendense mee te modelleer.

Eerstens behoort die langtermynkostemodel aanpasbaar te wees vir die verandering in die hoeveelheid padkonstruksie soos bepaal word deur ekonomiese of ander behoeftes. Ekonomiese ontwikkeling kan byvoorbeeld noodsaak dat meer en verbeterde padinfrastruktuur benodig word of dat ontwikkeling versadigings punt bereik het in sekere streke of provinsies en derhalwe verminder moet word. Tweedens behoort die model ook aanpasbaar te wees om 'n verskuiwing van projektypes te reflekteer. Dit is reeds genoem dat die koste

verbonde aan die daarstelling van nuwe infrastruktuur teenoor die instandhouding en verbetering daarvan verskil van mekaar, omdat daar verskillende aktiwiteite by betrokke is wat verskillende hulpbronne verbruik met verskille in die risiko's. Die model behoort derhalwe voorsiening te kan maak vir 'n verskuiwing in fokus van sê bou van nuwe infrastruktuur na die instandhouding daarvan en moontlike latere verbeteringswerk.

Die ligging van die konstruksiewerk speel ook 'n belangrike rol ten opsigte van die gemiddelde koste. In sommige streke is die gemiddelde padkonstruksiekoste per kilometer laer of hoër as in ander vir verskeie redes. Hierdie redes kan wissel van algemene geologiese toestande, topografie, neerslag, ondersteunende infrastruktuur, ens. Tydens die langtermynkosteberekening is daar egter te min inligting om hierdie aspekte voldoende individueel te kwantifiseer en derhalwe moet van 'n hoogs veralgemeende invloedfaktor gebruik gemaak word, wat wel beskikbaar is, om hierdie aspekte in aanmerking te neem.

Die twee invloedfaktore wat dus gebruik word om die projekeienskappe mee te modelleer is die omvang van werk wat gedoen moet word, gemeet in area, en die ligging waar dit gedoen moet word. In *Figuur 6-9* word die invloedmodel getoon.



Die derde invloedfaktor, konstruksie-omgewing, verwys na die omgewing waarbinne getender word en is 'n samevoeging van *Herbsman (1983)* se "influence factors" en *Wilmont (2003)* se kontrakteursomgewing. Die konstruksie-omgewing se invloedfaktore, soos die projektype se invloedfaktore, moet veralgemeen genoeg wees om breë tendense wat in die bedryf voorkom, te weerspieël. Twee konkrete aspekte is geïdentifiseer, nl. tendervolume en tegniese veranderinge, asook twee ander faktore wat effens meer vaag is, nl. prestasie en subjektiewe faktore, *Figuur 6-10*.

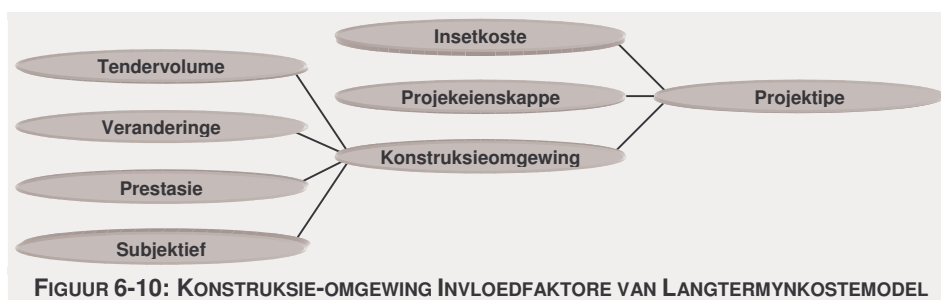
Die tendervolume invloedfaktor verwys na die tendens wat voorkom tussen tendervolume en getenderde pryse [*Herbsman (1984)*]. Wanneer daar 'n afname is in die jaarlikse totale bedrag waarvoor getender kan word, is dit gevind dat die getenderde pryse ook afneem. Die teenoorgestelde is ook gevind. Indien die jaarlikse totale bedrag waarvoor getender kan word begin styg, is daar 'n toename in die pryse wat getender word. Die mate van kompetisie is dus omgekeerd eweredig aan die tendervolume, 'n aspek wat ook bevestig is deur die CSRA-databasisbestuurders.



Die tegniese verandering invloedfaktor verwys na veranderinge in konstruksietegniese, ontwerpstandaarde en spesifikasies vir 'n gegewe projekte wat kostebesparings of verhoogde kostes kan veroorsaak. Hierdie aspek is deur *Wilmont (2003)* geïdentifiseer as een van die faktore wat 'n noemenswaardige impak kan hê op totale projektkostes. 'n Sprekende voorbeeld hiervan in Suid-Afrika is vaardigheidsoordrag wat verpligtend geword het in die standaardspesifikasies en tot 'n gevolglike styging in koste gelei het, alhoewel minder so as wat 'n oorskakeling van gemeganiseerde konstruksiemetodes na meer arbeidsintensiewe konstruksiemetodes gehad het om werkloosheid te verminder.

Die prestasie invloedfaktor verwys na die vermoë van die raadgevende ingenieur, kontrakteur en subkontraakteur om 'n padkonstruksieprojek binne die verwagte begroting te voltooi. Dit kan geargumenteer word dat hierdie faktor nie relevant is tot hierdie hoëvlak kosteberekeninge nie, omdat dit meer projekspesifiek is teenoor 'n groep projekte waar daar 'n balans behoort te wees tussen projekte wat binne begroting afgehandel word en die wat hul begrotings oorskry het. Globale tendense kan egter veroorsaak dat meer/minder projekte binne begroting voltooi word, soos wanneer daar 'n groot aantal kontrakteurs met min ervaring tot die mark toetree.

Die subjektiewe faktore waarna *Herbsman (1984)* verwys kan enersyds in die model ingebou word om voorsiening te maak vir politieke, ekonomiese en sosiale faktore en andersyds weggelaat word, omdat langtermyn kwantifisering van hierdie faktore feitlik onmoontlik is. Op eie bodem is die geweldige verswakking en die daaropvolgende versterking van die Suid-Afrikaanse Rand teenoor ander wêreldgeldeenhede 'n sprekende voorbeeld van onverwagte subjektiewe faktore wat 'n dramatiese styging in die koste van konstruksietoerusting gehad het.



## 6.6 KONSEPTUELE KOSTEMODEL INVLOEDFAKTORE

Die invloedfaktore wat die sewe kostekomponente in die konseptuele model beïnvloed kan in twee groepe verdeel word, nl. kwantitatiewe en kwalitatiewe faktore.

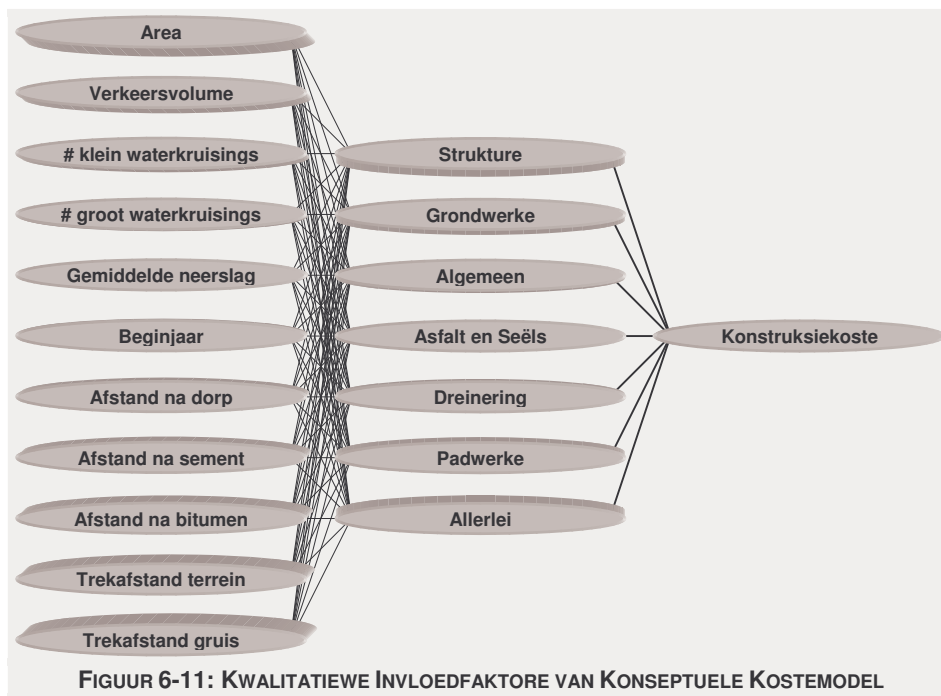
### 6.6.1 KWANTITATIEWE FAKTORE

Kwantitatiewe faktore is faktore waaraan daar 'n fisiese betekenis of afmeting gegee kan word. 'n Tipiese voorbeeld van 'n kwantitatiewe beskrywing is die afstande van die pad, die gemiddelde neerslag en aantal waterkruisings. Vir die doel van hierdie studie kan die kwantitatiewe faktore in die algemeen gesien word as faktore wat gemiddeldes van die hoeveelhede wat benodig word om 'n projek te voltooi, beskryf, teenoor kwalitatiewe faktore wat die afwyking vanaf die gemiddeld van die hoeveelhede beskryf.

Die kwantitatiewe faktore wat geïdentifiseer is [Akinci (1998), Adeli (1998), El-Rayes (2001), Hegazy (1998), Roy (2001), SANRAL (2002)] wat die konstruksiekoste van 'n pad kan beïnvloed word vervolgens bespreek.

- Area: die hoeveelhede van die eerste vlak invloedfaktore word bepaal deur die omvang van die projek.
- Verkeersvolume: belastings wat die pad moet dra bepaal duursaamheid van ontwerp met gepaardgaande toename in hulpbronne wat tydens konstruksie gebruik moet word.
- Aantal klein waterkruisings: bepaal die aktiwiteite en volume hulpbronne wat benodig word ten opsigte van duikers.
- Aantal groot waterkruisings: bepaal die aktiwiteite en volume hulpbronne wat benodig word ten opsigte van brûe.
- Gemiddelde neerslag: die grootte van dreineringsstelsels en moontlike probleme ten opsigte van fondamenttoestande.
- Beginjaar: die kostes van die hulpbronne word bepaal deur die effek wat inflasie daarop kan hê.
- Afstand na naaste groot dorp: bepalende invloed op vervoerkoste en verblyfreëlings wat vir die konstruksiespan getref moet word.
- Afstand na naaste sementbron: bepaal vervoerkostes van sement asook bergingsgeriewe.
- Afstand na naaste bitumenbron: bepaal vervoerkostes van bitumen asook bergingsgeriewe.
- Gemiddelde trekafstand op terrein: bepaal die vervoerkostes vir uitgrawings en opvullings.
- Gemiddelde trekafstand na gruisgroef: bepaal die vervoerkostes betrokke by die verskuiwing van gruis.

Die model wat voorgestel word om die verwantskappe tussen die eerste vlak invloedfaktore wat die konstruksiekostekomponente voorstel en die tweede vlak invloedfaktore wat die kwantitatiewe faktore beskryf, word in *Figuur 6-11* getoon. Hierdie model stel al die moontlike kombinasies van invloedfaktore op die hoëvlak invloedfaktore voor (sommiges mag onrealisties wees). Die sterkte van invloed van die kwantitatiewe invloedfaktor op die konstruksiekostekomponente wissel egter. Sommige kostekomponente word minimaal beïnvloed deur van die kwantitatiewe invloedfaktore terwyl ander weer feitlik in geheel afhanklik is daarvan. Hierdie sterktes en die gepaardgaande onsekerhede van die voorspellings kan bepaal word met die metodes wat in *Hoofstuk 4* bespreek is.



## 6.6.2 KWALITATIEWE FAKTORE

Kwalitatiewe faktore is faktore waaraan daar nie fisiese betekenisvolle waardes gekoppel kan word nie maar eerder eienskappe wat beskryf word. 'n Voorbeeld van 'n tipiese kwalitatiewe beskrywing is: "Die neerslag in hierdie gebied is hoog" (neerslag is natuurlik ook kwantifiseerbaar) of "Die beweegruiimte op terrein is beperk". Die kwalitatiewe invloedfaktore moet in hierdie model die kompleksiteit en die moeilikheidsgraad van 'n projek beskryf.

Die kompleksiteit en moeilikheidsgraad tydens die uitvoering van 'n projek bepaal die afwyking vanaf die gemiddelde waardes van die hulpbronne, soos materiaal, brandstof, toerusting en arbeid, wat benut word, omdat meer of minder van hierdie hulpbronne benodig word om take te voltooi. Sommige projektipes is meer kompleks as ander, omdat daar heelwat meer aktiwiteite by betrokke is en meer hulpbronne geveerg word om die take te voltooi.

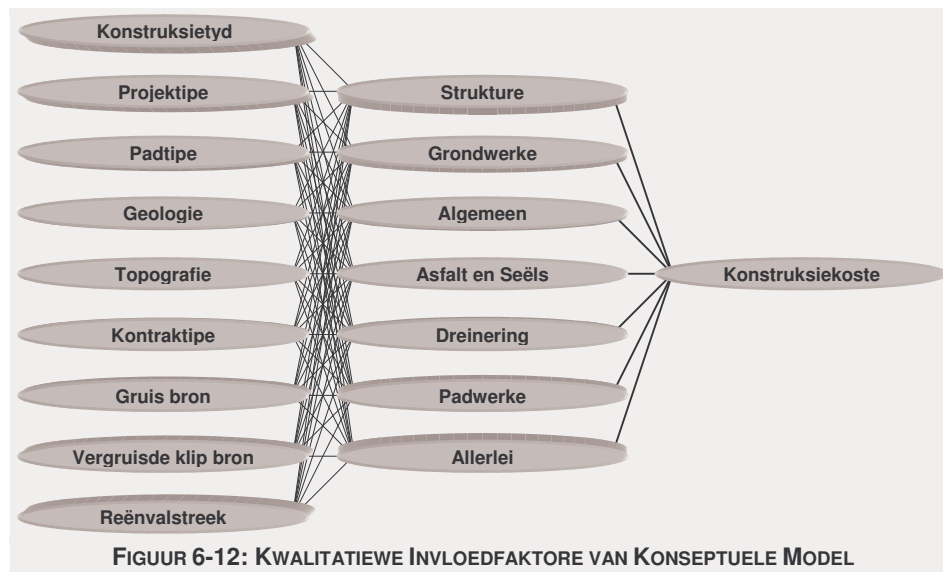
'n Padrehabilitasieprojek, waar daar slegs herstelwerk aangebring moet word, sal minder hulpbronne verbruik as die algehele herkonstruksie van 'n pad, alhoewel die area dieselfde kan wees. Dit word egter nêrens in die voorafgaande kwantitatiewe faktore bepaal nie. Omgewingsaspekte kan weer die moeilikheidsgraad beïnvloed van die aktiwiteite wat uitgevoer moet word wat ook kan lei tot 'n vermeerdering in die verbruik van hulpbronne. Hoë neerslag of bergagtige omgewing is enkele van die voorbeelde wat, onderskeidelik, die tipe dreineringsstelsel beïnvloed wat benodig word of hoeveelheid uitgrawings en opvullings wat gedoen moet word.

Die kwalitatiewe faktore wat geïdentifiseer is [Adeli (1998), El-Rayes (2001), Hegazy (1998), Roy (2001), SANRAL (2002)] wat die konstruksiekoste van 'n pad kan beïnvloed word vervolgens bespreek.

- Konstruksietyd: addisionele koste gekoppel aan projekte wat vinnig afgehandel moet word.
- Projektype: projektypes het 'n unieke samestelling van aktiwiteite wat uitgevoer moet word.
- Padtype: bepaal tot watter standaard die pad gebou of herstel moet word asook die aktiwiteite wat daarmee gepaardgaan.
- Topografie: beïnvloed die horisontale en vertikale belyning en dus die uitgrawings en opvullings.
- Geologie: bepaal die gemak van uitgrawingswerk en die kompleksiteit van die ontwerp vir die fondamenttoestande.
- Kontraktipe: risiko-oordrag met gepaardgaande kosteverskuiwing tussen kliënt en kontrakteur word bepaal deur die kontraktipe.
- Bron van gruis: aktiwiteit wat gepaardgaan met die gebruik van kommersiële bronne teenoor leengroewe word bepaal deur watter bron gekies word.
- Bron van vergruisde klip: aktiwiteit wat gepaardgaan met die gebruik van kommersiële bronne teenoor leengroewe word bepaal deur watter bron gekies word.
- Reënvalstreek: die intensiteitsverskille tussen winter- en somerreënvalstreke bepaal die tipe dreineringsstelsel en ook die verdragings wat verwag kan word.

Die model wat voorgestel word om die verwantskappe tussen die eerste vlak invloedfaktore wat die konstruksiekostekomponente voorstel en die tweede vlak

invloedfaktore wat die kwalitatiewe faktore beskryf, word in *Figuur 6-12* getoon. Weereens word al die teoreties moontlike invloede getoon, alhoewel sommige nie realisties is nie.



## 6.7 DETAILKOSTEMODEL INVLOEDFAKTORE

Om insig te ontwikkel in die invloedfaktore wat die eenheidtariewe van items beïnvloed, is dit nodig om onderskeid te maak tussen die verskillende perspektiewe wat die twee rolspelers, nl. die kontrakteur en die raadgewende ingenieur, van 'n konstruksieprojek ten opsigte van kosteberekeninge het.

### 6.7.1 KONTRAKTEURPERSPEKTIEF

Die inligting wat deur die kontrakteur gebruik word om sy getenderde eenheidtariewe te bepaal word vervolgens bespreek aan die hand van 'n generiese invloedsmodel en uitgebrei in daaropvolgende afdelings.

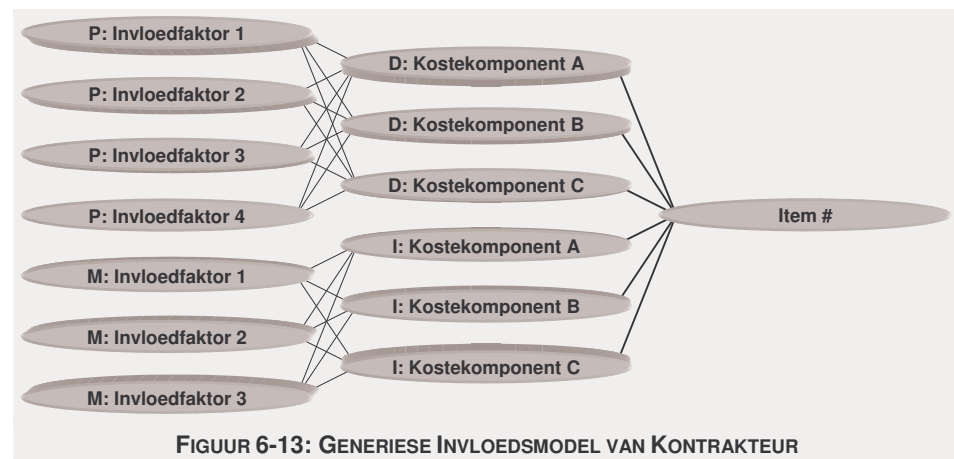
#### 6.7.1.1 GENERIESE INVLOEDSMODEL

Wanneer die kontrakteur tender vir 'n projek het hy spesifieke inligting beskikbaar oor die projek waarvoor getender word, asook kennis van sy maatskappy se finansiële toestand, konstruksie ondervinding, logistiese en strategiese raamwerk en die ekonomiese toestand waarbinne getender word. Die eenheidtarief van 'n item kan derhalwe beskou word as 'n samestelling van 'n verskeidenheid van kostekomponente, soos byvoorbeeld: arbeid; materiaal; toerusting; brandstof; algemene kostes ("over-heads"); wins en gebeurlikheidskoste. Hierdie kostekomponente word beïnvloed deur die kontrakteur se maatskappy- en projekstrategie.

Die maatskappystrategie word bepaal deur faktore soos: sterkte van kompetisie; die huidige ekonomiese marktoestand; die beskikbaarheid van

toekomstige werk; strategiese waarde; ondervinding; risiko's weens projekteienskappe; ens. [Fayek (1989)]. Die projekstrategie word bepaal deur: arbeidspanverdelings; produktiwiteit van arbeid; klimaat; projekligging; ens. [Bell (1987)].

'n Generiese invloedmodel word dus vir die kontrakteur voorgestel om onderskeid te maak tussen die item, die kostekomponente waaruit die item saamgestel is en die invloedfaktore wat die kostekomponente beïnvloed. Die invloedmodel kan diagrammaties soos in *Figuur 6-13* voorgestel word. Die kostekomponente in hierdie model van die kontrakteur kan gegroepeer word tussen direkte (materiaal, brandstof, toerusting en arbeid) en indirekte kostes (algemene, wins en gebeurlikheid). Die direkte kostes, D, is afkomstig vanaf die projek se uitvoering wat bepaal word deur projekinvloedfaktore, P, terwyl die indirekte kostes, I, bepaal word deur die maatskappy se strategiese invloedfaktore, M, weens die bestuursbeleid.



### 6.7.1.2 PROJEK

Die vier direkte kostekomponente, nl. arbeid, materiaal, toerusting en brandstof, word bepaal deur projekinvloedfaktore. Die faktore wat hierdie kostes kan beïnvloed is: volumes wat verskuif moet word; tipe materiaal wat verwerk moet word; aanwending van en tipe toerusting; werkspan strukturering; arbeidsproduktiwiteit; koste van materiaal; kompleksiteit; toerusting se funksionering; trekafstande; weersomstandighede; bewegbaarheid van toerusting op terrein; beskikbaarheid van water en die periode waarbinne konstruksie voltooi moet word [Easa (1988), Peurifoy (1989), Semple (1994)].

Sommige aspekte soos weersomstandighede het weer 'n indirekte effek op faktore soos arbeidsproduktiwiteit en funksionering van toerusting, terwyl ligging soos hoogte bo seevlak of onherbergsame terrein ook die funksionering van toerusting of die bewegbaarheid van toerusting beïnvloed.

Weens die presiese spesifisering van items wat in die hoeveelheidslys onder beskouing voorkom, is sommige van hierdie faktore egter nie van toepassing nie. Duidelike onderskeid word byvoorbeeld gemaak tussen die verskillende hardhede van materiaal wat verwerk moet word, so ook ten opsigte van bronne van materiaal, hetsy kommersieel of groewe, en trekafstande. Dit is dus nie nodig om al hierdie faktore as invloedfaktore te beskou wat die kostes van 'n item kan beïnvloed nie, omdat onderskeid al reeds gemaak word op grond van die item se spesifikasie.

---

### 6.7.1.3 STRATEGIES

---

Die drie indirekte kostekomponente, nl. algemene kostes ("over-heads"), wins en gebeurlikheidskoste, word bepaal deur die strategie van die konstruksie-maatskappy. Die strategiese invloedfaktore wat hierdie komponente beïnvloed kan intern aan die maatskappy wees of ekstern. Interne maatskappy faktore kan wees: behoefte aan werk, beskikbaarheid van maatskappy hulpbronne, ondervinding in 'n streek, beproef nuwe veld van konstruksie, strukturering van 'n tenderaanbod. Eksterne maatskappy faktore kan wees: sterk kompetisie, ekonomiese toestand van konstruksiebedryf, opbou van reputasie en projek verwante risiko's [Fayek (1989), Williams (1994)].

Dit kan verwag word dat feitlik al die faktore konsekwent op al die items toegepas sal word, omdat hierdie faktor ekstern tot die projek en afhanklik van die maatskappy se strategie is. Dit kan egter gebeur dat sommige items se gebeurlikheidskoste varieer weens risiko's wat afkomstig is weens die projek. Die winsopslag op sommige items kan ook wissel as die kontrakteur besef dat hoeveelhede verkeerd bereken is en dus deur 'n strukturering van die tender op sommige eenheidtariewe van items onder en op ander oor te kwoteer.

---

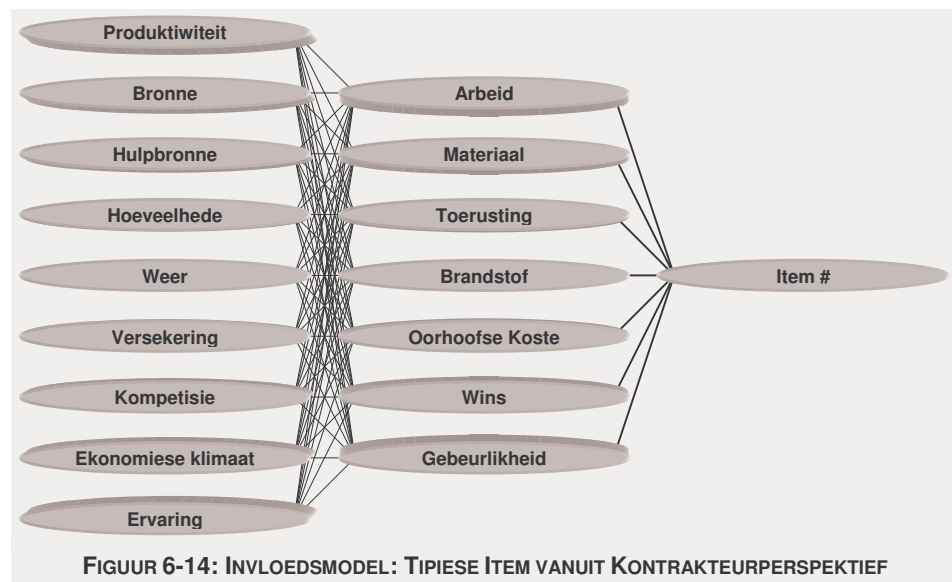
### 6.7.1.4 ALGEMENE INVLOEDSMODEL

---

Feitlik al die items in die spesifikasie is saamgestel uit die sewe bogenoemde kostekomponente, nl. arbeid, materiaal, toerusting, brandstof, algemene kostes, gebeurlikheidskoste en wins. Die faktore en die bydrae van elk van hierdie faktore wat die kostes van die kostekomponente bepaal wissel egter van item tot item. Slegs sommige van die projekfaktore is vir die raadgewende ingenieur bekend en feitlik geen van die strategiese faktore soos hierbo uiteengesit is beskikbaar vir die raadgewende ingenieur nie, derhalwe word die invloedsmodel vir die kontrakteur nie verder in diepte uitgebrei nie.

'n Veralgemeende diagrammatiese invloedsmodel vir die 18-items word voorgestel, soos in *Figuur 6-14*. Al die kostekomponente word getoon, maar slegs 'n paar van die faktore wat dit kan beïnvloed, asook die verwantskappe daartussen. Dit word vanuit die kontrakteurperspektief ontwikkel.

---



## 6.7.2 RAADGEWENDE INGENIEURPERSPEKTIEF

Die inligting wat deur die raadgewende ingenieur gebruik word om beraamde eenheidtariewe te bepaal word vervolgens bespreek aan die hand van 'n generiese invloedsmodel en uitgebrei in daaropvolgende afdelings.

### 6.7.2.1 GENERIESE INVLOEDSMODEL

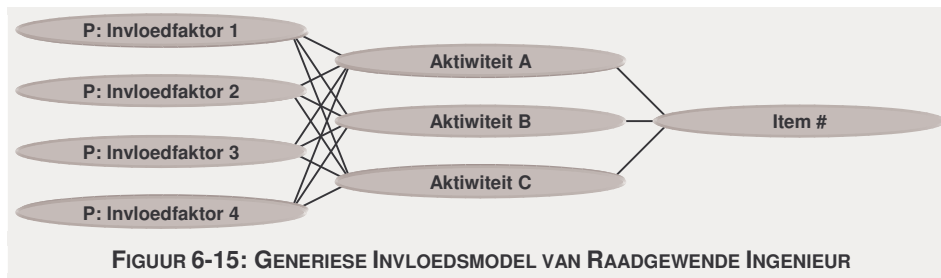
Die raadgewende ingenieur het net projekspesifieke inligting beskikbaar vir sy huidige projek, asook 'n verskeidenheid van historiese projekte waarop kostebereamings gebaseer kan word. Elk van hierdie historiese projekte is voltooi deur die uitvoering van 'n aantal take wat bekend is aan die raadgewende ingenieur, omdat hierdie take in gestandaardiseerde spesifikasies voorgeskryf word.

Eenheidtariewe van items kan derhalwe beskou word as 'n aantal aktiwiteite wat voltooi moet word, byvoorbeeld: voorsiening; plasing en verwerking van materiaal; ens. Hierdie aktiwiteit word egter beïnvloed deur projekspesifieke faktore soos die projekte, ligging en klimaat. Eenheidtariewe van vorige projekte moet dus deur die raadgewende ingenieur aangepas word om voorsiening te maak vir verskille wat voorkom tussen die huidige projek en ander historiese projekte.

Vir die raadgewende ingenieur word daar ook 'n generiese invloedsmodel voorgestel. Hierdie invloedsmodel bestaan ook uit die itemkomponent, maar verskil van die kontrakteur se invloedsmodel deurdat dit aktiwiteitskomponente het in plaas van 'n kostekomponent waaruit die itemkomponent se koste saamgestel is. Die invloedsdiagram vir die raadgewende ingenieur kan diagrammaties, soos in *Figuur 6-15*, voorgestel word. Die item se koste word bepaal deur die kostes van die aktiwiteit soos voorsiening, plasing en verwer-



king. Hierdie aktiwiteite se kostes word egter slegs bepaal deur projekinvloedsfaktore.



### 6.7.2.2 PROJEK

Die aktiwiteite wat die eenheidtarief by elke item bepaal is bekend aan die raadgewende ingenieur wanneer die kosteberamings uitgevoer word, omdat dit duidelik in die CSRA-spesifikasies gedefinieer word. By sommige items moet materiaal gestroop en verwyder word, terwyl by ander items die materiaal gestroop, verwyder, verwerk geplaas en gekompakteer word. Die aktiwiteite wissel dus tussen verskillende items en so dan ook die faktore wat dit kan beïnvloed. Hierdie projekfaktore is aan die ingenieur bekend vir die projek waarvoor die kosteberekening gedoen word. Dit kan gebruik word om historiese projekte se eenheidtariewe te rekonsilieer met sy projek indien die historiese projekte se projekfaktore ook bekend sou wees.

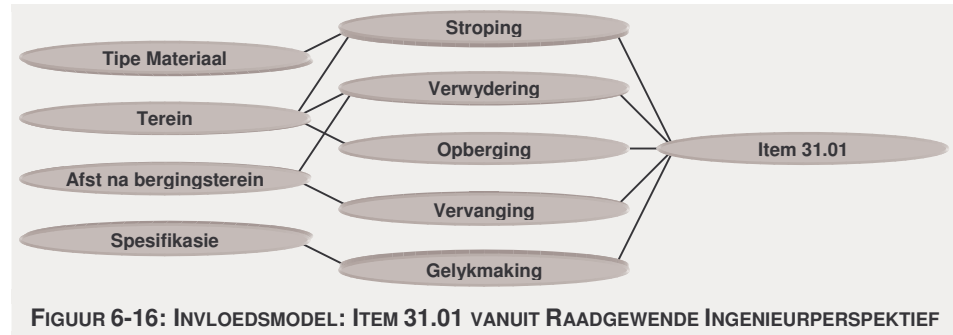
Die tipiese aktiwiteite wat voorkom kan in ses groepe gegroepeer word, nl. uitgrawing, vervoer, verwerking, kompaktering, toetsing en onderhoud van voltooide werke. Elke van hierdie groepe het tipiese faktore wat dit beïnvloed [Easa (1988), Gates (1980), Hegazy (1998), Peurifoy (1989)]:

- Uitgrawing: grondtoestande; ruimte beskikbaar op terrein, volume.
- Vervoer: toeganklikheid, volume.
- Verwerking: ruimte beskikbaar, klimaat, spesifikasies, water.
- Kompaktering: materiaal, ruimte beskikbaar, klimaat, spesifikasies, water.
- Toetsing: fasiliteit op terrein.
- Onderhoud: klimaat, verkeer.

### 6.7.2.3 INVLOEDSMODELLE

Die diagrammatiese invloedsmodelle is vir elk van die 17-items van die *Series 3000* wat in *Afdeling 5.1* bepaal is vanuit die raadgewende ingenieurperspektief ontwikkel. 'n Voorbeeld van een van die items, *Item 31.01* ("Verwydering van

oormaat bolaag per m<sup>3</sup>), se invloedsmoedelle word in *Figuur 6-16* getoon terwyl die res in *Aanhangsel P* getoon word.



## 6.8 INVLOEDFAKTORE

Met die invloedfaktore bekend is dit nodig om te identifiseer watter van die faktore 'n gereelde noemenswaardige impak het en dus onderskep moet word. Aangesien hierdie faktore in die kosteberekening gebruik moet word is dit ook nodig om te bepaal hoe en waar dit in die kostedatabasis gestoor gaan word, asook hoe die data oor die invloedfaktore aangewend kan word.

### 6.8.1 IDENTIFISERING

In hierdie afdeling word daar gekyk na hoekom slegs invloedfaktore van die raadgewende ingenieurperspektief onderskep moet word. Die leidrade wat die eenheidtariefkorrelasies bied word ondersoek en 'n metode word dan voorgestel hoe om die beduidendste invloedfaktore te identifiseer.

#### 6.8.1.1 RAADGEWENDE INGENIEURPERSPEKTIEF

Die tipe data wat binne die databasis gestoor moet word, om die eenheidtariewe van items beter te kwalifiseer, is die data wat die gebruiker van die databasis ook tot sy beskikking oor sy eie projek moet hê om die historiese projekte se eenheidtariewe met sy projek te konsilieer. Databasisse wat eenheidtariewe van historiese projekte stoor word grotendeels deur raadgewende ingenieurs gebruik vir kosteberekening. Dit is dus belangrik om in ag te neem watter invloedfaktore se data die raadgewende ingenieur tot sy beskikking het wanneer kosteberekening gedoen word.

Vir kosteberekening wat deur 'n raadgewende ingenieur uitgevoer word, sal die interne strategiese maatskappy faktore vanuit die kontrakteurperspektief nooit bekend wees nie, terwyl die eksterne strategiese maatskappy faktore wel tot 'n mate deur die raadgewende ingenieur gekwantifiseer kan word as daar byvoorbeeld na tendervolume in voorafgaande jare gekyk word [Herbsman (1983)]. Die faktore wat die projekkostekomponente vanuit die kontrakteurperspektief beïnvloed is ook nie bekend aan die raadgewende ingenieur nie. Tog is

daar aspekte soos klimaat, seisoen en terreintoestand, wat wel bekend is aan die raadgewende ingenieur, omdat dit deel vorm van die raadgewende ingenieurperspektief. Die invloedfaktore wat voorkom in die invloedsmoedelle van die raadgewende ingenieurperspektief is dus die enigste inligting wat aan hom bekend sal wees en waarmee die projekkostekomponent van die kontrakteurperspektief beskryf kan word. Dit is dan ook die tipe projekinligting wat behoort onderskep en gestoor te word in databasisse van historiese projekte.

Die onderskepping van al die faktore vir elke item in 'n projek blyk op die oog af omslagtig en onprakties te wees in terme van die hoeveelheid werk wat gedoen moet word. As aangeneem word 'n projek kan tot soveel as 600-items hê en elke item tot 10-faktore moet 6000 faktore onderskep word. Die impak van faktore is ook nie noodwendig dieselfde vir elke projek nie. Dit is dus nodig om te identifiseer watter invloedfaktor die grootste impak het sodat 'n meer gefokusde benadering gevolg kan word ten opsigte van die onderskepping van historiese projekinvloedfaktore vir alle projekte.

---

### 6.8.1.2 KORRELASIE-LEIDRAAD

---

Dit is reeds in *Afdeling 5.5* bespreek dat daar noemenswaardige sterk korrelasies voorkom tussen die verskillende eenheidtariewe van items. *Ang (1975)* wys egter daarop dat hierdie verwantskap nie beteken dat daar 'n kousale effek tussen twee veranderlikes bestaan nie (een veranderlike is nie afhanklike van die ander nie), maar dat die twee veranderlikes (of dan eenheidtariewe) afhanklik kan wees van 'n ander veranderlike of veranderlikes.

As daar na die korrelasie,  $c_{XY}$ , tussen twee ewekansige veranderlikes,  $X$  en  $Y$ , gekyk word dan lyk die verband soos volg:

$$c_{XY} = \frac{\sum_{i=1}^n (x_i - \mu_X)(y_i - \mu_Y)}{n}$$

waar  
 $n$  = aantal waarnemings ..... *Vgl. 6-1*  
 $x_i, y_i$  = ewekansige veranderlikes in  $X, Y$   
 $\mu_X, \mu_Y$  = gemiddelde van bevolking  $X, Y$   
 $X, Y$  = bevolkings met  $n$  elemente

Dit is duidelik dat hoe groter die verskil tussen  $x_i$  en die gemiddelde van die populasie,  $\mu_X$ , is hoe groter die produk sal wees en dienooreenkomstig vir  $y_i$  en  $Y$ . As die twee gesamentlik in dieselfde rigting afwyk (albei verskille negatief of positief), dan sal die produk 'n noemenswaardige bydrae tot die som van die produkte maak. As dit vir al die waarnemings sou plaasvind dan sal die som van die produkte relatief groot wees, wat 'n sterk positiewe korrelasie aandui; verskille in die teken van die onderskeie afwykings sal lei tot sterk negatiewe korrelasie as die afwykings groot is.

---

Hierdie gelyktydige afwyking vanaf die gemiddelde, as daar nie 'n kousale effek tussen die twee populasies bestaan nie, impliseer dat die afwyking veroorsaak word deur een of ander eksterne bron of invloed. Die invloeddiagramme wat hierbo bespreek is, dui juis hierdie gedeelde bronne aan wat die afwykings kan veroorsaak. Om te bepaal watter invloedfaktore 'n noemenswaardige impak het op die eenheidtariewe van 'n projek en dus behoort onderskep te word, kan daar gekyk word na die sterkte van korrelasie wat die invloedfaktore veroorsaak.

---

### 6.8.1.3 IDENTIFISERINGSMETODE

---

Die sterkte waarmee 'n invloedfaktor die koste van 'n aktiwiteit van 'n item beïnvloed, bepaal hoe sterk die verband tussen hierdie twee veranderlikes is. Indien die verband tussen die twee veranderlikes swak is, sal 'n afwyking vanaf die gemiddelde van die invloedfaktor min kere gereflekteer word deur 'n soortgelyke afwyking van die item se eenheidtarief. 'n Sterk verband tussen die twee veranderlikes bestaan egter as afwykings vanaf die gemiddeld van die invloedfaktor in 'n soortgelyke afwyking van die eenheidtarief gereflekteer word.

Invloedfaktore beïnvloed meer as een item se aktiwiteitskoste. Indien twee of meer items se aktiwiteit sterk deur dieselfde invloedfaktor beïnvloed word, sal die items se eenheidtariewe dieselfde afwykings van die invloedfaktor reflekteer. Die korrelasie tussen die twee eenheidtariewe sal dus sterk wees. Die omgekeerde geld ook; indien slegs die een item se aktiwiteitskoste deur die invloedfaktor beïnvloed word, maar die ander items word nie, dan sal die eenheidtariewe van die items nie dieselfde afwykings toon nie. Die korrelasie tussen die items se eenheidtariewe is dus swak.

Aangesien 'n item nie net uit een aktiwiteit bestaan met een invloedfaktor nie, maar 'n verskeidenheid van aktiwiteite elk met gedeelde invloedfaktore is die bepaling van die beduidendste invloedfaktor nie 'n triviale oefening nie. 'n Tegniek word dus benodig om die beduidendste invloedfaktore wat onderskep moet word te identifiseer. In die volgende afdeling is so 'n tegniek ontwikkel. Tydens die (her)ontwikkeling van 'n databasis hoef hierdie oefening net eenmalig uitgevoer te word om die invloedfaktore te vind. Die metode kan egter ook voordelig gebruik word by elke nuwe projek om insig te ontwikkel in die korrelasiemeganismes wat bestaan tussen die eenheidtariewe van die projek en watter invloedfaktore die sterkste invloed kan uitoefen op die projekkosies.

---

### 6.8.2 BEPALING VAN BEDUIDENDSTE INVLOEDFAKTORE

---

Die metode word aan die hand van nege stappe verduidelik. Uittreksels van die sigblad wat ontwikkel is om hierdie metode te implementeer, word hier getoon. In *Aanhangsel Q* kan 'n meer volledige bespreking daarvan gevind word.

---

- *Stap 1* – Identifisering van beduidende items: Die items wat 'n relatief groot bydrae op 'n gereelde basis tot die totale koste van projekte lewer moet geïdentifiseer word. Die metode wat in *Afdeling 5.1* beskryf word om beduidende items te identifiseer kan netso gebruik word, maar dit moet effens uitgebrei word om 'n groter aantal beduidende items te identifiseer as die 44-items wat vir die ontledings in *Hoofstuk 5* gebruik is. Die versameling van items kan vergroot word deur die kriteria aan te pas, soos in *Figuur 5-3* getoon word. Verkieslik moet die bydrae kriteria van 5% verlaag word na 1% in plaas van die voorkoms kriteria van 1% te verlaag, om te verseker dat items wat gereeld voorkom ingesluit word eerder as items wat net 'n groot impak het en nie gereeld voorkom nie.
- *Stap 2* – Identifisering van invloedfaktore en aktiwiteite: Die invloedfaktore wat elk van die aktiwiteite van die beduidende items kan beïnvloed moet bepaal word. Die invloedfaktore kan bepaal word volgens die metode wat in *Afdeling 6.7.2* verduidelik word, deur die aktiwiteite van elke item te identifiseer en dan die invloedfaktore van hierdie aktiwiteite te bepaal. Die aktiwiteite waaruit elke item bestaan kan direk vanuit die CSRA-spesifikasie verkry word soos in *Afdeling 6.7.2.1* genoem word. Elke aktiwiteit se invloedfaktore moet omvattend bepaal word, selfs al is die invloedfaktor se bydrae gering moet dit ingesluit word. Dieselfde invloedfaktore gaan egter altyd dieselfde aktiwiteite beïnvloed, derhalwe is dit slegs nodig om eenmalig die invloedfaktore van elke aktiwiteit te bepaal.
- *Stap 3* – Beoordeling van aktiwiteite se itemkostebydrae: Met elke item se aktiwiteite bekend is dit nodig om te bepaal wat die bydrae van elke aktiwiteit tot die koste van die item is. Die ideaal is dat hierdie kostebydraes kwantitatief bepaal word van inligting wat vanaf die kontrakteurs verkry word. Indien dit egter nie moontlik is nie kan dit ook kwalitatief bepaal word. Hiervoor kan van die 4-punt kostebydrae skaal, wat in *Tabel 6-1* getoon word, gebruik gemaak word.

**TABEL 6-1: KOSTEBYDRAE SKAAL**

Bydrae Skaal	Definisie
0	Geen
1	Klein
2	Gemiddeld
3	Groot

Die kwalitatiewe metode werk soos volg. Aan elke aktiwiteit word daar 'n punt op die kostebydrae skaal toegeken wat die kostebydrae van die aktiwiteit tot die item se koste reflekteer. Die sigblad verwerk hierdie punte en bepaal die relatiewe bydrae van elke aktiwiteit. In *Figuur 6-17* word 'n denkbeeldige voorbeeld getoon waar daar 12-items is waarvoor daar

elk die kostebydrae van 10-aktiwiteite bepaal moet word. In die figuur stel die groen area die geskatte bydrae van elke aktiwiteit voor terwyl die bruin area die berekende relatiewe bydraes toon.

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10
E1	3	0	3	0	0	0	0	0	0	0	50%	0%	50%	0%	0%	0%	0%	0%	0%	0%
E2	0	2	3	2	2	0	1	1	0	0	0%	18%	27%	18%	18%	0%	9%	9%	0%	0%
E3	3	1	3	1	0	2	1	0	2	1	21%	7%	21%	7%	0%	14%	7%	0%	14%	7%
E4	3	0	2	0	0	0	0	0	0	0	60%	0%	40%	0%	0%	0%	0%	0%	0%	0%
E5	1	2	0	3	3	1	1	2	2	3	6%	11%	0%	17%	17%	6%	6%	11%	11%	17%
E6	0	3	1	1	3	3	1	0	1	2	0%	20%	7%	7%	20%	20%	7%	0%	7%	13%
E7	2	3	2	2	1	1	2	0	2	3	11%	17%	11%	11%	6%	6%	11%	0%	11%	17%
E8	0	1	2	3	0	0	1	0	3	3	0%	8%	15%	23%	0%	0%	8%	0%	23%	23%
E9	3	0	1	2	0	0	0	0	0	2	38%	0%	13%	25%	0%	0%	0%	0%	0%	25%
E10	2	2	0	0	1	0	0	3	2	2	17%	17%	0%	0%	8%	0%	0%	25%	17%	17%
E11	0	0	3	2	1	0	3	1	1	2	0%	0%	23%	15%	8%	0%	23%	8%	8%	15%
E12	2	3	2	0	1	1	1	2	3	2	12%	18%	12%	0%	6%	6%	6%	12%	18%	12%

FIGUUR 6-17: BEPALING VAN AKTIWITEITE SE KOSTEBYDRAE PER ITEM

- *Stap 4* – Beoordeling van invloedfaktorsterkte op aktiwiteite: Die sterkte van die invloed van elke aktiwiteit se invloedfaktore op die aktiwiteit moet bepaal word. In die geval van korrelasies tussen eenheidtariewe, *Afdeling 5.5*, is dit gevind dat daar slegs positiewe korrelasie voorkom. Dit mag dus voorkom asof daar geen omgekeerde verbande bestaan nie, alhoewel dit moontlik mag wees by individuele invloedfaktore. Dit is dus nodig om onderskeid te maak tussen invloedfaktore wat 'n positiewe en 'n omgekeerde verband met die aktiwiteite het. Die 7-punt skaal wat in *Tabel 6-2* getoon word kan hiervoor gebruik word.

TABEL 6-2: INVLOEDFAKTORSTERKTE SKAAL

Sterkte Skaal	Definisie
-3	Sterk negatief
-2	Gemiddeld negatief
-1	Swak negatief
0	Geen invloed
1	Swak positief
2	Gemiddeld positief
3	Sterk positief

In *Figuur 6-18* word die geskatte sterkte van 15 denkbeeldige invloedfaktore op die 10 aktiwiteite, waarvan die bydrae tot die onderskeie items se kostes in die vorige stap bepaal is, getoon.

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15
A1	0	2	1	0	3	-1	0	0	1	1	1	3	2	1	2
A2	2	0	3	0	3	-1	3	0	3	0	0	0	2	1	2
A3	0	1	2	0	3	-1	2	3	1	1	2	3	3	3	2
A4	0	2	1	2	3	-2	3	0	1	2	3	1	3	0	1
A5	1	0	3	1	1	0	1	2	1	1	0	1	2	2	1
A6	1	0	0	3	2	-1	1	3	2	1	3	2	3	3	0
A7	2	3	0	0	0	-3	1	3	0	0	1	1	3	1	1
A8	1	1	0	3	2	-2	0	3	1	0	0	3	2	2	2
A9	2	0	2	2	1	-1	1	0	2	1	0	2	1	0	2
A10	1	3	3	2	1	0	1	1	1	2	3	0	3	3	0

FIGUUR 6-18: GESKATTE INVLOEDFAKTORSTERKTE OP AKTIWITEITE

- *Stap 5* – Berekening van invloedfaktorsterkte op items: Die sterkte van die invloed van elke invloedfaktor op die items moet nou bepaal word. Dit kan gedoen word deur die matriksproduk tussen die relatiewe aktiwiteit-bydrae matriks, *Figuur 6-17*, en die geskatte invloedfaktorsterkte matriks, *Figuur 6-18*, te bepaal. In *Figuur 6-19* word die resultaat van die berekening getoon.

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15
E1	0.0	0.5	2.5	0.0	3.0	0.0	1.0	1.5	1.0	0.0	2.5	1.5	1.0	2.0	2.0
E2	1.0	0.1	0.9	1.6	1.1	-0.9	0.9	0.6	1.1	1.7	0.3	0.4	2.0	0.3	0.6
E3	0.6	0.3	1.5	0.7	2.6	-0.4	0.9	0.9	1.3	0.3	1.8	1.3	0.9	1.2	1.6
E4	0.0	0.4	2.6	0.0	3.0	0.2	0.8	1.2	1.2	0.0	2.4	1.8	1.2	1.8	2.0
E5	0.8	0.0	0.3	1.3	1.3	-0.6	0.8	0.5	1.3	1.5	0.9	0.6	1.4	0.1	0.5
E6	1.1	0.1	0.7	1.1	1.7	-0.5	1.7	0.6	1.3	1.4	0.8	0.7	1.6	1.0	0.7
E7	0.9	0.0	0.9	0.9	1.7	-0.3	1.3	0.0	1.9	0.9	1.9	0.9	1.4	0.3	0.6
E8	1.1	0.2	0.4	1.4	1.9	-1.0	1.2	0.5	1.4	0.8	1.4	0.5	0.8	0.5	0.9
E9	0.7	0.2	1.8	0.0	2.7	0.3	1.3	0.5	2.0	0.0	2.5	1.5	1.0	1.0	1.3
E10	0.8	0.0	1.2	0.2	2.0	0.4	1.2	0.4	2.0	0.6	2.0	1.2	1.2	0.4	0.8
E11	1.2	0.3	0.5	0.9	1.4	-0.6	0.9	0.8	1.5	0.7	2.0	1.1	0.6	0.7	0.6
E12	0.3	0.3	1.7	0.2	2.0	0.0	0.7	1.3	1.2	0.5	2.2	1.5	1.0	1.3	1.3

FIGUUR 6-19: INVLOEDFAKTORSTERKTE OP ITEMS

- *Stap 6* – Korrektheid van item invloedfaktorsterkte: Om te bepaal of die invloedfaktorsterkte korrek beraam is, kan die sterkte van verband tussen die verskillende items, gebaseer op die item invloedfaktorsterkte, *Figuur 6-19*, bereken word en vergelyk word met die werklike korrelasies wat tussen die items voorkom. Die liniêre sterkte van verband,  $s$ , tussen twee items  $X$  en  $Y$ , gebaseer op die invloedfaktorsterkte, kan as volg gevind word:

$$s_{XY} = \sum_{i=1}^n \frac{T}{3} \sqrt{ABS(x_i/\#_X \times y_i/\#_Y)}$$

waar

- $x_i, y_i$  = invloedfaktorsterkte van item  $X, Y$
- $\#_X, \#_Y$  = aantal invloedsfaktor van item  $X, Y$  ..... Vgl. 6-2
- $T$  =  $\begin{cases} 1, & \text{as } x_i \times y_i \geq 0 \\ -1, & \text{as } x_i \times y_i < 0 \end{cases}$
- $n$  = aantal faktore

In *Figuur 6-20* word die resultate van die sterkte-van-verband (tussen eenheidtariewe van items) berekeninge onder die diagonaal van die matriks getoon. Die werklike korrelasie van die historiese eenheidtariewe,

wat met die metodes in *Afdeling 5.5* bepaal is, kan bo die diagonaal van hierdie matriks gevind word.

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12
E1	1.0	0.8	0.7	0.5	0.6	0.6	0.5	0.6	0.7	0.6	0.6	0.6
E2	0.5	1.0	0.7	0.6	0.5	0.5	0.4	0.2	0.6	0.5	0.5	0.5
E3	0.5	0.5	1.0	0.4	0.7	0.6	0.6	0.3	0.5	0.4	0.4	0.4
E4	0.6	0.5	0.5	1.0	0.4	0.4	0.5	0.3	0.5	0.4	0.4	0.4
E5	0.5	0.5	0.5	0.5	1.0	0.7	0.5	0.4	0.6	0.5	0.5	0.5
E6	0.5	0.5	0.5	0.5	0.5	1.0	0.7	0.6	0.8	0.8	0.8	0.8
E7	0.5	0.5	0.5	0.5	0.5	0.5	1.0	0.5	0.6	0.7	0.7	0.7
E8	0.5	0.5	0.5	0.5	0.5	0.5	0.5	1.0	0.8	0.3	0.4	0.8
E9	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	1.0	0.7	0.8	0.2
E10	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	1.0	0.4	0.8
E11	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	1.0	0.2
E12	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	1.0

**FIGUUR 6-20: BEREKENDE STERKTE-VAN-VERBAND**

- *Stap 7 – Afwyking vanaf korrelasies:* Die verskil tussen die sterkte-van-verband berekeninge en die korrelasie wat bepaal is, kan nou gevind word, *Figuur 6-21*. Indien 'n verskil van kleiner as 0.1 tussen die werklike en voorspelde korrelasies gevind word kan daar oorgegaan word na *Stap 9*, andersins moet na *Stap 8* gegaan word.

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12
E1	0.0											
E2	0.3	0.0										
E3	0.2	0.2	0.0									
E4	0.1	0.1	0.0	0.0								
E5	0.1	0.1	0.2	0.0	0.0							
E6	0.2	0.0	0.1	0.0	0.2	0.0						
E7	0.0	0.0	0.1	0.0	0.0	0.2	0.0					
E8	0.1	0.3	0.2	0.1	0.1	0.1	0.0	0.0				
E9	0.2	0.1	0.0	0.0	0.1	0.3	0.1	0.3	0.0			
E10	0.1	0.0	0.0	0.0	0.0	0.4	0.2	0.2	0.2	0.0		
E11	0.1	0.0	0.1	0.1	0.0	0.3	0.2	0.1	0.3	0.1	0.0	
E12	0.1	0.0	0.1	0.0	0.0	0.3	0.2	0.3	0.3	0.3	0.3	0.0

**FIGUUR 6-21: VERSKIL TUSSEN BEREKENDE VERBANDSTERKTE EN KORRELASIE**

- *Stap 8 – Kostebydrae en Invloedfaktorsterkte aanpassing:* Indien dit gevind word dat daar 'n noemenswaardige verskil tussen die sterkte van verband en die korrelasies is kan die invoere tot die berekeninge, nl. kostebydraes van aktiwiteite en die invloedfaktorsterktes op aktiwiteite, *Figuur 6-17* en *Figuur 6-18*, iteratief aangepas word.
- *Stap 9 – Bepaling van beduidendste invloedfaktore:* Die beduidendste invloedfaktore kan bepaal word deur die individuele impakte daarvan te bepaal op die totale gemiddelde koste van die beduidende items en dan te orden volgens grootste impak. Dit kan gedoen word deur die gemiddelde koste (produk van gemiddelde eenheidstarief en hoeveelheid) van die beduidende items wat in *Stap 1* geïdentifiseer is, te bepaal. Elke invloedfaktor se impak kan gevind word deur die somproduk van die iteminvloedfaktorsterktes van die invloedfaktor onder beskouing, *Figuur 6-19*, en die ooreenstemmende beduidende itemkoste te bepaal, *Figuur 6-22*.



	Itemkoste		F1		Produk
E1	R 2,967.00	X	0.0	=	R -
E2	R 7,106.00		1.0		R 7,106.00
E3	R 6,030.00		0.6		R 3,618.00
E4	R 5,224.00		0.0		R -
E5	R 3,298.00		0.8		R 2,748.33
E6	R 8,378.00		1.1		R 9,495.07
E7	R 1,900.00		0.9		R 1,628.57
E8	R 7,443.00		1.1		R 8,119.64
E9	R 5,593.00		0.7		R 3,728.67
E10	R 2,314.00		0.8		R 1,851.20
E11	R 9,318.00		1.2		R 10,751.54
E12	R 1,237.00		0.3		R 412.33
			Somproduk =	R 49,459.35	

FIGUUR 6-22: SOMPRODUK VAN ITEMKOSTES EN INVLOEDFAKTORSTERKTE

Die somprodukte van elke invloedfaktor kan nou genormaliseer word om die relatiewe bydrae van elk te bepaal vir 'n gemiddelde projek, *Figuur 6-23*.

F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15
5%	1%	7%	6%	13%	3%	7%	5%	9%	5%	11%	7%	8%	6%	7%

FIGUUR 6-23: RELATIEWE IMPAKTE VAN INVLOEDFAKTORE

Met die relatiewe impak van elke invloedfaktor op 'n gemiddelde projek bekend, kan die beduidendste invloedfaktore nou gekies word vir onderskepping. As daar byvoorbeeld besluit word om die top agt te kies vir die denkbeeldige voorbeeld dan sal die keuse op invloedfaktor F5, F11, F9, F13, F3, F7, F12 en F15 val.

### 6.8.3 NUWE PROJEKTOEPASSING

Die tweede handige toepassing van hierdie metode is by nuwe projekte se detailkosteberekening voor die tenderfase. Wanneer die raadgewende ingenieur die kosteberekening voltooi het en die koste (produk van hoeveelheid en eenheidtarief) van elke item is bekend, kan bepaal word watter invloedfaktore die grootste impakte kan hê op die totale koste van die projek.

Deur slegs *Stap 9* in die vorige afdeling te herhaal, maar vir hierdie geval die nuwe projek se itemkoste te gebruik kan die rangorde gevind word van die beduidende invloedfaktor met die grootste impak op die spesifieke nuwe projek. Hierdie insig kan van onskatbare nut wees om te verseker dat die invloedfaktore wat variasie kan veroorsaak wel in ag geneem is en korrek gekwantifiseer is tydens die beraming van die eenheidtariewe.

### 6.8.4 VOORSTELLING EN UITBREIDING VAN DATABASISMODEL

Die verwantskappe wat bestaan tussen 'n item, aktiwiteite en invloedfaktore kan voorgestel word met versamelingsteorie. Die item  $i$ , is 'n element in die versameling van items  $I$ , in die standaardspesifikasie. Elke item bestaan uit 'n aantal aktiwiteite. Die aktiwiteit  $a$ , is 'n element in die versameling aktiwiteite  $A$ . Elke

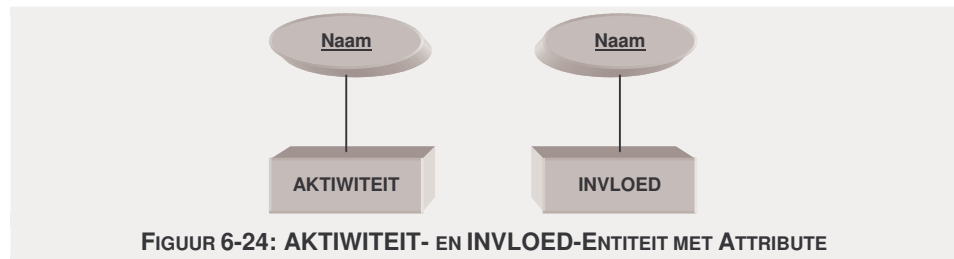
element in die versameling aktiwiteite  $A$ , word beïnvloed deur 'n aantal invloedfaktore. Die invloedfaktor  $f$ , is 'n element van die versameling invloedfaktore  $F$ .

Die relasie  $IA$ , tussen die versameling items en aktiwiteite, en die relasie  $AF$  tussen die versameling aktiwiteite en invloedfaktore kan dan as volg gedefinieer word:

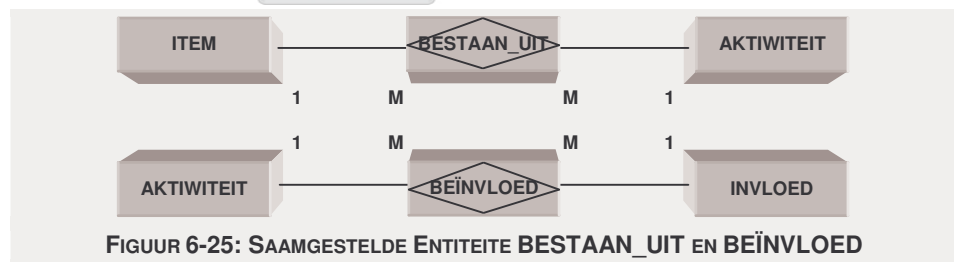
$$IA = \{(i, a) | i \in I, a \in A \text{ bestaan uit}\}$$

$$AF = \{(a, f) | a \in A, f \in F \text{ beïnvloed deur}\} \dots\dots\dots \text{Vgl. 6-3}$$

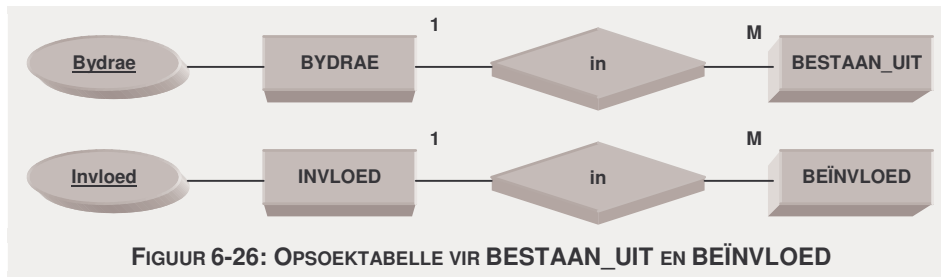
Die databasismodel kan maklik uitgebrei word om hierdie relasies te stoor. Twee nuwe tabelle (entiteite) word benodig om die versamelings van aktiwiteite en invloedfaktore te stoor (die ITEM-tabel bestaan reeds). In *Figuur 6-27* word die twee nuwe entiteite, nl. AKTIWITEIT en INVLOED, met hul attribute in 'n EV-diagram, *Afdeling 3.3.3*, getoon.



Die relasies, *Vgl. 6-2*, wat bestaan tussen hierdie entiteite kan met die twee samegestelde entiteite, nl. BESTAAN\_UIT en BEÏNVLOED, voorgestel word. Hierdie entiteite met hul verwantskappe word in *Figuur 6-25* getoon.



Die kostebydrae (*Stap 3*) en die invloedfaktorsterkte (*Stap 4*) kan onderskeidelik by die samegestelde relasie BESTAAN\_UIT en BEÏNVLOED gestoor word. Om egter te verseker dat hierdie gewigte nie besoedel word met onrealistiese waardes nie, word twee opsoektabelle, soos die in *Afdeling 3.5.1.4*, bygevoeg. Die twee opsoektabelle word in *Figuur 6-26* met die EV-diagram voorgestel.

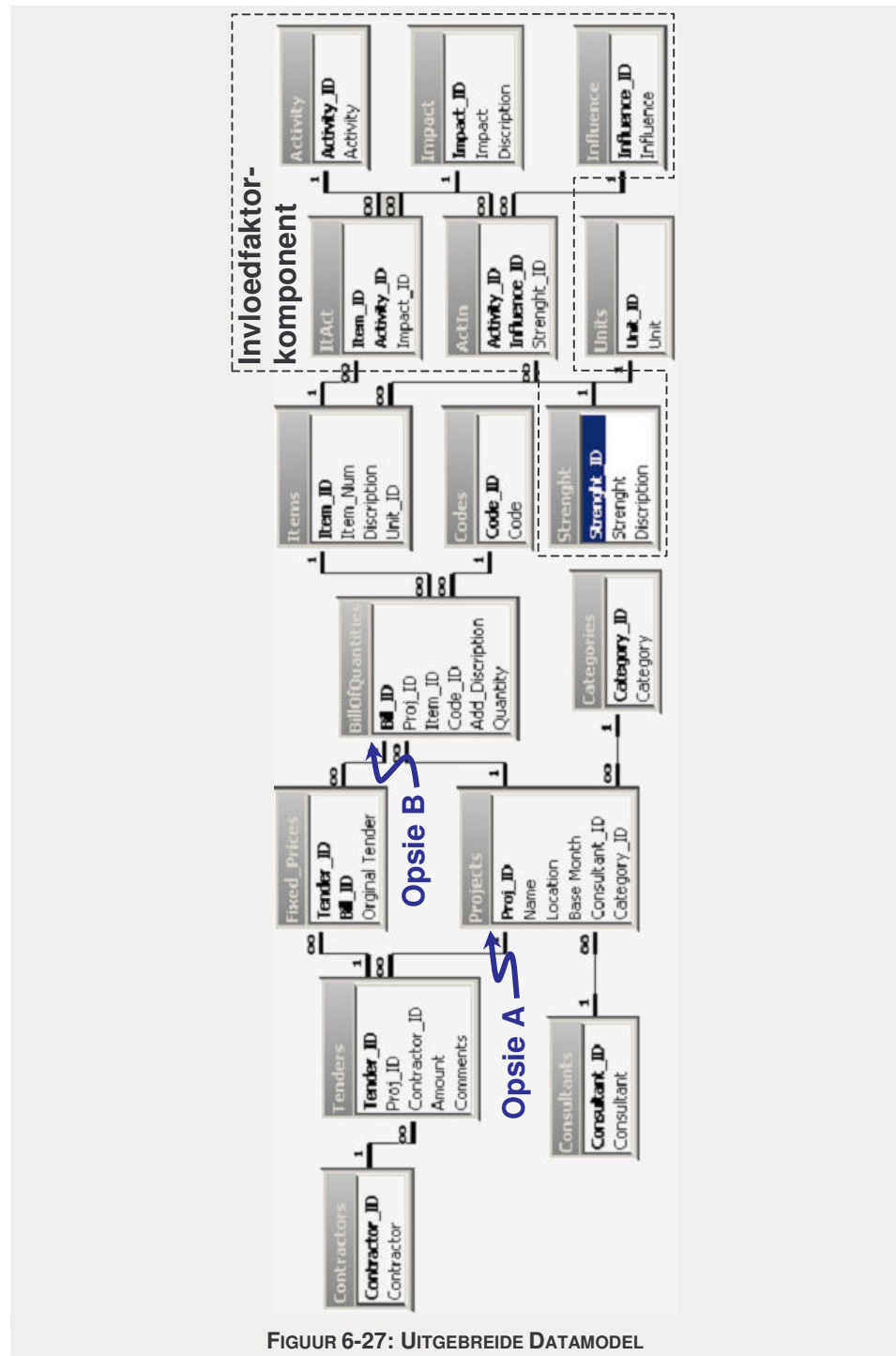


Hierdie entiteite is almal triviaal in BCNV, *Afdeling 3.3.4*, derhalwe word daar nie 'n formele toets uitgevoer om die onderskeie normaalvorms te toets nie. Die uitgebreide datamodel wat geïmplementeer is, *Figuur 6-27*, is gebaseer op die prototipe databasis, *Figuur 3-32*. Hierdie nuwe uitbreiding is onafhanklik van die uitbreidings wat in *Figuur 3-34* getoon word en kan met of sonder dit geïmplementeer word.

As die beduidende faktore geïdentifiseer is volgens die metode wat in die vorige afdeling bespreek is, is dit nodig dat die invloedfaktore se waardes gestoor moet word. Twee moontlikhede bestaan hoe om die projekfaktore in te sluit nl., dit kan gekoppel word aan die projek-tabel of dit kan gekoppel word aan die hoeveelheidslys-tabel, soos getoon in *Figuur 6-27*.

Die voordeel van eersgenoemde is dat veralgemeende projekfaktordata hier geplaas kan word en dat min data versamel hoef te word. Die nadeel is egter dat indien daar verskillende projekfaktore voorkom by verskillende items, daar nie onderskeid gemaak sal kan word tussen watter items deur die ooreenkomstige projekfaktore beïnvloed word nie. Byvoorbeeld indien daar wisselende terreintoestande voorkom tussen twee items in dieselfde projek sou slegs 'n enkele veralgemeende waarde as terreinfaktor gestoor kon word. Indien daar egter besluit word om die projekfaktor te stoor in die hoeveelheidslys-tabel, sou dit moontlik wees om nie net vir die dominante items invloedfaktore te stoor nie, maar vir al die items in elke projek se hoeveelheidslys. Dit sal egter heelwat addisionele versameling van inligting vereis.

'n Kombinasie van die twee moontlikhede behoort eerder oorweeg te word, sodat projekfaktore wat klein wissellings toon in die projek-tabel geplaas word, terwyl faktore wat groot wisselings tussen items toon in die hoeveelheidslys-tabel geplaas moet word.



FIGUUR 6-27: UITGEBREIDE DATAMODEL

### 6.8.5 AANWENDING

Indien hierdie projektfaktore versamel sou word, sal dit moontlik raak om gevorderde analise tegnieke soos neurale netwerke te begin gebruik om voorspellings van eenheidtariewe van items te maak. As voorbeeld word 'n neurale netwerk getoon wat in *MS Excel* geïmplementeer is. Die voorbeeld en data is afkomstig uit 'n joernaalartikel deur *Hegazy (1998)*.

TABEL 6–3: INVOERDATA VIR NEURALE NETWERK

Data	Projek Tipe	Projek Bestek	Jaar	Seisoen	Ligging	Lengte	Afstand	Kapasiteit	Water	Grond Toestand	Werklike Koste
Proj 1	3	3	1991	2	3	1	4.75	1	0	7	0.151
Proj 2	1	1	1990	2	3	9	0.10	1	1	7	0.335
Proj 3	2	2	1991	2	3	8	7.90	1	0	8	0.579
Proj 4	2	3	1991	3	3	12	1.00	1	1	9	0.376
Proj 5	2	2	1990	2	2	10	0.40	1	0	8	0.339
Proj 6	2	2	1991	2	3	9	6.40	1	0	8	0.612
Proj 7	2	2	1990	2	3	19	7.90	1	0	8	1.080
Proj 8	2	2	1990	2	2	20	0.87	1	0	8	0.673
Proj 9	2	2	1991	2	3	3	11.20	1	1	8	0.678
Proj 10	2	3	1990	2	3	4	1.00	1	0	4	0.129
Proj 11	2	2	1990	1	3	22	13.50	1	1	8	1.210
Proj 12	1	1	1991	2	3	17	0.10	1	1	5	0.220
Proj 13	2	2	1990	2	3	16	40.00	2	1	8	1.738
Proj 14	2	2	1991	2	3	20	0.90	1	0	8	0.637
Proj 15	2	2	1991	2	3	12	7.55	1	0	0	0.608
Proj 16	2	2	1989	1	3	36	11.60	1	1	0	0.717
Proj 17	2	2	1990	2	2	12	9.00	1	0	8	0.959
Proj 18	2	2	1991	2	3	11	8.40	1	0	8	0.710

Die neurale netwerk is gebaseer op die voorwaartsvoerende multigelaagde netwerke wat in *Afdeling 4.6* bespreek word. As invoernodusse word 10 projek kriteria gebruik en een afvoernodus wat die koste van die projek is. Die invoer data word getoon in *Tabel 6–3*.

TABEL 6–4: AFVOER VAN NEURALE NETWERK

Data	NN Berekende	Werklike Koste	% Fout
Proj 1	0.15	0.15	0.00
Proj 2	0.34	0.34	0.50
Proj 3	0.58	0.58	0.50
Proj 4	0.37	0.38	0.50
Proj 5	0.34	0.34	0.50
Proj 6	0.61	0.61	0.50
Proj 7	1.09	1.08	0.50
Proj 8	0.67	0.67	0.50
Proj 9	0.68	0.68	0.50
Proj 10	0.13	0.13	0.04
Proj 11	1.22	1.21	0.50
Proj 12	0.22	0.22	0.01
Proj 13	1.74	1.74	0.04
Proj 14	0.64	0.64	0.50
Proj 15	0.61	0.61	0.00
Proj 16	0.71	0.72	1.00
Proj 17	0.96	0.96	0.00
Proj 18	0.70	0.71	1.00

Nadat die neurale netwerk “geleer” is op die data kan daar nou met behulp daarvan voorspellings gemaak word. Die resultate word in *Tabel 6–4* getoon. Netso sal dit ook moontlik word om die data wat afkomstig is van projekfaktore wat onderskep word te analiseer en voorspellings daarmee te maak vir 'n gegewe bestaande projek.

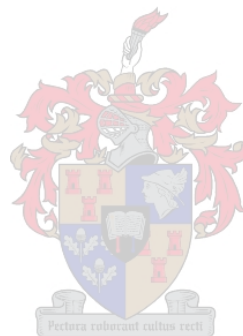
Die vraag ontstaan egter hoe die onsekerheid in terme van die akkuraatheid wat met hierdie tipe analise gepaardgaan realisties gemodelleer kan word. Dit is alreeds in *Afdeling 4.6* genoem dat die akkuraatheid waarmee die model geleer word as moontlike aanduiding kan dien. Die onsekerheid waarmee elke eenheidtariewe voorspel word kan dan in die LSKF-model gebruik word om die globale onsekerheid verbonde aan die beramings te bepaal.

Die onsekerheid wat egter bepaal word – gebaseer op die akkuraatheid waarmee die model geleer is – neem egter nie die onsekerheid in ag van die foute wat gemaak kan word met die kwantifisering van die invloedfaktore van die nuwe projek deur die raadgevende ingenieurs nie. Die neurale netwerk

model kan met ander woorde 'n voorspelling maak van die eenheidstaref van 'n item gegewe die invloedfaktore waarmee dit gevoer is en dit sal 'n antwoord lewer met die onsekerheid wat gepaardgaan met die akkuraatheid waarmee dit geleer het, maar hierdie fout-waarde neem nie in ag dat die raadgewers verkeerde invloedfaktore kon ingevoer het nie.

Hierdie “oorkoepelende” onsekerheid kan bepaal word deur die beraamde waardes van die eenheidstarewe van die raadgewende ingenieurs en die werklike getenderde waardes met mekaar te vergelyk. Die verskille wat tussen hierdie waardes gekry word, is die werklike fout wat gemaak word tydens voorspellings en behoort in die probabilistiese modelle gebruik te word.

---



## HOOFSTUK 7

---

### ONTWIKKELING VAN 'N NUWE DATABRON

---

In die voorafgaande hoofstukke is verskillende aspekte bespreek wat gevind is gebruik kan word in die verdere inkrementele ontwikkeling van 'n bestaande padkonstruksie kostedatabasis om die inligting daarin beter te kan benut vir besluitnemingsdoeleindes. Die metodes wat gevolg is, is egter nie beperk tot die padkonstruksie omgewing en die gegewe vlak van ontwikkeling van 'n bestaande databasis nie, maar is generiese metodes wat in enige ander konstruksiekosteberekening omgewing gebruik kan word.

In hierdie hoofstuk word die omgewing van damkonstruksie gebruik om te toon hoe die metode op 'n nuwe omgewing toegepas kan word. Daar word vlugtig gekyk na die verskille tussen die twee omgewings en die behoeftes wat kan bestaan om eenheidtariefberamings te verbeter. Die metodes wat gebruik word om hierdie behoeftes aan te spreek word bespreek en daar word dan getoon hoe om die leemtes wat by nuwe omgewings kan bestaan ten opsigte van standaardspesifikasies te akkommodeer.

---

#### 7.1 ONTWIKKELINGSVLAK

---

Die oorwegende verskil tussen die twee omgewings is die vlak van ontwikkeling van 'n sentrale elektroniese bron van koste-inligting weens die verskil in die struktuur van die tipiese kliënt-ingenieur-kontrakteur verhouding.

---

##### 7.1.1 STAPSGEWYSE ONTWIKKELING

---

Die CSRA-databasis wat tot dusver in hierdie studie gebruik is, is die uitvloeisel van omstandighede wat gelei het tot die ontwikkeling van standaardspesifika-

---

sies vir padkonstruksie in Suid-Afrika. Hierdie databasis wat gebaseer is op die CSRA-spesifikasies, verteenwoordig al 'n gevorderde vlak van ontwikkeling in die stelselmatige ontwikkelingsproses van so 'n databasis vir besluitsteundoeleindes. Die stapsgewyse ontwikkeling van standaardspesifikasie het 'n groot rol hierin gespeel.

Voor 1987 is drie verskillende standaardspesifikasies vir pad- en brugwerke deur die destydse vyf padbou instansies in Suid-Afrika gebruik. Die Nasionale Vervoerkommissie en die Paaiedepartemente van Transvaal en Oranje-Vrystaat het al drie dieselfde spesifikasie gebruik, terwyl die Paaiedepartemente van die Kaapprovinsie en Natal elk hul eie spesifikasies gebruik het. Die SABS spesifikasies wat as 'n oorkoepelende dokument moes dien was te lomp om van enige praktiese nut te wees. Weens die verwarring wat die drie verskillende dokumente onder raadgewende ingenieurs en kontrakteurs veroorsaak het, is daar deur die Komitee van Staatspadowerhede besluit dat 'n standaardspesifikasie opgestel moes word wat landswyd gebruik kon word. Die produk wat deur die aangestelde komitee in 1987 gelewer is, was die CSRA-Standaardspesifikasies vir Pad- en Brugwerke (CSRA-spesifikasies), wat grotendeels gebaseer was op die Nasionale Vervoerkommissie se Standaardspesifikasies. [Nel (1987)]

Met die totstandkoming van die CSRA-spesifikasies op nasionale vlak, was daar een standaardspesifikasie wat landswyd gebruik is. Hierdie spesifikasies moes deur al die ontwerpers gebruik word, nadat die ontwerp voltooi is, om die konstruksiemetodes te spesifiseer en die hoeveelhedslyste hiervolgens op te stel. Getenderde eenheidtariewe van dieselfde item in verskillende projekte het dus landswyd presies dieselfde aktiwiteite ingesluit. Hierdie standaardisering het noodwendig die geleentheid gebied om sentraal eenheidtariewe van alle padbouprojekte te stoor en sodoende tot die ontstaan van die CSRA-databasis gelei. Eenheidtariewe van die standaarditems is sedert 1987 in hierdie databasis onderskep van projekte oor die hele land.

Weens verskille in die struktuur van die omgewings, *Afdeling 7.1.2*, het daar nog nie standaardspesifikasies in die damkonstruksie omgewing ontwikkel wat veralgemeen genoeg is om vir alle tipe damkonstruksieprojekte in Suid-Afrika gebruik te kan word nie. Alhoewel daar deur die jare spesifikasies ontwikkel is, soos ook plaasgevind het in die eerste fases van ontwikkeling van die padkonstruksie spesifikasies, is hierdie spesifikasie op 'n *ad hoc* wyse aangewend by damkonstruksieprojekte en aangepas na gelang van die behoeftes by elke projek [Kroon (2003)]. Items wat herhaaldelik voorkom wat uniek gedefinieer is, is dus moeilik om te identifiseer.



'n Databasis van historiese eenheidtariewe van standaarditems bestaan dus nie, omdat die ontwikkeling van so 'n databasis bemoelik word deur die voorkoms van oorwegend projekspesifieke items. Kosteberekening word derhalwe deur ervare personeel behartig met die nodige ondervinding en kennis oor die historiese projekte waarop die beraamde eenheidtariewe gebaseer word. Die nodige aanpassings tussen die eenheidtariewe van verskillende projekte word gebaseer op die kosteberamer se persepsie en vermoë om die faktore wat die kostes by 'n projek kan beïnvloed te identifiseer en te kwantifiseer.

---

### 7.1.2 STRUKTUUR VAN OMGEWING

---

In die padkonstruksie voorbeeld was daar vyf verskillende padbou instansies, een op nasionale vlak en die ander vier op provinsiale vlak, onder die koepel van die Komitee van Staatspadowerhede wat veranderwoordelik was vir die ontwikkeling van die land se padvervoerinfrastruktuur. Hierdie instansies het opgetree as die kliënt ('n baie ervare kliënt), terwyl van die dienste van raadgewende ingenieurs en kontrakteurs gebruik gemaak is vir die ontwerp en bou van paaie. Die probleme wat dit teweeg gebring het vir die raadgewende ingenieurs en kontrakteurs weens die verskillende spesifikasies van die verskillende padbou instansies het 'n behoefte laat ontstaan vir gestandaardiseerde spesifikasies. Die standaardspesifikasie vir paaie en brûe het sodoende tot stand gekom om hierdie probleme aan te spreek, soos reeds bespreek is.

Die infrastruktuur om water op te gaar en te versprei in Suid-Afrika, is tradisioneel deur die Departement van Waterwese en Bosbou (DWB) en ander staatsdepartemente wat dit vooraf gegaan het, op nasionale vlak ontwikkel en onderhou. Die beplanning, ontwerp en konstruksie van staatswaterskemas is feitlik in geheel deur hierdie departement behartig, omdat dit so 'n gespesialiseerde tegniese gebied is. Die kliënt, raadgewende ingenieur en kontrakteur was dus binne dieselfde liggaam met die gepaardgaande voordele van 'n tipiese ontwerp-bou kontrak [Gordon (2001)].

Een van die grootste voordele van die ontwerp-bou kontrakte is dat die volume van tegniese dokumentasie (konstruksietekeninge en spesifikasies) aansienlik beperk word, omdat daar nie 'n tenderproses vir die konstruksiefase bestaan waarvoor die eindproduk tot op 'n detailvlak gespesifiseer moet word nie. Sterk gedefinieerde standaardspesifikasies het dus nooit ontwikkel nie, omdat die probleme wat in die padkonstruksie omgewing voorgekom het nie binne hierdie omgewing kon voorkom nie of dan ten minste in 'n aansienlike mindere mate voorgekom het. Die afwesigheid van sterk gedefinieerde standaardspesifikasie met geassosieerde items het sodoende veroorsaak dat dit nie moontlik was om 'n standaard hoeveelheidlys vir alle damkonstruksieprojekte op te stel nie. 'n Databasis vir onderskepping van eenheidtariewe van items sou

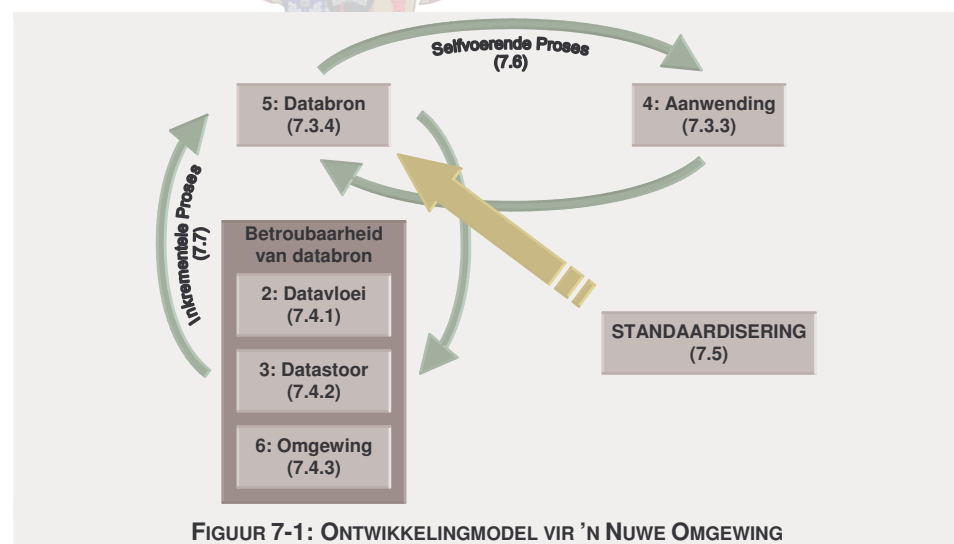
---

dis nie van veel nut gewees het nie, omdat items se spesifikasie kon wissel van projek tot projek. Verder meer was die konstruksieafdeling nog altyd gerieflik naby gewees om eenheidtariewe vanaf hulpbronne vlak (arbeid, toerusting, brandstof en materiaal) te bepaal.

Hierdie tradisionele struktuur van die departement is egter besig om te verander en dit is moontlik dat in die afsienbare toekoms (binne die volgende vyf jaar) daar beweeg sal word na 'n struktuur wat soortgelyk is aan die wat tans in die padkonstruksie omgewing gevind word. Dit sal heelwaarskynlik teweeg bring dat die probleme wat ervaar is deur die raadgewende ingenieurs en die kontrakteurs in die padbou omgewing weens 'n leemte in standaardspesifikasies en -items ook in hierdie omgewing sal voorkom.

## 7.2 ONTWIKKELINGSTRUKTUUR

Die verdere besprekings in hierdie hoofstuk is oorwegend gebaseer op die aspekte wat elk individueel in die voorafgaande hoofstukke in diepte gedefinieer en ontwikkel is. Hier word dit nou saamgevoeg en toegepas in die dambou omgewing om 'n inligtingstelsel vir risiko-gebaseerde kosteberekeninge op 'n geïntegreerde wyse te ontwikkel. Die res van die hoofstuk is gestruktureer rondom die model wat in *Figuur 7-1* getoon word. In die figuur word die verband met die afdelings in hierdie hoofstuk en die ooreenstemmende voorafgaande hoofstukke getoon waar dit van toepassing is.



## 7.3 BEHOEFTES

In die padkonstruksie omgewing is die tipiese behoeftes wat voorkom ten opsigte van konstruksieprojekte geïdentifiseer. Hier word dit veralgemeen en dan van toepassing op damkonstruksie gemaak.

### **7.3.1 PROJEKTIPES**

---

Siviele ingenieursprojekte kan tipies in vier generiese groepe verdeel word, nl. nuwe projekte, opgraderings/verbeterings, rehabilitasie/herstelwerk en instandhoudingwerk. Elk van hierdie projektipies bevat unieke aktiwiteite met gepaardgaande onsekerhede. Tydens kosteberekeninge kan en behoort hierdie unieke onsekerhede vasgevang te word om die risiko verbonde aan die projekte te kwantifiseer. In die damkonstruksie omgewing kan hierdie generiese projektipies ook geïdentifiseer word.

Eersgenoemde groep van projekte is voor die handliggend die konstruksie van nuwe opgaar-, vloedabsorpsie- en uitkeerdamme (stuwalle). Hierdie projektipies het unieke onsekerhede ten opsigte van die ondersteunende infrastruktuur, rivierverleggings en onbekende geologiese toestande. Hierteenoor is die opgradering of verbetering van bestaande infrastruktuur soos die vergroting van stoor- of vloeditlaatkapasiteit of verbeterings aan die uitlaatwerke, ens. voorbeelde van die tweede groep van projektipies met die unieke onsekerhede wat hiermee geassosieer kan word. Kwaliteit van bestaande boumateriaal, beperkte geometriese inligting, ens. is in hierdie geval tipiese bronne van onsekerheid.

Rehabilitasie- of herstelwerkprojekte word weer genoodsaak deur onveilige toestande wat aan die lig kom tydens damveiligheids-evaluasies of wanneer abnormale toestande skade aan van die komponente van die damstruktuur veroorsaak. Weens die diversiteit van rehabilitasiemoontlikhede waartussen gekies kan word om damveiligheidsprobleme aan te spreek is dit nodig om in ag te neem dat daar groter onsekerhede by sommige rehabilitasietegnieke is as by ander ten opsigte van die werklike kostes [Andersen (2001)]. Hierteenoor is die gewone instandhoudingswerk wat by feitlik enige dam kan voorkom en wissel van die elementêre verwydering van plantmateriaal tot die herboor van dreinerings- en drukverligtinggate. Hierdie projektype word weer gekenmerk deur baie min risikobronne.

Kosteberekeninge by verskillende projektipies behoort hierdie verskillende risiko's wat geassosieer word met die projektype in ag te neem. 'n Databron wat ontwikkel word moet dus voorsiening maak om sulke verskille in onsekerheid te kan reflekteer in die inligting wat dit lewer.

---

### **7.3.2 VLAK VAN DETAIL EN VOLLEDIGHEID**

---

Die drie generiese vlakke van kosteberekeninge, nl. langtermyn, konseptueel en detailkosteberekeninge, wat in konstruksieprojekte voorkom is in *Afdeling 4.3* bespreek en is ook geïdentifiseer in die padkonstruksie omgewing,

---

*Afdeling 6.4.* Hierdie vlakke van kosteberekeninge kan ook netso maklik vir die verskillende projektipos in die damkonstruksie omgewing geïdentifiseer word.

Gepaardgaande met die verskillende vlakke van kosteberekeninge is die akkuraathede hiervan soos in *Afdeling 4.3* bespreek is. Hierdie akkuraatheid, gebaseer op historiese projekte, is slegs gemiddelde waardes van afwykings wat by projekte verwag kan word. Dit is egter in die padkonstruksie omgewing gevind dat die onsekerhede wat voorkom by die verskillende projektipos van mekaar verskil vir die redes soos reeds bespreek in *Afdeling 6.4.1* en *6.6.2*. vir langtermyn en konseptuele kosteberekeninge en dat dit ook veroorsaak dat by detailkosteberekeninge nie altyd die korrekte hoeveelhede van items ingesluit word nie. Gegewe die verskille in risikobronne wat kan voorkom by die verskillende projektipos in damkonstruksie kan soortgelyke verskille ook in hierdie omgewing verwag word.

Die enigste manier om hierdie onsekerheidsverskille te kan kwantifiseer is om historiese detailkosteberekeninge as basis te gebruik vir die bepaling van die hoër vlakke van kosteberekeninge. Dit kan wel geargumenteer word dat hoër vlakke van kosteberekeninge gebaseer kan word op slegs 'n lessenaarstudie volumeberekening van 'n klein seleksie van die grootste kostebronne en dat die res van die kostes as 'n persentasie hiervan bereken kan word, m.a.w. gebaseer op die 80-20-reël. Dit impliseer egter dat die kosteberekening omvattend genoeg gedoen word om die grootste bydrae kostes in te sluit en die onsekerhede wat kan voorkom tussen werklike en beraamde hoeveelhede ook ingesluit is, soos dit kan wissel van projektype tot projektype.

Ervare ingenieurs behoort met redelike vertroue te kan voorspel wat die beïndrukkendste kostes vir 'n tipiese projek is. Onsekerhede gekoppel aan voorspellings is egter moeiliker om te bepaal, want in meeste gevalle word daar nie historiese inligting op 'n gestruktureerde wyse versamel om voorspellings wat gemaak is te toets en die ingenieur se beraamings te kalibreer nie. Die grootste probleem lê egter by die beskikbaarheid van ervare ingenieurs om kostebereamings te maak. Weens die hoë omset van ingenieurs wat die industrie op die oomblik kenmerk, is die bron van ervare ingenieurs in spesialisiteitsrigtings besig om te krimp.

---

### 7.3.3 KWANTIFISERING VAN RISIKO

---

Deterministiese ontwerpstandaarde was tradisioneel gebruik vir ontwerp doeleindes, maar weens verskeie redes is dit besig om te verander. Veranderende damveiligheidsstandaarde en stygende kostes van inkrementele veiligheidsverbeteringe het 'n behoefte laat ontstaan vir alternatiewe analise tegnieke [Hartford (2000)]. Verder meer wil die gemeenskappe wat geïmpak word deur besluite groter deursigtigheid hê, omdat daar 'n besef is dat daar in enige

---

konstruksieprojek onsekerhede bestaan [McDonald (2000)]. Risiko-analise bied die raamwerk vir 'n sistematiese toepassing van ingenieuroordeel en die gebruikmaking van statistiek van data [Høeg (1998)].

Pad- en damkonstruksie is albei kapitaalintensiewe projektipes wat tradisioneel deur die staat gefinansier is, maar nou meer en meer deur die privaatsektor gefinansier word. Tolpaaie wat met BOT (“Build Operate Transfer”) kontrakte gebou word is tipiese voorbeelde in die padkonstruksie omgewing. Voorbeelde van nuwe damkonstruksies en verhogings van bestaande staatsdamme wat deur die privaatsektor gefinansier word, word ook meer gereeld aangetref in sogenaamde “Public Private Partnership” ooreenkomste tussen die staat en die privaatsektor. Hierdie nuwe kliënte is meer ingestel op geoptimeerde ontwerpe waarby alle risiko's, hetsy ontwerp, kosteberaming, instandhouding en die omgewingsimpak in ag geneem word.

Netsoos dit onvoldoende is om deterministiese standaarde te gebruik in ontwerpe, netso is dit ook onvoldoende om deterministiese kosteberekeninge te maak indien onsekerhede kan voorkom in eenheidtariewe en volumes van materiaal. Die kwantitatiewe risiko-analise soos die LSKF- en MCS-tegniek, wat in *Hoofstuk 4* vir padkonstruksie getoon is, kan gebruik word om nie net die deterministiese kostes te bepaal nie, maar ook om die risiko's verbonde aan die kosteberamings vir verskillende alternatiewe in damkonstruksieprojekte te bepaal.

In die multi-kriteria probabilistiese besluitnemingsmodel wat deur *Oosthuizen (1985)* vir die prioritering van aktiwiteite in die damveiligheidsprogram vir staatsdamme ontwikkel is (en ook veralgemeen is om vir nuwe damkonstruksieprojekte gebruik te kan word), is die konstruksiekoste een van die invoere tot die model. Die invloedmodel van hierdie multi-kriteria probabilistiese besluitnemingsmodel is soortgelyk aan die invloedmodel wat vir padkonstruksie getoon is in *Figuur 6-3*. Die voor-en-nadele en die kostes verbonde aan die verskillende alternatiewe, saam met die voorkeure van die besluitnemers wat in die utiliteitsfunksie geoptimeer moet word, word binne sekere beperkings oorweeg. Historiese data om onsekerhede ten opsigte van konstruksiekoste mee te kwantifiseer is die voorkeur bron van inligting.

---

### **7.3.4 BETROUBARE BRON**

---

Goeie risikobestuur benodig goeie risiko-inligting – informasie wat deur risiko-analise geskep is. Risiko-analise is egter meer as slegs die skatting van risiko's, want dit is altyd moontlik om skattings te maak, maar geen hoeveelheid Monte Carlo-simulasie gebaseer op empiriese statistiese beskrywings kan opmaak vir ontoereikende teorie, eksperimentele data, empiriese bewys en gesonde oordeel nie. Prosedures wat robuustheid en naspeuring verseker word benodig,

---

omdat dit nie meer voldoende is om besluite en oordele te lewer sonder enige verduidelikings nie [Hartford (2000)]. Deur dus bloot data te ontleed en regressie tegnieke toe te pas om statistiese beskrywings van invoer parameters tot probabilistiese modelle te kry, is onvoldoende.

In die geval van kosteberekening data is dit nodig om te verseker dat daar 'n betroubare databron beskikbaar is en dat die data deeglike ontleed word om verklarings te vind vir afwykings. Tydens die ontleding van die padkonstruksie data is dit gevind dat 'n groot deel van die veranderlikheid van eenheidtariefdata epistemies van aard is. Die aspekte wat aangespreek behoort te word om epistemiese veranderlikheid by eenheidtariewe van padkonstruksie data te verlaag, is reeds in die voorafgaande hoofstukke bespreek. Hierdie aspekte is netso van toepassing by die damkonstruksie omgewing.

---

## 7.4 ONTWIKKELINGASPEKTE

---

Tydens die herontwikkeling van 'n betroubare databron is daar in die padkonstruksie omgewing gevind dat drie aspekte, nl. lae datavloei, ontoereikende databasismodelle en swak karakterisering van die omgewing, aanleiding kan gee tot epistemiese veranderlikheid. By die ontwikkeling van 'n nuwe databron, soos vir damkonstruksie, behoort hierdie aspekte ook in ag geneem te word.

---

### 7.4.1 DATAVLOEI

---

Die vloei van data ten opsigte van 'n inligtingstelsel is reeds in *Afdeling 1.3* bespreek. Datavloei verwys na die oordrag van data tussen verskillende aktiwiteite wat tydens die leeftyd van 'n projek uitgevoer word. In die padkonstruksie omgewing was daar 'n reeds bestaande datavloei gewees. Hierdie datavloei was egter nie sonder probleme nie. Van die probleme wat geïdentifiseer was, was die herhaaldelike heronderskepping van feitlik dieselfde data, stadige tempo waarteen inligting na die databasis gevloei het en die omslagtige datavloei. Hierdie probleme het veroorsaak dat daar foute in die data ontstaan het, data verouderd geraak het en dat nie al die data wat geskep is na die databasis gevloei het nie.

By die ontwikkeling van 'n nuwe stelsel kan prosesmodellering wat in *Hoofstuk 2* bespreek is gebruik word om hierdie probleme aan te spreek en te voorkom. Die doel hiervan is om op 'n logies geordende manier die interaksie van mense, aktiwiteite, data en toerusting op 'n oorhoofse wyse te verstaan, sodat die datavloei waarin belanggestel word geanaliseer kan word. Die ontleding van die datavloei en indien nodig, die verbetering daarvan is noodsaaklik om te verseker dat die data wat geïdentifiseer word in die later afdelings, wel onderskepbaar is op 'n effektiewe wyse sodat die produktiwiteit, kwaliteit en kwantiteit verbeter kan word, indien nodig.

---

Die damkonstruksie omgewing verskil van die padkonstruksie omgewing ten opsigte van die kompleksiteit daarvan. Die padkonstruksie omgewing hou oorwegend verband met tipiese siviele aktiwiteite wat die meeste hulpbronne tydens konstruksie gebruik. Damkonstruksie se kompleksiteit word veroorsaak deur uiteenlopende ingenieursaktiwiteite wat verbandhou met instrumentering, meganiese en elektriese toerusting en natuurlik konstruksie. Weens die verskillende tipes kostedata wat benodig word vir hierdie aktiwiteite, is daar nie net 'n enkele bron van waar hierdie kostedata in die DWB onderskep kan word nie. Hierdie inligting per projek is versprei oor verskillende afdelings elk met verskillende metodes hoe om die kostes te bereken en te stoor, hetsy in digitale of hardekopie formaat.

'n Duidelike prosesmodel van 'n tipiese damprojek word dus benodig om aan te toon waar orals kosteberekeninge uitgevoer word en wat die insette tot hierdie kosteberekeninge binne die DWB is. Indien so 'n model beskikbaar is wat oorhoofs die verskillende aktiwiteite beskryf kan daar bepaal word watter inligting beskikbaar is, in watter formaat hierdie inligting bestaan en hoe dit kan afgevoer word na 'n kostedatabasis. Tydens die ontwikkeling van so prosesmodel kan die *Generiese Ontwerp en Konstruksie Prosesprotokol* wat in *Afdeling 2.1.4* geïdentifiseer is as basis gebruik word, saam met die prosesmodel wat vir die padkonstruksie omgewing opgestel is, *Aanhangsel A*, om 'n prosesmodel vir die damkonstruksie omgewing soos dit tans lyk, op te stel.

Die komplekse datavloei wat hier kan voorkom sal noodsaak dat die model ontleed word vir volledigheid en korrektheid. Die tegnieke wat in *Hoofstuk 2* bespreek is om te toets vir volledigheid en korrektheid kan met groot sukses in so 'n omgewing aangewend word om te bepaal of die prosesmodel korrek opgestel is.

---

## 7.4.2 DATABASIS

---

Die funksie van 'n databasis ten opsigte van 'n inligtingstelsel is reeds in *Afdeling 1.3* bespreek. Die databasis se funksie is om die historiese data op 'n konsekwente en effektiewe wyse digitaal te stoor, sodat korrekte data beskikbaar is vir latere toepassings. Tydens die ontleding van die CSRA-databasis is enkele probleme geïdentifiseer met die bestaande databasis. Hierdie probleme het te make met die ontwerp van die datamodel vir die stoor van eenheidtariewe en onvoldoende fout korrigerende wat kan lei tot verhoogde veranderlikheid in data. Hierteenoor is gevind dat omvattend gedefinieerde standaarditems een van die redes is hoekom die databasis suksesvol kon funksioneer.

Die probleme wat met die datamodel geïdentifiseer is, word veroorsaak deur die onnodige duplisering van dieselfde data in die databasis. Nie net het dit onnodige omslagtige invoere na die databasis veroorsaak nie, maar het dit ook

---

data-inkonsekwentheid veroorsaak. Alhoewel dit gevind is dat hierdie probleem nie 'n noemenswaardige bydrae tot verhoogde veranderlikheid gelewer het nie, maar wel aanleiding kon gee tot verkeerde soekresultate in die databasis, behoort hierdie aspek deeglik aangespreek te word by die ontwikkeling van 'n damkonstruksie kostedatabasis.

Foutidentifisering wat veronderstel is om te verseker dat data wat in die databasis geplaas word nie foutief is ten opsigte van die brondata nie, is wel gedeeltelik in die CSRA-databasis geïmplementeer. Dit was egter nie omvattend genoeg gedoen om te verseker dat alle foute voorkom is nie. Die aard van die berekeninge waarvoor die data van die databasis gebruik was, was egter nie so sensitief vir hierdie foute gewees nie. Met die verbeterde ontledingstegnieke wat egter moontlik raak en waarvoor die damkonstruksie databasis se data gebruik gaan word, kan hierdie foute 'n noemenswaardige impak op die resultate hê, *Afdeling 4.5.3*. Omvattende foutidentifisering moet dus ingebou word in die prosesse om hierdie probleem aan te spreek.

Die basiese datamodel wat ontwikkel is vir die padkonstruksie omgewing kan feitlik netso gebruik word vir die ontwikkeling van die kostedatabasis vir die damkonstruksie omgewing. Die vyf relasies (tabelle) wat geïdentifiseer is in *Hoofstuk 3*, nl. PROJEK, ITEM, L\_V\_H, TENDER en GET\_EP, wat onderskeidelik die projekinligting, standaarditems, die lys van hoeveelhede, die totale getenderde bedrag en die getenderde eenheidpryse stoor, is die absolute minimum tabelle wat gebruik kan word om die eenheidtariefdata van getenderde projekte te stoor. Hierdie tabelle is so ten opsigte van mekaar uitgelê om slegs beheerde duplisering in die databasis toe te laat en sodoende te verseker dat die integriteit van die data behoue bly tydens opdaterings.

Inkonsekwentheid kan ontstaan indien daar nie voorsiening gemaak word om invoere na die databasis op 'n geordende wyse te laat plaasvind nie. Hierdie probleem is tydens die herontwikkeling van die CSRA-databasis aangespreek en behoort ook tydens die ontwikkeling van die damkonstruksie kostedatabasis in ag geneem te word. Die byvoeging van vyf verdere opsoektabelle by die basiese model was voldoende gewees om hierdie probleem op te los en kan ook netso gebruik word by die damkonstruksie databasis. Die model wat dus vir die damkonstruksie omgewing gebruik kan word om die probleme te voorkom wat in die CSRA-databasis gevind is lyk soos die model wat voorgestel word in *Figuur 3-31*.

Die voorgestelde CSRA-databasismodel is egter vir 'n tipiese eenheidtarief projek opgestel waar daar streng gedefinieerde items voorkom en al die inligting hieraan verbonde afkomstig is van slegs twee verskillende tipes dokument, nl. die tenderverslag en die projekvoltooiingsverslag. Min uitsonderings kom dus



voor. Vir hierdie nuwe toepassing sal die databasis verskillende tipes inligting moet stoor vanaf verskillende bronne waar die items nie so gestandaardiseer is soos by die CSRA-databasis nie. Hierdie aspek word verder in *Afdeling 7.5* bespreek.

---

### 7.4.3 KARAKTERISERING VAN OMGEWING

---

Die aanbevelings in die voorafgaande twee afdelings kan bydra tot 'n verlaging in die veranderlikheid van die data. Dit het egter geblyk dat daar nog 'n derde bron is wat aanleiding kan gee tot veranderlikheid. Hierdie derde bron het sy oorsprong in die tekort aan beskrywende inligting wat kan verklaar hoekom daar verskille tussen eenheidartiewe van projekte voorkom, *Hoofstuk 6*. Dit is gevind dat meer omvattende projekinligting kan help om hierdie probleem aan te spreek. Dit is dus nodig om die onderskepping van data uit te brei om sodoende te poog om hierdie veranderlikheid beter te verklaar.

Om egter vir elk van die verskillende tipe kostes al die moontlik invloedfaktore te onderskep wat die veranderlikheid beter kan verklaar kan lei tot die onderskepping van faktore wat nie 'n substansiële bydrae tot die verlaging van die globale veranderlikheid van die kosteberekening van die projek maak nie. Die faktore wat dus gekies word, moet die faktore wees wat die omvattendste die projekomgewing in terme van die totale projekkostes karakteriseer.

Aangesien dit egter nie bekend is watter faktore die grootste bydrae gaan lewer nie en daar ook in meeste gevalle nie historiese data beskikbaar is om dit te kan bepaal nie, is dit nodig om van die ervaring van die kosteberamers gebruik te maak. Die metode wat in *Afdeling 6.8* bespreek is om beraamde gewigte van die kosteberamers aan die verskillende invloedfaktore toe te ken, kan gebruik word. Indien dit vir al die moontlike geïdentifiseerde kosteversamings van al die verskillende tipes koste gedoen word kan daar oorhoofs bepaal word watter invloedfaktore die grootste moontlike impak het op die totale kostes van damprojekte. Hierdie invloedfaktor se eienskappe moet vir nuwe projekte onderskep word.

Die onderskepping van invloedfaktore noodsaak egter dat daar standaarditems met standaardspesifikasies is, sodat dit duidelik is watter aktiwiteite ingesluit word in die kostes van 'n gegewe item. Indien die aktiwiteite wat ingesluit word by 'n item verskil van projek tot projek sal dit veroorsaak dat die faktore wat die items beïnvloed ook wissel van projek tot projek. Netsoos in die geval van die ontwikkeling van die databasis vir damkonstruksie in die vorige afdeling, is dit dus nodig om eers standaarditems te identifiseer voordat daar na die onderskepping van invloedfaktore gekyk kan word. Die invloedfaktore bied egter die moontlikheid om die rigiditeit wat nodig is by standaarditems effens te verslap soos in *Afdeling 7.5* bespreek word.

---

## 7.5 GEBREK AAN STANDAARDITEMS

Bogenoemde generiese riglyne kan gebruik word om die kwaliteit van historiese kostedata te verbeter. 'n Metode moet egter gevind word om die leemte wat die tekort aan standaarditems veroorsaak aan te spreek anders sal die verbeteringe van min waarde wees om epistemiese veranderlikheid te verlaag by detailkosteberekeninge.

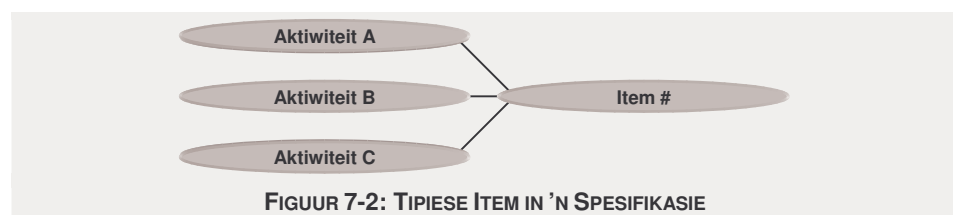
### 7.5.1 ITEMS TEENoor AKTIWITEITE

Twee moontlike alternatiewe om die tekort aan standaarditems op te los, is afkomstig van die berekeningmodelle wat in *Afdeling 6.7* bespreek is. Eenheid-tarief kan gedefinieer word op aktiwiteitvlak soos vanuit die kontrakteurperspektief of op itemvlak (saamgevoegde aktiwiteitkoste) soos vanuit die ingenieurperspektief. Elk van hierdie modelle het voor- en nadele.

#### 7.5.1.1 INGENIEURPERSPEKTIEF

Die kleinste eenheid van koste vanuit die ingenieurperspektief (en ook die ontwerpafdeling by DWB) tydens detail kosteberekeninge is die item. Tydens die ontwerp bepaal die ingenieur onder andere wat die materiaal eienskappe en afmetings van 'n komponent in die struktuur moet wees. Die eienskappe en afmetings van die komponent word vasgevang in die spesifikasie en word as 'n item in die hoeveelheidslys getoon. Die item bestaan egter nie net uit die materiaal wat voorgeskryf word nie, maar sluit ook die geassosieerde aktiwiteite in om die materiaal te laat voldoen aan die spesifikasie.

In die geval van die CSRA-spesifikasies sal 'n item soos die gruisstutlaag, *Item 34.01(a)*, die lewering, plasing, verspreiding en kompaktering insluit, netsoos die itemspesifikasie van byvoorbeeld die kleikern by damkonstruksie ook van hierdie aktiwiteite sal insluit. In die spesifikasie van die item word verdere aspekte wat die kostebydrae van elke aktiwiteit kan beïnvloed, ook presies uiteengesit. Die trekafstand van die vervoer aktiwiteit, die mate waartoe gekompakteer moet word of die metode hoe gekompakteer moet word, is voorbeelde van moontlike sulke voorskrifte wat variasie in die koste van die aktiwiteite kan veroorsaak. 'n Tipiese item kan voorgestel word met die model soos in *Figuur 7-2*.

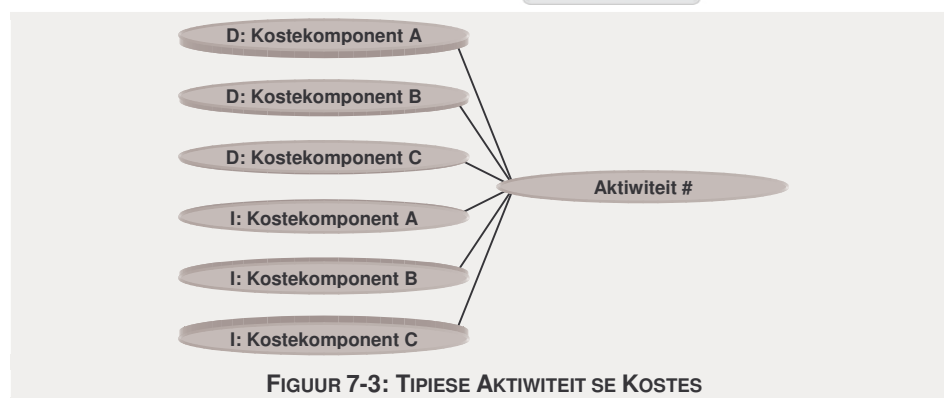


FIGUUR 7-2: TIPIESE ITEM IN 'N SPESIFIKASIE

Die voordeel van itemgebaseerde kosteberekening is dat die materiaal en geassosieerde aktiwiteitkoste van 'n komponent alreeds saamgevoeg is soos dit ooreenstem met die spesifikasie. Die eenheidstarief van 'n item vorm dus 'n logiese eenheid van betaling. In die damkonstruksie omgewing waar daar egter nie standaardspesifikasies is nie, bestaan daar ook nie standaarditems wat gebaseer is op standaard aktiwiteite nie. Dit veroorsaak dat aktiwiteite wat ingesluit word by die eenheidtariewe van items kan wissel van projek tot projek soos die spesifikasies wissel met gepaardgaande variasie in die eenheidtariewe. Die opstel van standaardspesifikasies om hierdie probleem aan te spreek is 'n moontlikheid, maar dit sal heelwaarskynlik 'n tydsame proses wees wat die ontwikkeling van 'n kostedatabasis sal vertraag, indien daar ooit gestandaardiseerde spesifikasies opgestel kan word vir die damkonstruksie omgewing.

### 7.5.1.2 KONTRAKTEURPERSPEKTIEF

Die perspektief van die kontrakteur (en ook die konstruksieafdeling by DWB [Du Buisson (2003)]) is vanuit aktiwiteitsgebaseerde kosteberekening wat gebaseer is op die hulpbronne wat 'n aktiwiteit gebruik. Die aktiwiteit is die tipiese aktiwiteit wat voorkom by 'n konstruksieprojek, soos byvoorbeeld uitgraving, lewering en kompaktering. Die hulpbronne is arbeid, toerusting, materiaal, brandstof wat gebruik word as ook die algemene kostes ("overheads"), wins (wat natuurlik nie by die staat van toepassing is nie) en die gebeurlikheidskoste. Eersgenoemde vier komponente is die direkte kostes en laasgenoemde drie die indirekte kostes. 'n Tipiese aktiwiteit kan voorgestel word met die model soos in *Figuur 7-3*.



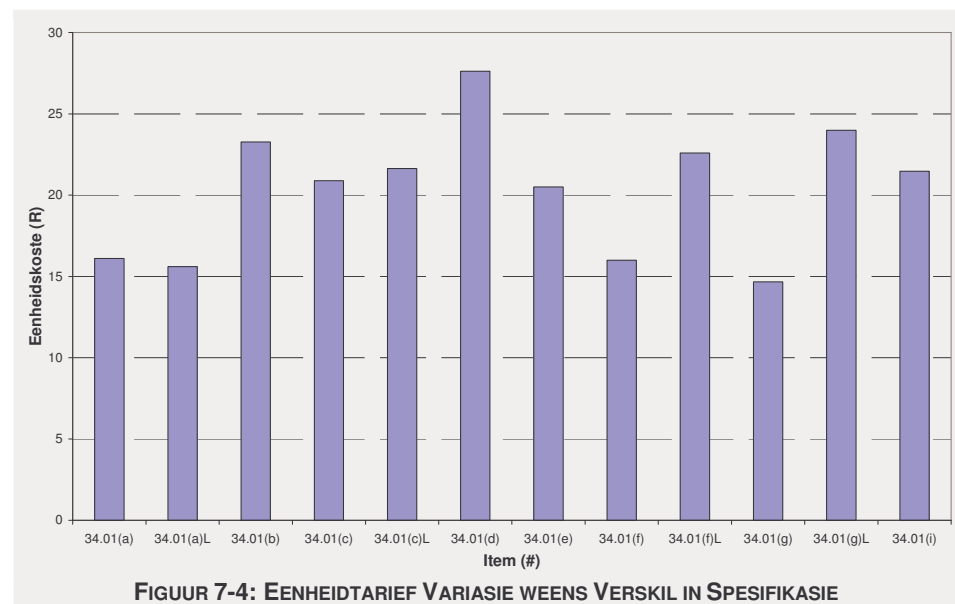
Die voordeel om aktiwiteite as basis van kosteberekening te gebruik is dat dit onafhanklik van enige spesifikasie is. Die versameling van aktiwiteite waarop gestandaardiseer word is die aktiwiteite wat in enige konstruksieprojek kan voorkom. Die eenheidtariewe, gebaseer op hulpbronnkoste, van aktiwiteite word dus gestoor in die databasis soos getender deur 'n kontrakteur (of beraam deur die konstruksieafdeling). Enige gespesifiseerde item se koste kan dus saamgestel word vanaf die aktiwiteite se kostes soos dit in 'n spesifikasies voorgeskryf word.

Die nadeel is egter dat daar 'n magdom verskillende aktiwiteite kan voorkom, veral as na die aantal moontlike variasies ten opsigte van 'n enkele aktiwiteit gekyk word. Die samevoeging van aktiwiteite soos voorgeskryf in die spesifikasies van 'n projek om items se eenheidtariewe te lewer moet ook nog steeds gedoen word. Hierdie samevoegings om items se kostes te bepaal sal ook in die databasis gestoor moet word, omdat betalings gewoonlik weens praktiese oorwegings gekoppel word aan die lewering van 'n gespesifiseerde item.

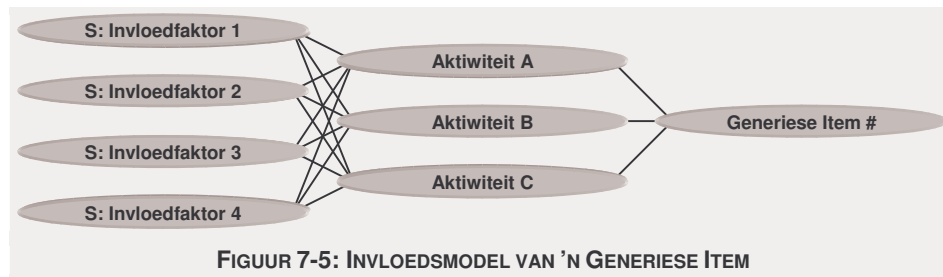
## 7.5.2 GENERIESE ITEMS

Die ingenieurperspektief van eenheidtariewe is die voorkeur metode, omdat dit aktiwiteite wat met mekaar verbandhou saam groepeer soos dit in die spesifikasies gevind word. Die realiteit is egter dat daar nie standaarditems kan wees indien daar nie standaardspesifikasies is nie. As alternatief moet daar eerder veralgemeende items gedefinieer word, onafhanklik van enige spesifikasie, maar ten minste met die aktiwiteite wat daarby ingesluit word, vasgelê. Hierdie items kan van die een tot die ander projek ten opsigte van die materiaalspesifikasie en ook die omvang van die aktiwiteite wat daarby ingesluit word, varieer.

Die effek wat hierdie variasies in spesifikasie op die item eenheidtarief het, kan in die CSRA-databasis se items gesien word. Die CSRA-databasis bestaan uit ongeveer 3000-standaarditems; baie van hierdie items is egter net variasies op dieselfde tema (item). Die kompaktering van gruis het byvoorbeeld verskillende onderafdelings van laagdiktes en digtheid. Hierdie variasies in aktiwiteite se kostes weens die spesifisering wat verskil tussen die “sub-items” veroorsaak variasies in die eenheidtariewe soos in *Figuur 7-4* vir *Item 31.01* gesien kan word.



Invloedfaktore kan weereens die moontlikheid bied om die variasie in eenheidtariewe wat kan ontstaan tussen verskillende projekte weens variasie in spesifikasies aan te spreek. Die minimum vereiste is egter dat daar gestandaardiseer moet word op die aktiwiteite wat ingesluit word by elke item. As die aktiwiteite vasgelê is vir elke item kan die invloedrykste aktiwiteitfaktore wat elke item beïnvloed bepaal word. Die generiese invloedsmodel sal dan weereens soos die model van die raadgevende ingenieur lyk, *Figuur 7-5*, maar die invloedfaktore wat die aktiwiteite beïnvloed sal oorwegend gebaseer wees op die spesifikasie van elke projek.



FIGUUR 7-5: INVLOEDSMODEL VAN 'N GENERIESE ITEM

Die databasismodel waarin hierdie inligting gelaai moet word is presies dieselfde as die model wat voorgestel is om die invloedfaktore en die gewigte daarvan in te stoor, *Figuur 6-27*.

### 7.5.3 HOËVLAK KOSTEBEREKENINGE

By die CSRA-databasis se uitbreiding van toepassingsmoontlikhede is die hoëvlakke van kosteberekeninge (konseptuele en langtermyn) gebaseer op onderskeidelik die samegevoegde kostes van ooreenstemmende items en totale koste van 'n gegewe projekteipe. Voorsiening vir verskille in projekte, wat die kostes kan beïnvloed, word gemaak deur projekte te karakteriseer met die invloedfaktor soos bespreek in *Afdelings 6.5* en *6.6*.

By die ontwikkeling van 'n databasis vir die damkonstruksie omgewing is een van die behoeftes om kosteberamings op verskillende vlakke van detail uit te voer. Dit is reeds genoem dat detail kostedata die voorkeur basis is om hierdie kosteberekeninge op te baseer. As die inligting op detailvlak onderskep word, word dit ook in hierdie geval moontlik om hoëvlak kosteberekeninge hierop te baseer deur slegs die verskillende generiese items in versamelings van ooreenstemmende items te groepeer.

Die probleme wat die leemte aan standaardspesifikasies veroorsaak ten opsigte van standaarditems is egter nie 'n probleem op die hoëvlak kosteberekeninge soos op die detailvlak eenheidtariefberekeninge nie. Die samevoeging van items se kostes in versamelings, soos die CSRA-databasis se itemkoste versamel is in *Series*, veroorsaak dat die detail omtrent individue-

le items vervaag en dat die veralgemeende faktore wat die versamelings beskryf meer bepalend word.

---

## 7.6 GROEI VAN DATABRON

---

In die damkonstruksie omgewing waar daar nie 'n bestaande databasis is nie kan die groei van die databron verdeel word, nadat die databasis geïmplementeer is, in 'n beginfase wanneer daar nie gestruktureerde data beskikbaar is nie en die selfvoerende fase, nadat die eerste data gelaai is.

---

### 7.6.1 BEGINFASE VAN DATABRONGROEI

---

Die derde fase in die lewensiklus van 'n databasis is die implementering en laai van data. Die implementering kan volgens die aanbevelings in *Afdeling 3.5* gedoen word. Om egter die databasis te kan laai met data moet daar eers bepaal word waarmee dit gelaai moet word.

---

#### 7.6.1.1 BEPALING VAN GENERIESE ITEMS

---

Een van die eerste stappe tydens die laai van die kostedatabasis is die eenmalige bevolking van die ITEM-tabel met standaarditems, soos in die CSRA-databasis se oordrag van bestaande data gedoen is. In die damkonstruksie omgewing waar daar nie 'n versameling van standaarditems bestaan nie, weens die tekort aan standaardspesifikasies, mag dit voorkom asof daar met 'n leë databasis begin moet word. Dit is egter nie die geval nie; daar is altyd data beskikbaar, hetsy in historiese formaat (papier of digitaal) of ervaring van kosteberamers waarop vorige beramings gedoen is. Hierdie inligting kan gebruik word om die “eerste asem” in die databasis te blaas.

Historiese projekte se hoeveelheidslyste kan as vertrekpunt gebruik word om generiese items te bepaal en die spesifikasies van hierdie projekte kan gebruik word om die aktiwiteite te vind as ook die invloedfaktore wat veranderlikheid vir elke aktiwiteit kan veroorsaak. Kosteberamers kan ook hier waardevolle insette lewer om hierdie lys van generiese items so omvattend moontlik te definieer asook die aktiwiteite en invloedfaktore, volgens die metode soos in *Afdeling 7.5.2* bespreek is. Indien hierdie generiese items, hul aktiwiteite en meegaande invloedfaktore bepaal is, kan die databasis hiermee gelaai word.

---

#### 7.6.1.2 SAAMSTEL VAN HISTORIESE INLIGTING

---

As die databasis eers gelaai is met die generiese items en die moontlike invloedfaktore vir elke aktiwiteit van die items, kan begin word om projekte se kostedata en invloedfaktore in die databasis te laai. Hier kan die databasisbestuurders egter besluit om onmiddellik te beweeg na die selfvoerende fase, *Afdeling 7.6.2*, van datagroei of om die databasis eers met historiese projekinligting te bevolk.

---

Eersgenoemde opsie sou beteken dat die eerste gebruikers nie waarde vanuit die databasis vir kosteberekeninge verkry nie, omdat daar geen inligting beskikbaar is om die kosteberekeninge op te baseer nie. Laasgenoemde opsie verg egter meer insette van die databasis ontwikkelaars om historiese projekinligting te versamel en te struktureer, sodat dit in die model van die databasis gestoor kan word. Hierdie opsie word egter verkies, omdat dit vir die eerste gebruikers van die databasis al resultate kan lewer, alhoewel met 'n hoë onsekerheid daaraan verbonde. Die mate van onsekerheid is egter afhanklik van die moeite wat gedoen word om die historiese bron te ontwikkel en te spesifiseer gebaseer op historiese projekte.

---

## 7.6.2 SELFVOERENDE FASE VAN DATABRONGROEI

---

Na die beginfase, waartydens die databasis gelaai word met generiese items en geassosieerde data, asook historiese projekdata, word daar beweeg na die operasionele funksionering van die databasis. Tydens hierdie fase word historiese data van die databasis gebruik in die aanwendingstegnieke, *Afdeling 4.4* en *4.5*, om die kostes van nuwe projekte te bepaal. Verskille wat voorkom ten opsigte van die projekomgewing/spesifikasie noodsaak egter dat die historiese data aangepas moet sodat dit vir die nuwe projek gebruik kan word, *Afdeling 4.6*. Om dit egter te kan doen moet die kosteberamer die nuwe projek karakteriseer (beskryf) aan die hand van die beduidendste invloedfaktore. Hierdie inligting data word saam met die historiese projekte se beskrywings gebruik om koste te beraam en dan probabilisties te kombineer, *Afdeling 4.7*.

Na voltooiing hiervan word die karakterisering inligting van die nuwe projek saam met die beraamde kostes soos bepaal deur die ontwerper en die kostes wat getender word deur die kontrakteur in die databasis gevoer. Sodoende ontstaan die selfvoerende fase waar die databasis met nuwe data groei soos die historiese data aangewend word vir nuwe projek kosteberamings.

---

## 7.7 VERDERE INKREMENTELE ONTWIKKELING

---

Databasisse behoort nie net aangepas te word as die behoeftes van die gebruikers verander nie, maar behoort ook te ontwikkel soos die inligting wat beskikbaar word oor die besprekingsruimte toeneem. Die CSRA-databasis kan as sprekende voorbeeld dien. Een van die belangrikste insigte wat uit die databasis verkry is, is dat die eenheidtariewe van sommige van die items wat gereeld noemenswaardige bydraes tot projekcostes lewer groot veranderlikheid toon en dat daar sterk korrelasies bestaan tussen van hierdie items se eenheidtariewe. By die herontwikkeling van die databasis behoort die oorsake wat kan aanleiding gee tot hierdie veranderlikheid ook aangespreek te word om kosteberekening te kan verbeter.

---

Inkrementele ontwikkeling aan die voorgestelde kostedatabasis vir damkonstruksie, soort gelyk aan bogenoemde, sal ook nodig wees soos meer inligting beskikbaar word en aspekte waarvan die ontwikkelaars nie bewus was nie, geïdentifiseer word. Een van die aspekte wat heelwaarskynlik duidelik na vore sal tree is die wyse waarop generiese items gedefinieer is, sommige items sal saamgevoeg kan word of nog verder verdeel word. Die enigste manier om dit egter te kan bepaal is as daar inligting op 'n gestruktureerde wyse onderskep word waarop daar in die toekoms weldeurdagte besluite geneem kan word gebaseer op historiese projekte.

---





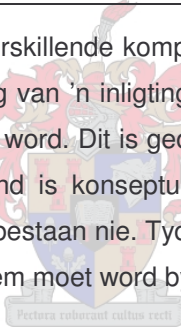
## HOOFSTUK 8

---

### ONTWIKKELING VAN 'N INLIGTINGSTELSEL

---

In die voorafgaande hoofstukke is die verskillende komponente geïdentifiseer en bespreek wat tydens die ontwikkeling van 'n inligtingstelsel vir kostebereamingsdoeleindes in ag geneem behoort te word. Dit is gedoen deur 'n bestaande databasis te ontleed en dit wat gevind is konseptueel toe te pas op 'n omgewing waar daar nog nie so 'n stelsel bestaan nie. Tydens die toepassing is nog aspekte geïdentifiseer wat in ag geneem moet word by die ontwikkeling van so 'n stelsel.



In hierdie hoofstuk word 'n generiese ontwikkelingmodel voorgestel wat al die aspekte integreer by die ontwikkeling/herontwikkeling van 'n koste-inligtingstelsel. Die verskillende komponente waaruit so 'n stelsel bestaan en hoe dit by mekaar inskakel word geskets. Die aspekte wat gevind is die betroubaarheid van die databron bepaal, word getoon en die gereedskap wat gebruik kan word om dit te verbeter word bespreek. Laastens word 'n geïntegreerde metode voorgestel wat gevolg kan word om 'n inligtingstelsel omvattend te ontwikkel.

---

#### 8.1 BEHOEFTE AAN VERBETERDE BESLUITNEMING

---

Verbeterende besluitneming ten opsigte van konstruksieprojekte word noodsaak deur faktore soos onder andere strenger omgewingswetgewing, groter deursigtigheid van besluite en kompetisie in die bedryf. Projekte se lewensvatbaarheid moet vandag gemeet word nie net aan inkomste en uitgawes nie, maar ook aan die impakte wat dit op die sosiale en natuurlike omgewing het. Hierdie impakte word al hoe meer in finansiële terme gekwantifiseer. Kliënte

---

vereis dat keuses ten opsigte van alternatiewe op ekonomiese gronde geneem word met gepaardgaande risiko's wat alternatiewe kan inhou vir optimale besluitneming. Verder veroorsaak 'n toenemende afname in finansiële bronne groter kompetisie en noodsaak verbeterde besluitneming.

Weens die behoefte aan verbeterde besluitneming, word verbeterde tegniese benodig. Hierdie verbeterde tegnieke is nie net beperk tot tegniese ontwikkeling wat meer ekonomiese ontwerpe lewer nie, maar ook ten opsigte van ekonomiese ontledings. Ekonomiese modelle wat al hoe meer omvattend is, word gebruik om die onsekerhede van die besigheidsmodel en die verskillende konstruksie alternatiewe te modelleer. Kosteberekening resultate moet derhalwe meer akkuraat en omvattend opgestel word, sodat konstruksieprojek risiko's hierby ingesluit kan word. Netsoos geoptimeerde strukture vandag ontwerp kan word, omdat verbeterde ontledingstegnieke beskikbaar is, moet die ekonomiese modelle wat die koste verbonde aan hierdie strukture beskryf ook verbeter word deur die tegnieke waarmee dit opgestel word te verbeter.

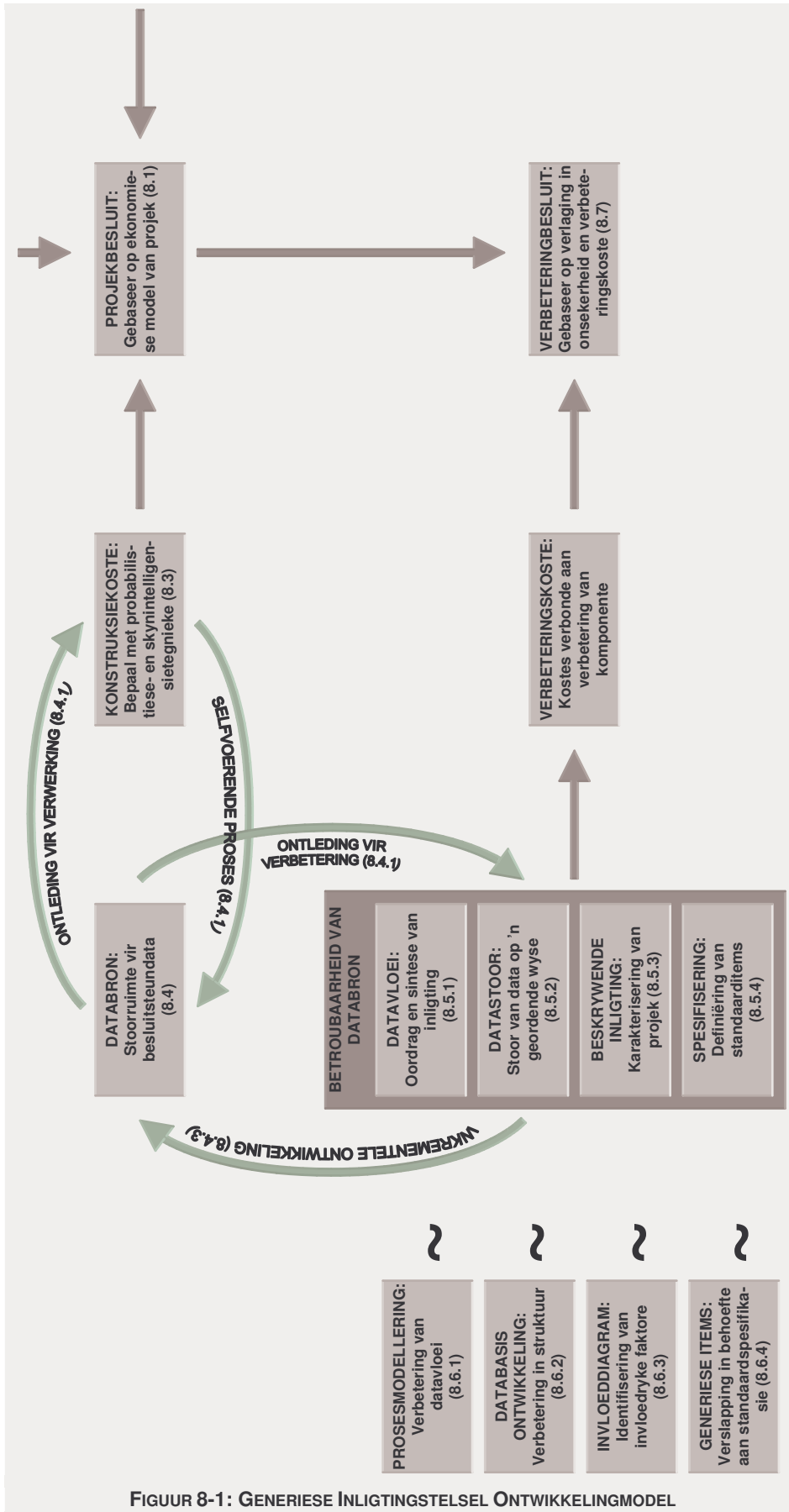
Onsekerhede wat voorkom in kosteberamingsmodelle kan met probabilistiese tegnieke gemodelleer word, maar naas die modelle wat benodig word om projekonsekerheid meer omvattend mee te modelleer word data benodig om hierdie onsekerhede te kwantifiseer. Die groot veranderlikheid wat egter in konstruksiekostedata voorkom veroorsaak dat probabilistiese tegnieke by hoë vlakke van betroubaarheid onrealistiese waardes lewer. Uitbreiding en verbetering van data-onderskepping is gevolglik 'n wesentliche element van 'n gevorderde besluitnemingsproses.

*Pectus roburant cultus recti*

## 8.2 GENERIESE INLIGTINGSTELSEL ONTWIKKELINGSMODEL

Die veranderlikheid wat in konstruksiekostedata gevind word, is oorwegend epistemies van aard, *Afdeling 5.6.2.3*. Dit kan aangespreek word met 'n verbetering in die inligtingstelsel wat die data onderskep en stoor. Die sistematiese ontwikkelingproses wat in hierdie studie gevolg is, vanaf 'n omgewing waar daar 'n reeds bestaande stelsel is wat geëvalueer kon word om probleme en moontlike verfynings te identifiseer tot 'n omgewing waar daar nog nie so 'n stelsel bestaan nie, het tot die ontwikkeling van 'n generiese inligtingstelsel ontwikkelingmodel gelei.

Die onderskeie komponente wat in hierdie generiese model geïdentifiseer is en die wyse hoe hul by mekaar inskakel word in *Figuur 8-1* getoon. Die noodsaaklikheid van betroubare probabilistiese inligting vir besluitnemingdoeleindes in ekonomiese modelle van projekte is reeds in die voorafgaande afdeling bespreek. Die res van die komponente word in die volgende afdelings saamgevat, soos dit in die voorafgaande gedeeltes van die studie geïdentifiseer is, en die rol daarvan in die generiese ontwikkelingmodel word bespreek.



FIGUR 8-1: GENEERIESE INLIGTINGSTELSEL ONTWIKKELINGSMODEL

---

## 8.3 KONSTRUKSIEKOSTE-INLIGTING

---

Die verskillende wyses waarop konstruksiekosteberekeninge uitgevoer kan word, is reeds in *Hoofstuk 4* bespreek. In hierdie afdeling word die voorkeur tegnieke genoem, wat gevind is met groot vrug aangewend kan word, om kosteberekening inligting vir besluitneming te lewer.

---

### 8.3.1 PROBABILISTIESE TEGNIEKE

---

Onsekerheid kom deurlopend voor in konstruksieprojekte. 'n Verskeidenheid van bronne, hetsy ruimtelik (klimaat, geologie, terreintoestand, ens.) of temporaal (inflasie, sterkte van die rand, koste van brandstof, ens.) veroorsaak die onsekerheid. Probabilistiese tegnieke bied die geleentheid om die onsekerhede wat by kosteberekening in die hoeveelhede en eenheidtariewe manifesteer weens die onderskeie bronne, in berekening te bring. Die globale onsekerheid geassosieer met die koste van 'n projek kan dus bepaal word. Van die twee probabilistiese tegnieke wat in hierdie studie bespreek is, nl. Monte Carlo-simulasie en Limietstaat Kostefunksie lewer laasgenoemde die mees insiggewende besluitsteuninligting. Die meer omvangryke benutting van data se statistiese eienskappe en die insigte wat dit op itemvlak bied, noodsaak egter die verbetering in kwaliteit van databronne om betroubare resultate te lewer.

---

### 8.3.2 SKYNINTELLIGENSIE-TEGNIEKE

---

Skynintelligensietegnieke bied die geleentheid om onsekerheid te verlaag met beskrywende inligting. Die tegnieke is toepaslik op die konstruksie omgewing waar veranderlikheid in kostedata weens verskille in projekomgewings voorkom. Dit is in hierdie studie gevind dat neurale netwerke die voorkeur tegniek blyk te wees om hierdie tipe ontledings mee uit te voer indien beskrywende data beskikbaar is om projekomgewings mee te karakteriseer. Netsoos dit egter ook belangrik is in probabilistiese tegnieke om inligting beskikbaar te hê om die statistieke van die onderskeie basiese veranderlikes op te baseer, so is dit ook essensieel dat daar beskrywende inligting beskikbaar is om die projekomgewing van historiese projekte mee te klassifiseer. Hierdie tipe inligting van projekte is in meeste gevalle nie beskikbaar nie.

---

## 8.4 DATABRON

---

Die suksesvolle groei van 'n organisasie is afhanklik van goeie besluite gebaseer op korrekte inligting. Tydens die beoefening van die kern besigheid van 'n organisasie word groot hoeveelhede data geskep deur die aktiwiteite wat uitgevoer word. Die data word egter nie onmiddellik, nadat dit geskep is, benodig vir besluitneming nie; daar kan 'n lang tydsverloop wees voordat dit aangewend word. Die data behoort dus op 'n selektiewe wyse onderskep en gestoor te word, sodat dit op 'n later tydstip beskikbaar is vir ontleding. Die volgende

---

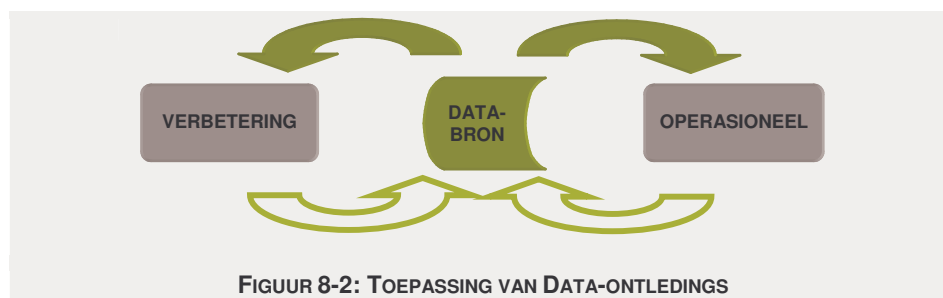
aspekte is van belang by die databron van die generiese ontwikkelingmodel, *Figuur 8-1*.

#### 8.4.1 DATA-ONTLEDING

Die data in die databron moet ontleed word om die geaardheid van die data te bepaal, *Hoofstuk 5*. Die inligting afkomstig hiervan kan tweeledig gebruik word, nl. vir operasionele besluitneming of vir verbetering aan die databron. Die operasionele funksie bestaan uit die lewering van inligting vir die kosteberekeningstegnieke wat gebruik word om kosteberamings mee te maak. Inligting soos die invloedrykste komponente in die kosteberekeninge kan bepaal word asook die basiese statistieke, korrelasies, statistiese verdelings en beskrywings van ruimtelike en temporale veranderlikheid. Hierdie meer omvangryke benutting van data ontstaan a.g.v. die probabilistiese en skynintelligensie-ontledingstegnieke wat gebruik word.

Aangesien die resultate van die probabilistiese ontledingstegnieke meer sensitief is vir onakkuraathede in die data, is dit nodig om die data se kwaliteit te verbeter. Data-ontledings vir verbetering aan die databron is soortgelyk as die van die operasionele funksie, maar in hierdie geval word die inligting gebruik om foute in die stelsel te identifiseer of om ontoereikende data-onderskepping in die stelsel te bepaal. Die word gedoen om onderskeidelik die verskillende aspekte wat die betroubaarheid van die data kan beïnvloed te verbeter en om die omvang van die data-onderskepping te optimeer.

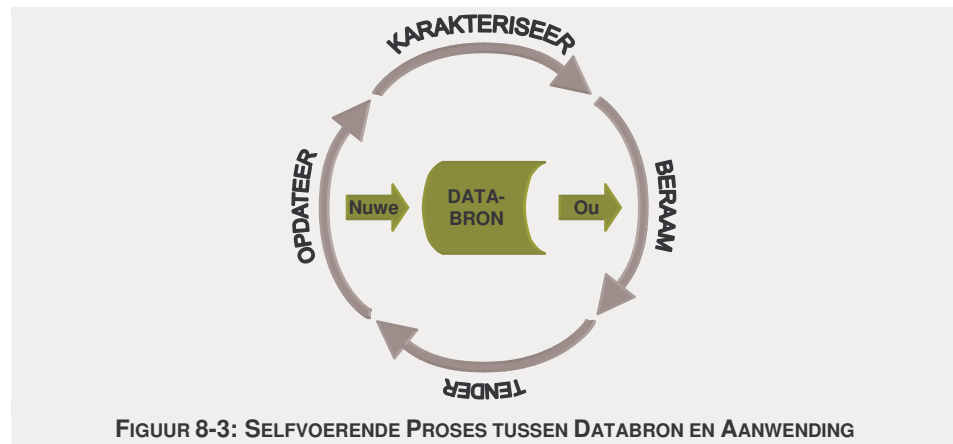
Die inligting wat van hierdie tweeledige toepassings verkry word, word onderskeidelik indirek weer na die databasis teruggevoer of aangewend om die databron te verbeter en meer betroubaar te maak, *Figuur 8-2*, sodoende ontstaan onderskeidelik 'n selfvoerende proses en inkrementele ontwikkeling.



#### 8.4.2 SELFVOERENDE PROSES

Die selfvoerende proses wat in hierdie studie voorgestel word vind plaas tydens die operasionele funksionering van die databasis tussen die kosteberekeninge en die databron, *Figuur 8-3*. Wanneer 'n gebruiker (raadgewende ingenieur) kosteberekeninge wil uitvoer het hy inligting oor sy eie projek beskikbaar. Hierdie inligting word gebruik om die projek te karakteriseer aan die hand van

invloedfaktore, *Afdeling 6.7.2*. Met hierdie inligting kan die historiese projekte se kostedata gekonsilieer word met behulp van skynintelligensietegnieke, soos neurale netwerke, om kostes wat spesifiek op die nuwe projek van toepassing is, te bepaal.



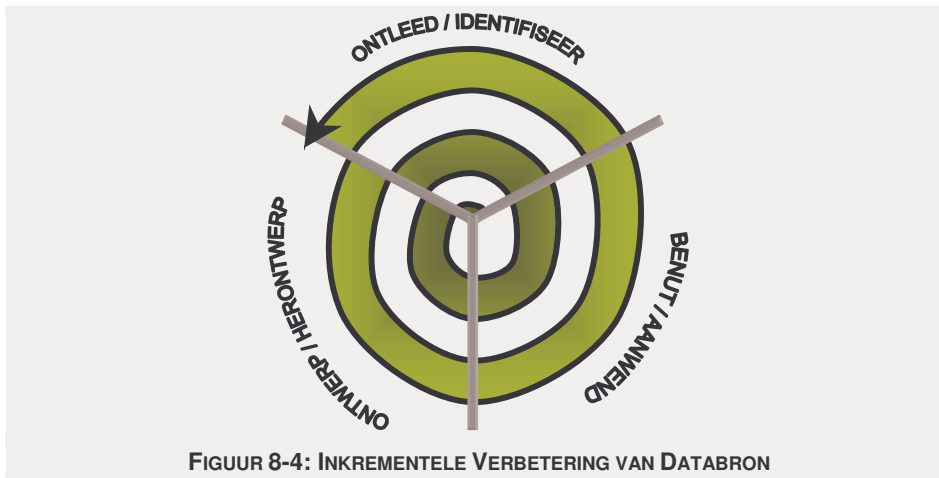
Tydens die tenderfase word die werklike eenheidtariefdata van die kontrakteurs beskikbaar. Hierdie inligting word saam met die beraamde eenheidtariewe van die raadgewende ingenieur, asook die gekwantifiseerde invloedfaktore wat in die aanvanklike kosteberekening gebruik is, saamgevat in die tipiese tenderverslag. Die data in hierdie verslag word dan afgevoer na die databasis en word sodoende deel van die historiese bron van projekte. Tydens opvolgende nuwe projekte is die vorige “nuwe” projek dus deel van die bron van historiese gekarakteriseerde projekdata wat gebruik kan word vir voorspellings.

Die selfvoerende proses van ’n databron is kritiek vir die suksesvolle funksionering van die databasis. Indien die gebruikers nie voordeel trek uit die akkurate karakterisering van die projekte nie kan dit lei tot swak beskrywende inligting of beperkte terugvloei van data na die databasis. Dit kan aanleiding gee tot ’n verlaging in die betroubaarheid van die databron. Die selfvoerende proses wat dus hier ontwikkel is, kan ’n substansiële bydrae in die verbetering van datakwaliteit meebring, omdat die korrekte beskrywing van nuwe projekte onmiddellik tot voordeel van die gebruiker is, maar ook vir latere gebruikers.

### 8.4.3 INKREMENTELE VERBETERING

’n Geïmplementeerde databasis word soms gesien as ’n finale produk, nie net ten opsigte van die databasis self nie, maar ook waarvoor dit aangewend word. Databasisontwikkeling eindig egter nie wanneer die databasis geïmplementeer is nie. Wanneer ’n databasis die eerste keer ontwerp word, is dit nie moontlik of geregverdig om al die eienskappe van die besprekingsruimte volledig te modelleer en aan al die behoeftes van die verbruikers te voldoen nie. Slegs wanneer die databron operasioneel gebruik word en datavloei vind na en vanaf die bron plaas kan dit ontleed word om te bepaal waar die tekortkominge voorkom. Na

die herontwerp van die databron moet die stelsel eers weer deur 'n operasionele fase beweeg voordat die volgende rondte van ontleding en die gevolglike herontwerp uitgevoer kan word, *Figuur 8-4*.



In die bestaande databasis wat ontleed is, is dit gevind dat die kwaliteit en die omvang van data-onderskepping voldoende is vir waarvoor die databasis oorspronklik ontwikkel is. Weens veranderde behoeftes, soos reeds genoem, is dit egter nodig om meer gesofistikeerde kosteberekeninge uit te voer. Hierdie berekeninge noodsaak dat die databron se kwaliteit en omvang egter uitgebrei moet word. Deur die data in die bestaende databasis te ontleed kon geïdentifiseer word watter aspekte vir die nuwe tipe kosteberekeninge verbeter moes word.

## 8.5 BETROUBAARHEID VAN BRON

In hierdie studie is gevind dat die betroubaarheid van die databron afhanklik is van vier hoof aspekte, nl. die datavloei, die datastoor, die beskikbaarheid van beskrywende inligting en die standardisering van dit wat onderskep word, *Figuur 8-1*.

### 8.5.1 DATAVLOEI

Datavloei verwys na die oordrag en sintese van data tussen die aktiwiteit waar dit ontstaan en die punt waar dit weer aangewend word. Tydens die datavloei kan foute ontstaan weens onakkuraathede gedurende oordragte, stadige of beperkte vloei van data en nog vele ander foute weens die vloei van data, soos in hierdie studie gevind is. Hierdie foute kan die betroubaarheid van die data nadelig beïnvloed, veral as die data gebruik moet word vir gevorderde ontledingdoeleindes.

---

## 8.5.2 DATASTOOR

---

Die vorm waarin data gestoor word bepaal hoe gereedlik beskikbaar en akkuraat hierdie data is. Die tradisionele medium van stoor was die liasseerkabinet gewees, maar weens toenemende volumes data en tegnologiese ontwikkeling het elektroniese stoorruimtes (databasisse) die voorkeur stooromgewing geword. Die doel van 'n databasis is om data op 'n geordende wyse te stoor, sodat die relevante data op 'n latere tydskop vinnig gevind en korrek opgeroep kan word vir ontledingdoeleindes. Dit is in hierdie studie gevind dat foute in databasisse kan ontstaan weens onnodige dupliseringe en onvoldoende foutidentifiseringsmeganismes.

---

## 8.5.3 BESKRYWENDE INLIGTING

---

Wanneer data temporale en ruimtelike veranderlikheid toon is dit nie voldoende om slegs die data te onderskep nie, maar beskrywende inligting word benodig om die data te kan klassifiseer. Indien daar nie beskrywende inligting hiervoor beskikbaar is nie, kan dit veroorsaak dat indien die data vir ontledingsdoeleindes aangewend word groot onsekerheid kan bestaan oor die toepaslikheid van die data in die omgewing waar dit aangewend word. Hierdie aspek is nie net geïdentifiseer by die databasis wat ondersoek is nie, maar kom ook in ander soortgelyke databasisse voor waar geen verklaring vir die hoë veranderlikheid van eenheidtariewe gevind kan word nie.

---

## 8.5.4 SPESIFISERING

---

Een van die oorsake van groot veranderlikhede in die data van 'n eenheidtarief-databasis kan toegeskryf word aan die swak definiëring van die items waarvan die eenheidtariewe onderskep word. Indien items nie streng gedefinieer is nie kan dit veroorsaak dat die kostekomponente waaruit die eenheidtarief saamgestel is van projek tot projek wissel. Dit gee aanleiding tot epistemiese veranderlikheid naas die temporale en ruimtelike veranderlikheid.

---

## 8.6 GEREEDSKAP BESKIKBAAR

---

Die betroubaarheid van die databron kan verbeter word deur die vier komponente in die voorafgaande afdeling te verbeter, *Figuur 8-1*. Hierdie komponente kan egter nie afsonderlik van mekaar ontwikkel word nie, maar behoort geïntegreerd gedoen te word.

---

### 8.6.1 PROSESMODELLERING

---

Prosesmodellering kan, in omgewings waar daar komplekse datavloei voorkom, gebruik word om op 'n gestruktureerde wyse al die aktiwiteite wat inligting skep en gebruik te identifiseer asook om die interaksie tussen hierdie aktiwiteite te bepaal. Die insigte wat dit bied op die datavloei wat binne hierdie prosesse

---



plaasvind kan gebruik word om die onderskepping van data te verbeter deur omslagtige aktiwiteite uit te skakel, oordragte van data te stroomlyn en onderskepping meer omvattend te maak. Nuttige gereedskap is in hierdie studie ontwikkel om vir volledigheid en korrektheid van sulke modelle te toets, *Afdeling 2.2*.

---

### **8.6.2 DATABASISONTWIKKELING**

---

Inkonsekwenthede in data ontstaan wanneer onnodige dupliserings van data voorkom. Databasis normalisering kan tydens die ontwerp van die datamodel gebruik word om te verseker dat slegs beheerde dupliserings voorkom, soos in hierdie studie gedoen is vir die ontwikkeling van 'n eenheidtariefdatabasismodel. 'n Korrek genormaliseerde databasis kan ook makliker uitgebrei word in latere fases van ontwikkeling om meer omvattende inligting te onderskep soos die behoeftes mag verander. Om te verseker dat oordragte na die databasis korrek plaasvind kan van die foutidentifiseringsmeganismes wat in hierdie studie bespreek is aangewend word om foute in die databasis tydens invoere te identifiseer en te voorkom.

---

### **8.6.3 INVLOEDFAKTORE**

---

Ruimtelike veranderlikheid kan verlaag word indien inligting beskikbaar is om die projekomgewing te karakteriseer. Die uitbreiding van data-onderskepping om ruimtelike veranderlikheid aan te spreek behoort egter op 'n gestruktureerde wyse gedoen te word om te voorkom dat nuttelose inligting onderskep word wat nie bydra tot die algehele verlaging in die veranderlikheid van 'n projek se kostes nie. Die metode wat in hierdie studie ontwikkel is, *Afdeling 6.8*, om die beduidendste invloedfaktore te bepaal en dit in verband met die korrelasie wat tussen eenheidtariewe voorkom te bring kan met groot vrug aangewend word om data-onderskepping te struktureer om te verseker dat die waardevolste projekomgewinginligting onderskep word.

---

### **8.6.4 GENERIESE ITEMS**

---

Standaardspesifikasie lei tot die ontwikkeling van standaarditems waar al die aktiwiteite van elke item en moontlike variasies vasgelê is. In 'n omgewing waar daar nie standaardspesifikasie is nie, kan kostes van items van projek tot projek verskil weens verskille in die aktiwiteite en die eienskappe van die aktiwiteite. Indien slegs die aktiwiteite van elke item egter vas gelê word, maar die moontlike variasies van die aktiwiteite word nie gedefinieer nie, kan generiese items gevorm word soos in hierdie studie voorgestel is, *Afdeling 7.5*. Die moontlike veranderlikheid wat veroorsaak word deur verskille in spesifikasies tussen projekte kan dan deur invloedfaktore verklaar word.

---

---

### 8.6.5 GEÏNTEGREERDE ONTWIKKELING

---

Die ontwikkeling van 'n inligtingstelsel kan nie afsonderlik plaasvind van die aanwendingstegnieke nie. Van die tegnieke wat in hierdie studie bespreek is vereis verhoogde kwaliteit en kwantiteit in brondata, omdat die tegnieke se resultate meer sensitief is vir onder ander veranderlikheid in die data. Hiervoor word meer omvattende en verbeterde data-onderskepping benodig wat die datakwaliteit kan verbeter en verklarings kan bied vir groot veranderlikheid.

Die aanwendingstegnieke bepaal egter nie net omvang en kwaliteit van data onderskepping in die inligtingstelsel nie, maar kan ook die verdere ontwikkelingmoontlikhede bepaal. Die selfvoerende proses wat ontstaan tydens data-aanwending stel inligting beskikbaar wat aangewend kan word tydens die herontwikkeling van die stelsel. Indien beperkte inligting beskikbaar is oor die omgewing waarbinne die inligtingstelsel funksioneer bemoeilik dit uitbreikbaarheid daarvan, *Afdeling 5.1*.

Die vier aspekte wat tydens die ontwikkeling/herontwikkeling van 'n inligtingstelsel aangespreek moet word kan ook nie in afsondering aangewend word om die betroubaarheid van die databron te verbeter nie. Selfs al word sommige van die komponente soos die databasis en die datavloei optimaal ontwikkel, maar daar word nie op 'n gestruktureerde wyse bepaal wat onderskep behoort te word nie sal die verbetering van eersgenoemde twee aspekte van geen nut wees nie, omdat die ontwikkeling beperk sal word deur die ontoepaslike inligting wat onderskep kan word.

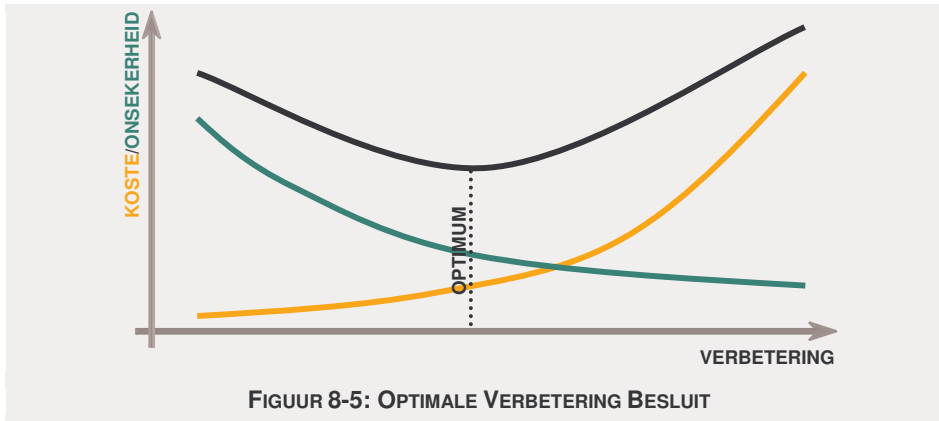
---

### 8.7 KOSTE VAN 'N VERBETERDE INLIGTINGSTELSEL

---

Die vier komponente wat bespreek is, kan gebruik word om die betroubaarheid van historiese data te verbeter vir gevorderde besluitondersteuningstegnieke. Aan elk van hierdie verbeteringmetodes is daar egter finansiële kostes verbonden, vir sommige meer as vir ander. Om te bepaal of dit lonend is om enige van die komponente te verbeter behoort die ontwikkelaar van so 'n stelsel vir elke van die komponente te besluit watter verlaging in onsekerheid te weeg gebring sal word vir die totale portefeulje van kosteberekeninge per jaar teenoor die kapitaal uitgawe en jaarlikse bedryfskoste van die verbetering/e aan die komponent/e. Deur die totale koste en onsekerheid te minimeer kan die optimale besluit ten opsigte van die omvang van die verbetering/e bepaal word, *Figuur 8-5*.

---

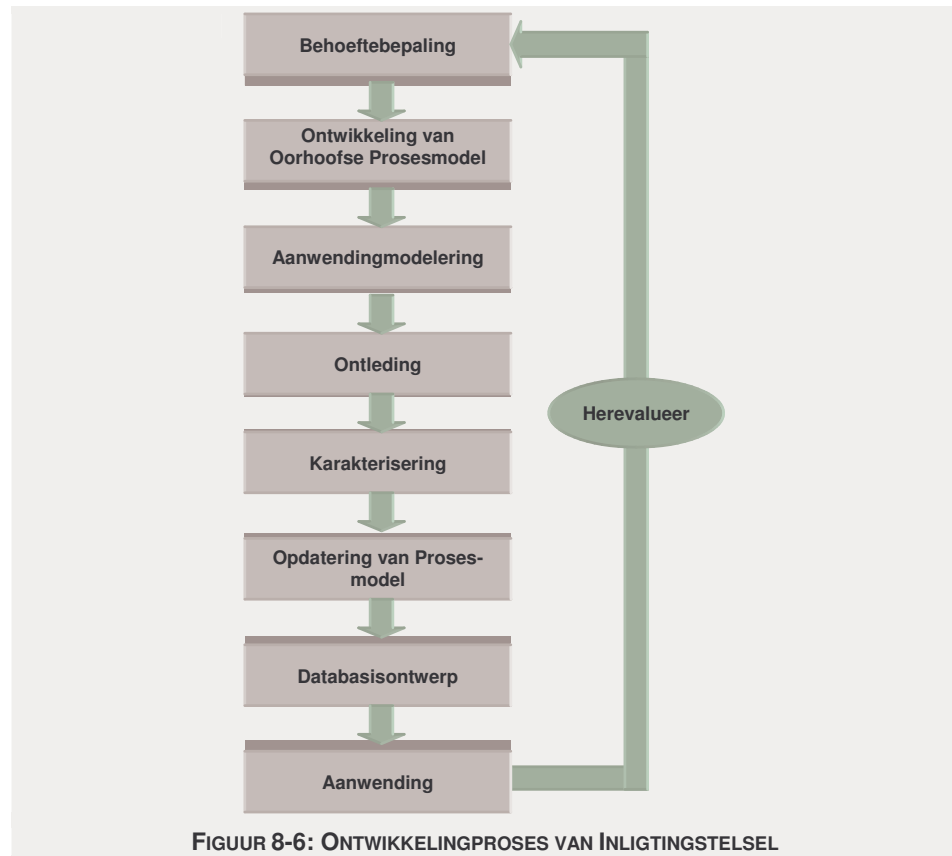


FIGUUR 8-5: OPTIMALE VERBETERING BESLUIT

## 8.8 ONTWIKKELINGPROSES

Die ontwikkelingproses wat hier bespreek word is 'n geïntegreerde proses wat van toepassing is op die generiese ontwikkelingmodel van nuwe/bestaande inligtingstelsels vir kosteberamings. Alhoewel die onderskeie stappe, *Figuur 8-6*, in 'n logiese volgorde bespreek word, is die werklike ontwikkeling van so 'n stelsel nie noodwendig so gestruktureerd nie. Gedurende die stappe is daar gereeld terugvoerings en aanpassings nodig aan van die modelle wat in die vorige stappe ontwikkel is soos die kennis van die besprekingsruimte toeneem en verdere behoeftes of tekortkominge geïdentifiseer word.





### 8.8.1 BEHOEFTEBEPALING

Die eerste stap in die ontwikkeling/herontwikkeling van enige stelsel is om te bepaal waarvoor dit aangewend gaan word. Vir die ontwikkeling van 'n inligtingstelsel vir kosteberekeninge moet bepaal word vir watter tipe projekte dit gebruik gaan word en watter vlak van detail benodig word. Die verskillende generiese tipe projekte en vlakke van detail wat kan voorkom is onderskeidelik in *Afdelings 7.3.1* en *7.3.2* in die voorbeeld oor die dambou bespreek en meer algemeen in *Afdelings 4.3* en *6.2*.

Die ontledingstegnieke wat gebruik gaan word het ook 'n bepaalde invloed op die ontwerp van die stelsel. Deterministiese tegnieke is minder sensitief vir die kwaliteit van data teenoor probabilistiese tegnieke waar dit die resultate noemenswaardig kan beïnvloed. Die ontwerp van die inligtingstelsel ten opsigte van omvangrykheid van data-onderskepping en foutidentifiseringsmeganismes word dus bepaal deur die behoeftes wat deur die ontledingstegnieke gestel word.

### 8.8.2 OORHOOFSE PROSESMODEL

Die behoeftes wat geplaas word op die inligtingstelsel bepaal vir watter besluite die inligting afkomstig uit die stelsel gebruik gaan word. Hierdie besluite vorm deel van die projek en bepaal watter roetes gevolg word in die projek. Om te kan aandui waar in die proses die inligting benodig word, word 'n oorhoofse

prosesmodel benodig om die aktiwiteite, persone en inligting hierby betrokke te toon en die effek wat die besluite gaan hê. Die oorhoofse model wat in *Afdeling 2.1* ontwikkel is, kan as basis gebruik word om die omgewingspesifieke oorhoofse model op te stel.

---

### **8.8.3 OPSTEL VAN AANWENDINGMODEL**

---

Vir die verskillende tipe kosteberekening wat geïdentifiseer is vir uitvoering, moet bepaal word wat die bronne van inligting is wat benodig word vir die berekening. Die aanwendingmodelle, *Afdeling 6.2*, kan as basis gebruik word om hierdie modelle mee op te stel. Die nuwe aanwendingmodel wat ontwikkel word, behoort so na as moontlik aan die ideale model wat in bogenoemde afdeling getoon is te wees, om die maksimum voordeel van so 'n ontwikkeling te kry. Aanpassings aan die prosesmodel moet aangebring word indien daar in die aanwendingmodel aspekte geïdentifiseer is wat nie alreeds ingesluit is in die oorhoofse prosesmodel nie.

---

### **8.8.4 ONTLEDING VAN DATABRONNE**

---

Dit is reeds genoem dat databronontleding 'n tweeledige rol vervul; eerstens vir die bepaling van inligting vir ontledingstegnieke en tweedens vir die ontwikkeling van die databron. Laasgenoemde fase is hier van toepassing. Indien data beskikbaar is by van die bronne wat in die vorige stap geïdentifiseer is, moet bepaal word watter van die data in die onderskeie versamelings die grootste invloed op die kosteberekening het. Hierdie invloedrykste data moet statisties ontleed word om te bepaal wat die veranderlikheid van die data is en wat die oorsprong van die veranderlikheid is, bv. foute, temporale of ontoereikende spesifikasie (item en/of projek), soos in *Hoofstuk 5* bespreek word. Vir die gevalle waar daar nie databronne beskikbaar is nie moet spesialiste se kennis gebruik word om beramings van die eienskappe van die data te maak.

---

### **8.8.5 INVLOEDFAKTORE**

---

Die inligting afkomstig uit die vorige afdeling bepaal hoe bestaande en nuwe databronne se spesifisering ontwikkel moet word.

---

#### **8.8.5.1 BESTAANDE DATABRONNE**

---

Indien dit gevind word dat temporale veranderlikheid of foute nie die oorsaak van die veranderlikheid van die data is nie en dit blyk 'n ontoereikende spesifikasie van die item of die projek te wees, is dit nodig om die eienskappe wat die veranderlikheid veroorsaak beter te beskryf. Hiervoor kan van invloedfaktore gebruik gemaak word. Die invloedfaktore wat bepaal word moet vanaf die raadgewende ingenieurperspektief ontwikkel word vir die redes wat vermeld word in *Afdeling 6.8.1.1*. Die bepaling van die beduidendste invloedfaktore kan

---

met die metode soos in *Afdeling 6.8.1.3* beskryf, gedoen word en geoptimeer word met die korrelasie inligting wat afkomstig is van die historiese data.

---

#### **8.8.5.2 NUWE DATABRONNE**

---

Indien daar nie bestaande databronne beskikbaar is vir van die inligting wat in die kosteberekening benodig word nie of die bestaande bronne is van so 'n aard dat die spesifikasie van die items nie voldoende gestandaardiseer is nie en dus aanleiding gee tot onaanvaarbare groot veranderlikheid, moet daar 'n gestandaardiseerde lys van generiese items ontwikkel word. Om die verskeie variasies van items te akkommodeer kan ook van invloedfaktore gebruik gemaak word om die bronne van veranderlikheid te karakteriseer van die onderskeie kostekomponente waaruit die standaarditem se kostes saamgestel is nadat die kostekomponente vir elke item vasgestel is. Die metode word in *Afdeling 7.5* bespreek.

---

#### **8.8.6 OPDATERING VAN PROSESMODEL**

---

Die onderskeie bronne van inligting wat benodig word om die kostes te bepaal en om die bronne van veranderlikheid te karakteriseer behoort teen hierdie punt vasgelê te wees. Die prosesmodel kan dus uitgebrei word om alle aktiwiteite te identifiseer waar die data geskep word en wie daarby betrokke is. Die datavloei moet dan binne hierdie prosesmodel bepaal word en geoptimeer word om omslagtige aktiwiteite uit te skakel wat kan lei tot foute en veroudering van data, *Hoofstuk 2*.

---

#### **8.8.7 DATABASISONTWIKKELING**

---

Die databasisontwikkeling bestaan uit drie fases, nl. die ontwikkeling van die databasismodel, die beginfase en die operasionele fase.

---

##### **8.8.7.1 DATABASISMODEL**

---

Die ontwerp van die databasismodel word bepaal deur die inligting wat geïdentifiseer is vir die kosteberekening en dus gestoor moet word in die databasis vir latere aanwending. Die onderskeie aspekte wat gepaardgaan met die ontwikkeling van 'n databasis om hierdie inligting effektief te stoor is volledig in *Hoofstuk 3* bespreek. As basis vir die databasismodel kan die generiese standaarditems model gebruik word, *Afdeling 3.5.2*, en die riglyne wat in die daaropvolgende logiese ontwerp bespreek word, asook die foutidentifiseringsmeganismes wat genoem word. Indien die model verder moet uitgebrei word om die regsdokumente, standaardspesifikasies, invloedfaktore (projek of item) te onderskep kan die modelle wat in *Afdelings 3.8* en *6.8.4* getoon word, gebruik word.

---

---

### 8.8.7.2 BEGINFASE VAN DATABRONGROEI

---

Die bestaande databronne wat geïdentifiseer is moet gebruik word om die databasis te bevolk. Indien die bestaande inligting se datamodel verskil van die nuwe model is 'n omskakelingproses nodig met die nodig kontroles om te verseker dat die data korrek oorgedra word en dat verwantskappe behoue bly, *Afdeling 3.6*. Vir die gevalle waar daar nie bestaande databasisse is nie en die bronne dus nog ontwikkel moet word, kan een van die opsies gevolg word soos in *Afdeling 7.6.1.2* bespreek word.

---

### 8.8.7.3 SELFVOERENDE FASE VAN DATABRONGROEI

---

Een van die belangrikste eienskappe van die inligtingstelsel is die selfvoerende proses wat voorkom, *Afdeling 8.4.2*. Wanneer die databasis oorgaan na die operasionele fase is dit nodig om te verseker dat inligting wat gebruik word om projekte te karakteriseer tydens die kosteberekeninge van nuwe projekte teruggevoer word na die databasis. Hiervoor is dit nodig dat die prosesmodel sodanig aangepas word om te verseker dat die komponent wat voorheen nie deel was van die kosteberamingsproses nie, wel ingesluit word en dat die onderskepping en terugvoering hiervan duidelik ontwikkel word in die datavloei na die databasis.

---

### 8.8.8 AANWENDINGSTEGNIEKE

---

Die kosteberamingstegnieke wat gebruik kan word is afhanklik van die kwaliteit en kwantiteit van die inligting wat onderskep word en word dus in 'n groot mate vasgelê deur die behoeftebepaling wat in die eerste fase uitgevoer is. Die verskillende kosteberamingstegnieke wat gebruik kan word, word in *Hoofstuk 4* bespreek. Die inligting afkomstig van die data-ontleding, tydens die operasionele fase van die databasis sal bepaal watter tegnieke aangewend moet word vir die kosteberamings. Dit kan slegs plaasvind as werklike data vanuit die databasisbesprekingsruimte die databasis tydens die operasionele fase begin bevolk.

---

### 8.8.9 HEREVALUERING VAN STELSEL

---

Die kwaliteit van die besluitneminginligting wat die inligtingstelsel lewer is eers moontlik om te bepaal wanneer die stelsel operasioneel is. Die statistiese ontleding van die data in die databasis kan leidrade bied tot die korrektheid van die inligting en of die gegewe vlak van data-onderskeppingsomvang voldoende is. Tekortkominge wat geïdentifiseer is kan aangespreek word met die metodes soos in *Afdeling 8.6* bespreek word om verfyninge aan die stelsel aan te bring deur die proses soos hier bespreek is, te herhaal.

---

---

## HOOFSTUK 9

---

### SAMEVATTING EN GEVOLGTREKKINGS

---

In hierdie studie, waar daar ondersoek ingestel is na die verbetering van historiese data as basis vir besluitsteuninligting, is dit gevind dat 'n verskeidenheid van aspekte die kwaliteite en kwantiteite van inligting kan beïnvloed wat beskikbaar gestel word vir besluitneming. Aspekte wat wissel vanaf die kwaliteit van die databronne wat gebruik word tot die ontledingstegnieke wat gevolg word om besluitsteuninligting te lewer is ondersoek.

In hierdie hoofstuk word die kerngedagtes wat in die voorafgaande hoofstukke geïdentifiseer en ontwikkel is, saamgevat. As vertrekpunt van die studie is ondersoek ingestel na die aspekte wat die kwaliteit van data in 'n bestaande databasis beïnvloed en hoe verbeterde ontledingstegnieke waarde kan toevoeg tot die data. Die aspekte wat geïdentifiseer is om die inligtingstelsel te verbeter is vervolgens in 'n nuwe omgewing, waar daar nog nie so 'n stelsel bestaan nie, getoets om sodoende 'n geïntegreerde proses te ontwikkel wat gevolg kan word by die ontwerp of herontwerp van 'n nuwe of bestaande inligtingstelsel.

---

#### 9.1 BESTAANDE DATABASIS

---

Tydens die ontleding van 'n bestaande kostedatabasis is daar na verskeie aspekte gekyk, nie net om die kwaliteit en kwantiteit van die data te verbeter nie, maar ook om te identifiseer hoe die inligting wat vanuit die databasis verkry word, beter aangewend kan word om besluitneming te ondersteun.

---



### 9.1.1 PROSESMODELLERING

---

- Prosesmodellering bied die geleentheid om die konstruksieprosesse wat bestaan uit aktiwiteite, mense, materiaal en/of data asook die transformasie daarvan te modelleer om sodoende die raamwerk daar te stel waarbinne datavloei ontleed kan word om die interaksie en sintese van data te bepaal. In die model wat vir die bestaande databasis opgestel is, is daar 'n aantal omslagtige prosesse gevind wat die kwaliteit van die data kon beïnvloed.
- 'n Verskeidenheid van hoëvlak modelle is al ontwikkel om die verskillende fases van konstruksieprojekte mee te modelleer, maar slegs die *Generiese Ontwerp en Konstruksie Prosesprotokol* is omvattend genoeg opgestel om al die fases in te sluit. Die prosesmodel wat in hierdie studie ontwikkel is om die onderskepping van konstruksiekostedata te toon stem feitlik presies ooreen met bogenoemde model, maar is meer gefokus op die kosteberekening prosesse.
- Die kompleksiteit van die laevlak modelle en die eenmaligheid van projekte, weens die fragmentasie van die siviele ingenieursomgewing en die aard van konstruksieprojekte, noodsaak dat prosesmodelle getoets moet word vir onder andere korrektheid en volledigheid. Die algoritmes en datastrukture wat beskikbaar is in versameling- en grafiekteorie is gebruik om funksionaliteite te ontwikkel wat gebruik kan word om te toets vir korrektheid en volledigheid van detail prosesmodelle om sodoende die ontwerp van sulke modelle te verbeter. Die rekenaararmatige implementasie van hierdie funksionaliteite het getoon dat dit wel moontlik is om vir hierdie aspekte te toets.

### 9.1.2 DATABASISONTWIKKELING

---

- Die ontwerp van 'n databasis kan die kwaliteit van die data in die databasis en die uitbreikbaarheid daarvan beïnvloed. Meer omvattende fout identifiseringsmeganismes is geïdentifiseer wat die kwaliteit van data wat in die databasis geplaas word, kan verbeter. 'n Herontwerpte datamodel is voorgestel wat voldoen aan die onderskeie normaalvorms om onnodige duplisering van data wat aanleiding gee tot inkonsekwentheid, te beperk.
  - Die vlak van spesifisering van items in 'n databasis met behulp van standaardspesifikasies het 'n bepalende invloed op die veranderlikheid van die data. Strenger spesifikasie gee aanleiding tot 'n verlaging in veranderlikheid, maar kan ook terselfdertyd aanleiding gee tot die ontstaan van yl bevolkte databasisse weens oorspesifisering.
-

- Die inligting afkomstig van kosteberekeninge tydens die verskillende fases van die konstruksieproses, soos onder andere die tenderverslag en die projekvoltooiingsverslag, kan maklik in databasismodelle wat op die relasionele databasisstruktuur gebaseer is, gestoor word. Die datamodel wat ontwikkel is, is omvattend genoeg om al die inligting wat met konstruksiekoste, spesifikasie (hetsy item of projek) en regsdokumente verbandhou, te kan stoor.
- Databasisontwikkeling is 'n iteratiewe proses wat nie net eenmalig uitgevoer word nie, maar wat van tyd tot tyd herhaal moet word soos meer inligting beskikbaar word oor die besprekingsruimte wat die databasis modelleer. Hierdie herontwikkeling kan egter nie afsonderlik van die ander komponente wat deeluitmaak van die inligtingstelsel, geskied nie.

---

### 9.1.3 KWANTITATIEWE RISIKO-ANALISE

---

- Die deterministiese kostefunksie, op komponent- of detailvlak, kan uitgebrei word om onsekerhede in die veranderlikes van die kostefunksie stochasties te modelleer. Die totale koste van verskillende alternatiewe van projekte of tenders kan sodoende met mekaar vergelyk word in die lig van die onsekerhede wat bestaan oor die werklike waardes van die basiese veranderlikes (hoeveelhede en eenheidtariewe).
  - Die tipe verdelings wat in die kostefunksie gebruik word om kostes stochasties mee te modelleer, beïnvloed die akkuraatheid van die probabilistiese kostefunksie. Dit is gevind dat die log-normaalverdeling 'n geskikte generiese verdeling is wat die eienskappe soos skeefheid en onbegrensdeheid besit om stochastiese kostes realisties mee te modelleer.
  - Korrelasiemeganismes wat bestaan tussen die basiese veranderlikes van die kostefunksie beïnvloed die resultate van die ontledings. Hierdie korrelasies behoort dus in ag geneem te word, want dit blyk dat ignorering hiervan lei tot die onderskatting van projekkkostes.
  - Die Limietstaat Kostefunksie- (LSKF) en Monte Carlo-simulasietegniek (MCS) bied die geleentheid om die kostefunksie se veranderlikes stochasties te modelleer en om die effek van korrelasies tussen hierdie veranderlikes in berekening te bring. Die twee metodes, alhoewel fundamenteel verskillend, lewer in essensie dieselfde resultate. Selfs die benaderde oplossingsmetodes van die LSKF-tegniek lewer dieselfde resultate as die makliker implementeerbare MCS-tegniek, maar met die voordeel dat dit parametries insig verleen in hoe risiko versprei is tussen die stochasties gemodelleerde veranderlikes van die kostefunksie.
-

- Databasisse wat historiese projekdata stoor kan aangewend word om die verdelings te bepaal wat die stochastiese veranderlikes in 'n probabilistiese model beskryf. Die veranderlikheid van die data weens epistemiese veranderlikheid veroorsaak egter dat die tegnieke uitermatige hoë projek-kostes lewer as hoë vlakke van betroubaarheid vereis word. Die nut van die probabilistiese tegnieke word dus beperk deur die beskikbaarheid van betroubare data.

---

#### 9.1.4 DATA-ONTLEDING

---

- Die ontleding van historiese data het 'n tweeledige toepassing, dit kan aangewend word in kwantitatiewe risiko-analise vir besluitsteundoeleindes soos in die voorafgaande afdeling, maar dit kan ook aangewend word om te identifiseer waar probleme in die inligtingstelsel voorkom en verbeter behoort te word.
  - Die verskynsel van 'n klein hoeveelheid veranderlikes met 'n groot impak kom duidelik na vore in die data wat ontleed is. 'n Noemenswaardige gedeelte van die totale projek-koste kan probabilisties gemodelleer word met 'n klein aantal stochastiese veranderlikes. Verdere uitbreidings aan die databasis om die veranderlikheid van die historiese data te verklaar kan dus gefokus word op hierdie veranderlikes.
  - Inflasiegebaseerde aanpassings van historiese eenheidtariefdata van padkonstruksie items verwyder nie altyd volledig die dalende of stygende tendens wat die historiese data toon nie en veroorsaak dat daar residuele temporale veranderlikheid voorkom. Indien 'n tendensaanpassing gedoen word, kan die veranderlikheid geminimeer word.
  - Orde grootte foute in die data het 'n noemenswaardige impak op die veranderlikheid van die datastel. Deur die verwydering van hierdie foute met filters kan 'n substansiële verlaging in die veranderlikheid te weeg gebring word. Hierdie tipe foute behoort egter voorkom te word indien daar goeie foutidentifiseringsmeganismes in die databasis is.
  - Sterk korrelasie kom voor tussen sommige items se eenheidtariewe en, alhoewel minder so, ook tussen hoeveelhede van items, maar geen noemenswaardige korrelasie kom tussen items se eenheidtariewe en hoeveelhede voor nie. Die sterkte van die korrelasiemeganisme wat bestaan tussen items se eenheidtariewe dui op een-of-ander gedeelde eksterne invloed, omdat daar nie 'n kousale verband tussen eenheidtariewe bestaan nie.
-

- Epistemiese veranderlikheid, weens onvolledige inligting, word oorwegend veroorsaak deur ruimtelike veranderlikheid nadat temporale en data veranderlikheid (weens foute in die data) verwyder is. Verbeterde data-onderskepping oor die omgewing waarbinne projekte plaasvind kan lei tot 'n verlaging in hierdie tipe veranderlikheid.

---

### 9.1.5 UITBREIDING VAN DATA-ONDERSKEPPING

---

- Die epistemiese veranderlikheid wat in die kostedata geïdentifiseer is kan om twee redes ontstaan, nl. 'n tekort aan spesifikasies wat die items streng definieer of 'n tekort aan projekbeskrywende inligting. Aangesien die items streng gedefinieer is in die bestaande databasis is dit slegs die tekort aan projekbeskrywende inligting wat die epistemiese veranderlikheid veroorsaak.
- 'n Tekort aan projekbeskrywende inligting op detailkosteberekeningvlak veroorsaak dat by die hoëvlak kosteberekening inligting ontbreek om die projekte te karakteriseer. Indien die inligtingonderskepping uitgebrei word, moet dit in ag geneem word dat van die inligting wat op detailvlak oor die projekomgewing onderskep word, ook van nut kan wees vir die hoër vlakke van kosteberekening. 'n Duidelike aanwendingsmodel is saamgestel om die nut van die verskillende tipes kostedata binne die verskillende tipes kosteberekening te toon.
- Die aspekte wat die kostes van projekte beïnvloed is met invloedsmodelle voorgestel om die verwantskappe wat bestaan tussen die invloedsfaktore en die komponente waaruit die kostes saamgestel is, te toon. Hierdie modelle is opgestel vir die verskillende vlakke van kosteberekening en in die geval van die detailkosteberekening ook vanuit die verskillende perspektiewe van die kontrakteur en die raadgewende ingenieur.
- Weens die magdom van invloedfaktore wat op detailvlak die kostes van items kan beïnvloed, is 'n metode ontwikkel wat gebaseer is op die korrelasies wat bestaan tussen die eenheidtariewe van items. Hierdie metode bied die moontlikheid om die beduidendste invloedfaktore te bepaal en sodoende die onderskepping van addisionele inligting en die databasis-uitbreiding te optimeer.

---

## 9.2 NUWE OMGEWING

---

Die aspekte wat in die padbou-omgewing geïdentifiseer en ontwikkel is, is vervolgens aangewend in 'n nuwe omgewing, nl. damkonstruksie, waar daar

---

nog nie so 'n stelsel bestaan nie en is dan veralgemeen om 'n generiese proses te vorm wat vir nuwe of bestaande inligtingstelsels gebruik kan word.

---

### 9.2.1 DAMKONSTRUKSIE

---

- Die dambou omgewing verskil van die padbou omgewing ten opsigte van die vlak van ontwikkeling van 'n koste-inligtingstelsel oorwegend weens die verskil in die struktuur van die twee omgewings. By die padbou omgewing, waar die tradisionele kliënt-ingenieur-kontraakteur verhouding voorkom, het die behoefte aan standaardspesifikasie veroorsaak dat 'n databasis met gestandaardiseerde items kon ontwikkel. Hierteenoor vorm die kliënt-ingenieur-kontraakteur 'n eenheid in die dambou omgewing. Die behoefte het dus nog nie so sterk ontwikkel vir gestandaardiseerde spesifikasies nie en gevolglik is die ontwikkelingsmoontlikheid van 'n gestandaardiseerde databasis van eenheidtariewe beperk.
- Die behoeftes wat in die padbou omgewing geïdentifiseer is ten opsigte van die onsekerhede wat voorkom by die verskillende vlakke van kosteberekeninge en verskillende tipes projekte is ook in die dambou omgewing gevind. Die kwantifisering hiervan is veral besig om in die dambou omgewing weens veranderde damveiligheidsstandaarde en vereistes van die publiek vir groter deursigtigheid by dambouprojekte, belangrik te word. Probabilistiese tegnieke is beskikbaar om die onderskeie onsekerhede wat kan voorkom te modelleer, maar daar is 'n behoefte aan goeie kwaliteit databronne.
- Die bronne wat epistemiese veranderlikheid kan veroorsaak in die kostedata van padprojekte is netso van toepassing by die dambou omgewing. Aspekte soos die datavloei, databasis ontwikkeling en uitbreiding van data-onderskepping moet ontleed en korrek ontwikkel word om te verseker dat die veranderlikheid wat hierdie bronne veroorsaak tot 'n minimum beperk word. Die voordele wat 'n geïntegreerde ontwikkeling van hierdie komponente te weeg kan bring word egter beperk deur die tekort aan standaarditems.
- Die behoefte aan standaardspesifikasies vir die bepaling van standaarditems kan omseil word met die definiëring van generiese items. Generiese items word ontwikkel vanaf die raadgewende ingenieurperspektief waar elke item asook die koste komponente wat daarby ingesluit word vasgelê word, asook die faktore wat die komponente kan beïnvloed. Die presiese eienskappe van elke faktor word egter nie soos in die geval van gespesifiseerde items vasgestel nie, maar kan vir elke projek gevarieer word. Hierdie variasie word op 'n gestruktureerde wyse in die databasis onder-

skep om verklaring te bied vir die afwykings wat kan voorkom tussen verskillende projekte se spesifikasies.

---

### 9.2.2 GENERIESE ONTWIKKELINGSMODEL

---

- Die generiese ontwikkelingmodel wat voorgestel word as basis vir die ontwikkeling/herontwikkeling van 'n nuwe/bestaande inligtingstelsel is saamgestel uit die onderskeie aspekte wat in die twee proefvoorbeelde geïdentifiseer is. Hierdie omvattende model integreer nie net die onderskeie individuele aspekte wat die kwaliteit en kwantiteit van die data in die databron kan beïnvloed nie, maar dit toon hoe 'n selfvoerende proses tussen die gebruiker en die databasis ontwikkel wat vir die kosteberamer van 'n nuwe projek en vir latere gebruikers waarde toevoeg tot berekening wat uitgevoer word.
- 'n Geïntegreerde ontwikkelingproses is op die generiese ontwikkelingmodel gebou wat die verbetering van die onderskeie aspekte in die model soos die aanwending van inligting, die datastoor, die datavloei, die omvang van data-onderskepping en die standardisering van items aanspreek. Hierdie stapsgewyse proses berus op 'n inkrementele verbetering van die stelsel soos inligting beskikbaar raak oor die besprekingsruimte en die behoeftes van die gebruikers verander.

---

### 9.3 GEVOLGTREKKINGS

---

Besluitneming wat gebaseer is op ekonomiese uitkomst is 'n inherente komponent van enige siviele ingenieursprojek. Historiese data is die voorkeurbron van inligting om hierdie besluite op te skoei, maar die betroubaarheid (kwaliteit) daarvan is dikwels onvoldoende weens die meer omvangryke databenutting van risiko-gebaseerde ontledingstegnieke. Hierdie tegnieke kan egter waardevolle inligting blootlê ten opsigte van onsekerhede wat voorkom in beraamde kostes, indien die kwaliteit en kwantiteit van data-onderskepping verbeter kan word.

Veranderde omstandighede intern sowel as ekstern tot die konstruksiebedryf noodsaak die ontwikkeling en verbetering van databronne om data meer koste effektief te onderskep, bestuur, manipuleer en analiseer. Die gereedskap wat die ontwikkelinge in die informasietegnologie omgewing bied, bied die geleentheid om omvattende en geïntegreerde databronne te ontwikkel wat die kwantiteit en kwaliteit van historiese databronne vir besluitnemingsdoeleindes kan verbeter. Van hierdie gereedskap is gebruik om drie hoofaspekte vir die verbetering van historiese databronne aan te spreek.

- **Standaarditems:** Die ontwikkeling van 'n kostedatabasis noodsaak die standardisering van die koste-items. In 'n omgewing waar daar reeds
-

bestaande standaardspesifikasies is, is items gewoonlik voldoende gestandaardiseer, maar wanneer dit nie bestaan nie gee dit aanleiding tot groot veranderlikheid in die data. Die metode van generiese items wat in hierdie studie ontwikkel is met behulp van invloedfaktore oorkom hierdie probleem deur die vereiste van standaardspesifikasie te verslap, maar tog op 'n gestruktureerde wyse die faktore wat die veranderlikheid kan veroorsaak te onderskep.

- **Inkrementele ontwikkeling:** Bronne van onsekerheid wat nie verband hou met konstruksieverwante onsekerhede nie kan teruggespeur word na tekortkominge in die wyse waarop historiese databronne bestuur word. Deur 'n inkrementele ontwikkeling waar die onderskeie aspekte wat in hierdie studie geïdentifiseer is soos die datavloei, die datastoor en die beskrywende inligting te verbeter soos inligting vanaf die stelsel beskikbaar word om hierdie bronne te identifiseer kan hierdie onsekerhede verminder word.
- **Selfvoerende proses:** Die karakterisering van die omgewing van nuwe projekte noodsaak addisionele insette van die kosteberamer wat die uitvoering hiervan kan beperk as dit slegs van waarde is vir latere kosteberaminge. Die selfvoerende proses wat egter hier voorgestel word bied 'n meganisme wat die akkurate karakterisering van die projek verseker. Akkurate karakteriseer van 'n nuwe projek deur die kosteberamer bied onmiddellike die voordeel van verbeterde beraamde kostes vir die kosteberamer. Dit bied die nodige motivering om sorgvuldig die karakterisering van 'n projek te doen.

Die aanwending en toepassing van hierdie drie aspekte tydens die ontwikkeling/herontwikkeling van kostedatabronne vir die siviele ingenieursomgewing kan 'n substansiële bydrae maak om op 'n gestruktureerde en optimale wyse die onderskepping, stoor, ontleding en aanwending van data te verbeter. Besluitneming by 'n gegewe vlak van detail en vir 'n spesifieke tipe projek kan sodoende verbeter word om vir die besluitnemer insig te gee in die werklike onsekerheid in die projek koste.

---

## 9.4 VERDER STUDIE

---

Die volgende aspekte is in die studie geïdentifiseer wat met vrug gebruik kan word as onderwerpe vir verdere studie:

- **Prosesmodellering** – Die aanwending van prosesmodelle in hierdie studie is oorwegend vir die ontwikkeling van die onderliggende datavloei modelle gebruik. Prosesmodelle bied egter die moontlikheid om ondersoek in te stel na optimalisering van die komplekse prosesse wat voorkom tydens kon-
-

struksieprojekte. Indien die bestaande argitektuur wat in hierdie studie gebruik is, uitgebrei word om tydsverloop te modelleer kan die toepassingmoontlikhede aansienlike uitgebrei word vir die optimering en verbetering van konstruksieprosesse. Die toepassingmoontlikhede bied 'n wye veld van navorsinggeleenthede.

- *Generiese Ontwerp en Konstruksie Prosesprotokol* – Die GOKPP bied die geleentheid om die basis van verskillende tipes konstruksieprojekte se siviele ingenieurswese prosesse te standaardiseer. Veral in die staatsdiensomgewing waar verskeie departemente betrokke is by verskillende tipes infrastruktuurontwikkeling kan dit van besondere nut wees vir die bestuur en monitering van projekte. Die aanpasbaarheid van die GOKPP en die waarde wat dit kan toevoeg tot bestuur en monitering moet egter getoetse word in die onderskeie omgewings, asook vir die verskillende tipes kliënt-ingenieur-kontrakteur strukture wat kan voorkom.
- *Globale projek databasismodel* – Die databasismodel wat hier ontwikkel is, is spesifiek vir die onderskepping van inligting afkomstig uit die onderskeie kosteberekeningfasies en gepaard gaande inligting soos regsdocumente en standaardspesifikasies. Kosteberekeningdata is egter 'n klein, alhoewel belangrik, onderafdeling van die globale versameling van projekdata wat al die tegniese, omgewings, regs- en finansiële aspekte insluit. Die ontwikkeling van databasismodelle wat generiese genoeg is om sulke inligting op 'n effektiewe wyse te stoor behoort ondersoek te word nie net vir projekadministrasie nie, maar ook om te kan dien as bronne van historiese inligting.
- *Invloedfaktore* – Die suksesvolle aanwending van neurale netwerke is al in vele omgewings getoon. Inligting word egter benodig om die omgewing waarvoor die neurale netwerk voorspellings moet maak te karakteriseer. Dit is egter nie altyd vooraf bekend watter van die omgewingseienskappe die grootste bydrae tot die geaardheid van die omgewing maak voordat inligting daarvoor onderskep is nie. Die metode wat in hierdie studie voorgestel word om invloedfaktore saam met werklike korrelasie te gebruik om sodoende die belangrikste invloedfaktore te bepaal bied 'n meganisme om dit te kan doen. Hierdie toepassingsmoontlikheid behoort verfyn en uitgebrei te word.



## HOOFSTUK 10

---

### VERWYSINGS

---

- [1] AACE (1997), *Cost Estimate Classification System*, AACE International Recommended Practise No. 17R-97, August.
- [2] AACE (1998), *Cost Estimate Classification System – As Applied in Engineering Procurement and Construction for the Process Industries*, AACE International Recommended Practise No. 18R-97, June.
- [3] Abudayyeh (1991), O. Y., "Design of Construction Industry Information Management Systems", *ASCE Journal of Construction Engineering and Management*, Vol. 117, No. 4, December, pp. 698-715.
- [4] Adams (1993), T. M., "Knowledge Representation and Processing in Relational Data Base", *ASCE Journal of Computing in Civil Engineering*, Vol. 7, No. 2, April, pp. 238-255.
- [5] Adeli (2001), H., Karim, A., *Construction Scheduling, Cost Optimisation and Management*, Spon Press, London.
- [6] Adeli (1998), H., Wu, M., "Regularization Neural Network for Construction Cost Estimation", *ASCE Journal of Construction Engineering and Management*, Vol. 124, No. 1, February, pp. 18-24.
- [7] Akinci (1998), B., Fischer, M., "Factors Affecting Contractors' Risk of Cost Overburden", *ASCE Journal of Management in Engineering*, Vol. 14, No. 1, January/February, pp. 67-76.
-

- [8] Alsugair (1998), A. M., Al-Qudrah, A. A., "Artificial Neural Network Approach for Pavement Maintenance", *ASCE Journal of Computing in Civil Engineering*, Vol. 12, No. 4, October, pp. 249-256.
- [9] Amekudzi (2000), A., McNeil, S., "Capturing Data and Model Uncertainties in Highway Performance Estimation", *ASCE Journal of Transportation Engineering*, Vol. 126, No. 6, November/December, pp. 455-463.
- [10] Andersen (1995), G. R., Torrey, V. H., "Function Based Condition Indexing for Embankment Dams", *ASCE Journal of Geotechnical Engineering*, August, pp. 579 – 588.
- [11] Andersen (2001), G. R., Chouinard, L. E., et al., "Risk Indexing Tool to Assist in Prioritizing Improvements to Embankment Dam Inventories", *ASCE Journal of Geotechnical and Geo-environmental Engineering*, April, pp. 325 – 334.
- [12] Ang (1975), A. H. S., Tang, W. H., *Probability Concept in Engineering Planning and Design*, Volume 1, John Wiley & Sons, New York.
- [13] Ang (1984), A. H. S., Tang, W. H., *Probability Concept in Engineering Planning and Design*, Volume 2, John Wiley & Sons, New York.
- [14] *ARIS (1991) Toolset*, Version 4.0, IDS Scheer AG, Germany.
- [15] Attalla (2001), M., Hegazy, T., "Predicting Cost Deviation in Reconstruction Projects: Artificial Neural Networks versus Regression", *ASCE Journal of Construction Engineering and Management*, Vol. 129, No. 4, July/August, pp. 405 – 411.
- [16] Back (2000a), W. E., Maxwell, D. A., Isidore, L. J., "Activity Based Costing as a Tool for process Improvement Evaluations", *ASCE Journal of Management in Engineering*, Vol. 16, No. 2, March/April, pp. 49-58.
- [17] Back (2000b), W. E., Moreau, K. A., "Cost and Schedule Impacts of Information Management on EPC Process", *ASCE Journal of Management in Engineering*, Vol. 16, No. 2, March/April, pp. 59-70.
- [18] Balena (1999), F., *Programming Microsoft Visual Basic 6.0*, Microsoft Press, Washington.
- [19] Bell (1987), L. C., Kaminsky, A., "Data Base for Preliminary Cost Estimating", *ASCE Journal of Transportation Engineering*, Vol. 113, No. 4, July, pp. 341-347.

- [20] Benyon (1990), D., *Information and Data Modelling*, Blackwell Scientific Publications, Oxford.
- [21] Bester (2001), A., Huhnt, W., "A Tendering Process Model for An Application of Reliability Analysis", *CIB W78 Conference Paper*, South Africa.
- [22] Birrell (1980), G. S., "Construction Cost Estimating in the Design Process", *ASCE Journal of the Construction Division*, Vol. 106, No. 4, December, pp. 551-566.
- [23] Boles (2000), W. W., Back, W. E., Fry, G. T., "Defining Triangular Probability Distributions from Historical Cost Data", *ASCE Journal of Construction Engineering and Management*, Vol. 126, No. 1, January/February, pp. 29-37.
- [24] Booch (1999), G., Rumbaugh, J., Jacobson, I., *The Unified Modelling Language User Guide*, Addison Wesley Publishing Co., Reading.
- [25] Bradley (1990), R. M., Powell, M. G., Soulsby, M. R., "Quantifying Variations in Project-Cost Estimates", *ASCE Journal of Management in Engineering*, Vol. 6, No. 1, January, pp. 99-106.
- [26] Budd (1994), T. A., *Classic Data Structures in C++*, Addison-Wesley Publishing Company, New York.
- [27] Carr (1989), R. I., "Cost-estimating Principles", *ASCE Journal of Construction Engineering and Management*, Vol. 115, No. 4, December, pp. 545-551.
- [28] Chao (1994), L., Skibniewski, M. J., "Estimating Construction Productivity: Neural Network-Based Approach", *ASCE Journal of Computing in Civil Engineering*, Vol. 8, No. 2, April, pp. 234-251.
- [29] Chen (1976) P., "The Entity-Relationship Model – Toward a Unified View of Data", *ACM Transactions on Databases Systems*, Vol. 1, No. 1 pp. 9-36, March.
- [30] Cleveland (1999), A. B., "Harvesting the Value of Information", *ASCE Journal of Management in Engineering*, Vol. 15, No. 4, July/August, pp. 37-42.
- [31] Codd (1970) E. F., "A Relational Model of Data for Large Shared Data Banks", *Communications of the ACM*, Vol. 13, No. 6, June.
- [32] Cooper (1998), R., Kagioglou, M., Aouad, G., Hinks, J., Sexton, M., Sheath, D., "The Development of a Generic Design and Construction
-

- Process", *European Conference, Product Data Technology (PDT) Days*, Building Research Establishment, March 1998, Watford, UK.
- [33] *Crystal Ball (2000)*, Decisioneering – Risk Analysis Software, Version 2000.
- [34] Date (1995), C. J., *An Introduction to Database Systems*, Addison-Wesley Publishing Company, Reading.
- [35] Du Buisson (2003), N., *Persoonlike Onderhoud*, Departement van Waterwese en Bosbou, Pretoria, Desember.
- [36] Easa (1988), S. M., "Earthwork Allocations with Linear Unit Costs", *ASCE Journal of Construction Engineering and Management*, Vol. 114, No. 4, December, pp. 641-655.
- [37] Easa (1987), S. M., "Earthwork Allocations with Non-constant Unit Costs", *ASCE Journal of Construction Engineering and Management*, Vol. 113, No. 1, March, pp. 34-50.
- [38] Elmasri (1994), R., Navathe, S. B., *Fundamentals of Database Systems*, 2nd Edition, Benjamin/Cummings Publishing Company, Redwood City.
- [39] El-Rayes (2001), K., Moselhi, O., "Impact of Rainfall on the Productivity of Highway Construction", *ASCE Journal of Construction Engineering and Management*, Vol. 127, No. 2, March/April, pp. 125-131.
- [40] Faghri (2000), A., "Transportation Project Cost Estimation using Artificial Neural Networks", *Engineering Valuation and Cost Analysis*, Vol. 2, pp. 465-480.
- [41] Fayek (1989), A., "Competitive Bidding Strategy Model and Software System for Bid Preparation", *ASCE Journal of Construction Engineering and Management*, Vol. 124, No. 1, January/February, pp. 1-10.
- [42] Froese (1995), T., "Models of Construction Process Information", *ASCE Journal of Computing in Civil Engineering*, March.
- [43] Gates (1980), M., Scarpa, A., "Criteria for the Selection of Construction Equipment", *ASCE Journal of the Construction Division*, Vol. 106, No. 2, June, pp. 207-219.
- [44] Giarratano (1994), J., Riley, G., *Expert Systems: Principles and Programming*, 2nd Edition, PWS Publishing Company, Boston.
-

- [45] Goh (1995), A. T. C., "Modelling Soil Correlations Using Neural Networks", *ASCE Journal of Computing in Civil Engineering*, Vol. 9, No. 4, October, pp. 275-278.
- [46] Gordon (2001), J. L., "Unit Price Contracts or Design Build?" *Hydropower and Dams*, Issue 5, pp. 99 – 102.
- [47] Gross (1998), J., Yellen, J., *Graph Theory and Its Applications*, CRC Press, Baco Raton.
- [48] Halpin (2001), T., *Information Modelling and Relational Databases*, Morgan Kaufmann, San Francisco.
- [49] Hartford (2000), D., "Dam Risk and the Owner's Dilemma", *International Water Power and Dam Construction*, September, pp. 18 – 23.
- [50] Hegazy (1998), T., Ayed, A., "Neural Network Model for Parametric Cost Estimation of Highway Projects", *ASCE Journal of Construction Engineering and Management*, Vol. 124, No. 3, May/June, pp. 210-218.
- [51] Herbsman (1984), Z., "Long-range Forecasting Highway Construction Costs", *ASCE Journal of Construction Engineering and Management*, Vol. 109, No. 4, December, pp. 423-434.
- [52] Høeg (1998), K., "New Dam Safety Legislation and the use of Risk Analyses", *Hydropower and Dams*, Issue 5, pp. 85 – 89.
- [53] Howard (1991), H. C., "Project-Specific Knowledge Bases in AEC Industry", *ASCE Journal of Computing in Civil Engineering*, Vol. 5, No. 1, January, pp. 25-41.
- [54] Ince (1995), D. C., *An Introduction to Discrete Mathematics, Formal System Specification and Z*, Oxford University Press, Oxford.
- [55] Juran (2003), J. M., "The Non-Pareto Principle; Mea Culpa", Internet artikel, [www.juran.com/pdf/SP7518.doc](http://www.juran.com/pdf/SP7518.doc).
- [56] Ker-Fox (1998), G. M., *The Limit State Cost Function as a Risk Management Tool for Construction Projects*, Masters Degree Thesis, University of Stellenbosch Printers, Stellenbosch.
- [57] Knight (2002), K., Fayek, A. R., "Use of Fuzzy Logic for Predicting Design Cost Overruns on Building Projects", *ASCE Journal of Construction Engineering and Management*, Vol. 128, No. 6, November/December, pp. 503 – 512.
-

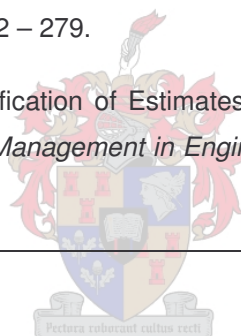
- [58] Kroon (2003), J., *Persoonlike Onderhoud*, Departement van Waterwese en Bosbou, Pretoria, Desember.
- [59] Levitt (1980), R. E., Logcher, R. D., and Ashley, D. B., "Allocating Risk and Incentive in Construction", *ASCE Journal of the Construction Division*, Vol. 106, No. 3, September, pp. 297-305.
- [60] Litwin (1997), P., Getz, K., Gilbert, M., *Access 97: Developer's Handbook*, 3rd Edition, Sybex Inc.
- [61] Manber (1989), U., *Introduction to Algorithms*, Addison Wesley, New York.
- [62] Mar (1989), B. W., Palmer, R. N., "Does Civil Engineering Need System Engineering?", *ASCE Journal of Professional Issues in Engineering*, Vol. 115, No. 1, January, pp. 45-52.
- [63] Mayberry (2000), J. P., *The Foundations of Mathematics in the Theory of Sets*, Cambridge University Press, Cambridge.
- [64] McDonald (2001), L. A., Hartford, D. N. D., et al., *Risk Assessment in Dam Safety Management – A Reconnaissance of Benefits, Methods, and Current Applications*, ICOLD Bulletin on Risk Assessment, August.
- [65] Mitropoulos (2000), P., Tatum, C. B., "Forces Driving Adoption of New Information Technologies", *ASCE Journal of Construction Engineering and Management*, Vol. 126, No. 5, September/October, pp. 340-348.
- [66] Nel (1987), M., *Padkonstruksie: 'n Ondersoek na Tenderpryse (1977-1986)*, Meestersgraad Tesis, Universiteit van Stellenbosch Drukkers, Stellenbosch.
- [67] Oberlender (2001), G. D., Trost, S. M., "Predicting Accuracy of Early Cost Estimates Based on Estimate Quality", *ASCE Journal of Construction Engineering and Management*, Vol. 127, No. 3, May/June, pp. 173-182.
- [68] O'Neil (2001), P., *Database Principles, Programming and Performance*, 2nd Edition, Morgan Kaufmann Publishers, San Francisco.
- [69] Oosthuizen (1985), C., *A Methodology for the Probabilistic Evaluation of Dams*, Department of Water Affairs and Forestry, Pretoria, July.
- [70] Pahl (2000), P. J., Damrath, R., *Mathematische Grundlagen der Ingenieurinformatik*, Springer Verlag, Berlin.

- [71] Palmer (1987), R. N., Tull, R. M., "Expert System Drought Management Planning", *ASCE Journal of Computing in Civil Engineering*, Vol. 1, No. 4, October, pp. 284-297.
- [72] Peurifoy (1989), R. L., Oberlender, G. D., *Estimating Construction Costs*, McGraw-Hill Inc., U.S.A.
- [73] Piff (1992), M., *Discrete Mathematics*, Cambridge University Press, Cambridge.
- [74] Porro (2001), B., "The Contribution of Insurers and Re-insurers to Risk Management", *Conference Report on International Conference on Safety, Risk and Reliability - Trends in Engineering*, Malta.
- [75] Rand (1995), *A Million Random Digits with 100000 Normal Deviates*, Rand Corporation.
- [76] Rasdorf (2000), W., Shuller, E., Poole, R., Abudayyeh, O., Robson, F., "Information Management at State Highway departments: Issues and Needs", *ASCE Journal of Transportation Engineering*, Vol. 126, No. 2, March/April, pp. 134-142.
- [77] Raftery (1994), J., *Risk Analysis in Project Management*, E & FN Spon, London.
- [78] Reiser (1992), M., Wirth, N., *Programming in Oberon*, Addison-Wesley Publishing Company, New York.
- [79] Render (2000), B., Stair, R. M., *Quantitative Analyses for Management*, 7th Edition, Prentice-Hall, New Jersey.
- [80] Rob (1997), P., Coronel, C., *Database Systems*, 3rd Edition, Thomson Publishing Company, New York.
- [81] Roy (2001), R., Forsberg, S., et al., "Risk Analyses of Parametric Cost Estimates within a Concurrent Engineering Environment", Cranfield University, Cranfield, United Kingdom.
- [82] Russel (1995), S., Norvig, P., *Artificial Intelligence*, Prentice-Hall, New Jersey.
- [83] Sanral (2002), Ongepubliseerde interne dokument.
- [84] Sanvido (1990), V., "Applying Computer-integrated manufacturing concepts to construction", *ASCE Journal of Construction Engineering and Management*, Vol.116, No. 2, June.

- [85] Scheer (1992), A. W., *Architecture of Integrated Information Systems*, Springer-Verlag, Berlin.
- [86] Semple (1994), C., Hartman, T., Jergeas, G., Construction "Claims and Disputes: Causes and Cost/Time Overruns", *ASCE Journal of Construction Engineering and Management*, Vol. 120, No. 4, December, pp. 785-795.
- [87] Severn (1995), C. G., Severn, S. R. T., "Avoiding the Database Quagmire", *ASCE Journal of Management in Engineering*, Vol. 11, No. 3, May/June, pp. 18-20.
- [88] Sheath (1996), D., Woolley, H., Cooper, R., Hinks, J., Aoud, G., "A Process for Change - The Development of a Generic Design and Construction Process Protocol", *Proceedings of the International Construction Information Technology Conference (InCIT '96)*, The Institution of Engineers - Australia, Sydney, April.
- [89] Smith (1996), A, "Cost Estimation Predictive Modelling: Regression versus Neural Network", *The Engineering Economist*, Vol. 41, No. 4, November.
- [90] Swamy (1992), M. N. S., Thulasiraman, K., *Graphs, Networks and Algorithms*, John Wiley & Sons, Inc., New York.
- [91] Tawfik (1997), M., Ibrahim, A., Fahmy, H., "Hysteresis Sensitive Neural Network for Modelling Rating Curves", *ASCE Journal of Computing in Civil Engineering*, Vol. 11, No. 3, July, pp. 206-212.
- [92] Teicholz (1993), P., " Forecasting Final cost and Budget of Construction Projects", *ASCE Journal of computing in Civil Engineering*, Vol. 7, No. 4, October, pp. 511-529.
- [93] Thompson (1992), P., Perry, J., *Engineering Construction Risks*, Thomas Telford Ltd., London.
- [94] Touran (1993), A., "Probabilistic Cost Estimating with Subjective Correlations", *ASCE Journal of Construction Engineering and Management*, Vol. 119, No. 1, March, pp. 58-71.
- [95] Touran (1997), A., Suphot, L., "Rank Correlations in Simulating Construction Costs", *ASCE Journal of Construction Engineering and Management*, Vol. 123, No. 3, September, pp. 297-301.



- [96] Touran (1992), A., Wiser, E. P., "Monte Carlo Technique with Correlated Random Variables", *ASCE Journal of Construction Engineering and Management*, Vol. 118, No. 2, June, pp. 258-272.
- [97] Trost (2003), S. M., Oberlender, G. D., "Predicting Accuracy of Early Cost Estimates using Factor Analysis and Multivariate Regression", *ASCE Journal of Construction Engineering and Management*, Vol. 129, No. 2, March/April, pp. 198 – 204.
- [98] Walpole (1993), R. E., Myers, R. H., *Probability and Statistics for Engineers and Scientists*, 5th Edition, Macmillan Publishing Company, New York.
- [99] Williams (1994), T. P., "Predicting Changes in Construction Cost Indexes Using Neural Networks", *ASCE Journal of Construction Engineering and Management*, Vol. 120, No. 2, June, pp. 306-320.
- [100] Wilmont (2003), C. G., Cheng, G., "Estimating Future Highway Construction Costs", *ASCE Journal of Construction Engineering and Management*, Vol. 129, No. 3, May/June, pp. 272 – 279.
- [101] Yeo (1990), K. T., "Risks, Classification of Estimates, and Contingency Management", *ASCE Journal of Management in Engineering*, Vol. 6, No. 4, October, pp. 458-470.

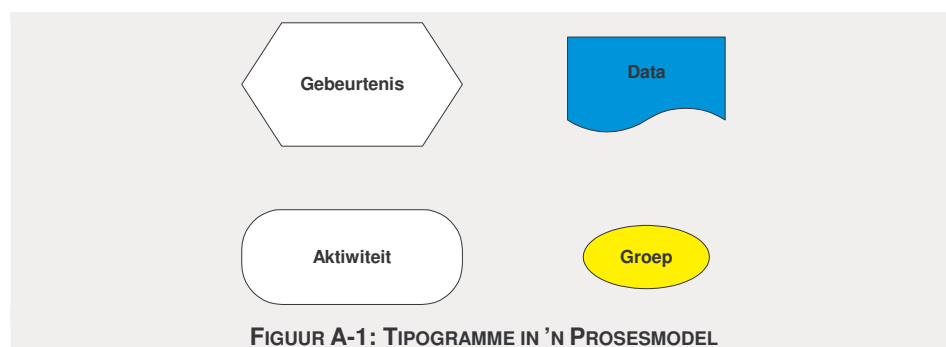


## AANHANGSEL A

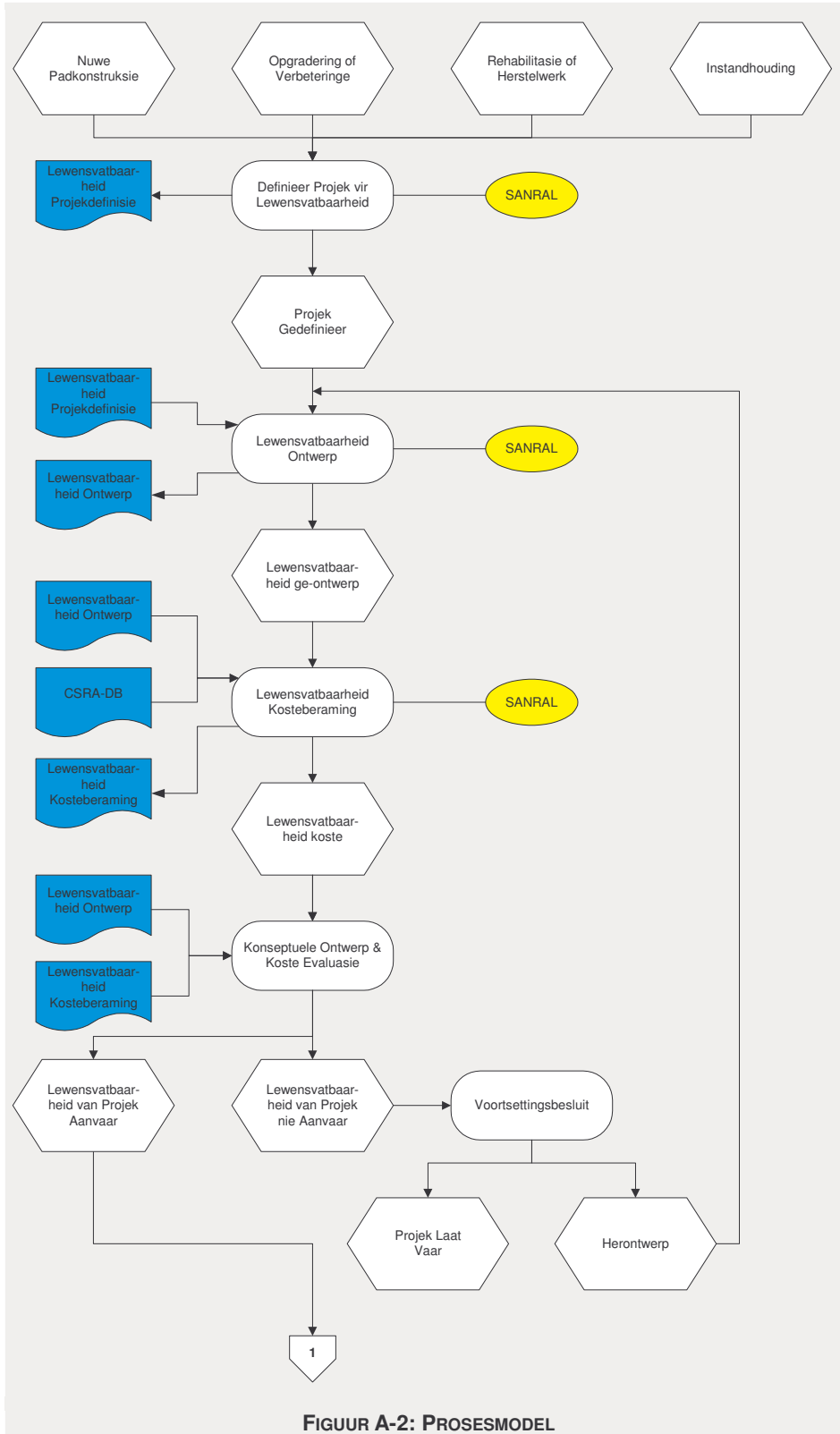
### PADKONSTRUKSIEPROSEMODEL

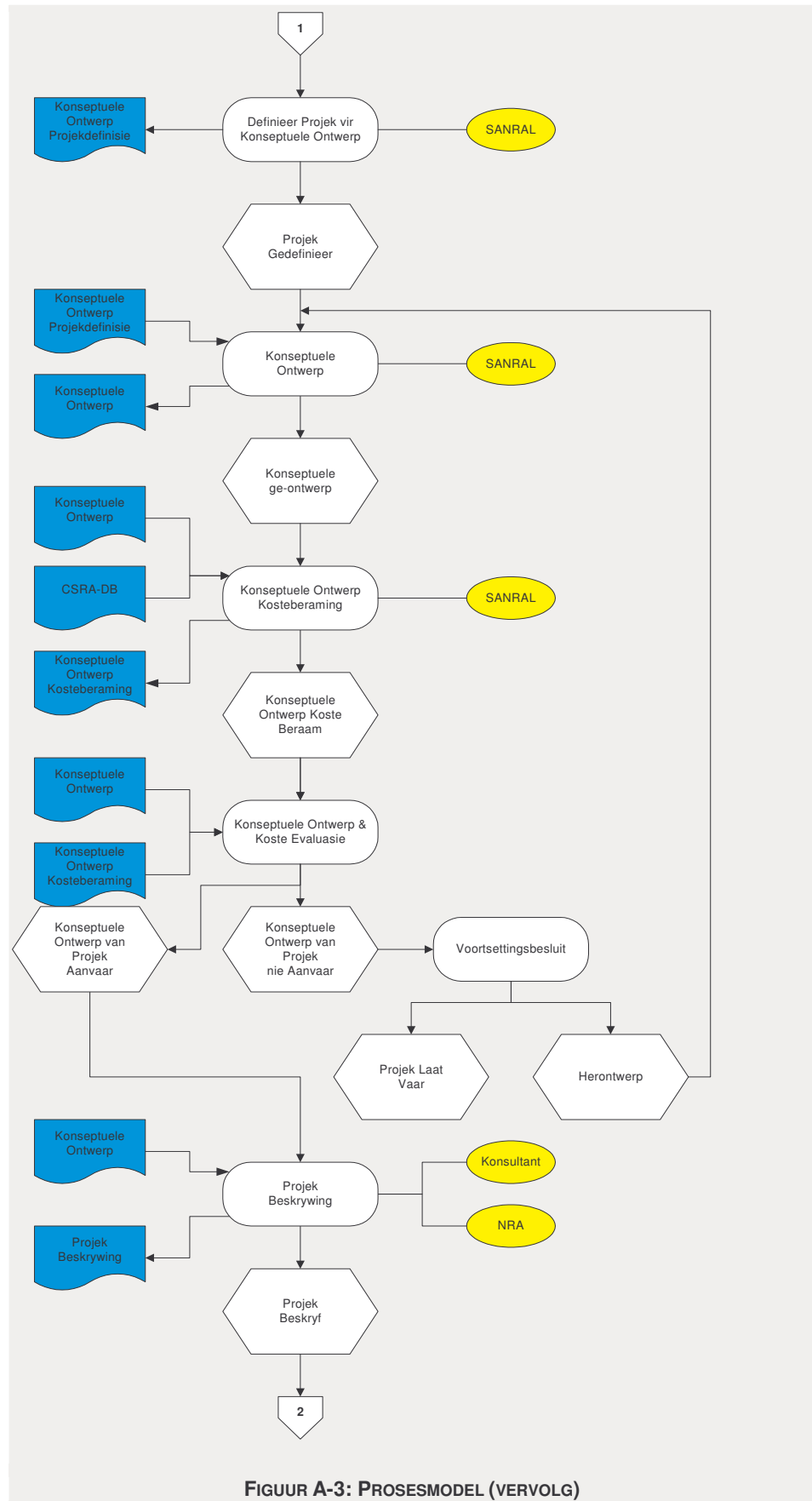
Hieronder volg 'n generiese prosesmodel van 'n padkonstruksieprojek met die fokus op kosteberekeninge. Die model toon hoe die aktiwiteite en gebeurtenisse op mekaar volg, wie daarby betrokke is en watter data gebruik of geskep word.

'n Aktiwiteit stel 'n werkstaak voor wat 'n periode het waarbinne dit verrig word. Gebeurtenisse word gebruik om toestande voor te stel wat 'n punt in die tydsraamwerk van die proses aandui. Groepe is mense of toerusting wat 'n taak verrig of monitor of daarby betrokke is. Vir elke aktiwiteit is daar slegs een groep wat verantwoordelik is daarvoor, maar meer as een groep kan betrokke wees. Data dui die inligting aan wat benodig word om 'n taak te kan uitvoer of ook die inligting wat geskep word tydens die uitvoering van 'n taak. Die onderskeid tussen gebruik en skep van data by 'n aktiwiteit word aangedui deur die rigting van die pyl. As data geskep word, is die pyl gerig weg van die aktiwiteit na die data toe en as data gebruik word is die pyl gerig na die aktiwiteit vanaf die data. Die tipogramme van die vier komponente word in *Figuur A-1* getoon.

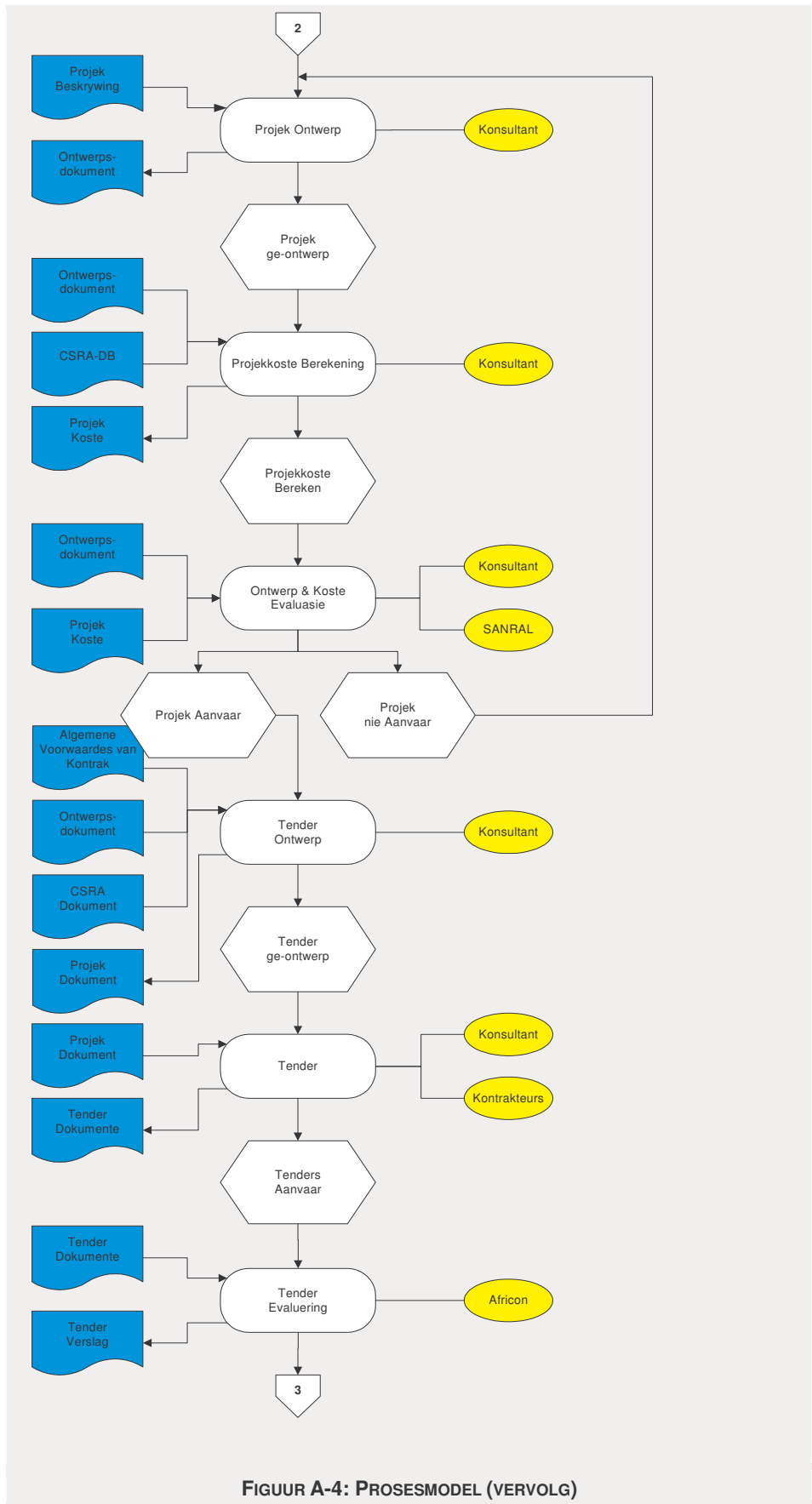


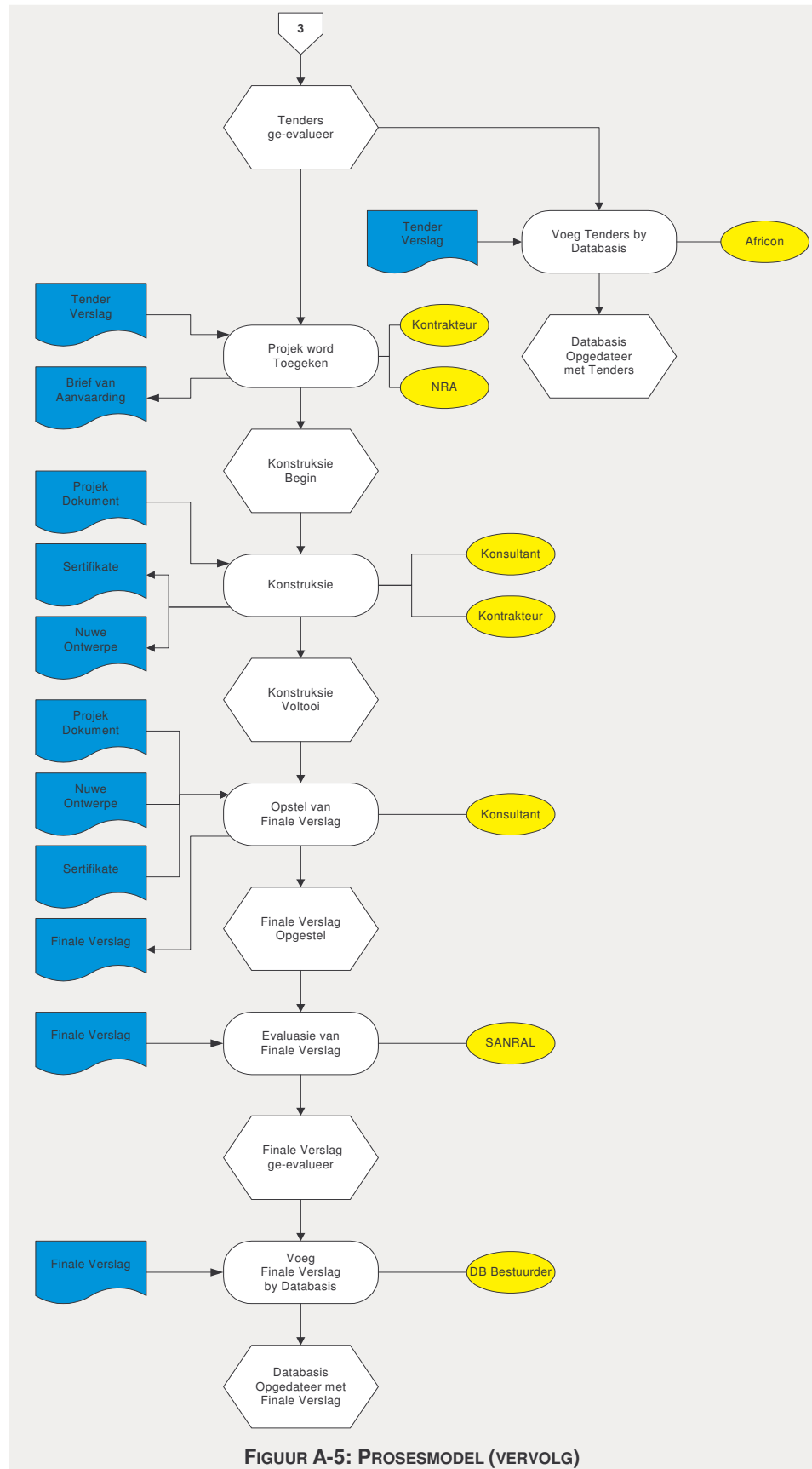
In *Figuur A-2* tot *A-5* word die prosesmodel getoon beginnend by 'n behoefte (gebeurtenis) aan een-of-ander padkonstruksieprojek om infrastruktuur daar te stel of te verbeter en eindig by die voltooiing van die projek wanneer die projekvoltooiingsverslag afgehandel is en die kosteberekeningsdatabasis daarmee opgedateer is.

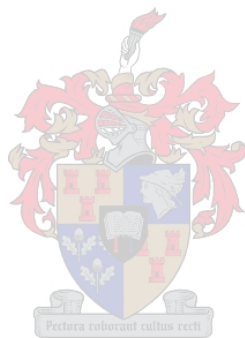




FIGUUR A-3: PROSEMODEL (VERVOLG)







---

## AANHANGSEL B

---

### IMPLEMENTERING VAN TOETSPROGRAM

---

Die program wat ontwikkel is om te toets vir volledigheid en korrektheid van prosesmodelle is in die objekgeoriënteerde programmeringstaal *JAVA* geskryf. Die program bestaan uit vyf klasse waarvan die objekte geskep word tydens uitvoering. Die klasdiagram, wat die interaksie tussen die verskillende klasse aandui word met *UML* ("Unified Modelling Language") in *Figuur B-1* getoon.

In bogenoemde figuur word die vyf klasse wat ontwikkel is se attribute en metodes gewys, terwyl agt van die standaard *JAVA* klasse wat deur die program gebruik word, ook getoon word. 'n Stippelpyl tussen twee klasse dui aan dat die een klas in die ander klas gebruik word, terwyl 'n ononderbroke pyl aandui dat een klas 'n ander klas se eienskappe oorerf.

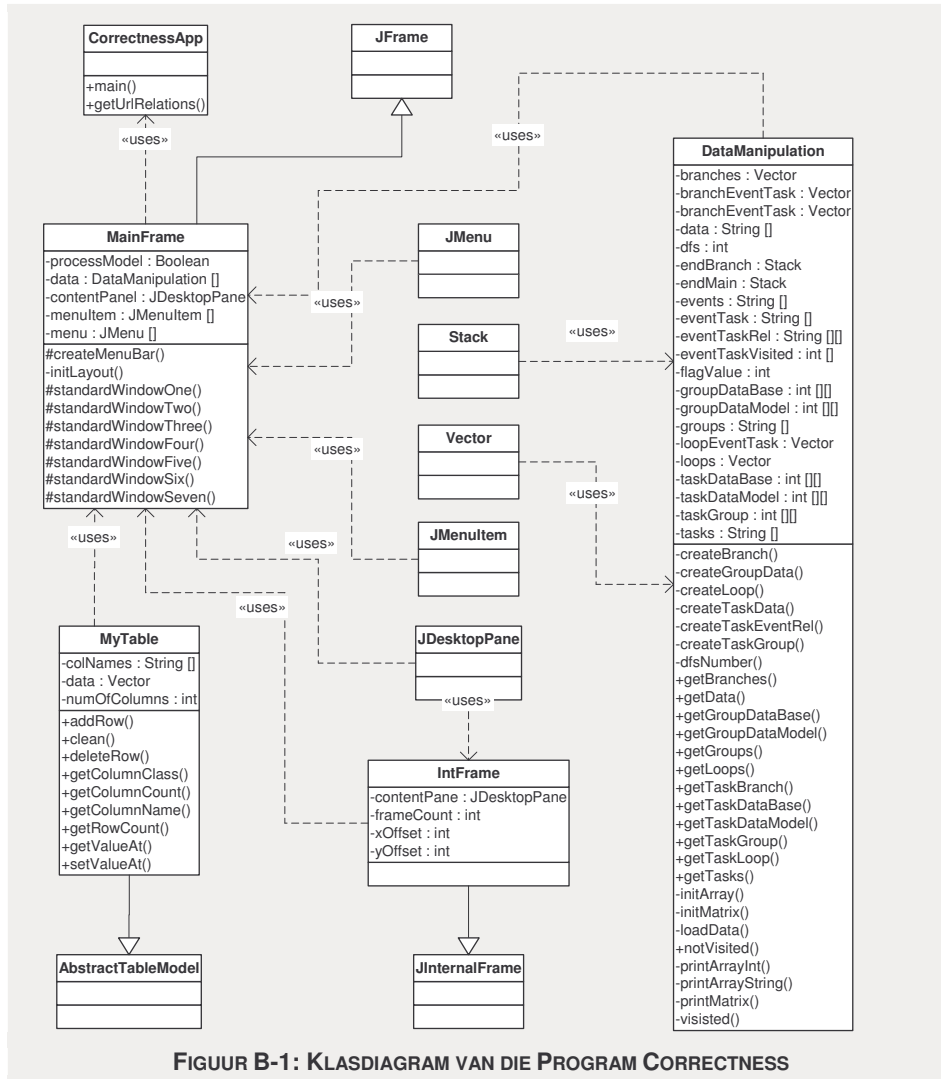
Die kode en die gekompileerde program kan op die KS gevind word in die gids "Aanhangsel B". Die verskillende klasse kan met die opdrag *Make Correctness.cmd* gekompileer word indien die *JAVA Virtuele Masjien* beskikbaar is. Die gekompileer kode kan dan geloop word met die opdrag *Correctness.cmd*.

Gedurende uitvoering van die program moet die beskrywing van die prosesmodel, wat die inligting van die aktiwiteite, gebeurtenisse, groepe en data stoor, asook die interaksie tussen hierdie groepe, gelaai word. Hierdie inligting word gestoor in 'n *MS Access* databasis, *SimbaModel.mdb*, wat ook in die gids "Aanhangsel B" gevind kan word. Vir die program om te kan skakel met die databasis is dit nodig om die databasis met die *ODBC*-funksionaliteit in *MS Windows* sigbaar te maak.

---



Die opstelling kan maklik gedoen word deur van die *ODBC Data Source Administrator* gebruik te maak. Die naam van die gebruikersdatabron wat in *ODBC* opgestel word, moet “*SimbaModel*” wees en dit moet verwys na bogenoemde databasis, *SimbaModel.mdb*, wat vanaf die KS na ’n hardeskyf oorgedra moet word. Daar is nie van gebruikersname of wagwoorde gebruik gemaak nie, dus is geen verdere verstellings by die *ODBC*-skerm nodig nie.



FIGUUR B-1: KLASDIAGRAM VAN DIE PROGRAM CORRECTNESS

Die kode van elk van die vyf klasse word vervolgens getoon. Die eerste klas wat beskou word is die klas *CorrectnessApp*, Java-klas B-1, wat die “main”-metode bevat en derhalwe gebruik word om die program uit te voer. Hierdie klas bewerkstellig die skakeling met die databasis waar die inligting van die prosesmodel, wat geanaliseer moet word, gestoor word en instansieer dan ’n objek van die klas wat die funksionaliteit van die gebruikskoppelvlak bevat.

Die klas *MainFrame* en *DataManipulation* is die twee hoofklasse van hierdie program. In die klas *DataManipulation* is die algoritmes wat in *Afdeling 2.2.2* bespreek word geïmplementeer. Die antwoorde wat verkry word vanaf hierdie klas word teruggestuur aan die klas *MainFrame* wat dit dan vertoon.

**JAVA-KLAS B-1: CORRECTNESSAPP.JAVA**

```
// Bauhaus Universität Weimar 2000

/**
 * CorrectnessApp.java
 *
 *
 * Created: Fri Dec 08 14:31:32 2000
 *
 * @author André Bester
 * @version 1.0.0
 */

public class CorrectnessApp {
    public static void main(String[] args) {
        try {
            Class.forName("sun.jdbc.odbc.JdbcOdbcDriver");
        }
        catch (Exception e) {
            System.out.println("Class.forName()");
            e.printStackTrace();
        }
        MainFrame frame = new MainFrame();
        frame.setVisible(true);
    }

    public static String getUrlRelations() {
        return "jdbc:odbc:SimbaModel";
    }
}
} // CorrectnessApp
```

Die klas *MainFrame* skep die gebruikskoppelvlak wat gebruik word tydens die interaksie van die gebruiker met die program, asook die funksionaliteit wat die instruksie vanaf die koppelvlak uitvoer, *Java-klas B-2*.

**JAVA-KLAS B-2: MAINFRAME.JAVA**

```
// Bauhaus Universität Weimar 2000

/**
 * RelationsFrame.java
 *
 *
 * Created: Wed Oct 18 17:24:08 2000
 *
 * @author André Bester
 * @version 1.0.0
 */

import java.sql.*;
```

**JAVA-KLAS B-3: MAINFRAME.JAVA (VERVOLG)**

```
import java.util.Vector;

import java.awt.*;
import java.awt.event.*;

import javax.swing.*;
import javax.swing.border.*;
import javax.swing.event.*;

public class MainFrame extends JFrame{
    private boolean    processModel;
    private DataManipulation data;
    private JDesktopPane  contentPanel;
    private JMenuItem []  menuItem;
    private JMenu []      menu;

    public MainFrame (){
        super("Correctness and Completeness");

        menuItem = new JMenuItem [25];
        menu     = new JMenu [6];
        processModel = true;

        InitLayout();

        // windows closing
        this.addWindowListener(new WindowAdapter()
        {
            public void windowClosing(WindowEvent e){
                System.exit(0);
            }
        });
    }

    private void InitLayout() {
        // set location and size
        this.setLocation(150,70);
        this.setSize(800,600);

        // set menu bar
        this.setJMenuBar(createMenuBar());

        // create contentPane and make settings
        contentPanel = new JDesktopPane();
        contentPanel.setBackground(null);
        contentPanel.setBorder(BorderFactory.createLoweredBevelBorder());

        // set contentPanel as contentPane and make settings
        this.setContentPane(contentPanel);
    }
}
```

---

**JAVA-KLAS B-4: MAINFRAME.JAVA (VERVOLG)**

```
protected JMenuBar createMenuBar() {
    JMenuBar menuBar = new JMenuBar();
    JMenu [] submenu = new JMenu [7];
    ButtonGroup group = new ButtonGroup();
    JRadioButtonMenuItem [] rbMenuItem = new JRadioButtonMenuItem [2];

    // Menu's
    menu[0] = new JMenu("File");
    menu[1] = new JMenu("Model");
    menu[2] = new JMenu("Task");
    menu[3] = new JMenu("Group");
    menu[4] = new JMenu("Data");
    menu[5] = new JMenu("Correctness");

    menu[1].setEnabled(false);
    menu[2].setEnabled(false);
    menu[3].setEnabled(false);
    menu[4].setEnabled(false);
    menu[5].setEnabled(false);

    // MenuItems
    menuItem[0] = new JMenuItem("Load");
    menuItem[24] = new JMenuItem("Exit");
    menuItem[1] = new JMenuItem("Groups");
    menuItem[2] = new JMenuItem("Data");
    menuItem[3] = new JMenuItem("Tasks in Circuit");
    menuItem[5] = new JMenuItem("Tasks in Loop");
    menuItem[23] = new JMenuItem("Tasks");
    menuItem[7] = new JMenuItem("Data");
    menuItem[8] = new JMenuItem("Used in");
    menuItem[9] = new JMenuItem("Change before");
    menuItem[10] = new JMenuItem("Change after");
    menuItem[11] = new JMenuItem("Used by");
    menuItem[12] = new JMenuItem("Change before");
    menuItem[13] = new JMenuItem("Change after");
    menuItem[14] = new JMenuItem("No group");
    menuItem[15] = new JMenuItem("More than one");
    menuItem[18] = new JMenuItem("No data");
    menuItem[19] = new JMenuItem("More than one");
    menuItem[20] = new JMenuItem("Sequence");
    menuItem[21] = new JMenuItem("Effecting tasks");
    menuItem[22] = new JMenuItem("Effecting groups");

    // submenu's
    submenu[2] = new JMenu("Tasks");
    submenu[3] = new JMenu("Groups");
    submenu[4] = new JMenu("Group");
    submenu[5] = new JMenu("Data");
    submenu[6] = new JMenu("Data Errors");
```

**JAVA-KLAS B-5: MAINFRAME.JAVA (VERVOLG)**

```
// radiobuttons
rbMenuItem[0] = new JRadioButtonMenuItem("Process");
rbMenuItem[1] = new JRadioButtonMenuItem("Database");

rbMenuItem[0].setSelected(true);

group.add(rbMenuItem[0]);
group.add(rbMenuItem[1]);

// add menuitems to submenu's
submenu[2].add(menuItem[8]);
submenu[2].add(menuItem[9]);
submenu[2].add(menuItem[10]);

submenu[3].add(menuItem[11]);
submenu[3].add(menuItem[12]);
submenu[3].add(menuItem[13]);

submenu[4].add(menuItem[14]);
submenu[4].add(menuItem[15]);

submenu[5].add(menuItem[18]);
submenu[5].add(menuItem[19]);

submenu[6].add(menuItem[21]);
submenu[6].add(menuItem[22]);

// add items to menu's
menu[0].add(menuItem[0]);
menu[0].add(menuItem[24]);

menu[1].add(rbMenuItem[0]);
menu[1].add(rbMenuItem[1]);

menu[2].add(menuItem[1]);
menu[2].add(menuItem[2]);
menu[2].add(menuItem[3]);
menu[2].add(menuItem[5]);

menu[3].add(menuItem[23]);
menu[3].add(menuItem[7]);

menu[4].add(submenu[2]);
menu[4].add(submenu[3]);

menu[5].add(submenu[4]);
menu[5].add(submenu[5]);
menu[5].add(menuItem[20]);
menu[5].add(submenu[6]);
```

---

**JAVA-KLAS B-6: MAINFRAME.JAVA (VERVOLG)**

```
// add menu's to menuBar
menuBar.add(menu[0]);
menuBar.add(menu[1]);
menuBar.add(menu[2]);
menuBar.add(menu[3]);
menuBar.add(menu[4]);
menuBar.add(menu[5]);

// menuItem File
menuItem[0].addActionListener(new ActionListener()
{
    public void actionPerformed(ActionEvent e) {
        JOptionPane.showInternalMessageDialog(contentPanel, "Reload for source data
changes to take effect!", "Warning", JOptionPane.WARNING_MESSAGE);
        data = new DataManipulation();
        if (!data.notVisited().equals(""))
            JOptionPane.showInternalMessageDialog(contentPanel, "Events or Tasks not connec-
ted:"+data.notVisited(), "Warning", JOptionPane.WARNING_MESSAGE);
        for(int i = 1; i != menu.length; i++) menu[i].setEnabled(true);
    }
});

// menuItem Exit
menuItem[24].addActionListener(new ActionListener()
{
    public void actionPerformed(ActionEvent e) {
        System.exit(0);
    }
});

// rbMenuItem Database
rbMenuItem[0].addActionListener(new ActionListener()
{
    public void actionPerformed(ActionEvent e){
        processModel=true;
    }
});

// rbMenuItem Model
rbMenuItem[1].addActionListener(new ActionListener()
{
    public void actionPerformed(ActionEvent e){
        processModel=false;
    }
});

// menuItem Group
menuItem[1].addActionListener(new ActionListener() {
    public void actionPerformed(ActionEvent e) {
```

### JAVA-KLAS B-7: MAINFRAME.JAVA (VERVOLG)

```

        menuItem[1].setEnabled(false);
        String [] colNames = {"Group"};
        standardWindowOne("Task-Groups", "Select task:", "Groups involved:", 1, colNames,
data.getTasks(), data.getGroups(), data.getTaskGroup(), 0);
    }
});

// menuItem Data
menuItem[2].addActionListener(new ActionListener() {
    public void actionPerformed(ActionEvent e) {
        menuItem[2].setEnabled(false);
        String [] colNames = {"Data", "Access"};
        standardWindowTwo("Task-Data", "Select task:", "Data used:", 2, colNames, da-
ta.getTasks(), data.getData(), data.getTaskDataModel(), data.getTaskDataBase(), 0);
    }
});

// subMenu Loop
menuItem[3].addActionListener(new ActionListener() {
    public void actionPerformed(ActionEvent e) {
        menuItem[3].setEnabled(false);
        String [] colNames = {"Main","Circuit"};
        standardWindowSeven("Task-Circuits", "Select circuit:", "Tasks/Events in circuit:", 3,
colNames, data.getLoops(), data.getTaskLoop());
    }
});

// subMenu Tasks in Branches
menuItem[5].addActionListener(new ActionListener() {
    public void actionPerformed(ActionEvent e) {
        menuItem[5].setEnabled(false);
        String [] colNames = {"Main","Second"};
        standardWindowSeven("Task-Paths", "Select path:", "Tasks/Events in path:", 5, colNa-
mes, data.getBranches(), data.getTaskBranch());
    }
});

// menuItem Task
menuItem[23].addActionListener(new ActionListener() {
    public void actionPerformed(ActionEvent e) {
        menuItem[23].setEnabled(false);
        String [] colNames = {"Task"};
        standardWindowOne("Group-Tasks", "Select group:", "Tasks executed:", 23, colNa-
mes, data.getGroups(), data.getTasks(), data.getTaskGroup(), 1);
    }
});

// menuItem Data
menuItem[7].addActionListener(new ActionListener() {
    public void actionPerformed(ActionEvent e) {

```

**JAVA-KLAS B-8: MAINFRAME.JAVA (VERVOLG)**

```

        menuItem[7].setEnabled(false);
        String [] colNames = {"Data","Access"};
        standardWindowTwo("Group-Data", "Select group:", "Data used:", 7, colNames, data.getGroups(), data.getData(), data.getGroupDataModel(), data.getGroupDataBase(), 0);
    }
});

// subMenu Used by
menuItem[8].addActionListener(new ActionListener() {
    public void actionPerformed(ActionEvent e) {
        menuItem[8].setEnabled(false);
        String [] colNames = {"Task","Access"};
        standardWindowTwo("Data-Tasks", "Select data:", "Used in:", 8, colNames, data.getData(), data.getTasks(), data.getTaskDataModel(), data.getTaskDataBase(), 1);
    }
});

// subMenu Change before
menuItem[9].addActionListener(new ActionListener() {
    public void actionPerformed(ActionEvent e) {
        menuItem[9].setEnabled(false);
        String [] colNames = {"Task","Access"};
        standardWindowThree("Data-Tasks", "Select data:", "Select task:", "Used before in:", 9, colNames, data.getData(), data.getTasks()); //additional parameters needs to be added for full implementation
    }
});

// subMenu Changs after
menuItem[10].addActionListener(new ActionListener() {
    public void actionPerformed(ActionEvent e) {
        menuItem[10].setEnabled(false);
        String [] colNames = {"Task","Access"};
        standardWindowThree("Data-Tasks", "Select data:", "Select task:", "Used after in:", 10, colNames, data.getData(), data.getTasks()); //additional parameters needs to be added for full implementation
    }
});

// subMenu Used by
menuItem[11].addActionListener(new ActionListener() {
    public void actionPerformed(ActionEvent e) {
        menuItem[11].setEnabled(false);
        String [] colNames = {"Group","Access"};
        standardWindowTwo("Data-Groups", "Select data:", "Used by:", 11, colNames, data.getData(), data.getGroups(), data.getGroupDataModel(), data.getGroupDataBase(), 1);
    }
});

// subMenu Change before

```



### JAVA-KLAS B-9: MAINFRAME.JAVA (VERVOLG)

```

menuItem[12].addActionListener(new ActionListener() {
    public void actionPerformed(ActionEvent e) {
        menuItem[12].setEnabled(false);
        String [] colNames = {"Group","Access"};
        standardWindowThree("Data-Groups", "Select data:", "Select group:", "Used before
by:", 12, colNames, data.getData(), data.getGroups()); //additional parameters needs to be
added for full implementation
    }
});

// subMenu Change after
menuItem[13].addActionListener(new ActionListener() {
    public void actionPerformed(ActionEvent e) {
        menuItem[13].setEnabled(false);
        String [] colNames = {"Group","Access"};
        standardWindowThree("Data-Groups", "Select data:", "Select group:", "Used after by:",
13, colNames, data.getData(), data.getGroups()); //additional parameters needs to be added
for full implementation
    }
});

// subMenu No group
menuItem[14].addActionListener(new ActionListener() {
    public void actionPerformed(ActionEvent e) {
        menuItem[14].setEnabled(false);
        String [] colNames = {"Task"};
        standardWindowFour("Correctness-Groups", "No group at tasks:", 14, colNames,
data.getTasks(), data.getTaskGroup(), 0);
    }
});

// subMenu More than one
menuItem[15].addActionListener(new ActionListener() {
    public void actionPerformed(ActionEvent e) {
        menuItem[15].setEnabled(false);
        String [] colNames = {"Task"};
        standardWindowFour("Correctness-Groups", "More than one group at tasks:", 15,
colNames, data.getTasks(), data.getTaskGroup(), 1);
    }
});

// subMenu No data
menuItem[18].addActionListener(new ActionListener() {
    public void actionPerformed(ActionEvent e) {
        menuItem[18].setEnabled(false);
        String [] colNames = {"Task"};
        standardWindowFour("Correctness-Data", "No data at tasks:", 18, colNames, da-
ta.getTasks(), data.getTaskDataModel(), 0);
    }
});

```

## JAVA-KLAS B-10: MAINFRAME.JAVA (VERVOLG)

```

// subMenu More than one
menuItem[19].addActionListener(new ActionListener() {
    public void actionPerformed(ActionEvent e) {
        menuItem[19].setEnabled(false);
        String [] colNames = {"Task"};
        standardWindowFour("Correctness-Data", "More than one data at tasks:", 19, colNames, data.getTasks(), data.getTaskDataModel(), 1);
    }
});

// MenuItem Sequence
menuItem[20].addActionListener(new ActionListener() {
    public void actionPerformed(ActionEvent e) {
        menuItem[20].setEnabled(false);
        String [] colNames = {"Data"};
        standardWindowFive("Correctness-Sequence", "Incorrect sequence of data access:", 20, colNames, data.getData()); //additional parameters needed for full implementation
    }
});

// subMenu Effecting tasks
menuItem[21].addActionListener(new ActionListener() {
    public void actionPerformed(ActionEvent e) {
        menuItem[21].setEnabled(false);
        String [] colNames = {"Task"};
        standardWindowSix("Correctness-Data Errors", "Select data:", "Tasks effected:", 21, colNames, data.getData(), data.getTasks()); //additional parameters needed for full implementation
    }
});

// subMenu Effecting groups
menuItem[22].addActionListener(new ActionListener() {
    public void actionPerformed(ActionEvent e) {
        menuItem[22].setEnabled(false);
        String [] colNames = {"Group"};
        standardWindowSix("Correctness-Data Errors", "Select data:", "Groups effected:", 22, colNames, data.getData(), data.getGroups()); //additional parameters needed for full implementation
    }
});

return menuBar;
}

protected void standardWindowOne(String wName, String cName, String tName, final int iMenu, String [] colNames, String [] comboValues, final String [] result, final int [][] relation, final int c) {
    JComponent [] comp = new JComponent [4];

```

**JAVA-KLAS B-11: MAINFRAME.JAVA (VERVOLG)**

```
JLabel txtSelect      = new JLabel(cName);
JLabel txtResult      = new JLabel(tName);
final JComboBox selectCombo = new JComboBox(comboValues);
final MyTable dataTable = new MyTable(colNames);
final JTable dataJTable = new JTable(dataTable);
JScrollPane dataScrPane = new JScrollPane(dataJTable);

// create compound border
Border line1 = BorderFactory.createLineBorder(Color.gray,1);
Border line2 = BorderFactory.createLineBorder(Color.lightGray,1);
Border line3 = BorderFactory.createLineBorder(Color.gray,1);
Border compound1 = BorderFactory.createCompoundBorder(line1, line2);
Border compound2 = BorderFactory.createCompoundBorder(compound1, line3);

// make settings
txtSelect.setForeground(Color.black);
selectCombo.setPreferredSize(new Dimension(150,25));
txtResult.setForeground(Color.black);
dataScrPane.setPreferredSize(new Dimension(280,103));
dataScrPane.setBorder(compound2);

// specify components
comp[0] = txtSelect;
comp[1] = selectCombo;
comp[2] = txtResult;
comp[3] = dataScrPane;

// instantiate internal frame and add
IntFrame frame = new IntFrame(wName, comp);
frame.setVisible(true);

contentPanel.add(frame);

// register listener for closing event
frame.addInternalFrameListener(new InternalFrameAdapter () {
    public void internalFrameClosed(InternalFrameEvent e) {
        menuItem[iMenu].setEnabled(true);
    }
});

// register listener for selectCombo
selectCombo.addActionListener(new ActionListener() {
    public void actionPerformed(ActionEvent e) {
        int value = selectCombo.getSelectedIndex();
        int j = 0;
        dataTable.clean();
        if (c == 0) {
            while (j != result.length) {
                if (relation[value][j]==1) dataTable.setValueAt(result[j],0,0);
                j++;
            }
        }
    }
});
```

## JAVA-KLAS B-12: MAINFRAME.JAVA (VERVOLG)

```

    }
    } else {
        while (j != result.length) {
            if (relation[j][value]==1) dataTable.setValueAt(result[j],0,0);
            j++;
        }
    }
    dataJTable.revalidate();
    dataJTable.repaint();
}
});

try {
    frame.setSelected(true);
} catch (java.beans.PropertyVetoException e) {}
}

protected void standardWindowTwo(String wName, String cName, String tName, final int
iMenu, String [] colNames, String [] comboValues, final String [] result, final int [][] relationModel,
final int [][] relationBase, final int c) {
    JComponent [] comp      = new JComponent [6];
    JLabel txtSelect        = new JLabel(cName);
    JLabel txtBuffer         = new JLabel(" ");
    JLabel txtResult         = new JLabel(tName);
    final JCheckBox modelChB = new JCheckBox("Process Model",processModel);
    final JComboBox selectCombo = new JComboBox(comboValues);
    final MyTable dataTable    = new MyTable(colNames);
    final JTable dataJTable    = new JTable(dataTable);
    JScrollPane dataScrPane   = new JScrollPane(dataJTable);

    // create compound border
    Border line1 = BorderFactory.createLineBorder(Color.gray,1);
    Border line2 = BorderFactory.createLineBorder(Color.lightGray,1);
    Border line3 = BorderFactory.createLineBorder(Color.gray,1);
    Border compound1 = BorderFactory.createCompoundBorder(line1, line2);
    Border compound2 = BorderFactory.createCompoundBorder(compound1, line3);

    // make settings
    txtSelect.setForeground(Color.black);
    selectCombo.setPreferredSize(new Dimension(150,25));
    txtResult.setForeground(Color.black);
    dataScrPane.setPreferredSize(new Dimension(280,80));
    dataScrPane.setBorder(compound2);
    modelChB.setBackground(Color.lightGray);

    // specify components
    comp[0] = txtSelect;
    comp[1] = selectCombo;
    comp[2] = txtBuffer;
    comp[3] = txtResult;

```

## JAVA-KLAS B-13: MAINFRAME.JAVA (VERVOLG)

```
comp[4] = dataScrPane;
comp[5] = modelChB;

// instantiate internal frame and add
IntFrame frame = new IntFrame(wName ,comp);
frame.setVisible(true);

contentPanel.add(frame);

// register listener for closing event
frame.addInternalFrameListener(new InternalFrameAdapter () {
    public void internalFrameClosed(InternalFrameEvent e) {
        menuItem[iMenu].setEnabled(true);
    }
});

// register listener for selectCombo
selectCombo.addActionListener(new ActionListener() {
    public void actionPerformed(ActionEvent e) {
        int value = selectCombo.getSelectedIndex();
        int j = 0;
        dataTable.clean();
        if (modelChB.isSelected()) {
            if (c == 0) {
                while (j != result.length) {
                    if (relationModel[value][j]==1) {
                        dataTable.setValueAt(result[j],0,0);
                        dataTable.setValueAt("Write",0,1);
                    } else if (relationModel[value][j]==-1) {
                        dataTable.setValueAt(result[j],0,0);
                        dataTable.setValueAt("Read",0,1);
                    } else if (relationModel[value][j]==2) {
                        dataTable.setValueAt(result[j],0,0);
                        dataTable.setValueAt("Write/Read",0,1);
                    }
                    j++;
                }
            } else if (c == 1) {
                while (j != result.length) {
                    if (relationModel[j][value]==1) {
                        dataTable.setValueAt(result[j],0,0);
                        dataTable.setValueAt("Write",0,1);
                    } else if (relationModel[j][value]==-1) {
                        dataTable.setValueAt(result[j],0,0);
                        dataTable.setValueAt("Read",0,1);
                    } else if (relationModel[j][value]==2) {
                        dataTable.setValueAt(result[j],0,0);
                        dataTable.setValueAt("Write/Read",0,1);
                    }
                    j++;
                }
            }
        }
    }
});
```

## JAVA-KLAS B-14: MAINFRAME.JAVA (VERVOLG)

```

    }
    }
} else if (!modelChB.isSelected()) {
    if (c == 0) {
        while (j != result.length) {
            if (relationBase[value][j]==1) {
                dataTable.setValueAt(result[j],0,0);
                dataTable.setValueAt("Write",0,1);
            } else if (relationBase[value][j]==-1) {
                dataTable.setValueAt(result[j],0,0);
                dataTable.setValueAt("Read",0,1);
            } else if (relationBase[value][j]==2) {
                dataTable.setValueAt(result[j],0,0);
                dataTable.setValueAt("Write/Read",0,1);
            }
            j++;
        }
    } else if (c == 1) {
        while (j != result.length) {
            if (relationBase[j][value]==1) {
                dataTable.setValueAt(result[j],0,0);
                dataTable.setValueAt("Write",0,1);
            } else if (relationBase[j][value]==-1) {
                dataTable.setValueAt(result[j],0,0);
                dataTable.setValueAt("Read",0,1);
            } else if (relationBase[j][value]==2) {
                dataTable.setValueAt(result[j],0,0);
                dataTable.setValueAt("Write/Read",0,1);
            }
            j++;
        }
    }
}
dataTable.revalidate();
dataTable.repaint();
}
});

// register listener for modelChB
modelChB.addItemListener(new ItemListener() {
    public void itemStateChanged(ItemEvent e) {
        int value = selectCombo.getSelectedIndex();
        int j = 0;
        dataTable.clean();
        if (modelChB.isSelected()) {
            if (c == 0) {
                while (j != result.length) {
                    if (relationModel[value][j]==1) {
                        dataTable.setValueAt(result[j],0,0);
                        dataTable.setValueAt("Write",0,1);
                    }
                }
            }
        }
    }
});

```

## JAVA-KLAS B-15: MAINFRAME.JAVA (VERVOLG)

```

    }else if (relationModel[value][j]==-1) {
        dataTable.setValueAt(result[j],0,0);
        dataTable.setValueAt("Read",0,1);
    } else if (relationModel[value][j]==2) {
        dataTable.setValueAt(result[j],0,0);
        dataTable.setValueAt("Write/Read",0,1);
    }
    j++;
}
} else if (c == 1) {
while (j != result.length) {
    if (relationModel[j][value]==1) {
        dataTable.setValueAt(result[j],0,0);
        dataTable.setValueAt("Write",0,1);
    } else if (relationModel[j][value]==-1) {
        dataTable.setValueAt(result[j],0,0);
        dataTable.setValueAt("Read",0,1);
    } else if (relationModel[j][value]==2) {
        dataTable.setValueAt(result[j],0,0);
        dataTable.setValueAt("Write/Read",0,1);
    }
    j++;
}
}
} else if (!modelChB.isSelected()) {
if (c == 0) {
while (j != result.length) {
    if (relationBase[value][j]==1) {
        dataTable.setValueAt(result[j],0,0);
        dataTable.setValueAt("Write",0,1);
    } else if (relationBase[value][j]==-1) {
        dataTable.setValueAt(result[j],0,0);
        dataTable.setValueAt("Read",0,1);
    } else if (relationBase[value][j]==2) {
        dataTable.setValueAt(result[j],0,0);
        dataTable.setValueAt("Write/Read",0,1);
    }
    j++;
}
} else if (c == 1) {
while (j != result.length) {
    if (relationBase[j][value]==1) {
        dataTable.setValueAt(result[j],0,0);
        dataTable.setValueAt("Write",0,1);
    } else if (relationBase[j][value]==-1) {
        dataTable.setValueAt(result[j],0,0);
        dataTable.setValueAt("Read",0,1);
    } else if (relationBase[j][value]==2) {
        dataTable.setValueAt(result[j],0,0);
        dataTable.setValueAt("Write/Read",0,1);
    }
}
}
}
}

```

**JAVA-KLAS B-16: MAINFRAME.JAVA (VERVOLG)**

```

        }
        j++;
    }
}
}
dataJTable.revalidate();
dataJTable.repaint();
}
});

try {
    frame.setSelected(true);
} catch (java.beans.PropertyVetoException e) {}
}

protected void standardWindowThree(String wName, String cName1, String cName2, String
tName, final int iMenu, String [] colNames, String [] comboValues1, String [] comboValues2) {
    JComponent [] comp    = new JComponent [8];
    JComboBox selectCombo1 = new JComboBox(comboValues1);
    JComboBox selectCombo2 = new JComboBox(comboValues2);
    JLabel txtSelect1      = new JLabel(cName1);
    JLabel txtResult       = new JLabel(tName);
    JLabel txtSelect2      = new JLabel(cName2);
    JLabel txtBuffer       = new JLabel(" ");
    JCheckBox modelChB     = new JCheckBox("Database Model",processModel);
    MyTable dataTable      = new MyTable(colNames);
    JTable dataJTable      = new JTable(dataTable);
    JScrollPane dataScrPane = new JScrollPane(dataJTable);

    // create compound border
    Border line1 = BorderFactory.createLineBorder(Color.gray,1);
    Border line2 = BorderFactory.createLineBorder(Color.lightGray,1);
    Border line3 = BorderFactory.createLineBorder(Color.gray,1);
    Border compound1 = BorderFactory.createCompoundBorder(line1, line2);
    Border compound2 = BorderFactory.createCompoundBorder(compound1, line3);

    // make settings
    txtSelect1.setForeground(Color.black);
    selectCombo1.setPreferredSize(new Dimension(150,25));
    txtSelect2.setForeground(Color.black);
    selectCombo2.setPreferredSize(new Dimension(150,25));
    txtResult.setForeground(Color.black);
    dataScrPane.setPreferredSize(new Dimension(280,50));
    dataScrPane.setBorder(compound2);
    modelChB.setBackground(Color.lightGray);

    // add components
    comp[0] = txtSelect1;
    comp[1] = selectCombo1;
    comp[2] = txtBuffer;

```



### JAVA-KLAS B-17: MAINFRAME.JAVA (VERVOLG)

```

comp[3] = txtSelect2;
comp[4] = selectCombo2;
comp[5] = txtResult;
comp[6] = dataScrPane;
comp[7] = modelChB;

// instantiate internal frame and add
IntFrame frame = new IntFrame(wName ,comp);
frame.setVisible(true);

contentPanel.add(frame);

// register listener for closing event
frame.addInternalFrameListener(new InternalFrameAdapter () {
    public void internalFrameClosed(InternalFrameEvent e) {
        menuItem[iMenu].setEnabled(true);
    }
});

try {
    frame.setSelected(true);
} catch (java.beans.PropertyVetoException e) {}

// for full implementation listener for selectCombo1 must be registered here

// for full implementation listener for selectCombo2 must be registered here

// for full implementation listener for modelChB must be registered here
}

protected void standardWindowFour(String wName, String tName, final int iMenu, String []
colNames, String [] result, int [][] relation, int c) {
    JComponent []comp    = new JComponent [2];
    JLabel txtResult      = new JLabel(tName);
    MyTable dataTable     = new MyTable(colNames);
    JTable dataJTable     = new JTable(dataTable);
    JScrollPane dataScrPane = new JScrollPane(dataJTable);

    // create compound border
    Border line1 = BorderFactory.createLineBorder(Color.gray,1);
    Border line2 = BorderFactory.createLineBorder(Color.lightGray,1);
    Border line3 = BorderFactory.createLineBorder(Color.gray,1);
    Border compound1 = BorderFactory.createCompoundBorder(line1, line2);
    Border compound2 = BorderFactory.createCompoundBorder(compound1, line3);

    // make settings
    txtResult.setForeground(Color.black);
    dataScrPane.setPreferredSize(new Dimension(280,133));
    dataScrPane.setBorder(compound2);

```

**JAVA-KLAS B-18: MAINFRAME.JAVA (VERVOLG)**

```

// add components
comp[0] = txtResult;
comp[1] = dataScrPane;

// instantiate internal frame and add
IntFrame frame = new IntFrame(wName ,comp);
frame.setVisible(true);

contentPanel.add(frame);

// calculate result
if (c == 0) {
    int j = 0;
    while (j != relation.length) {
        int k = 0;
        while ((k!=relation[0].length)&&(relation[j][k]!=1)&&(relation[j][k]!=-1)) k++;
        if (k==relation[0].length) dataTable.setValueAt(result[j],0,0);
        j++;
    }
} else if (c == 1) {
    int j = 0;
    while (j != relation.length) {
        int k = 0;
        int sum = 0;
        while (k != relation[0].length) {
            if (relation[j][k]==1||relation[j][k]==-1) sum++;
            k++;
        }
        if (sum > 1) dataTable.setValueAt(result[j],0,0);
        j++;
    }
}
dataTable.revalidate();
dataTable.repaint();

// register listener for closing event
frame.addInternalFrameListener(new InternalFrameAdapter () {
    public void internalFrameClosed(InternalFrameEvent e) {
        menuItem[iMenu].setEnabled(true);
    }
});

try {
    frame.setSelected(true);
} catch (java.beans.PropertyVetoException e) {}
}

protected void standardWindowFive(String wName, String tName, final int iMenu, String []
colNames, final String [] result) {
    JComponent []comp    = new JComponent [3];

```

**JAVA-KLAS B-19: MAINFRAME.JAVA (VERVOLG)**

```
JLabel txtResult      = new JLabel(tName);
JCheckBox modelChB    = new JCheckBox("Database Model",processModel);
MyTable dataTable    = new MyTable(colNames);
JTable dataJTable     = new JTable(dataTable);
JScrollPane dataScrPane = new JScrollPane(dataJTable);

// create compound border
Border line1 = BorderFactory.createLineBorder(Color.gray,1);
Border line2 = BorderFactory.createLineBorder(Color.lightGray,1);
Border line3 = BorderFactory.createLineBorder(Color.gray,1);
Border compound1 = BorderFactory.createCompoundBorder(line1, line2);
Border compound2 = BorderFactory.createCompoundBorder(compound1, line3);

// make settings
txtResult.setForeground(Color.black);
dataScrPane.setPreferredSize(new Dimension(280,110));
dataScrPane.setBorder(compound2);
modelChB.setBackground(Color.lightGray);

// add components
comp[0] = txtResult;
comp[1] = dataScrPane;
comp[2] = modelChB;

// instantiate frame
IntFrame frame = new IntFrame(wName ,comp);
frame.setVisible(true);

contentPanel.add(frame);

// register listener for closing event
frame.addInternalFrameListener(new InternalFrameAdapter () {
    public void internalFrameClosed(InternalFrameEvent e) {
        menuItem[iMenu].setEnabled(true);
    }
});

// for full implementation listener for modelChB must be registered here

try {
    frame.setSelected(true);
} catch (java.beans.PropertyVetoException e) {}
}

protected void standardWindowSix(String wName, String cName, String tName, final int
iMenu, String [] colNames, String [] comboValues, final String [] result) {
    JComponent [] comp      = new JComponent [4];
    JLabel txtSelect      = new JLabel(cName);
    JLabel txtResult      = new JLabel(tName);
    final JComboBox selectCombo = new JComboBox(comboValues);
```

**JAVA-KLAS B-20: MAINFRAME.JAVA (VERVOLG)**

```

final MyTable dataTable = new MyTable(colNames);
final JTable dataJTable = new JTable(dataTable);
JScrollPane dataScrPane = new JScrollPane(dataJTable);

// create compound border
Border line1 = BorderFactory.createLineBorder(Color.gray,1);
Border line2 = BorderFactory.createLineBorder(Color.lightGray,1);
Border line3 = BorderFactory.createLineBorder(Color.gray,1);
Border compound1 = BorderFactory.createCompoundBorder(line1, line2);
Border compound2 = BorderFactory.createCompoundBorder(compound1, line3);

// make settings
txtSelect.setForeground(Color.black);
selectCombo.setPreferredSize(new Dimension(150,25));
txtResult.setForeground(Color.black);
dataScrPane.setPreferredSize(new Dimension(280,103));
dataScrPane.setBorder(compound2);

// specify components
comp[0] = txtSelect;
comp[1] = selectCombo;
comp[2] = txtResult;
comp[3] = dataScrPane;

// instantiate internal frame and add
IntFrame frame = new IntFrame(wName, comp);
frame.setVisible(true);

contentPanel.add(frame);

// register listener for closing event
frame.addInternalFrameListener(new InternalFrameAdapter () {
    public void internalFrameClosed(InternalFrameEvent e) {
        menuItem[iMenu].setEnabled(true);
    }
});

// for full implementation listener for selectCombo must be registered here

try {
    frame.setSelected(true);
} catch (java.beans.PropertyVetoException e) {}
}

protected void standardWindowSeven(String wName, String cName, String tName, final int
iMenu, String [] colNames, String [] comboValues, final String [][] relation) {
    JComponent [] comp = new JComponent [4];
    JLabel txtSelect = new JLabel(cName);
    JLabel txtResult = new JLabel(tName);
    final JComboBox selectCombo = new JComboBox(comboValues);

```

**JAVA-KLAS B-21: MAINFRAME.JAVA (VERVOLG)**

```
final MyTable dataTable = new MyTable(colNames);
final JTable dataJTable = new JTable(dataTable);
JScrollPane dataScrPane = new JScrollPane(dataJTable);

// create compound border
Border line1 = BorderFactory.createLineBorder(Color.gray,1);
Border line2 = BorderFactory.createLineBorder(Color.lightGray,1);
Border line3 = BorderFactory.createLineBorder(Color.gray,1);
Border compound1 = BorderFactory.createCompoundBorder(line1, line2);
Border compound2 = BorderFactory.createCompoundBorder(compound1, line3);

// make settings
txtSelect.setForeground(Color.black);
selectCombo.setPreferredSize(new Dimension(175,25));
txtResult.setForeground(Color.black);
dataScrPane.setPreferredSize(new Dimension(280,103));
dataScrPane.setBorder(compound2);

// specify components
comp[0] = txtSelect;
comp[1] = selectCombo;
comp[2] = txtResult;
comp[3] = dataScrPane;

// instantiate internal frame and add
IntFrame frame = new IntFrame(wName, comp);
frame.setVisible(true);

contentPanel.add(frame);

// register listener for closing event
frame.addInternalFrameListener(new InternalFrameAdapter () {
    public void internalFrameClosed(InternalFrameEvent e) {
        menuItem[iMenu].setEnabled(true);
    }
});

// register listener for selectCombo
selectCombo.addActionListener(new ActionListener() {
    public void actionPerformed(ActionEvent e) {
        int value = selectCombo.getSelectedIndex();
        int j = 0;
        dataTable.clean();
        for (int i = 0; i != relation[value*2].length; i++)
            dataTable.setValueAt(relation[value*2][i],0,0);
        for (int i = 0; i != relation[value*2+1].length; i++)
            dataTable.setValueAt(relation[value*2+1][i],0,1);
        dataJTable.revalidate();
        dataJTable.repaint();
    }
});
```

**JAVA-KLAS B-22: MAINFRAME.JAVA (VERVOLG)**

```

});

try {
    frame.setSelected(true);
} catch (java.beans.PropertyVetoException e) {}
}
} // MainFrame

```

In die klas `DataManipulation` word die funksionaliteit geïmplementeer wat die toetse van volledigheid en korrektheid uitvoer, *Java-klas B-23*.

**JAVA-KLAS B-23: DATAMANIPULATION.JAVA**

```

// Bauhaus Universität Weimar 2000

/**
 * dataManipulation.java
 *
 *
 * Created: Mon Dec 11 14:38:14 2000
 *
 * @author André Bester
 * @version 1.0.0
 */

// NB - the algorithms implemented here is based on the assumption that
// there are no duplicates in tasks, groups, events, data and also in
// RelationTaskEvent and RelationEventTask

import java.sql.*;
import java.util.Stack;
import java.util.Vector;

public class DataManipulation {
    private String [][] eventTaskRel;
    private String [] tasks;
    private String [] groups;
    private String [] events;
    private String [] data;
    private String [] eventTask;

    private int [][] taskGroup;
    private int [][] taskDataModel;
    private int [][] taskDataBase;
    private int [][] groupDataModel;
    private int [][] groupDataBase;
    private int [] eventTaskVisited;

    private int flagValue = 0;

```

**JAVA-KLAS B-24: DATAMANIPULATION.JAVA (VERVOLG)**

```
private int dfs = 0;

private Stack endMain = new Stack();
private Stack endBranch = new Stack();

private Vector loops = new Vector ();
private Vector branches = new Vector ();
private Vector loopEventTask = new Vector ();
private Vector branchEventTask = new Vector ();

public DataManipulation (){
    // load data from database
    loadData("Name", "ProcessTasks");
    loadData("Name", "Person");
    loadData("Name", "ProcessEvents");
    loadData("Set", "ModelAttributes");

    // instantiate arrays
    eventTask = new String [tasks.length+events.length];
    eventTaskVisited = new int [tasks.length+events.length];

    // instantiate matrices
    taskGroup = new int [tasks.length][groups.length];
    taskDataModel = new int [tasks.length][data.length];
    taskDataBase = new int [tasks.length][data.length];
    groupDataModel = new int [groups.length][data.length];
    groupDataBase = new int [groups.length][data.length];

    // initialize array
    initArray(eventTaskVisited);

    // initialize matrices
    initMatrix(taskGroup);
    initMatrix(taskDataModel);
    initMatrix(taskDataBase);
    initMatrix(groupDataModel);
    initMatrix(groupDataBase);

    // generate values in matrices
    createTaskGroup();
    createTaskData();
    createGroupData();
    createTaskEventRel();
    dfsNumber("Start");

    // uncomment to see generated matrices
    // System.out.println("taskGroup:");
    // printMatrix(taskGroup);
    // System.out.println("events:");
    // printArrayString(events);
```

**JAVA-KLAS B-25: DATAMANIPULATION.JAVA (VERVOLG)**

```
// System.out.println("tasks:");
// printArrayString(tasks);
// System.out.println("taskDataModel:");
// printMatrix(taskDataModel);
// System.out.println("taskDataBase:");
// printMatrix(taskDataBase);
// System.out.println("groupDataModel:");
// printMatrix(groupDataModel);
// System.out.println("groupDataBase:");
// printMatrix(groupDataBase);
// System.out.println("taskLoop:");
// printMatrix(taskLoop);
// System.out.println("taskBranch:");
// printMatrix(taskBranch);
// System.out.println("eventTask:");
// printArrayString(eventTask);
// System.out.println("eventTaskVisited:");
// printArrayInt(eventTaskVisited);
}

public String [] getTasks() {
    return tasks;
}

public String [] getGroups() {
    return groups;
}

public String [] getData() {
    return data;
}

public int [][] getTaskGroup() {
    return taskGroup;
}

public int [][] getTaskDataModel() {
    return taskDataModel;
}

public int [][] getTaskDataBase() {
    return taskDataBase;
}

public int [][] getGroupDataModel() {
    return groupDataModel;
}

public int [][] getGroupDataBase() {
    return groupDataBase;
}
```



## JAVA-KLAS B-26: DATAMANIPULATION.JAVA (VERVOLG)

```
}

public String [] getLoops() {
    String [] getloops = new String [loops.size()];
    for (int i = 0; i != loops.size(); i++)
        getloops[i] = (String) loops.get(i);
    return getloops;
}

public String [] getBranches() {
    String [] getbranches = new String [branches.size()];
    for (int i = 0; i != branches.size(); i++)
        getbranches[i] = (String) branches.get(i);
    return getbranches;
}

public String [][] getTaskLoop() {
    String [][] getEventTask = new String [loopEventTask.size()][];
    for (int i = 0; i != loopEventTask.size(); i++) {
        getEventTask[i] = new String [((MyString)loopEventTask.get(i)).getLength()];
        getEventTask[i] = ((MyString)loopEventTask.get(i)).getString();
    }
    return getEventTask;
}

public String [][] getTaskBranch() {
    String [][] getEventTask = new String [branchEventTask.size()][];
    for (int i = 0; i != branchEventTask.size(); i++) {
        getEventTask[i] = new String [((MyString)branchEventTask.get(i)).getLength()];
        getEventTask[i] = ((MyString)branchEventTask.get(i)).getString();
    }
    return getEventTask;
}

public String notVisited() {
    String notVisited = "";
    int i = 0;
    while (i != eventTaskVisited.length) {
        if (eventTaskVisited[i] == 0) {
            if (i < events.length) {
                notVisited += " "+events[i];
            } else if ((i-events.length) < (tasks.length)) {
                notVisited += " "+tasks[i-(events.length)];
            }
        }
        i++;
    }
    return notVisited;
}
```

**JAVA-KLAS B-27: DATAMANIPULATION.JAVA (VERVOLG)**

```
private void loadData(String colName, String table) {
    try {
        int i = 0;
        int length = 0;
        String s1;

        Connection con = DriverManager.getConnection(CorrectnessApp.getUrlRelations(),"","");
        Statement stat =
con.createStatement(ResultSet.TYPE_SCROLL_INSENSITIVE,ResultSet.CONCUR_READ_O
NLY);
        ResultSet rs = stat.executeQuery("SELECT "+colName+" FROM "+table+" ORDER BY
"+colName+";");

        // find the length of the resultset
        rs.last();
        length = rs.getRow();
        rs.beforeFirst();

        // load resultset in to array
        if (table.equals("ProcessTasks")) {
            tasks = new String [length];
            while (rs.next()){
                s1 = rs.getString(1);
                tasks[i] = s1;
                i++;
            }
        } else if (table.equals("Person")) {
            groups = new String [length];
            while (rs.next()){
                s1 = rs.getString(1);
                groups[i] = s1;
                i++;
            }
        } else if (table.equals("ProcessEvents")) {
            events = new String [length];
            while (rs.next()){
                s1 = rs.getString(1);
                events[i] = s1;
                i++;
            }
        } else if (table.equals("ModelAttributes")) {
            data = new String [length];
            while (rs.next()){
                s1 = rs.getString(1);
                data[i] = s1;
                i++;
            }
        } else System.out.println("Correct Table not loaded in loadData()");
        rs.close();
        stat.close();
    }
}
```

**JAVA-KLAS B-28: DATAMANIPULATION.JAVA (VERVOLG)**

```
        con.close();
    }
    catch (SQLException except) {
        System.out.println("Error in loadData()");
        System.out.println(except.toString());
    }
}

private void initArray(int [] a) {
    for (int i = 0; i != a.length; i++) a[i] = 0;
}

private void initMatrix(int [][] a) {
    for (int i = 0; i != a.length; i++)
        for (int j = 0; j != a[0].length; j++)
            a[i][j] = 0;
}

private void createTaskGroup() {
    try {
        int i = 0;
        int j = 0;
        int k = 0;
        int l = 0;
        int count = 0;
        int length = 0;
        String [][] tempTaskGroup;
        String s1, s2;

        Connection con = DriverManager.getConnection(CorrectnessApp.getUrlRelations(), "", "");
        Statement stat =
con.createStatement(ResultSet.TYPE_SCROLL_INSENSITIVE,ResultSet.CONCUR_READ_O
NLY);
        ResultSet rs = stat.executeQuery("SELECT Task, Person FROM RelationPersonTask
ORDER BY Task, Person;");

        // find the length of the resultset
        rs.last();
        length = rs.getRow();
        rs.beforeFirst();

        // load resultset into temp array
        tempTaskGroup = new String [length][2];
        while (rs.next()){
            s1 = rs.getString(1);
            s2 = rs.getString(2);
            tempTaskGroup[i][0] = s1;
            tempTaskGroup[i][1] = s2;
            i++;
        }
    }
}
```

**JAVA-KLAS B-29: DATAMANIPULATION.JAVA (VERVOLG)**

```

rs.close();
stat.close();
con.close();

// write relations into matrix taskGroup
while (j != length) {
    k = 0;
    // iterate down tasks, if a corresponding task in tempTaskGroup is found, stop
    while ((i != tasks.length && (!tasks[i].equals(tempTaskGroup[j][0]))) i++);
    // iterate down groups, if a corresponding group in tempTaskGroup is found, stop
    while ((k != groups.length && (!groups[k].equals(tempTaskGroup[j][1]))) k++);
    // if the end of tasks or groups was not reached i,k indicates
    // a relation between tasks and groups, therefore write it in taskGroup
    if ((i != tasks.length && k != groups.length) {
        taskGroup[i][k] = 1;
        count++;
    }
    j++;
}

// check to ensure that all relations was written into taskGroup
if (count != length) System.out.println("All task-group relations not loaded");
}
catch (SQLException except) {
    System.out.println("Error in createTaskGroup()");
    System.out.println(except.toString());
}
}

private void createTaskData() {
    try {
        int i = 0;
        int j = 0;
        int k = 0;
        int l = 0;
        int count = 0;
        int length = 0;
        String [][] tempTaskData;
        String s1, s2, s3;

        Connection con = DriverManager.getConnection(CorrectnessApp.getUrlRelations(), "", "");
        Statement stat =
con.createStatement(ResultSet.TYPE_SCROLL_INSENSITIVE, ResultSet.CONCUR_READ_ONLY);
        ResultSet rs = stat.executeQuery("SELECT Task, Set, Semantics FROM RelationTaskSet
ORDER BY Task, Set;");

        // find the length of the resultset
        rs.last();
        length = rs.getRow();

```

### JAVA-KLAS B-30: DATAMANIPULATION.JAVA (VERVOLG)

```

rs.beforeFirst();

// load resultset into temp array
tempTaskData = new String [length][3];
while (rs.next()){
    s1 = rs.getString(1);
    s2 = rs.getString(2);
    s3 = rs.getString(3);
    tempTaskData[i][0] = s1;
    tempTaskData[i][1] = s2;
    tempTaskData[i][2] = s3;
    i++;
}

// write relations into matrix taskDataModel, taskDataBase
while (j != length) {
    k = 0;
    // iterate down tasks, if a corresponding task in tempTaskData is found, stop
    while (!(tasks.length && !tasks[i].equals(tempTaskData[j][0]))) i++;
    // iterate down data, if a corresponding data in tempTaskData is found, stop
    while (k != data.length && !data[k].equals(tempTaskData[j][1])) k++;
    // if the end of tasks or data was not reached i, k indicates a relation
    // between tasks and data, therefore, depending on access write in taskDataModel/Base
    // 1 indicates read, -1 write, 2 read and write
    if (!(tasks.length && k != data.length) {
        if (tempTaskData[j][2].equals("Create")) {
            if ((taskDataModel[i][k] == -1) || (taskDataModel[i][k] == 2)) taskDataModel[i][k] = 2;
            else taskDataModel[i][k] = 1;
            count++;
        } else if (tempTaskData[j][2].equals("Write") || tempTaskData[j][2].equals("Delete") || tempTaskData[j][2].equals("Read"))
        {
            if ((taskDataModel[i][k] == 1) || (taskDataModel[i][k] == 2)) taskDataModel[i][k] = 2;
            else taskDataModel[i][k] = -1;
            count++;
        } else System.out.println("Unexpected Access in createTaskData()");
        if (tempTaskData[j][2].equals("Create") || tempTaskData[j][2].equals("Write") || tempTaskData[j][2].equals("Delete")
        ) {
            if ((taskDataBase[i][k] == -1) || (taskDataBase[i][k] == 2)) taskDataBase[i][k] = 2;
            else taskDataBase[i][k] = 1;
            count++;
        } else if (tempTaskData[j][2].equals("Read")) {
            if ((taskDataBase[i][k] == 1) || (taskDataBase[i][k] == 2)) taskDataBase[i][k] = 2;
            else taskDataBase[i][k] = -1;
            count++;
        } else System.out.println("Unexpected Access in createTaskData()");
        }
    }
    j++;
}

```

### JAVA-KLAS B–31: DATAMANIPULATION.JAVA (VERVOLG)

```

// check to ensure that all relations was written into taskDataBase and taskDataModel
if (count/2 != length) System.out.println("All task-data relations not loaded");
}
catch (SQLException except) {
    System.out.println("Error in createTaskDataRead()");
    System.out.println(except.toString());
}
}

private void createGroupData() {
    int [][] tempGroupTask = new int [groups.length][tasks.length];

    // finding the transpose of taskGroup
    for (int i = 0; i != tasks.length; i++)
        for (int j = 0; j != groups.length; j++)
            tempGroupTask[j][i] = taskGroup[i][j];

    // finding the "matrix product" between tempGroupTask and taskDataModel
    // the three for-loops are for the typical "matrix product"
    // the if asserts that we have a read and write or if it is already
    // read/write situation for the specific group and data
    // the else if prevents overwrites
    for (int i = 0; i != data.length; i++)
        for (int j = 0; j != groups.length; j++)
            for (int k = 0; k != tasks.length; k++)
                if ((groupDataModel[j][i] != 0)&&(tempGroupTask[j][k]*taskDataModel[k][i] !=
0)&&(groupDataModel[j][i] != tempGroupTask[j][k]*taskDataModel[k][i])|(groupDataModel[j][i] ==
2)) {
                    groupDataModel[j][i] = 2;
                } else if (tempGroupTask[j][k]*taskDataModel[k][i] != 0) {
                    groupDataModel[j][i] = tempGroupTask[j][k]*taskDataModel[k][i];
                }

    // finding the "matrix product" between tempGroupTask and taskDataBase
    // the three for-loops are for the typical "matrix product"
    // the if asserts that we have a read and write or if it is already
    // read/write situation for the specific group and data
    // the else if prevents overwrites
    for (int i = 0; i != data.length; i++)
        for (int j = 0; j != groups.length; j++)
            for (int k = 0; k != tasks.length; k++)
                if ((groupDataBase[j][i] != 0)&&(tempGroupTask[j][k]*taskDataBase[k][i] !=
0)&&(groupDataBase[j][i] != tempGroupTask[j][k]*taskDataBase[k][i])|(groupDataBase[j][i] ==
2)) {
                    groupDataBase[j][i] = 2;
                } else if (tempGroupTask[j][k]*taskDataBase[k][i] != 0) {
                    groupDataBase[j][i] = tempGroupTask[j][k]*taskDataBase[k][i];
                }
    }
}

```

## JAVA-KLAS B-32: DATAMANIPULATION.JAVA (VERVOLG)

```
private void createTaskEventRel() {
    try {
        int i = 0;
        int j = 0;
        int k = 0;
        int l = 0;
        int count = 0;
        int length = 0;
        String [][] tempRel;
        String s1, s2;

        Connection con = DriverManager.getConnection(CorrectnessApp.getUrlRelations(),"","");
        Statement stat =
con.createStatement(ResultSet.TYPE_SCROLL_INSENSITIVE,ResultSet.CONCUR_READ_O
NLY);
        ResultSet rs = stat.executeQuery("SELECT Task, Event FROM RelationTaskEvent ORDER
BY Task, Event;");

        // find the length of the resultset of the first query
        rs.last();
        length = rs.getRow();
        rs.beforeFirst();

        // load resultset of the first query
        tempRel = new String [length][2];
        while (rs.next()){
            s1 = rs.getString(1);
            s2 = rs.getString(2);
            tempRel[i][0] = s1;
            tempRel[i][1] = s2;
            i++;
        }

        rs = stat.executeQuery("SELECT Event, Task FROM RelationEventTask ORDER BY Event,
Task;");

        // find the length of the resultset of the second query
        rs.last();
        length = rs.getRow();
        rs.beforeFirst();

        // load resultset of the second query
        eventTaskRel = new String [length+tempRel.length][2];
        while (rs.next()){
            s1 = rs.getString(1);
            s2 = rs.getString(2);
            eventTaskRel[j][0] = s1;
            eventTaskRel[j][1] = s2;
            j++;
        }
    }
}
```

---

**JAVA-KLAS B–33: DATAMANIPULATION.JAVA (VERVOLG)**

```

    }
    rs.close();
    stat.close();
    con.close();

    while (j != eventTaskRel.length) {
        eventTaskRel[j][0] = tempRel[k][0];
        eventTaskRel[j][1] = tempRel[k][1];
        j++;
        k++;
    }
}
catch (SQLException except) {
    System.out.println("Error in createTaskEvent()");
    System.out.println(except.toString());
}
}

private void createBranch(int mainBeg, int mainEnd, int branchBeg, int branchEnd) {
    branches.add(eventTask[mainBeg]+"-"+eventTask[mainEnd]);

    String [] main = new String [mainEnd-mainBeg-1];
    String [] branch = new String [branchEnd-branchBeg+1];

    for (int i = (mainBeg + 1); i != mainEnd; i++) {
        main[i-(mainBeg+1)] = eventTask[i];
    }

    for (int i = branchBeg; i != (branchEnd+1); i++) {
        branch[i-branchBeg] = eventTask[i];
    }

    MyString myMain = new MyString(main);
    MyString myBranch = new MyString(branch);

    branchEventTask.add(myMain);
    branchEventTask.add(myBranch);
}

private void createLoop(int mainBeg, int mainEnd, int loopBeg, int loopEnd) {
    loops.add(eventTask[mainEnd]+"-"+eventTask[mainBeg]);

    String [] main = new String [mainBeg-mainEnd-1];
    String [] loop = new String [loopEnd-loopBeg+1];

    for (int i = (mainEnd + 1); i != mainBeg; i++) {
        main[i-(mainEnd+1)] = eventTask[i];
    }

    for (int i = loopBeg; i != (loopEnd+1); i++) {

```

---



**JAVA-KLAS B-34: DATAMANIPULATION.JAVA (VERVOLG)**

```
loop[i-loopBeg] = eventTask[i];
}

MyString myMain = new MyString(main);
MyString myLoop = new MyString(loop);

loopEventTask.add(myMain);
loopEventTask.add(myLoop);
}

private void dfsNumber(String et) {
    int count = 0;
    int position = 0;
    int beg = 0;

    if (!visited(et)) {
        eventTask[dfs] = et;
        dfs++;

        int i = 0;
        while ((i != eventTaskRel.length)&&!(eventTaskRel[i][0].equals(et))) i++;
        while ((i != eventTaskRel.length)&&(eventTaskRel[i][0].equals(et))) {
            // count number of branches at event
            count++;
            // remember where the first branches was found
            if (count==1) position = i;
            i++;
        }

        // remember DFS number of branching point
        beg = dfs-1;

        for (int j = 0; j != count; j++) {
            // remember DFS number for each branch beginning
            int next = dfs;

            dfsNumber(eventTaskRel[position+j][1]);

            if ((j > 0)&&!endMain.empty()) {
                if (beg < ((Integer)endMain.peek()).intValue())
                    createBranch(beg, ((Integer)endMain.pop()).intValue(), next, ((Integer)endBranch.pop()).intValue());
                else
                    createLoop(beg, ((Integer)endMain.pop()).intValue(), next, ((Integer)endBranch.pop()).intValue());
            }
        }
    }
}
```

---

**JAVA-KLAS B-35: DATAMANIPULATION.JAVA (VERVOLG)**

```

private boolean visited(String et) {
    try {
        int i = 0;
        int j = 0;
        int k = 0;

        while ((i != events.length)&&!et.equals(events[i])) i++;
        if (i == events.length) while ((j != tasks.length)&&!et.equals(tasks[j])) j++;
        if (eventTaskVisited[i+j] != 0) {
            eventTaskVisited[i+j]++;
            while (!et.equals(eventTask[k])) k++;
            // remember DFS number of end of first branch
            endMain.push(new Integer(k));
            endBranch.push(new Integer(dfs-1));
            //flagValue = k;
            return true;
        } else {
            eventTaskVisited[i+j]++;
            return false;
        }
    }
    catch (ArrayIndexOutOfBoundsException except) {
        System.out.println("Error in visited" );
        System.out.println(except.toString());
        return true;
    }
}

private void printMatrix(int [][] a) {
    for (int i = 0; i != a.length; i++){
        for (int j = 0; j != a[0].length; j++) System.out.print(a[i][j]);
        System.out.println();
    }
}

private void printArrayString(String [] a) {
    for (int i = 0; i != a.length; i++){
        System.out.print("Pos: "+i+ " - "+a[i]);
        System.out.println();
    }
}

private void printArrayInt(int [] a) {
    for (int i = 0; i != a.length; i++){
        System.out.print("Pos: "+i+ " - "+a[i]);
        System.out.println();
    }
}

} // DataManipulation

```

In die klas MyTable word die funksionaliteit ingebou om 'n generiese tabel met enige aantal kolomme en rye te skep, *Java-klas B-36*. Hierdie klas word gebruik deur die klas Mainframe om resultate te vertoon in objekte van die klas IntFrame.

```
JAVA-KLAS B-36: MYTABLE.JAVA

// Bauhaus Universität Weimar 2000

/**
 * MyTable.java
 *
 *
 * Created: Thu Oct 19 11:30:00 2000
 *
 * @author André Bester
 * @version 1.0.0
 */

import java.util.*;

import javax.swing.table.*;

public class MyTable extends AbstractTableModel{
    private int numOfColumns;
    private Vector [] data;
    private String [] colNames;

    public MyTable (String [] names){
        this.numOfColumns = names.length;

        data = new Vector [numOfColumns];
        colNames = new String [numOfColumns];

        for (int i = 0; i != numOfColumns; i++) {
            data[i] = new Vector();
            this.colNames[i] = names[i];
        }
    }

    public int getColumnCount() {
        return numOfColumns;
    }

    public int getRowCount() {
        int max = 0;

        for (int i = 0; i != numOfColumns; i++)
            max = Math.max(max, data[i].size());

        return max;
    }
}
```

**JAVA-KLAS B-37: MYTABLE.JAVA (VERVOLG)**

```

}

public Object getValueAt(int row, int col) {
    if (row < data[col].size())
        return data[col].get(row);
    else
        return "";
}

public Class getColumnClass(int c) {
    return getValueAt(0, c).getClass();
}

public String getColumnName(int col) {
    return colNames[col];
}

public void setValueAt (Object aValue, int row, int col){
    data[col].add(aValue);
}

public void addRow(Object [] rowData) {
    for (int i = 0; i != numOfColumns; i++) setValueAt(rowData[i], 0, i);
}

public void deleteRow(int row) {
    for (int i = 0; i != numOfColumns; i++) data[i].remove(row);
}

public void clean() {
    for (int i = 0; i != numOfColumns; i++) data[i].removeAllElements();
}
} // MyTable

```

Die klas `IntFrame` word in die klas `MainFrame` gebruik om die resultate te vertoon, *Java-klas B-38*. Dit is 'n generiese raam wat die objekte wat dit ontvang uitlê en vertoon.

**JAVA-KLAS B-38: INTFRAME.JAVA**

```

// Bauhaus Universität Weimar 2000

/**
 * IntFrame.java
 *
 *
 * Created: Thu Dec 07 08:40:06 2000
 *
 * @author André Bester
 * @version 1.0.0

```

**JAVA-KLAS B-39: INTFRAME.JAVA (VERVOLG)**

```
*/  
  
import java.awt.*;  
  
import javax.swing.*;  
import javax.swing.JInternalFrame;  
  
public class IntFrame extends JInternalFrame{  
    static    int frameCount = 0;  
    static final int xOffset = 30;  
    static final int yOffset = 30;  
  
    private JDesktopPane contentPanel = new JDesktopPane();  
  
    public IntFrame (String name, JComponent [] contents) {  
        super(name, false, true, false, true);  
  
        // make settings and add components  
        contentPanel.setBackground(Color.lightGray);  
        contentPanel.setLayout(new FlowLayout(FlowLayout.LEFT));  
  
        for (int i=0; i != contents.length; i++) contentPanel.add(contents[i]);  
  
        // set contentPanel as contentPane and make settings  
        this.setLocation(xOffset*frameCount, yOffset*frameCount);  
        this.setSize(300,200);  
        this.setContentPane(contentPanel);  
  
        // increase number of frames  
        frameCount++;  
    }  
} // IntFrame
```

---

## AANHANGSEL C

---

### DATA-OORDRAG

---

Hier word stapsgewys verduidelik hoe die rou data verwerk is totdat dit in die voorgestelde databasismodel geplaas is. Die proses bestaan uit twee hoof gedeeltes: eerstens die oordrag van die rou data, wat in teksleërs ontvang is, na die bestaande databasismodel en tweedens die oordrag van die data vanuit die bestaande databasis na die voorgestelde databasismodel.

#### **STAP 0 – SKEP VAN DATABASISMODELLE:**

Die datamodel gee struktuur aan die data binne die databasis. Vir die oorsake-lingsprosesse is van twee databasismodelle gebruik wat in *MS Access* geskep is:

- Die bestaande databasismodel, *Prototype Old CSRA DB.mdb*, is geskep en kan op meegaande kompakte skyf (KS) in die gids “\Databasis\Prototipe” gevind word;
- Die voorgestelde databasismodel, *Prototype Normalised CSRA DB.mdb*, is geskep en ook in bogenoemde gids geplaas.

Beide hierdie twee databasisse het net die struktuur van die genoemde databasisse en geen data is binne dit gelaai.

#### **STAP A1 – LAAI VAN ROU DATA:**

- Die roudata is ontvang in die vorm van nege digitale teksleërs. Hierdie leërs kan gevind word in die gids “\Databasis\Gelaai\Data”.
-

- Die drie teksleërs wat elk 'n "c" in die uitbreiding het, bevat die definisies van die konstruksie items. Hierdie leërs is elk in drie afsonderlike *MS Excel*-sigblaaie gelaai en onder bogenoemde gids gestoor. Die naam van die teksleërs is behou, maar "*Item*" is vooraan gevoeg.
- Die drie teksleërs wat elk 'n "d" in die uitbreiding het, bevat die tenderdata. Hierdie leërs is elk in drie afsonderlike *MS Excel*-sigblaaie gelaai. Dit is ook in bogenoemde gids gestoor. Die naam van die teksleërs is behou, maar "*Tenders*" is vooraan gevoeg.
- Die drie teksleërs wat elk 'n "n" in die uitbreiding het, bevat die projekinligting. Hierdie leërs is ook in drie afsonderlike *MS Excel*-sigblaaie gelaai en is ook in bogenoemde gids gestoor. Die naam van die teksleërs is behou, maar "*Projekte*" is vooraan gevoeg.

#### **STAP A2 – LAAI VAN DIE BESTAANDE DATABASISMODEL MET DATA:**

'n Databasis met dieselfde struktuur as die bestaande databasis, *Transition Old Database.mdb*, is geskep en is geplaas in die gids "\Databasis\Gelaai\Data". In hierdie databasis is die drie tabelle ITEMS, TENDERS en PROJSETS hernoem na ITEMS COMBINED, TENDERS COMBINED en PROJECTS COMBINED. Drie kopieë van die ITEMS COMBINED-tabel is geskep met die name ITEMS1, ITEMS2, ITEMS3, netso is dit ook vir TENDERS COMBINED en PROJECTS COMBINED gedoen. Hierdie nege tabelle is toe gelaai met die data.

- Die kolomme van die nege *MS Excel*-sigblaaie, van *Stap A1* is geformateer, sodat die datatipes ooreenstem met die datatipes in die bestaande databasismodel.
  - Die drie item-sigblaaie is in ITEMS1, ITEMS2 en ITEMS3 gelaai, so die drie tender-sigblaaie in TENDERS1, TENDERS2 en TENDERS3 en so ook die drie projek-sigblaaie in PROJECTS1, PROJECTS2 en PROJECTS3.
  - Die drie tabel van elke groep is dan gelaai in die ooreenstemmende COMBINED-tabelle. Foute wat weens duplisering ontstaan het is uit gefilter tydens hierdie proses en word getoon in die tabelle PASTE ERRORS.
  - Hierdie databasis is netso oorgedra na die gids "\Databasis\Gelaai" en die naam is verander na *Loaded Old CSRA DB.mdb*. Die nege tabelle waar die *MS Excel*-leërs in gelaai is, is verwyder en die name van die tabelle ITEMS COMBINED, TENDERS COMBINED en PROJECTS COMBINED se name is terug verander na ITEMS, TENDERS en PROJECTS.
-

Hierdie databasis het die struktuur van die oorspronklike databasis en is nou gelaai met die data soos afkomstig uit die oorspronklike databasis.

### **STAP B: OORDAG VAN DATA TUSSEN OU EN NUWE MODEL**

Vervolgens is 'n kopie van die voorgestelde databasismodel geskep, sodat dit kon dien as 'n oorgansdatabasis waarin al die nodige aanpassings gemaak kon word om voorsiening te maak vir die oorskakelingsproses. Die databasis se naam was *Transitional Modified Normalised CSRA DB.mdb* en kan in die gids “\Databasis\Gelaai\Genormaliseer” gevind word. Die drie tabelle van *Loaded Old CSRA DB.mdb* is ook na hierdie databasis oorgedra en hernoem na ORIGINAL TENDERS, ORIGINAL PROJECTS en ORIGINAL ITEMS, om data-oordrag te vergemaklik tussen die bestaande databasis se data en die voorgestelde databasis.

'n MS Excel-sigblad, *Transition.xls*, is ook geskep om datamanipulasie tussen die twee verskillende databasismodelle te vergemaklik. Dit kan in die gids “\Databasis\Gelaai\Genormaliseer” gevind word. In hierdie sigblad is sewe werkblaaië geskep met die name *Projects, Categories, Consultants, Tenders, Codes, Items en Units* waarheen geselekteerde data vanuit die bestaande databasis oorgedra, gefilter en dan na die voorgestelde databasismodel oorgedra is. By elke van hierdie sewe werkblaaië is daar nommers aan die bokant van die werkblaaië geplaas wat ooreenstem met van die stappe wat gevolg is om die data op te stel, sodat dit kon oorgedra word na die voorgestelde databasismodel. Hierdie nommers word tussen hakkies aangedui in die stappe wat vervolgens verduidelik word.

Om onderskeid te kan maak tussen dit databasis se tabelle en kolomme, asook die sigblad se werkblaaië, is van die volgende notasie gebruik gemaak. Tabelle se naam word met hoofletters aangedui, kolomme met klein letters behalwe die eerste letter van die naam en werkblaaië van die sigblad in kursiewe skrif.

### **STAP B1 – OORSKAKELING VAN DATA VAN ORIGINAL PROJECTS NA VOORGESTELDE PROJECTS:**

- Die ORIGINAL PROJECTS is na die *Projects*-werkblad in die sigblad gekopieer.
- Vir elke konstruksieprojek is daar gewoonlik twee tenders (in sommige gevalle drie), soos in die databasismodel ontwikkeling genoem is, wat tot onnodig duplisering geleidelik het. Daar was onderskeid gemaak tussen die tenders deur aan die einde van elke projeknaam 'n beskrywing by te voeg, bv. “\Tend1” of “\Tend2”. Hierdie benaming is egter nie konsekwente gedoen, bv. “\Tend1” is soms as “\Tend 1”, “TEND 1”, “TENDER1” en



“\Tender 1” geskryf. Om die dupliserings te kon verwyder was dit nodig gewees om die verskillende vorms van die naam te verwyder deur van ’n gewone “Edit\Replace” funksie in *MS Excel* gebruik te maak en slegs boek te hou watter vorm van skryfwyse verwyder is. (Aangedui as 1 op werkblad.)

- ’n Visual Basic-makro is geloop om die dupliserings van projeknaam en projekinligting te verwyder. Die kode word in *Vb C-1* getoon.
- Kolomme is ingevoeg tussen die bestaande data in die werkblad om voorsiening te maak vir *Consultant\_ID* en *Catogory\_ID*. Dit is gevul met 1’e om te verseker dat die integriteitsreël wat forseer data alle inskrywing in *PROJECTS* data het, nie oortree is as die data oorgedra word nie. (Aangedui as 2 & 3 op werkblad.)
- Die *Amount*-kolom is verwyder, omdat dit gestoor is in *TENDER*. (Aangedui as 4 op werkblad.)
- Hierdie werkblad data was nou gereed om oorgedra te word na *PROJECTS* in die databasis.

#### VB C-1: TENDER UITKENNING

```
Sub Macro1()
'
' Macro1 Macro
' Macro recorded 11.19.00 by André Bester
'
' Keyboard Shortcut: Ctrl+a
'
While Not (ActiveCell.Value = "")
If ActiveCell.Value = ActiveCell.Offset(-1, 0).Value Then
    ActiveCell.Rows("1:1").EntireRow.Select
    Selection.Delete Shift:=xlUp
    ActiveCell.Offset(1, 1).Range("A1").Select
Else: ActiveCell.Offset(1, 0).Select
End If

End Sub
```

#### **STAP B2 – OORSKAKELING VAN CATEGORY-KOLOM IN ORIGINAL PROJECTS NA VOORGESTELDE CATEGORIES:**

- ’n *SQL*-navraag, *Sql C-1*, is uitgevoer in die *ORIGINAL PROJECTS*-tabel van die databasis om die unieke kategorieë te bepaal.
- Hierdie data is na die *Catogory*-werkblad van die sigblad gekopieer. (Aangedui as 2 op werkblad.)

- Vier kategorieë wat verkeerdlik ingeskryf is, kon geïdentifiseer word. (Aangedui as 3 op werkblad.)

**SQL C-1**

```
SELECT Projects.Category, Count(Projects.Category) AS CountOfCategory FROM Projects
GROUP BY Projects.Category;
```

**SQL C-2**

```
UPDATE Projects SET Projects.Category = "ROAD CNSTR" WHERE (((Pro-
jects.Category)="ROAD CNSTR"));
```

**SQL C-3**

```
UPDATE Projects SET Projects.Category = "ROAD CNSTR" WHERE (((Pro-
jects.Category)="ROAD CNSTR."));
```

**SQL C-4**

```
UPDATE Projects SET Projects.Category = "ROAD REHAB" WHERE (((Pro-
jects.Category)="ROAD REHAB          TEND 1"));
```

**SQL C-5**

```
UPDATE Projects SET Projects.Category = "CNSTR" WHERE (((Pro-
jects.Category)="CNSTR"));
```

**SQL C-6**

```
UPDATE Projects INNER JOIN Categories ON Projects.Category = Categories.Category SET
Projects.Category_ID = [Categories].[Category_ID] WHERE (((Catego-
ries.Category)=[Projects].[Category]));
```

- Vier SQL-opdaterings, *Sql C-2*, *Sql C-3*, *Sql C-4* en *Sql C-5*, is in die in transformasie databasis uit gevoer om die korreksie aan te bring.
- Die gekorrigeerde werkblad data is in CATEGORIES van die databasis geplaas.
- 'n SQL-opdatering, *Sql C-6*, is uit gevoer om die Category\_ID in PROJECTS konsekwent te kry met die Category\_ID in CATEGORIES van die databasis.
- Die Category-kolom in PROJECTS is verwyder aangesien die Category\_ID dit nou uniek geïdentifiseer het.

**STAP B3 – OORSKAKELING VAN DIE CONSULTANT-KOLOM IN DIE ORIGINAL PROJECTS NA DIE VOORGESTELDE CONSULTANTS:**

- 'n SQL-navraag, *Sql C-1*, is uitgevoer in die ORIGINAL PROJECTS van die databasis om die unieke konsultante te bepaal en hierdie data is na die *Consultant*-werkblad van die sigblad gekopieer. (Aangedui as 2 op werkblad.)
- Konsultante wat verkeerdlik ingeskryf was, is geïdentifiseer. (Aangedui as 3 op werkblad.)

- Aangesien hier heelwat foute was, maar dit nie veel van 'n effek sou hê op die latere berekeninge nie, is daar slegs twee regstellings as voorbeeld gemaak. Die twee SQL-opdaterings, *Sql C-2* en *Sql C-3*, is in die databasis doen.
- Die gekorrigeerde werkblad data is in CONSULTANTS van die databasis geplaas.
- 'n SQL-opdatering, *Sql C-4*, is uit gevoer om Consultant\_ID in die PROJECTS konsekwent te kry met die Consultant\_ID in CONSULTANT van die databasis.

**SQL C-1**

```
SELECT Projects.Consultant, Count(Projects.Consultant) AS CountOfConsultant FROM Projects  
GROUP BY Projects.Consultant;
```

**SQL C-2**

```
UPDATE Projects SET Projects.Consultant = "SS" WHERE (((Projects.Consultant)="SSCOTT"));
```

**SQL C-3**

```
UPDATE Projects SET Projects.Consultant = "SS" WHERE (((Projects.Consultant)="STEWART  
SCOTT"));
```

**SQL C-4**

```
UPDATE Projects INNER JOIN Categories ON Projects.Category = Categories.Category SET  
Projects.Category_ID = [Categories].[Category_ID] WHERE (((Categories.Category)=[Projects].[Category]));
```

- Die Consultant-kolom in PROJECTS is verwyder aangesien die Consultant\_ID dit nou uniek kon identifiseer.



**STAP B4 – OPDATERING VAN CONTRACTORS-TABEL:**

Om onderskeid te kan maak tussen die verskillende tenders van elke projek is daar gebruik gemaak van denkbeeldige name met die doel om die orde van die tenders voor te stel. Die kontrakteursname wat gebruik is, nl. "Tend1", "Tend2" en "Tend3", verwys ooreenkomstig na die laagste, tweede laagste en derde laagste tender.

- In die CONTRACTORS van die databasis is die volgende drie inskrywings bygevoeg: "Tend1", "Tend2" en "Tend3".

**STAP B5 – OORDRAG VAN TENDERBEDRAE EN KOMMENTAAR IN DIE ORIGINAL PROJECTS NA DIE VOORGESTELDE TENDERS:**

- 'n SQL-navraag, *Sql C-1*, in ORIGINAL PROJECTS van Amount, Comments en Contract No. is uit gevoer en gekopieer na die Tenders-werkblad. (Aangedui as 2 op werkblad.)
-

- 'n Kopie van die Contract No.-kolom is direk langs dit gemaak en die verskillende tenders is uit gefilter soos bespreek in die *Stap B1*. (Aangedui as 3 op werkblad.)
- 'n *MS Excel*-funksie wat kon bepaal watter tender dit is, is gebruik. (Aangedui as 4 op werkblad.)
- By hierdie werkblad is die kolomme **Proj\_ID** en **Contractor\_ID** bygevoeg en gevul met 1'e om integriteitsreël na te kom. (Aangedui as 5 op werkblad.)
- Die werkblad data is na **TENDERS** in die databasis gekopieer.
- 'n *SQL*-opdatering, *Sql C-2*, is uit gevoer om die **Proj\_ID** in **TENDERS** te laat ooreenstem met die in **PROJECTS**.
- Drie *SQL*-opdaterings, *Sql C-3*, *Sql C-4* en *Sql C-5*, is uit gevoer om die **Contractor\_ID** in **TENDERS** te laat ooreenstem met die **ID** in **CONTRACTORS** van die databasis.

**SQL C-1**

```
SELECT [Original Projects].Amount, [Original Projects].Comment, [Original Projects].[Contract No]
FROM [Original Projects];
```

**SQL C-2**

```
UPDATE Projects INNER JOIN Tenders ON Projects.Name = Tenders.Name SET Tenders.Proj_ID = [Projects].[Proj_ID] WHERE (((Projects.Name)=Tenders.[Name]));
```

**SQL C-3**

```
UPDATE Tenders, Contractors SET Tenders.Contractor_ID = [Contractors].[Contractor_ID]
WHERE (((Tenders.[Tend No])=1) AND ((Contractors.Contractor)="Tend1"));
```

**SQL C-4**

```
UPDATE Tenders, Contractors SET Tenders.Contractor_ID = [Contractors].[Contractor_ID]
WHERE (((Tenders.[Tend No])=2) AND ((Contractors.Contractor)="Tend2"));
```

**SQL C-5**

```
UPDATE Tenders, Contractors SET Tenders.Contractor_ID = [Contractors].[Contractor_ID]
WHERE (((Tenders.[Tend No])=3) AND ((Contractors.Contractor)="Tend3"));
```

**STAP B6 – OPDATERING VAN CODES:**

- 'n *SQL*-navraag, *Sql C-1*, is uitgevoer in **ORIGINAL TENDERS** van die databasis om die kodes uniek te bepaal en is in die *Code*-werkblad geplaas. (Aangedui as 2 op werkblad.)
- Hierdie werkblad data is netso in **CODES** van die databasis geplaas, omdat daar geen klaarblyklike duplisering was nie. Dit was egter opvallende dat daar 'n groot verskeidenheid van kodes was.

**Sql C-1**

```
SELECT [Original Tenders].Code, Count([Original Tenders].Code) AS CountOfCode FROM [Original Tenders] GROUP BY [Original Tenders].Code;
```

**STAP B7 – OPDATERING VAN UNITS:**

- 'n SQL-navraag, *Sql C-1*, is uitgevoer in ORIGINAL ITEMS van die transformasie databasis om die unieke eenhede te bepaal en is in die *Units*-werkblad geplak. (Aangedui as 2 op werkblad.)
- Twee duplisering is geïdentifiseer. (Aangedui as 3 op werkblad.)
- Aangesien hierdie foute nie 'n effek sou hê op die berekening nie, is dit netso in UNITS van die databasis geplaas.

**Sql C-1**

```
SELECT [Original Items].Unit, Count([Original Items].Unit) AS CountOfUnit FROM [Original Items] GROUP BY [Original Items].Unit;
```

**STAP B8 – OPDATERING VAN ITEMS:**

- Die ORIGINAL ITEMS-tabel is na die *Items*-werkblad gekopieer. (Aangedui as 2 op werkblad.)
- Die kolom Unit\_ID is bygevoeg en gevul met 1'e. (Aangedui as 3 op werkblad.)
- Hierdie data van die werkblad is in ITEMS van die databasis geplak.
- 'n SQL-opdatering, *Sql C-1*, is uit gevoer om Unit\_ID van ITEMS te laat ooreenstem met die Unit\_ID van UNITS.

**Sql C-1**

```
UPDATE Units INNER JOIN Items ON Units.Unit = Items.Unit SET Items.Unit_ID = [Units].[Unit_ID] WHERE (((Items.Unit)=[Units].[Unit]));
```

**STAP B9 – OPDATERING VAN BILLOFQUANTITIES:**

- Die kolomme Contract No, Code, Original Tender, Escalated Tender en % of Tender is by die BILLOFQUANTITIES by gevoeg.
  - 'n SQL-byvoeging, *Sql C-1*, is uit gevoer om die data van ORIGINAL TENDER in BILLOFQUANTITIES te voeg.
-

**SQL C-1**

```
INSERT INTO BillOfQuantities ( [CSRA Code], [Contract No], Code, Add_Discription, Quantity,
[Escalated Tender], [Orginal Tender], Percentage ) SELECT [Orginal Tenders].[CSRA Code],
[Orginal Tenders].[Contract No], [Orginal Tenders].Code, [Orginal Tenders].[Add Disc], [Orginal
Tenders].Quantity, [Orginal Tenders].[Escalated Tender], [Orginal Tenders].[Orginal Tender],
[Orginal Tenders].[% of Ten] FROM [Orginal Tenders];
```

**SQL C-2**

```
UPDATE Tenders INNER JOIN BillOfQuantities ON Tenders.[Contract No] = BillOfQuantities.[Contract No] SET BillOfQuantities.Proj_ID = [Tenders].[Proj_ID] WHERE
(((BillOfQuantities.[Contract No])=[Tenders].[Contract No]));
```

**SQL C-3**

```
UPDATE Items INNER JOIN BillOfQuantities ON Items.Item_Num = BillOfQuantities.[CSRA
Code] SET BillOfQuantities.Item_ID = [Items].[Item_ID] WHERE (((BillOfQuantities.[CSRA Co-
de])=[Items].[Item_Num]));
```

**SQL C-4**

```
UPDATE Codes INNER JOIN BillOfQuantities ON Codes.Code = BillOfQuantities.Code SET
BillOfQuantities.Code_ID = [Codes].[Code_ID] WHERE (((BillOfQuantities.Code)=[Codes].[Code]));
```

- Drie SQL-opdaterings, *Sql C-2*, *Sql C-3* en *Sql C-4*, uit gevoer om Proj\_ID, Item\_ID en Code\_ID te laat ooreenstem met die ID's soos in die ooreenstemmende PROJECTS, ITEMS en CODES.

Om verder met die opdatering van BILLOFQUANTITIES te kon voortgaan was dit eers nodig gewees om ook FIXED PRICES te ontwikkel.

**STAP B10 – OPDATERING VAN FIXED PRICES:**

- Die kolomme Contract No, Proj\_ID, Item\_ID en Code\_ID is by FIXED PRICES gevoeg.
- Die verwysingsintegriteit en die primêre sleutels van die tabel is verwyder.
- 'n SQL-byvoeging, *Sql C-1*, is uit gevoer om data van ORIGINAL TENDERS te kry in hierdie tabel.

Die verder ontwikkeling van die tabel BILLOFQUANTITIES moes nou voltooi word, voordat daar verder gegaan kon word met die uitbreiding van hierdie tabel.

**SQL C-1**

```
INSERT INTO Fixed_Prices ( [Orginal Tender], [Escalated Tender], Percentage, [Contract no],
Proj_ID, Item_ID, Code_ID ) SELECT BillOfQuantities.[Orginal Tender], BillOfQuantities.[Escalated Tender],
BillOfQuantities.Percentage, BillOfQuantities.[Contract No], BillOfQuantities.Proj_ID,
BillOfQuantities.Item_ID, BillOfQuantities.Code_ID FROM BillOfQuantities;
```

**STAP B9 (VERVOLG)– OPDATERING VAN BILLOFQUANTITIES:**

- 'n Kopie van die struktuur van BOQ met die naam BOQ-copy is gemaak.
- Die primêre sleutel is verander na die kolomme Proj\_ID, Item\_ID en Code\_ID en met 'n SQL-byvoeging, *Sql C-1*, is die data van BOQ by hierdie tabel gevoeg. Hierdie was 'n elegant manier om te verseker dat elke projek se hoeveelheidslys slegs eenkeer gestoor word al was daar meer as een tender, omdat die primêre sleutel beperking verseker het dat dieselfde kombinasie van Proj\_ID, Item\_ID en Code\_ID slegs eenkeer kon voorkom.
- Die primêre sleutels en integriteitsreëls op die kolomme Tender\_ID-kolom is herdefinieer.

**Sql C-1**

```
INSERT INTO [BillOfQuantities - Copy] SELECT BillOfQuantities.* FROM BillOfQuantities;
```

**STAP B10 (VERVOLG)– OPDATERING VAN FIXED PRICES:**

- 'n SQL-opdatering, *Sql C-1* en *Sql C-2*, is gedeon om die Tender\_ID en Bill\_ID van hierdie tabel te laat ooreenstem met onderskeidelik die TENDER en BILOFQUANTITIES.
- Die samegestelde primêre sleutel is weer aangeskakel, asook die verwysingsintegriteit op die Tender\_ID en Bill\_ID-kolomme.

**Sql C-1**

```
UPDATE Tenders INNER JOIN Fixed_Prices ON Tenders.[Contract No] = Fixed_Prices.[Contract no] SET Fixed_Prices.Tender_ID = [Tenders].[Tender_ID] WHERE (((Fixed_Prices.[Contract no])=[Tenders].[Contract No]));
```

**Sql C-2**

```
UPDATE Fixed_Prices INNER JOIN [BillOfQuantities - Copy] ON (Fixed_Prices.Proj_ID = [BillOfQuantities - Copy].Proj_ID) AND (Fixed_Prices.Code_ID = [BillOfQuantities - Copy].Code_ID) AND (Fixed_Prices.Item_ID = [BillOfQuantities - Copy].Item_ID) SET Fixed_Prices.Bill_ID = [BillOfQuantities - Copy].[Bill_ID] WHERE (((Fixed_Prices.Proj_ID)=[BillOfQuantities - Copy].[Proj_ID]) AND ((Fixed_Prices.Code_ID)=[BillOfQuantities - Copy].[Code_ID]) AND ((Fixed_Prices.Item_ID)=[BillOfQuantities - Copy].[Item_ID]));
```

**STAP B11 – AFRONDING AAN VOORGESTELDE DB:**

- 'n Kopie van oorgangsdatabasis is gemaak.
  - Die kolomme Contract No, Name, Tend No van TENDERS is verwyder.
  - Die tabelle ORIGINAL TENDER, ORIGINAL ITEMS en ORIGINAL PROJECTS is verwyder.
-

- In ITEMS is die Unit-kolom verwyder.
- In PROJECTS is die kolomme Contract No, Category, Consultant verwyder.
- BILLOFQUANTITIES is verwyder en Copy of BILLOFQUANTITIES is hernoem na BILLOFQUANTITIES.
- In FIXED PRICES is die kolomme Contract No, Prof\_ID, Code\_ID, Item\_ID verwyder.
- In BILLOFQUANTITIES is die kolomme CSRA Code, Contract No, Code, Original Tender, Escalated Tender, % of Tender verwyder.
- Alle navrae en opdaterings wat gestoor was, is verwyder in die databasis.
- Die databasis is gekompakteer.

Met die voltooiing van hierdie stap was die voorgestelde databasis gelaai met die bestaande data en al die onnodige dupliserings wat ontstaan het tydens die oordrag is verwyder. Die databasis *Loaded Modified Normalised CSRA DB.mdb* kan gevind word in gids “\Databasis\Gelaai”.







## AANHANGSEL D

---

### TOETSING VAN OORGEDRAGDE DATA

---

Tydens die oordrag van data tussen die bestaande en die voorgestelde databasis kon dit moontlik gebeur het dat sommige verwantskappe in die data verlore gegaan het. Om te verseker dat dit nie plaasgevind het nie en dat data-oordrag korrek gedoen is, was dit nodig om die data van die voorgestelde databasis te toets teen die van die bestaande databasis. Verder meer was dit ook nodig om te bevestig dat sogenaamde informasie aansigte (“views”) van die bestaande databasis in die voorgestelde databasis herskep kon word.

Drie toetse is derhalwe uitgevoer om te bepaal of dit die geval was. Dit is gedoen deur van *SQL*-navrae gebruik te maak en die resultate te toets in 'n *MS Excel*-sigblad, *Testing.xls* wat gevind kan word in die gids “\Databasis\Gelaai\Genormaliseer”. Hierdie sigblad bestaan uit drie werkblaaie, nl. *Projects*, *Items* en *Totals*, wat in die volgende drie toets gebruik is.

#### **TOETS 1:**

In hierdie toets is bepaal of die projekte se inligting korrek saamgestel kon word vanaf die voorgestelde databasis, sodat dit ooreenstem met die bestaande databasis se projekinligting.

- Die *PROJECTS* van *Loaded Old DB.mdb* is na *Projects*-werkblad van bogenoemde sigblad gekopieer.
  - 'n *SQL*-navraag, *Sql D-1*, in die voorgestelde DB wat geskryf is om dieselfde struktuur aan die data te gee soos in die *PROJECTS*-tabel van
-

die bestaande databasis is uitgevoer en die resultaat hiervan is langs die data van die projekte wat alreeds op die werkblad geplak is, geplaas.

- Die data van die twee bronne in die sigblad is met 'n *MS Excel*-funksie getoets om verskille aan te dui.

#### Sql D-1

```
SELECT Projects.Name, Contractors.Contractor, Projects.Location, Projects.[Base Month],
Tenders.Amount, Consultants.Consultant, Categories.Category FROM (Consultants INNER JOIN
(Categories INNER JOIN Projects ON Categories.Category_ID = Projects.Category_ID) ON
Consultants.Consultant_ID = Projects.Consultant_ID) INNER JOIN (Contractors INNER JOIN
Tenders ON Contractors.Contractor_ID = Tenders.Contractor_ID) ON Projects.Proj_ID = Ten-
ders.Proj_ID ORDER BY Projects.Name, Contractors.Contractor;
```

Enkele verskille het wel tussen die twee bronne se data voorgekom, maar by nadere ondersoek het dit geblyk dat hierdie verskille ontstaan het weens korreksies wat aangebring is tydens die oordrag proses. Geen foute is dus gemaak nie.

#### TOETS 2:

In hierdie toets is bepaal of die items se inligting korrek kon saamgestel word vanaf die voorgestelde databasis, sodat dit ooreenstem met die bestaande databasis. Die drie stappe stem presies oor een met die van Toets 1, behalwe vir die *SQL*-navraag, *Sql D-1*, wat verander het.

#### Sql D-1

```
SELECT Items.Item_Num, Items.Discription, Units.Unit FROM Units INNER JOIN Items ON
Units.Unit_ID = Items.Unit_ID ORDER BY Items.Item_Num;
```

Uit die vergelyking wat in die werkblad uitgevoer is, het dit geblyk dat daar geen verskille voorgekom het nie.

#### TOETS 3:

By die voorafgaande twee toets is twee van die drie tabelle van die bestaande databasis getoets, nl. **PROJEK** en **ITEM**. Die ideaal sou wees om ook te toets of die **TENDER**-tabel se data korrek versprei is tussen die voorgestelde databasis se tabelle. Weens die groot hoeveelheid data was dit egter nie moontlik om direk so 'n toets te kon uitvoer nie. As alternatief is daar eerder gekyk of die opsommende inligting per projek tussen die twee databasis ooreen gestem het. Dit is gedoen deur 'n *SQL*-navraag, *Sql D-1*, uit te voer wat die volgende gedoen het:

- die som van die oorspronklike eenheidtariewe per projek is bereken;
- die som van die aangepasde eenheidtariewe per projek is bereken;

- die som van die hoeveelhede van items per projek is bereken;
- die som van die produk tussen 'n item se eenheidtarief en hoeveelheid per projek is bereken.

**SQL D-1**

```
SELECT Projects.Name, Contractors.Contractor, Sum(Fixed_Prices.[Orginal Tender]) AS
[SumOfOrginal Tender], Sum(Fixed_Prices.[Escalated Tender]) AS [SumOfEscalated Tender],
Sum(BillOfQuantities.Quantity) AS SumOfQuantity,
Sum([BillOfQuantities][Quantity]*[Fixed_Prices][Orginal Tender]) AS Total FROM (Projects
INNER JOIN (BillOfQuantities INNER JOIN Fixed_Prices ON BillOfQuantities.Bill_ID =
Fixed_Prices.Bill_ID) ON Projects.Proj_ID = BillOfQuantities.Proj_ID) INNER JOIN (Contractors
INNER JOIN Tenders ON Contractors.Contractor_ID = Tenders.Contractor_ID) ON (Ten-
ders.Tender_ID = Fixed_Prices.Tender_ID) AND (Projects.Proj_ID = Tenders.Proj_ID) GROUP
BY Projects.Name, Contractors.Contractor ORDER BY Projects.Name, Contractors.Contractor;
```

Hierdie data afkomstig uit die voorgestelde databasis is vergelyk met data afkomstig uit die bestaande databasis wat ook met 'n *SQL*-navraag, *Sql D-2*, bepaal is.

**SQL D-2**

```
SELECT Tenders.[Contract No], Sum(Tenders.[Orginal Tender]) AS [SumOfOrginal Tender],
Sum(Tenders.[Escalated Tender]) AS [SumOfEscalated Tender], Sum(Tenders.Quantity) AS
SumOfQuantity, Sum([Tenders][Orginal Tender]*[Tenders][Quantity]) AS Expr1 FROM Tenders
GROUP BY Tenders.[Contract No];
```

Die vergelyking is in die *Totals*-werkblad van bogenoemde sigblad gedoen. Hieruit het dit geblyk dat daar verskille by die hoeveelhede voorgekom het weens inkonsekwentheid. Hierdie inkonsekwentheid spruit voort uit die samevoeging van die twee (soms drie) Hoeveelheidslyste van die verskillende tenders, waarin daar inkonsekwentheid van die oorspronklike data voorgekom het weens vingerfoute wat tydens invoering gemaak is. Dit is nie verder ondersoek nie.



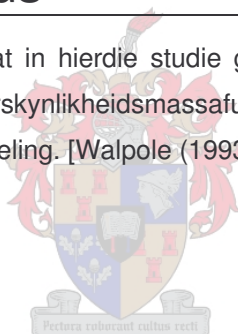
## AANHANGSEL E

---

### VERDELINGS

---

Die verdelings wat in hierdie studie gebruik is word vervolgens getoon. Elke verdeling se waarskynlikheidsmassafunksie word getoon as 'n grafiek van die vorm van die verdeling. [Walpole (1993), Ang (1975)]



**GUMBELVERDELING**

Die Gumbelverdeling (of ekstreemwaarde) word in die algemeen gebruik om die grootste waarde van 'n respons te beskryf oor 'n bepaalde periode. Die Gumbelverdeling se waarskynlikheidsmassafunksie, met die parameters modus,  $m$ , en skaal,  $\alpha$ , word gegee deur:

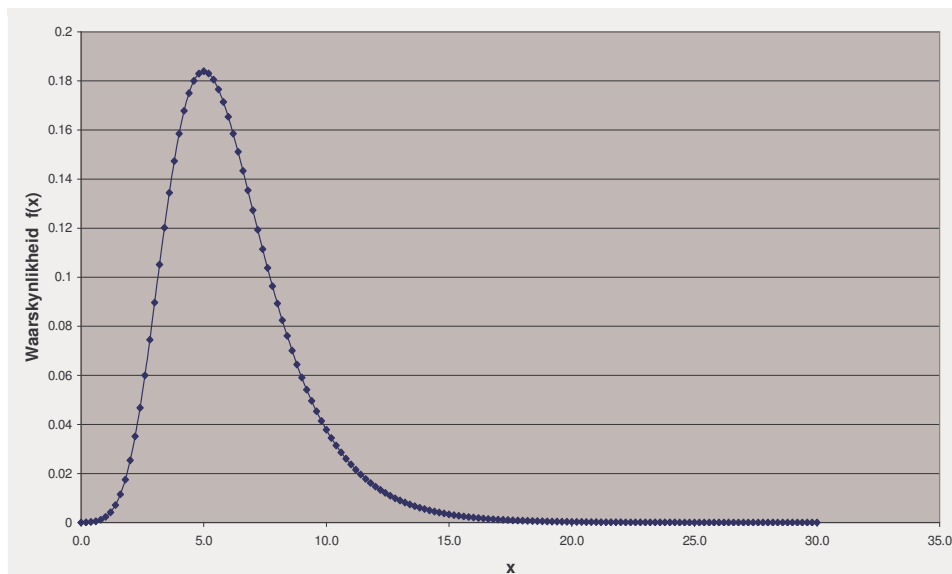
$$f(x) = \frac{1}{\alpha} z e^{-z} \quad \text{vir } -\infty < x < \infty$$

waar:

$$z = e^{-\left(\frac{x-m}{\alpha}\right)}$$

Die skaal,  $\alpha$ , is 'n funksie van die variansie,  $\sigma^2$ , en word gegee deur:

$$\alpha = \sqrt{\frac{6\sigma^2}{\pi^2}}$$



**FIGUUR E-1: GUMBELVERDELING**

**LOG-NORMAALVERDELING**

Die waarskynlikheidsmassafunksie van die log-normaalverdeling word gegee deur:

$$f(x) = \frac{1}{x\xi\sqrt{2\pi s}} e^{-\frac{1}{2}\left(\frac{\ln(x)-\lambda}{\xi}\right)^2} \quad \text{vir } 0 \leq x < \infty$$

Die parameters van die funksie is die log-gemiddelde,  $\lambda$ , en die log-standaardafwyking,  $\xi$ , van  $\ln(x)$ , met:

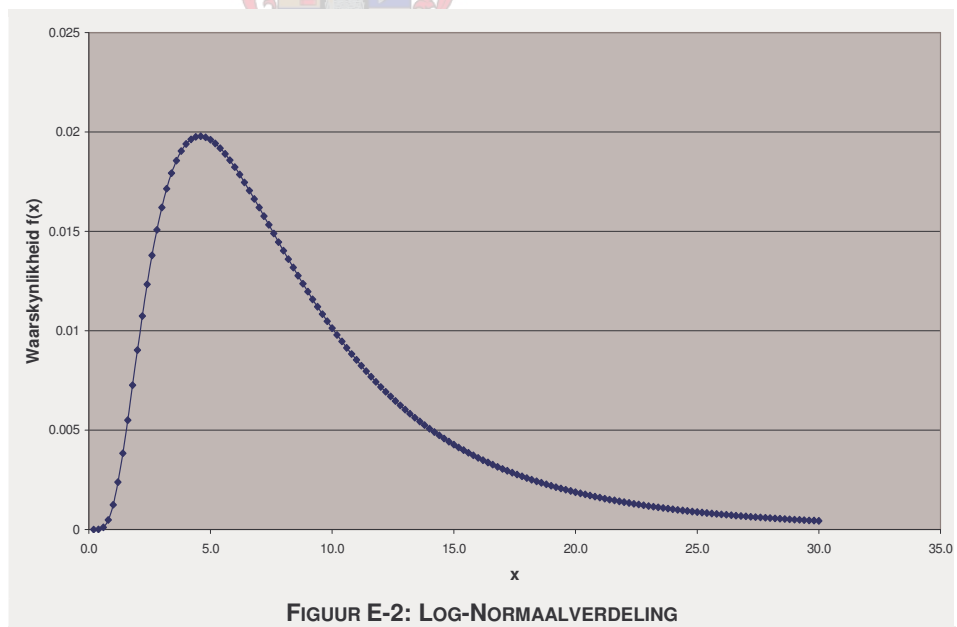
$$\lambda = E(\ln(x))$$

$$\xi = \sqrt{\text{Var}(\ln(x))}$$

Die parameters kan bepaal word vanaf die gemiddelde,  $\mu$ , en die standaardafwyking,  $\sigma$ :

$$\xi^2 = \ln\left(1 + \frac{\sigma^2}{\mu^2}\right)$$

$$\lambda = \ln(\mu) - \frac{1}{2}\xi^2$$





**DRIEHOEKVERDELING**

Die driehoekverdeling beskryf 'n situasie waar 'n minimum, maksimum en waarskynlike waarde bekend is. Die waarskynlikheidsmassafunksie word gegee deur:

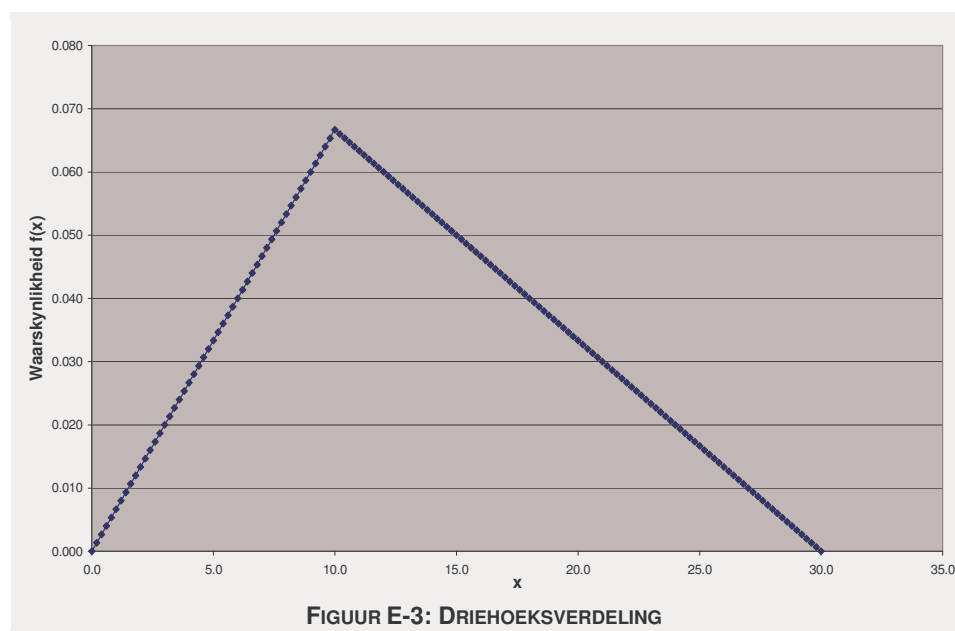
$$f(x) = \frac{h(x - \text{Min})}{\text{Waarskynlikste} - \text{Min}} \quad \text{vir} \quad \text{Min} < x < \text{Waarskynlikste}$$

of

$$f(x) = \frac{h(x - \text{Waarskynlikste})}{\text{Waarskynlikste} - \text{Maks}} + h \quad \text{vir} \quad \text{Waarskynlikste} < x < \text{Maks}$$

waar

$$h = \frac{2}{\text{Maks} - \text{Min}}$$



**GAMMAVERDELING**

Die Gammafunksie word gedefinieer met:

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx \quad \text{vir } \alpha > 0$$

Die kontinue willekeurige veranderlike,  $x$ , beskik oor 'n Gammaverdeling, met parameters  $\alpha$  en  $\beta$ , indien sy waarskynlikheidsmassafunksie gegee word deur:

$$f(x) = \left( \frac{x^{\alpha-1} e^{-x/\beta}}{\beta^{\alpha} \Gamma(\alpha)} \right) \quad \text{vir } x > 0$$

en

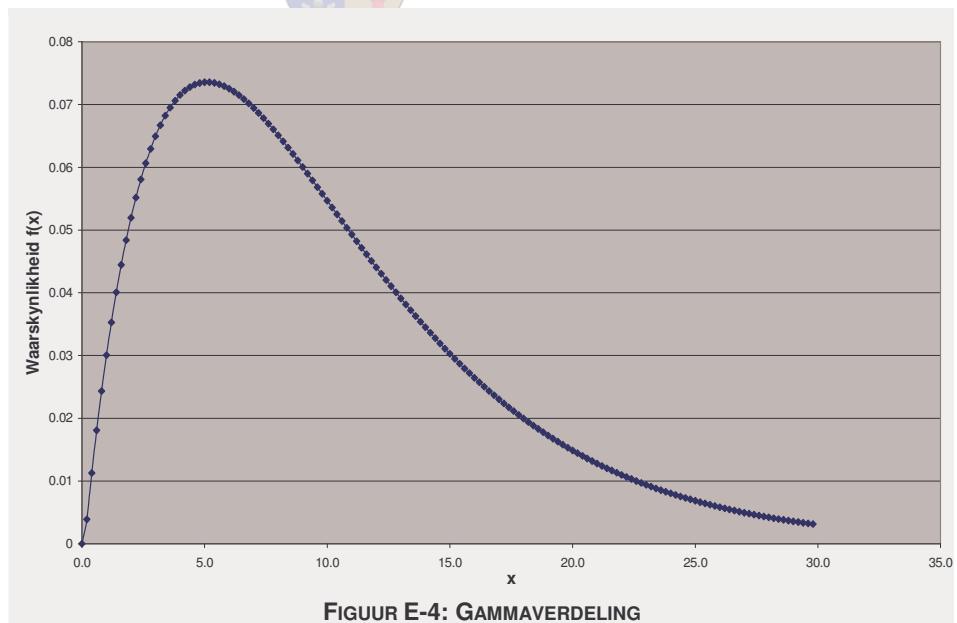
$$f(x) = 0 \quad \text{as } x \leq 0$$

Die gemiddelde,  $\mu$ , en variansie,  $\sigma^2$ , in terme van die parameters van die massafunksie word gegee deur:

$$\mu = \alpha\beta$$

en

$$\sigma^2 = \alpha\beta^2$$



**BETAVERDELING**

Die waarskynlikheidsmassafunksie van die Betaverdeling word gegee deur:

$$f(x) = \frac{1}{B(q,r)} \frac{(x-a)^{q-1} (b-x)^{r-1}}{(b-a)^{q+r-1}} \quad \text{vir } a \leq x \leq b$$

en  $f(x) = 0$  andersins.

Die parameters van die funksie is  $q$  en  $r$ , terwyl  $B(q,r)$  die Betafunksie is:

$$B(q,r) = \int_0^1 x^{q-1} (1-x)^{r-1} dx$$

Die Betafunksie is ook verwant aan die Gammafunksie:

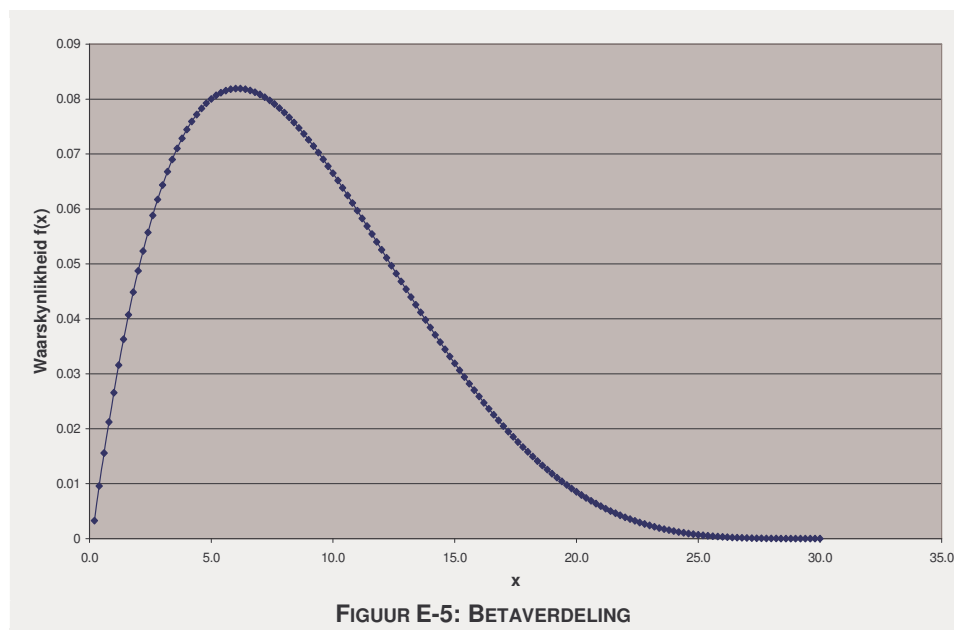
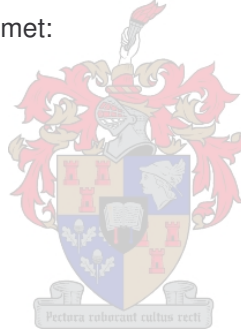
$$B(q,r) = \frac{\Gamma(q)\Gamma(r)}{\Gamma(q+r)}$$

Die gemiddelde en variansie as 'n funksie van die Betaverdeling parameters  $q$  en  $r$ , en die grense  $a$  en  $b$ , word beskryf met:

$$\mu = a + \frac{q}{q+r}(b-a)$$

en

$$\sigma^2 = \frac{qr}{(q+r)^2(q+r+1)}(b-a)^2$$



**CHI-KWADRAATVERDELING**

Die Chi-Kwadraatverdeling is 'n spesiale geval van die Gammaverdeling, en word beskryf deur:

$$f(x) = \frac{1}{2^{v/2} \Gamma(v/2)} x^{v/2-1} e^{-x/2} \quad \text{vir } x > 0$$

en

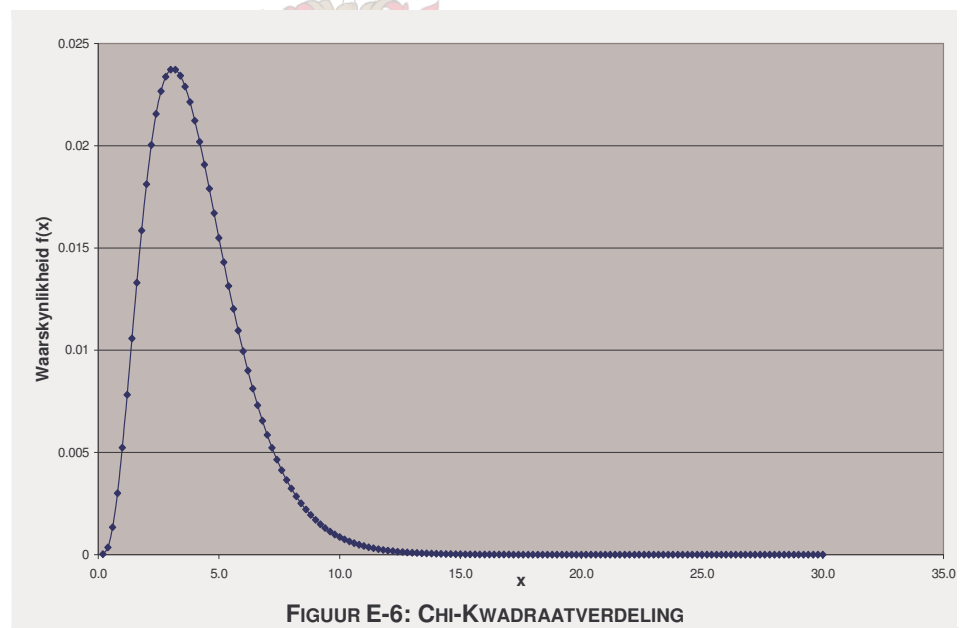
$$f(x) = 0 \quad \text{andersins.}$$

Die verdeling beskik oor 'n enkele parameter, die grade van vryheid,  $\nu$ . Die gemiddelde,  $\mu$ , en variansie,  $\sigma^2$ , van die Chi-Kwadraatverdeling word gegee deur:

$$\mu = \nu$$

en

$$\sigma^2 = 2\nu$$





## AANHANGSEL F

### MCS-TEGNIK

Die MCS-tegniek word gedemonstreer aan die hand van 'n denkbeeldige projek wat in *Hoofstuk 5* bespreek word. Drie simulasiemodelle word saamgestel om die effek van die verskillende verdelings op die resultate te toon. Die data vir die simulaties is verkry uit die CSRA-databasis wat in die gids “\Databasis” gestoor is. Dit word as volg gedoen:

#### STAP 1 – DATA:



Vir die berekeninge is die eenheidtariewe en hoeveelhede se data van die beduidende items benodig. Die eenheidtariewe moes as stochastiese veranderlikes gemodelleer word en dus moes al die eenheidtariewe van elke beduidende item bepaal word, teenoor die hoeveelhede waar slegs die gemiddeldes van elke item benodig is.

- Die eenheidtariefdata is met 'n *SQL*-navraag, *Sql F-1*, bepaal wat die tendensaangepaste data van elke item teenoor die projek waar dit voorkom teruggestuur het.

#### Sql F-1

```
TRANSFORM Avg([Aangepaste Data].NewEscalatedTender) AS AvgOfNewEscalatedTender
SELECT [Aangepaste Data].Name
FROM [Aangepaste Data]
WHERE ((([Aangepaste Data].Item_Num)>="16.02" And ([Aangepaste Data].Item_Num)<>"42.02(a)(xiii)" And ([Aangepaste Data].Item_Num)<"66.15"))
GROUP BY [Aangepaste Data].Name
PIVOT [Aangepaste Data].Item_Num;
```

- Die hoeveelhede is met 'n SQL-navraag, *Sql F-2*, bepaal. Die navraag het die gemiddelde hoeveelhede van elk van die beduidende items bereken.

#### SQL F-2

```
SELECT [Geselekteerde Items].[Item Num], Codes.Code, Avg(BillOfQuantities.Quantity) AS
AvgOfQuantity
FROM (Items INNER JOIN (Codes INNER JOIN BillOfQuantities ON Codes.Code_ID = BillOf-
Quantities.Code_ID) ON Items.Item_ID = BillOfQuantities.Item_ID) INNER JOIN [Geselekteerde
Items] ON Items.Item_Num = [Geselekteerde Items].[Item Num]
GROUP BY [Geselekteerde Items].[Item Num], Codes.Code
HAVING ((([Geselekteerde Items].[Item Num])>="16.02" And ([Geselekteerde Items].[Item
Num])<>"42.02(a)(xiii)" And ([Geselekteerde Items].[Item Num])<"66.15") AND ((Co-
des.Code)="A"));
```

- Albei hierdie stelle data is na drie verskillende *MS Excel*-sigblaaie geko- pieer wat in die gids “\Aanhangsel F” op meegaande KS is.

#### **STAP 2 – BEREKENINGE:**

- Die verdelings wat op elke item se eenheidtariefdata die beste pas is eers met *Batch Fit* van *Crystal Ball* bepaal. In die sigblad, *MC-simulasie verdelings outomaties.xls*, is die bes passende verdeling met *Batch Fit* gekies. In die sigblad, *MC-simulasie verdelings normaal.xls* is *Batch Fit* gebruik om slegs bes passende normaalverdelings te pas en in die sig- blad, *MC-simulasie verdelings log nom.xls* is slegs bes passende log- normaalverdelings aanvaar.
- Die koste funksie is in elke sigblad opgestel deur die somproduk van die eenheidtariewe en hoeveelhede te bepaal met die *MS Excel*-funksie “Sumproduct”.
- Hierdie funksie is gedefinieer in elke sigblad as die sogenaamde voor- spellings sel vir *Crystal Ball*.
- Crystal Ball* is opgestel dat dit vir elke sigblad dieselfde 20 000 ewekan- sigte getalle gebruik tydens die simulasies.
- Die simulasies is vir elke sigblad uitgevoer met en sonder korrelasie.

#### **STAP 3 – RESULTATE:**

Die resultate van die berekening kan op elke sigblad in die *Report*- en *Report (with correl)*-werkblad gesien word. Die data is gekombineer en word getoon in *Figuur 4-3* en *Figuur 4-4*.

---

## AANHANGSEL G

---

### LSKF-TEGNIK

---

Die LSKF-tegniek word hier geïmplementeer vir verskillende vorms van die kostefunksie, nl. die liniêre vorm waar die veranderlikes onafhanklik van mekaar is, die liniêre vorm waar die veranderlikes afhanklik van mekaar is en die nie-liniêre vorm waar die veranderlikes ook afhanklik van mekaar is. Die data vir die simulasies is verkry uit die CSRA-databasis wat in die gids “\Databasis” gestoor is.



#### **STAP 1 – DATA:**

- Die gemiddelde en standaardafwykings wat in *Tabel 5-4* vir die tendens-aangepaste data getoon word van die beduidende items, is na die *MS Excel*-sigblad, *LSKF.xls*, in die gids “\Aanhangsel G” op meegaande KS gekopieer.
  - Die hoeveelhede is met ’n *SQL*-navraag, *Sql G-1*, bepaal. Die navraag het die gemiddelde hoeveelhede van elk van die beduidende items bereken, dit is ook oorgedra na die sigblad.
  - Vir die gevalle waar dit aangeneem is dat die veranderlikes nie onafhanklik is nie, was die korrelasie tussen die veranderlikes benodig. Dit is verkry vanaf twee bronne. In *Aanhangsel O* is die Pearson-korrelasies verkry wat tussen die eenheidtariewe en die hoeveelhede voorkom. Die “Rank”-korrelasies tussen die eenheidtariewe is verkry van die MCS-tegniek in *Aanhangsel F*.
-



**Sql G-1**

```

SELECT [Geselekteerde Items].[Item Num], Codes.Code, Avg(BillOfQuantities.Quantity) AS
AvgOfQuantity
FROM (Items INNER JOIN (Codes INNER JOIN BillOfQuantities ON Codes.Code_ID = BillOf-
Quantities.Code_ID) ON Items.Item_ID = BillOfQuantities.Item_ID) INNER JOIN [Geselekteerde
Items] ON Items.Item_Num = [Geselekteerde Items].[Item Num]
GROUP BY [Geselekteerde Items].[Item Num], Codes.Code
HAVING ((([Geselekteerde Items].[Item Num])>="16.02" And ([Geselekteerde Items].[Item
Num])<>"42.02(a)(xiii)" And ([Geselekteerde Items].[Item Num])<"66.15") AND ((Co-
des.Code)="A"));

```

**STAP 2 – BEREKENINGE:**

Die verskillende vorms van die LSKF is geïmplementeer in die sigblad op verskillende werkblaaie:

- In die *Sonder korrelasie*-werkblad is die liniêre vorm sonder enige afhanklikhede opgestel.
- In die *Met korrelasie (Ranked)*-werkblad is die liniêre vorm met afhanklikhede tussen die eenheidtariewe opgestel. Die korrelasie wat met die MCS-tegniek bepaal is, is hier gebruik.
- In die *Met korrelasie (Pearson)*-werkblad is die liniêre vorm met afhanklikhede tussen die eenheidtariewe opgestel. Die korrelasie wat in *Aanhangsel O* bepaal is, is hier gebruik.
- In die *Nie-liniêr*-werkblad is die nie-liniêre vorm opgestel. Hierdie opstelling veroorsaak sogenaamde rekursiewe verwysings en derhalwe is van *MS Excel* se "Iteration"-funksionaliteit gebruik gemaak om die werkblad te balanseer.
- In die *Nie-liniêr (by gem)*-werkblad is die nie-liniêre vorm opgestel, dit was nie nodig om hier te itereer nie, want die falingspunte is beskou by die gemiddelde waardes.

**STAP 3 – RESULTATE:**

Die resultate van die berekeninge is verwerk in vier grafieke wat in *Figuur 4-6*, *Figuur 4-7*, *Figuur 4-8* en *Figuur 4-9* getoon word en ook op die sigblad gevind kan word.

---

## AANHANGSEL H

---

### NEURALE NETWERK

---

Verwys na die meegaande KS vir die sigblad van die neurale netwerk. Die stappe in die sigblad stem ooreen met die stappe soos verduidelik in *Afdeling 4.6.1*.

---





## AANHANGSEL I

---

### **KUMULATIEWE BYDRAE TOT PROJEKKOSTE**

---

Dit is bekend dat in konstruksieprojekte 'n klein aantal items 'n groot bydrae tot die totale koste van 'n projek lewer. Dit bring mee dat slegs 'n beperkte aantal beduidende items in diepte ondersoek hoef te word vir besluitnemingsdoelendes sonder 'n noemenswaardige afname in die akkuraatheid waarmee risiko's gemodelleer word.

In hierdie aanhangsel word bepaal of die verskynsel van 'n klein aantal items wat 'n groot bydrae tot die totale kostes van 'n projek maak ook vir die data wat in die CSRA-databasis gestoor word van toepassing is.

#### **STAP 1 – BEPALING VAN DATA:**

Die data is met die volgende stappe uit die databasis wat in die gids “\Databasis” is, verkry:

- Die eenheidtarief en hoeveelheid van elke item per projek is bepaal. Dit is uitgevoer met 'n *SQL*-navraag, *Sql I-1*, wat die items per projek georden het van hoogste tot laagste koste.
-

#### SQL I-1

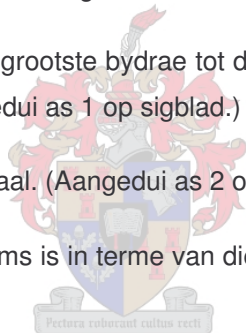
```
SELECT Projects.Name, Contractors.Contractor, Items.Item_Num, Codes.Code, BillOfQuantities.Quantity, Fixed_Prices.[Original Tender], [BillOfQuantities][Quantity]*[Fixed_Prices][Original Tender] AS Expr1, Tenders.Amount
FROM (Projects INNER JOIN (Items INNER JOIN (Codes INNER JOIN (BillOfQuantities INNER JOIN Fixed_Prices ON BillOfQuantities.Bill_ID = Fixed_Prices.Bill_ID) ON Codes.Code_ID = BillOfQuantities.Code_ID) ON Items.Item_ID = BillOfQuantities.Item_ID) ON Projects.Proj_ID = BillOfQuantities.Proj_ID) INNER JOIN (Contractors INNER JOIN Tenders ON Contractors.Contractor_ID = Tenders.Contractor_ID) ON (Tenders.Tender_ID = Fixed_Prices.Tender_ID) AND (Projects.Proj_ID = Tenders.Proj_ID)
GROUP BY Projects.Name, Contractors.Contractor, Items.Item_Num, Codes.Code, BillOfQuantities.Quantity, Fixed_Prices.[Original Tender], Tenders.Amount
HAVING (((Projects.Name)<>"NW682/95") AND ((Contractors.Contractor)="Tend1"))
ORDER BY Projects.Name, [BillOfQuantities][Quantity]*[Fixed_Prices][Original Tender] DESC;
```

- Hierdie data is oorgedra na 'n *MS Excel*-sigblad, *Persentasie bydrae.xls*, in die gids "Aanhangsel I" op meegaande KS.

#### **STAP 2 – BEREKENINGE:**

Die berekeninge is vervolgens in die sigblad uitgevoer:

- 'n Rangorde posisie in terme van grootste bydrae tot die koste per projek is aan die items toegeken. (Aangedui as 1 op sigblad.)
- Die aantal items per projek is bepaal. (Aangedui as 2 op sigblad.)
- Die kumulatiewe koste van die items is in terme van die rangorde bepaal. (Aangedui as 3 op sigblad.)
- Die totale koste per projek is bepaal. (Aangedui as 4 op sigblad.)
- Die persentasie van die aantal items en die dienooreenkomstige kumulatiewe bydrae van die items per projek is bereken. (Aangedui as 5 & 6 op sigblad.)



#### **STAP 3 – RESULTATE**

Die resultate is in 'n grafiek saamgevat en word getoon in *Figuur 5-2*.

---

## AANHANGSEL J

### IDENTIFISERING VAN BEDUIDENDE ITEMS

Die identifisering van beduidende items wat meer as 1% bydrae lewer tot die totale koste van die projekte en wat by meer as 5% van die projekte in die databasis voorkom, word hier bepaal. Die data is afkomstig van die CSRA-databasis wat in die gids "Databasis" gestoor is.

#### **STAP 1 – VERDELING IN SERIES:**

- Die data in die databasis is verdeel tussen die nege *Series* deur van die 'n *SQL*-navraag gebruik te maak wat dit tussen die eerste en tweede tender verdeel het en ook tussen die nege *Series*. In *Sql J-1* word die navraag getoon om dit vir vir die eerste tender te bepaal.

#### **SQL J-1**

```
SELECT Series.Series, Count(Fixed_Prices.[Percentage of Tender]) AS [CountOfPercentage of
Tender]
FROM (Contractors INNER JOIN Tenders ON Contractors.Contractor_ID = Ten-
ders.Contractor_ID) INNER JOIN (Series INNER JOIN (Items INNER JOIN (BillOfQuantities
INNER JOIN Fixed_Prices ON BillOfQuantities.Bill_ID = Fixed_Prices.Bill_ID) ON Items.Item_ID
= BillOfQuantities.Item_ID) ON Series.Series_ID = Items.Series_ID) ON Tenders.Tender_ID =
Fixed_Prices.Tender_ID
WHERE (((Contractors.Contractor)="Tend1"))
GROUP BY Series.Series;
```

#### **STAP 2 – BEPALING VAN ITEMS WAT 1%-BYDRAE LEWER:**

- Die data afkomstig uit *Stap 1* is gebruik om die projekitems wat 'n bydrae meer as 1% tot die totale koste van 'n projek maak, te bepaal met die navraag, *Sql J-1*.

#### SQL J-1

```
SELECT Series.Series, Count(Fixed_Prices.[Percentage of Tender]) AS [CountOfPercentage of Tender]
FROM Series INNER JOIN ((Projects INNER JOIN (Items INNER JOIN (BillOfQuantities INNER JOIN Fixed_Prices ON BillOfQuantities.Bill_ID = Fixed_Prices.Bill_ID) ON Items.Item_ID = BillOfQuantities.Item_ID) ON Projects.Proj_ID = BillOfQuantities.Proj_ID) INNER JOIN (Contractors INNER JOIN Tenders ON Contractors.Contractor_ID = Tenders.Contractor_ID) ON (Tenders.Tender_ID = Fixed_Prices.Tender_ID) AND (Projects.Proj_ID = Tenders.Proj_ID)) ON Series.Series_ID = Items.Series_ID
WHERE (((Contractors.Contractor)="Tend1") AND ((Fixed_Prices.[Percentage of Tender])>1))
GROUP BY Series.Series;
```

### **STAP 3 – BEPALING VAN ITEMS MET 'N VOORKOMS VAN MEER AS 5%:**

- Die data afkomstig uit *Stap 2* is gebruik om te bepaal watter van die items by meer as 5% van die projekte voorkom met navraag *Sql J-1*.

#### SQL J-1

```
SELECT Series.Series, Count([1% bydrae 5% voorkoms Tend1].Item_Num) AS CountOfItem_Num
FROM [1% bydrae 5% voorkoms Tend1] INNER JOIN (Series INNER JOIN Items ON Series.Series_ID = Items.Series_ID) ON [1% bydrae 5% voorkoms Tend1].Item_Num = Items.Item_Num
GROUP BY Series.Series;
```



### **STAP 4 – VOORKOMS IN TENDER 1 EN 2**

- Om te bepaal watter items in albei van die tenders voorkom is die inligting afkomstig uit *Stap 3* van tender een en twee gebruik en die navraag *Sql J-1* is hierop uitgevoer.

#### SQL J-1

```
SELECT [1% bydrae 5% voorkoms Tend1].Item_Num, [1% bydrae 5% voorkoms Tend1].Discription, Units.Unit
FROM Units INNER JOIN (Items INNER JOIN ([1% bydrae 5% voorkoms Tend1] INNER JOIN [1% bydrae 5% voorkoms Tend2] ON [1% bydrae 5% voorkoms Tend1].Item_Num = [1% bydrae 5% voorkoms Tend2].Item_Num) ON Items.Item_Num = [1% bydrae 5% voorkoms Tend2].Item_Num) ON Units.Unit_ID = Items.Unit_ID
ORDER BY [1% bydrae 5% voorkoms Tend1].Item_Num DESC;
```

### **STAP 5 – RESULTATE:**

- Die resultate van die soektogte word in *Tabel 5-2* en *Tabel 5-3* getoon.

## AANHANGSEL K

---

### BEPALING VAN UITSKIETER PROJEKTE

---

Om die uitskieter projekte te bepaal waarin daar 'n noemenswaardige aantal bo en onder uitskieter eenheidariewe van die beduidende items voorkom is daar van vyf *SQL*-navrae gebruik gemaak. Die data wat hiervoor gebruik is, is afkomstig van die CSRA-databasis wat in die gids “\Databasis” gevind kan word. Die data is egter eers aangepas om enige “orde grootte”-uitskieters en tydafhanklikhede te verwyder, voordat met die ontleding begin is.



Pectora roburant cultus recti

In die eerste twee navrae, *Sql K-1* en *Sql K-2*, is die items en hul projekte wat met meer as een standaardafwyking onderskeidelik na onder en bo uitgewyk het, geïdentifiseer. Die aantal items wat vir elke van hierdie gevalle gevind is, word getoon in *Tabel 5-4*.

---



**SQL K-1**

```

SELECT [Aangepaste Data].Name, [Aangepaste Data].[Base Month], [Aangepaste Data].Item_Num, [Aangepaste Data].[Original Tender], [Aangepaste Data].NewEscalatedTender, [Gem en Var van Geslk data].AvgOfNewEscalatedTender, [Gem en Var van Geslk data].StDevOfNewEscalatedTender
FROM [Aangepaste Data] INNER JOIN [Gem en Var van Geslk data] ON [Aangepaste Data].Item_Num = [Gem en Var van Geslk data].Item_Num
WHERE ((([Aangepaste Data].Item_Num)>"16") AND ((If([Aangepaste Data].[NewEscalatedTender]<=[Gem en Var van Geslk data].[AvgOfNewEscalatedTender]-1*[Gem en Var van Geslk data].[StDevOfNewEscalatedTender]),1,0)=1))
ORDER BY [Aangepaste Data].Name, [Aangepaste Data].Item_Num;

```

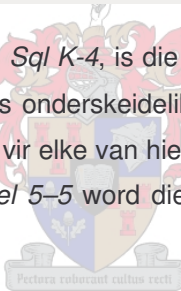
**SQL K-2**

```

SELECT [Aangepaste Data].Name, [Aangepaste Data].[Base Month], [Aangepaste Data].Item_Num, [Aangepaste Data].[Original Tender], [Aangepaste Data].NewEscalatedTender, [Gem en Var van Geslk data].AvgOfNewEscalatedTender, [Gem en Var van Geslk data].StDevOfNewEscalatedTender
FROM [Aangepaste Data] INNER JOIN [Gem en Var van Geslk data] ON [Aangepaste Data].Item_Num = [Gem en Var van Geslk data].Item_Num
WHERE ((([Aangepaste Data].Item_Num)>"16") AND ((If([Aangepaste Data].[NewEscalatedTender]>=[Gem en Var van Geslk data].[AvgOfNewEscalatedTender]+1*[Gem en Var van Geslk data].[StDevOfNewEscalatedTender]),1,0)=1))
ORDER BY [Aangepaste Data].Name, [Aangepaste Data].Item_Num;

```

In die volgend twee navrae, *Sql K-3* en *Sql K-4*, is die projekte wat meer as vyf van hierdie uitskieter beduidende items onderskeidelik na onder en bo per projek het, bepaal. Die aantal projekte wat vir elke van hierdie gevalle gevind is, word ook getoon in *Tabel 5-4* en in *Tabel 5-5* word die projekte getoon wat geïdentifiseer is.



Ongeveer 'n derde van hierdie projekte se brondokumente is by die databasisbestuurders opgespoor. Die dokumente is ondersoek om te bepaal of enige inligting gevind kon word om te verklaar waarom daar soveel van hierdie uitskieter eenheidtariewe in die projekte voorgekom het en of dit foute is wat gemaak was. Dit is bevestig dat die inligting korrek was, maar geen verklaring kon gevind word vir die groot afwykings nie.

**SQL K-3 VOORKOMS VAN BO BEGRENSEDE PROJEK**

```

SELECT [Voorkoms van uitskieter items (Bo begrens)].Name, Count([Voorkoms van uitskieter items (Bo begrens)].Name) AS CountOfName
FROM [Voorkoms van uitskieter items (Bo begrens)]
GROUP BY [Voorkoms van uitskieter items (Bo begrens)].Name
HAVING (((Count([Voorkoms van uitskieter items (Bo begrens)].Name))>=5));

```

**SQL K-4 VOORKOMS VAN ONDER BEGRENSEDE PROJEK**

```

SELECT [Voorkoms van uitskieter items (Onder begrens)].Name, Count([Voorkoms van uitskieter items (Onder begrens)].Name) AS CountOfName
FROM [Voorkoms van uitskieter items (Onder begrens)]
GROUP BY [Voorkoms van uitskieter items (Onder begrens)].Name
HAVING (((Count([Voorkoms van uitskieter items (Onder begrens)].Name))>=5));

```

Laastens is bepaal of daar enige oorvleueling van hierdie bo en onder uitskieter items in dieselfde projek voorkom, *Sql K-5*. Dit is gevind dat daar geen oorvleueling is nie. Die resultaat van hierdie navraag word ook in *Tabel 5-4* getoon.

**SQL K-5 BEIDE PROJEKTE**

```
SELECT [Voorkoms van uitskieter projekte (Bo begrens)].Name  
FROM [Voorkoms van uitskieter projekte (Bo begrens)] INNER JOIN [Voorkoms van uitskieter  
projekte (Onder begrens)] ON [Voorkoms van uitskieter projekte (Bo begrens)].Name = [Voor-  
koms van uitskieter projekte (Onder begrens)].Name;
```





## AANHANGSEL L

---

### INFLASIE-EFFEK

---

Om die effek van inflasie op die basiese statistieke van data te toon en ook die effek daarvan op die statistiese verdelings van die data is van *MS Excel* en *Crystall Ball* gebruik gemaak. Dit is in drie stappe vir die uniform- en normaalverdeelde data gedoen; vervolgens word die uniformverdeelde data gebruik om die metode te bespreek. Die *MS Excel*-sigblad, *Effek van inflasie op verdelings.xls*, waar in die berekeninge gedoen is kan in die gids “Aanhangsel L” op meegaande KS gevind word.

#### **STAP 1 – SKEP VAN DATA:**

- Tien jaar se datums, met 'n twee week interval vanaf 1990 is geskep. (Aangedui as 1 op sigblad.)
- Ewekansige getalle wat uniformverdeel is, met 'n gemiddeld van 75 en 'n standaardafwyking van 15 (20% van die gemiddeld) is met *MS Excel* se ewekansige getal generator geskep. (Aangedui as 2 op sigblad.)
- Die data is met die inflasie-formule, *Vgl. 5-1*, aangepas vir 'n inflasie van 10% en 'n deflasie van 10%. (Aangedui as 3 & 4 op sigblad.)

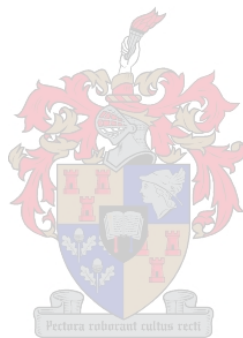
#### **STAP 2 – ONTLEDING:**

- Die gemiddelde en standaardafwyking van die drie stelle data (oorspronklik, 10% inflasie en 10% deflasie) is bereken met *MS Excel* se ingeboude funksies. (Aangedui as 5 op sigblad.)
-

- Die statistiese verdelings wat die beste pas om elke van die stelle data is met *Crystal Ball* se *Batch Fit*-funksie bepaal. (Aangedui as 6 op sigblad.)

**STAP 3 – RESULTATE:**

- Die resultate van die basiese statistiese analyses word in *Tabel 5-6* getoon en die resultate van die passings in *Figuur 5-7* vir uniformverdeelde data.
- 



## AANHANGSEL M

---

### INFLASIE-AANPASSINGS

---

Weens inflasie is dit nodig om die historiese eenheidtariewe van items aan te pas. Aanpassings is in die CSRA-databasis uitgevoer deur die databasisbestuurders en derhalwe word dit verwag dat stygende (of dalende) tendense nie sal voorkom nie. In hierdie aanhangsel word verduidelik hoe ondersoek ingestel is om te bepaal of dit wel so is. Waar dit gevind is dat dit nie die geval is nie, is die data aangepas om die tendense te verwyder. Die data is afkomstig van die CSRA-databasis wat in die gids “\Databasis” gestoor is.

#### **STAP 1 – DATA:**

- Die historiese projek eenheidtariewe, die begin datum van die projek wat ooreenstem met die eenheidtarief en die inflasie-aangepaste data van elk van die 44-items is met 'n *SQL*-navraag, *Sql M-1*, vanuit die databasis verkry. Die data is gegroeppeer volgens die itemnommers en gerangskik volgens 'n stygende begin datum.
  - Hierdie data is oorgedra na 'n *MS Excel*-sigblad, *Inflasie en statistieke.xls*, in die gids “\Aanhangsel M” op meegaande KS.
-

**Sql M-1**

```

SELECT [Geselekteerde Items].[Item Num], Count(Codes.Code) AS CountOfCode, Pro-
jects.[Base Month], Projects.Name, Count(Projects.Name) AS CountOfName,
Avg(BillOfQuantities.Quantity) AS AvgOfQuantity, Avg(Fixed_Prices.[Orginal Tender]) AS
[AvgOfOrginal Tender], Avg(Fixed_Prices.[Escalated Tender]) AS [AvgOfEscalated Tender]
FROM (Projects INNER JOIN ((Items INNER JOIN (Codes INNER JOIN (BillOfQuantities INNER
JOIN Fixed_Prices ON BillOfQuantities.Bill_ID = Fixed_Prices.Bill_ID) ON Codes.Code_ID =
BillOfQuantities.Code_ID) ON Items.Item_ID = BillOfQuantities.Item_ID) INNER JOIN [Geselek-
teerde Items] ON Items.Item_Num = [Geselekteerde Items].[Item Num]) ON Projects.Proj_ID =
BillOfQuantities.Proj_ID) INNER JOIN (Contractors INNER JOIN Tenders ON Con-
tractors.Contractor_ID = Tenders.Contractor_ID) ON (Tenders.Tender_ID =
Fixed_Prices.Tender_ID) AND (Projects.Proj_ID = Tenders.Proj_ID)
GROUP BY [Geselekteerde Items].[Item Num], Projects.[Base Month], Projects.Name
HAVING (((Avg(Fixed_Prices.[Orginal Tender]))<>0))
ORDER BY [Geselekteerde Items].[Item Num], Projects.[Base Month], Projects.Name;

```

**STAP 2 – BEREKENINGE:**

Vervolgens word die handmatige algoritme bespreek wat gevolg is om die berekeninge uit te voer.

- Die getenderde eenheidstarief en inflasie-aangepaste data is teenoor die begin datum van die ooreenstemmende projek vir elke item geplot. (Aangedui as 1 op sigblad.)
- Om te bepaal of elke item se eenheidstarief 'n stygende of dalende neiging toon is daar van liniêre regressie gebruik gemaak om 'n bespassede lyn deur die historiese data van elke item te pas met *MS Excel* se funksie "Add Trendline".
- Dieselfde tegnieke is gevolg met die inflasie-aangepaste eenheidstariewe soos dit vanuit die databasis verkry is.
- Aangesien die inflasie-aangepaste data nog steeds 'n stygende of dalende tendens getoon het, is elke item se data aangepas met 'n samegestelde rentefunksie wat die huidige waarde bereken. (Aangedui as 2 op sigblad.)
- 'n Regressielyn is ook deur hierdie data van elke item gepas en die helling is bepaal met *MS Excel* se "Linest"-funksie. Die inflasiekoers,  $i$ , is dan aangepas tot die regressielyn presies horisontaal was en daar dus nie meer 'n stygende of dalend tendens voorgekom het nie. Hierdie gedeelte van die berekening is geoutomatiseer met die "Goal Seek"-funksionaliteit wat in *MS Excel* beskikbaar is. (Aangedui as 3 op sigblad.)

- Die gemiddelde waarde en standaardafwyking van hierdie aangepaste data is vir elke item bepaal met die funksie “Average” en “Stdev”. (Aangedui as 4 op sigblad.)

Aangesien dit tydrowend is om dit vir elke item uit te voer, is die proses geautomatiseer deur van 'n *Visual Basic*-makro, *Vb M-1*, gebruik te maak wat geskryf is om bogenoemde algoritme uit te voer.

#### VB M-1: INFLASIEBEREKENINGE

```

Sub Macro1()
'
' Macro1 Macro
' Macro recorded 2001/05/15 by A Bester
'
' Keyboard Shortcut: Ctrl+w
'
    Range("I2:J10000").Clear
    Range("I2").Select
    j = 2
    c = 3

    While Not (ActiveCell.Offset(0, -8).Value = "")
    c = c + 1
    j = j + i
    i = 0
    cell = ActiveCell.Offset(i, -8).Value
    yearfrac = "=RC[-2]*(1+R" & Trim(Str(j)) & "C10)^YEARFRAC(RC[-6],""96/10/01"",1)"
    ActiveCell.Offset(i, 0).FormulaR1C1 = yearfrac
    ActiveCell.Offset(i, 0).Range("A1").Copy

    While (ActiveCell.Offset(i, -8).Value = cell)
    i = i + 1
    Wend

    test = "A1:A" & Trim(Str(i))
    Application.CutCopyMode = False
    Selection.AutoFill Destination:=ActiveCell.Range(test), Type:=xlFillDefault

    LinEst = "=LINEST(R[-1]C[-1]:R[" & Trim(Str(i - 2)) & "]C[-1],R[-1]C[-7]:R[" & Trim(Str(i - 2)) & "]C[-7])"
    ActiveCell.Offset(1, 1).FormulaR1C1 = LinEst

    ActiveCell.Offset(1, 1).GoalSeek Goal:=0, ChangingCell:=ActiveCell.Offset(0, 1).Range("A1")

    avera = "=AVERAGE(R[-2]C[-1]:R[" & Trim(Str(i - 3)) & "]C[-1])"
    ActiveCell.Offset(2, 1).FormulaR1C1 = avera

    std = "=STDEV(R[-3]C[-1]:R[" & Trim(Str(i - 4)) & "]C[-1])"
    ActiveCell.Offset(3, 1).FormulaR1C1 = std

```



**Vb M-2: INFLASIEBEREKENINGE (VERVOLG)**

```

ActiveCell.Offset(0, 1).Range("A1:A4").Copy
ActiveCell.Offset(c - j, 5).PasteSpecial xlPasteValues, xlPasteSpecialOperationNone, False,
True

ActiveCell.Offset(j - c, -13).Copy
ActiveCell.Offset(0, -1).PasteSpecial

ActiveCell.Offset(j - c + i, -4).Range("A1").Select
Wend

Range("I2").Select
End Sub

```

Die grafieke se opstelling en passing van die verskillende tendenslyne is ook met 'n *Visual Basic*-makro, *Vb M-3*, uitgevoer.

**Vb M-3: GRAFIEK-OPSTELLING**

```

Sub Macro2()
'
' Macro2 Macro
' Macro recorded 2001/05/21 by A Bester
'
'
'
Range("A2").Select

i = 0
j = 2

While Not (ActiveCell.Value = "")
j = j + i
i = 0
cell = ActiveCell.Value

While (ActiveCell.Offset(i, 0).Value = cell)
i = i + 1
Wend

'ActiveCell.Offset(14, -3).Range("A1").Activate
Charts.Add
ActiveChart.ChartType = xlXYScatter
sel = "C" & Trim(Str(j)) & ":C" & Trim(Str(j + i - 1)) & ",G" & Trim(Str(j)) & ":I" & Trim(Str(j + i - 1))
ActiveChart.SetSourceData Source:=Sheets("Data").Range(sel), PlotBy:=xlColumns
ActiveChart.Location Where:=xlLocationAsNewSheet

' Resize graph and move legend
ActiveChart.PlotArea.Select

```

**VB M-4: GRAFIEK-OPSTELLING (VERVOLG)**

```

Selection.Top = 1
Selection.Height = 436
ActiveChart.Legend.Select
Selection.Left = 371
Selection.Top = 154
ActiveChart.PlotArea.Select
Selection.Width = 718

With ActiveChart
    .HasTitle = False
    .Axes(xlCategory, xlPrimary).HasTitle = True
    .Axes(xlCategory, xlPrimary).AxisTitle.Characters.Text = "Begindatum"
    .Axes(xlValue, xlPrimary).HasTitle = True
    .Axes(xlValue, xlPrimary).AxisTitle.Characters.Text = "Eenheidkoste"
End With

With ActiveChart.Axes(xlCategory)
    .HasMajorGridlines = True
    .HasMinorGridlines = False
End With

With ActiveChart.Axes(xlValue)
    .HasMajorGridlines = True
    .HasMinorGridlines = False
End With

With ActiveChart.Axes(xlCategory)
    .MinimumScale = 32143
    .MaximumScale = 35431
    .MinorUnitIsAuto = True
    .MajorUnit = 365.25
    .Crosses = xlAutomatic
    .ReversePlotOrder = False
    .ScaleType = xlLinear
    .DisplayUnit = xlNone
End With

ActiveChart.SeriesCollection(1).Name = ""Oorspronklike Data""
ActiveChart.SeriesCollection(2).Name = ""Indeks Aangepaste Data""
ActiveChart.SeriesCollection(3).Name = ""Tendense Aangepaste Data""

ActiveChart.SeriesCollection(1).Trendlines.Add(Type:=xlLinear, Forward:=0, _
    Backward:=0, DisplayEquation:=False, DisplayRSquared:=False, Name:= _
    "Oorspronklik").Select
ActiveChart.SeriesCollection(2).Trendlines.Add(Type:=xlLinear, Forward:=0, _
    Backward:=0, DisplayEquation:=False, DisplayRSquared:=False, Name:= _
    "Indeks Aangepas").Select
ActiveChart.SeriesCollection(3).Trendlines.Add(Type:=xlLinear, Forward:=0, _
    Backward:=0, DisplayEquation:=False, DisplayRSquared:=False, Name:= _
    "Tendense Aangepas").Select

```

### Vb M-5: GRAFIEK-OPSTELLING (VERVOLG)

```
ActiveChart.SeriesCollection(1).Trendlines(1).Select
With Selection.Border
    .ColorIndex = 25
    .Weight = xlMedium
    .LineStyle = xlContinuous
End With
ActiveChart.SeriesCollection(2).Trendlines(1).Select
With Selection.Border
    .ColorIndex = 26
    .Weight = xlMedium
    .LineStyle = xlContinuous
End With
ActiveChart.SeriesCollection(3).Trendlines(1).Select
With Selection.Border
    .ColorIndex = 27
    .Weight = xlMedium
    .LineStyle = xlContinuous
End With

ActiveChart.HasLegend = True

chartName = ActiveChart.Name

Sheets(chartName).Name = cell
Sheets("Data").Select

ActiveCell.Offset(i, 0).Select
Wend
End Sub
```

### **STAP 3 – RESULTATE:**

Die resultate van die berekening kan gevind word in hierdie sigblad, voorbeelde van die grafieke word getoon in *Figuur 5-11* en *Figuur 5-12*.

---

---

## AANHANGSEL N

### VERWERKING VAN RESULTATE

Die gemiddelde en standaardafwyking van elke item gebaseer op die: oorspronklike data, inflasie-aangepaste data en tendensaangepaste data is bereken. Die data is afkomstig van die CSRA-databasis wat in die gids “Databasis” gestoor is.

#### **STAP 1 – DATA EN BEREKENINGE:**

- Die data wat in *Aanhangsel M* gebruik is om die inflasiebepaling te doen is oorgedra na 'n tabel in die CSRA-databasis met die naam **A-ANGEPASTE DATA**;
- 'n *SQL*-navraag, *Sql N-1*, wat die gemiddelde en standaardafwyking van elke item bepaal het vir die: oorspronklike, inflasie-aangepaste en tendensaangepaste data is uitgevoer op hierdie tabel;
- Die data is oorgedra na 'n *MS Excel*-sigblad, *Berekeninge.xls*, wat in die gids “Aanhangsel N” op die KS is.

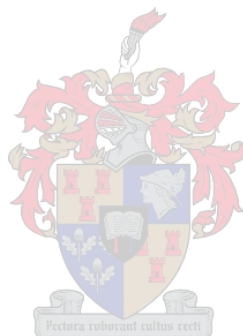
#### **SQL N-1**

```
SELECT [Aangepaste Data].Item_Num, Avg([Aangepaste Data].[Orginal Tender]) AS [AvgOfOrginal Tender], StDev([Aangepaste Data].[Orginal Tender]) AS [StDevOfOrginal Tender], Avg([Aangepaste Data].[Escalated Tender]) AS [AvgOfEscalated Tender], StDev([Aangepaste Data].[Escalated Tender]) AS [StDevOfEscalated Tender], Avg([Aangepaste Data].NewEscalatedTender) AS AvgOfNewEscalatedTender, StDev([Aangepaste Data].NewEscalatedTender) AS StDevOfNewEscalatedTender
FROM [Aangepaste Data]
GROUP BY [Aangepaste Data].Item_Num;
```

**STAP 3 – RESULTATE:**

Die gemiddelde en standaardafwyking van elke item is bepaal met die funksie wat beskikbaar is in *MS Excel*. Die tendensaangepaste inflasie vir elke item word in *Figuur 5-13* getoon en die resultate van die statistiese ontledings word in *Tabel 5-8* getoon.

---



## AANHANGSEL O

### KORRELASIE BEREKENINGE

Die korrelasies tussen die items is bepaal vir korrelasie tussen die eenheidtariewe van die items, tussen die hoeveelhede van die items en ook tussen die eenheidtariewe en die hoeveelhede. Die data is afkomstig van die CSRA-databasis wat in die gids "Databasis" gestoor is.

#### STAP 1 – DATA:

- Die data wat vir die korrelasie berekening tussen die eenheidtariewe gebruik is, is bepaal deur 'n *SQL*-navraag, *Sql O-1*, uit te voer. Dit het die eenheidtariewe van die beduidende items bepaal, teenoor die projekte waarin dit voorgekom het.

#### SQL O-1

```
TRANSFORM Avg([Aangepaste Data].[Orginal Tender]) AS [AvgOfOrginal Tender]
SELECT [Aangepaste Data].Name
FROM [Aangepaste Data]
WHERE ((([Aangepaste Data].Item_Num)>="16.01" And ([Aangepaste Data].Item_Num)<>"33.01(a)(i)L" And ([Aangepaste Data].Item_Num)<>"61.15"))
GROUP BY [Aangepaste Data].Name
PIVOT [Aangepaste Data].Item_Num;
```

- Die data wat verkry is, is na 'n *MS Excel*-sigblad, *Kovariansie.xls*, wat in die gids "Aanhangsel O" op die KS gevind kan word, verskuif.
- Die data wat vir die korrelasie berekening tussen die hoeveelhede gebruik is, is ook bepaal met 'n *SQL*-navraag, *Sql O-2*. Dit het die

hoeveelhede van die beduidende items bepaal, teenoor die projekte waarin dit voorgekom het.

#### SQL O-2

```

TRANSFORM Avg(BillOfQuantities.Quantity) AS AvgOfQuantity
SELECT Projects.Name
FROM (Projects INNER JOIN ((Items INNER JOIN (BillOfQuantities INNER JOIN Fixed_Prices
ON BillOfQuantities.Bill_ID = Fixed_Prices.Bill_ID) ON Items.Item_ID = BillOfQuantities.Item_ID)
INNER JOIN [Geselekteerde Items] ON Items.Item_Num = [Geselekteerde Items].[Item Num])
ON Projects.Proj_ID = BillOfQuantities.Proj_ID) INNER JOIN Tenders ON (Tenders.Tender_ID =
Fixed_Prices.Tender_ID) AND (Projects.Proj_ID = Tenders.Proj_ID)
WHERE ((([Geselekteerde Items].[Item Num])>="16.01" And ([Geselekteerde Items].[Item
Num])<>"33.01(a)(i)L" And ([Geselekteerde Items].[Item Num])<>"61.15") AND
((Fixed_Prices.[Orginal Tender])<>0))
GROUP BY Projects.Name
ORDER BY Projects.Name, [Geselekteerde Items].[Item Num]
PIVOT [Geselekteerde Items].[Item Num];

```

- Die data wat verkry is, is ook na die *MS Excel*-sigblad *Kovariansie.xls* wat in die gids “\Aanhangsel O” op die KS gevind kan word, verskuif.

#### **STAP 2 – BEREKENINGE:**

- Die korrelasies tussen twee items is bereken met 'n *MS Excel*-funksie “Pearson” wat die Pearson-korrelasie bepaal tussen twee datastelle.
- In sommige gevalle was daar baie min datapunte gewees om korrelasie op te kon bereken. 'n Filter is ontwikkel wat gebruik kon word om te verseker dat daar ten minste 'n minimum aantal pare datapunte was waarop die berekening kon uitgevoer word. Vir die berekening is 'n minimum van 10-pare punte voorgeskryf.
- 'n Verdere filter om vir die gebruiker makliker sterk korrelasies te kan bepaal is ingebou wat slegs korrelasie binne sekere grense vertoon.

Weens die herhalende geaardheid om 'n funksie en die filters wat saam met dit gaan vir elke kombinasie van pare items per sel te definieer is 'n *Visual Basic*-makro geskryf wat die proses geoutomatiseer het. Die kode hiervan word getoon in *Vb O-1*.

#### **STAP 3 – RESULTATE:**

Die resultate van die berekening kan gesien word op die sigblad. Uittreksels van die resultate word getoon in *Figuur 5-18*, *Figuur 5-19* en *Figuur 5-20*.

## VB O-1: GRAFIEK-OPSTELLING

```

Sub Macro1 ()
'
' Macro1 Macro
' Macro recorded 2001/05/24 by A Bester
'
'
    k = 258           'number of spaces to move down to incidence matrix
    l = 245           'number of data sets-2

    Range("B2:AW49").Clear

    Range("B2").Select

    i = 0

    While Not (ActiveCell.Offset(i, -1) = "")
        j = 0
        While Not (ActiveCell.Offset(-1, j) = "")
            corFor = "=IF(SUMPRODUCT(Units!R[" & Trim(Str(-i + k)) & "]C[" & Trim(Str(-j + i)) & "]:R[" &
Trim(Str(l - i + k)) & "]C[" & Trim(Str(-j + i)) & "],Units!R[" & Trim(Str(-i + k)) & "]C[0]:R[" & Trim(Str(l
- i + k)) & "]C[0]>=R[" & Trim(Str(-i + 52)) & "]C[" & Trim(Str(-j)) & "],PEARSON(Units!R[" &
Trim(Str(-i)) & "]C[" & Trim(Str(-j + i)) & "]:R[" & Trim(Str(l - i)) & "]C[" & Trim(Str(-j + i)) &
"],Units!R[" & Trim(Str(i)) & "]C[0]:R[" & Trim(Str(l - i)) & "]C[0],0)"
            ActiveCell.Offset(i, j).FormulaR1C1 = corFor
            j = j + 1
        Wend
        i = i + 1
    Wend

    Range("B2:AW49").Select
    Selection.FormatConditions.Add Type:=xlCellValue, Operator:=xlBetween, For-
mula1:="=$B$56", Formula2:="=$B$57"
    Selection.FormatConditions(1).Font.ColorIndex = 2
    Selection.FormatConditions.Add Type:=xlCellValue, Operator:=xlNotBetween, For-
mula1:="=$C$56", Formula2:="=$C$57"
    Selection.FormatConditions(2).Font.ColorIndex = 2
    Selection.NumberFormat = "0.0"
    With Selection
        .HorizontalAlignment = xlCenter
        .VerticalAlignment = xlBottom
        .WrapText = False
        .Orientation = 0
        .AddIndent = False
        .ShrinkToFit = False
        .MergeCells = False
    End With

```

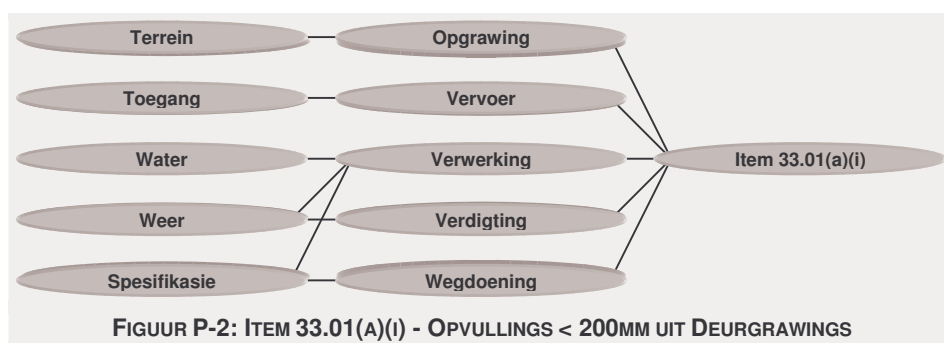
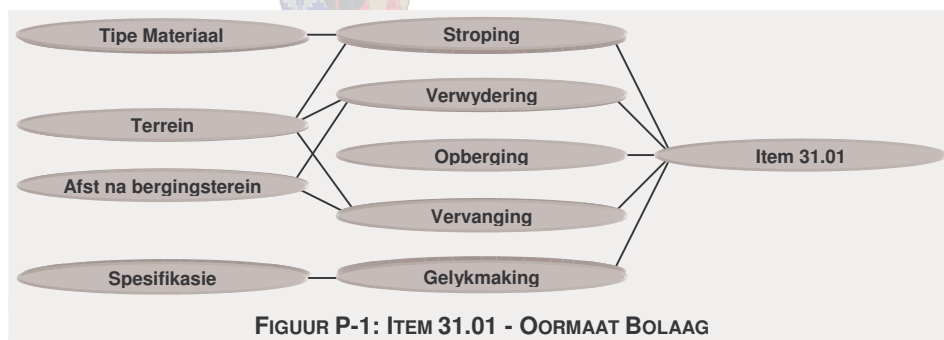


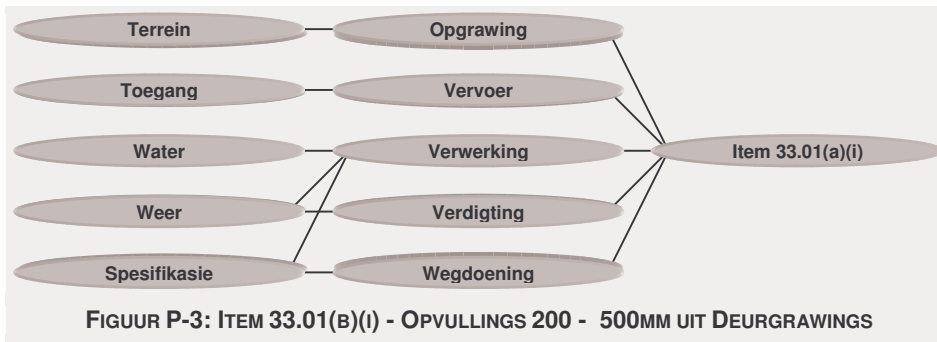


## AANHANGSEL P

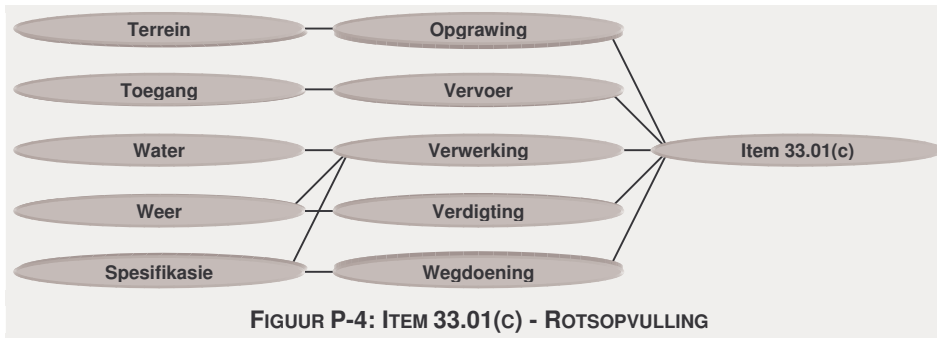
### INVLOEDDIAGRAMME

Die invloedmodelle wat ontwikkel is vir die 17-items wat voorkom in die *Serie 3000* word hier getoon. Die modelle toon vir elke item die aktiwiteite soos dit in die standaardspesifikasie gevind word. Verder word die invloedfaktore wat hierdie aktiwiteite beïnvloed ook getoon, asook die verwantskappe daartussen.

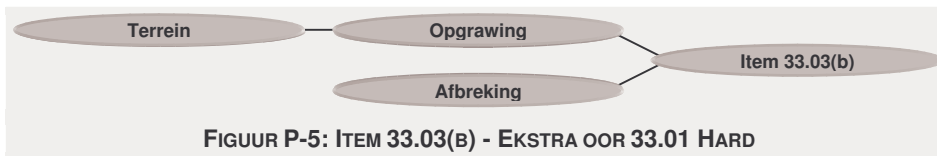




FIGUUR P-3: ITEM 33.01(B)(I) - OPVULLINGS 200 - 500MM UIT DEURGRAWINGS



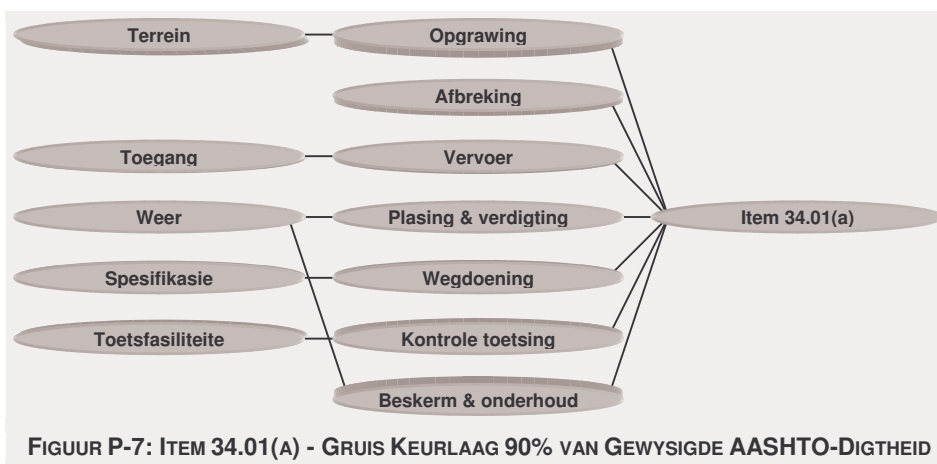
FIGUUR P-4: ITEM 33.01(C) - ROTSOPVULLING



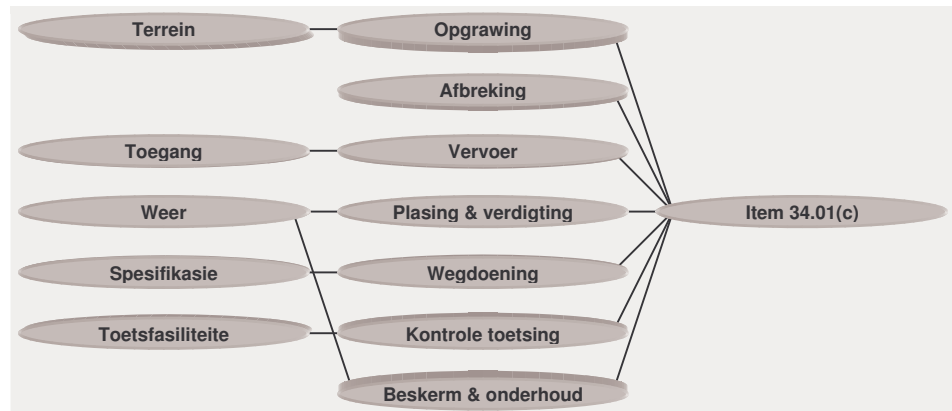
FIGUUR P-5: ITEM 33.03(B) - EKSTRA OOR 33.01 HARD



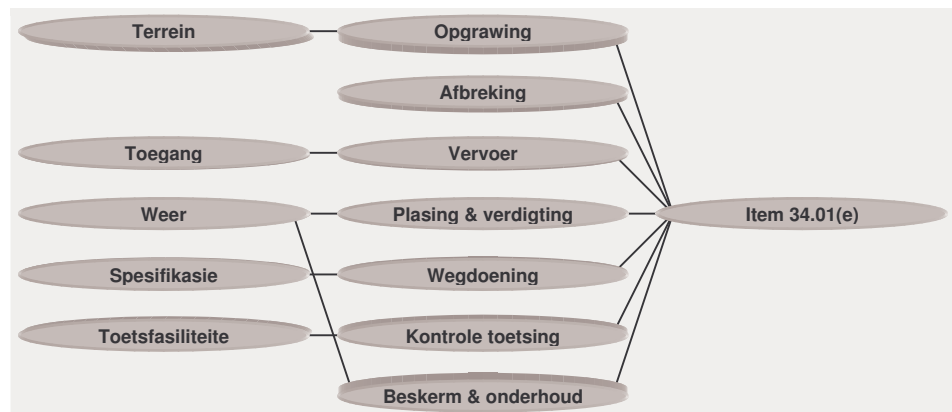
FIGUUR P-6: ITEM 33.04(A) - WEGGOOI VAN DEURGRAWINGSMATERIAAL



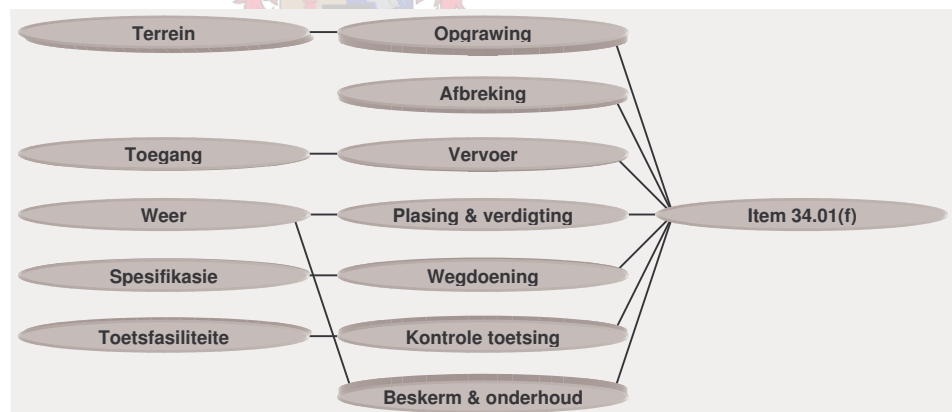
FIGUUR P-7: ITEM 34.01(A) - GRUIS KEURLAAG 90% VAN GEWYSIGDE AASHTO-DIGTHEID



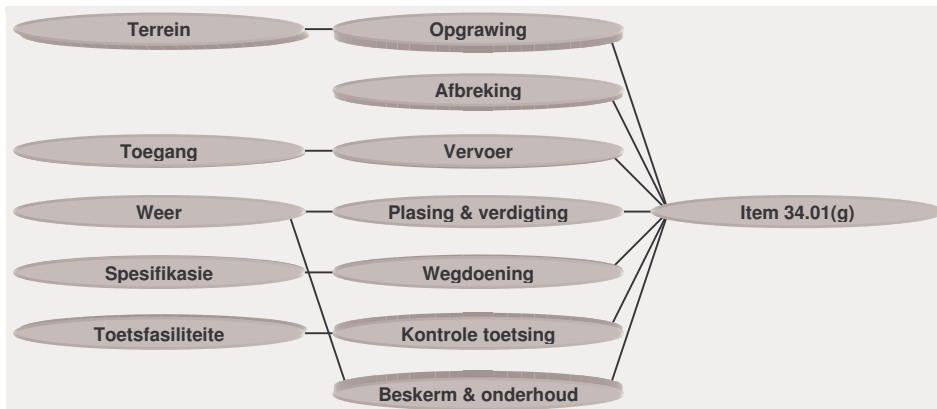
FIGUUR P-8: ITEM 34.01(C) GRUIS STUTLAAG 95% VAN GEWYSIGDE AASHTO-DIGTHEID



FIGUUR P-9: ITEM 34.01(E) GRUIS KROONLAAG 97% VAN GEWYSIGDE AASHTO-DIGTHEID



FIGUUR P-10: ITEM 34.01(F) - GRUIS SKOUERS 93% VAN GEWYSIGDE AASHTO-DIGTHEID



FIGUUR P-11: ITEM 34.01(G) GRUIS SLYTLAAG 93% VAN GEWYSIGDE AASHTO-DIGTHEID



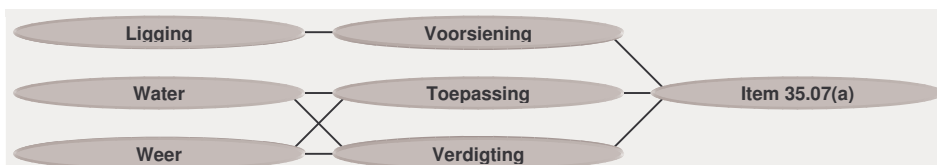
FIGUUR P-12: ITEM 35.02(A) - STABILISERINGSMIDDEL PORTLANDSEMENT



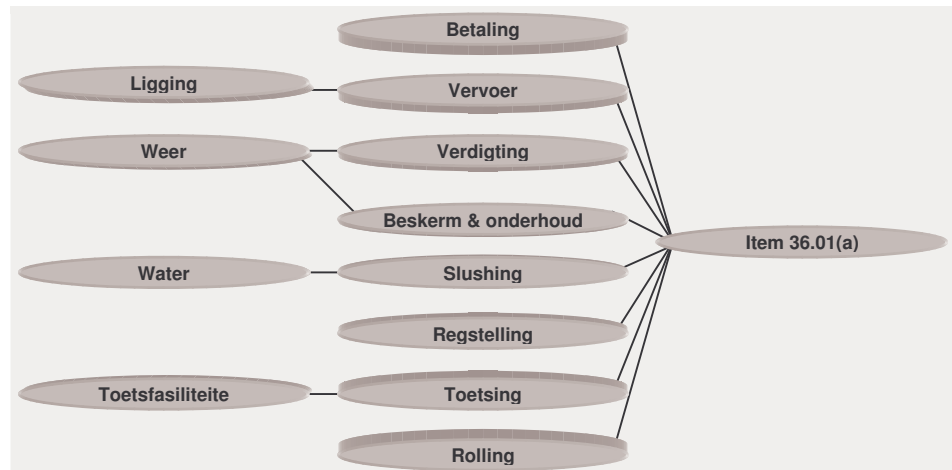
FIGUUR P-13: ITEM 35.02(B) - STABILISERINGSMIDDEL HOOGOONDSLAKSEMENT



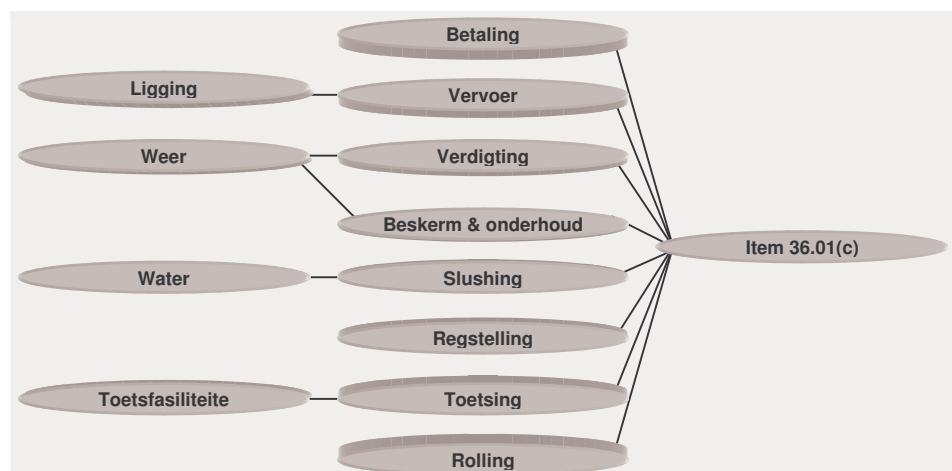
FIGUUR P-14: ITEM 35.02(C) - STABILISERINGSMIDDEL GEBLUSTE KALK



FIGUUR P-15: ITEM 35.07(A) - ANIONIESE BITUMINEUSE STABILISERINGSMIDDEL

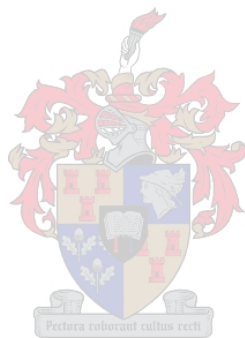


FIGUUR P-16: ITEM 36.01(A) - G1 GEBREEKTE KLIP KROONLAAG KOMMERSIEEL



FIGUUR P-17: ITEM 36.01(C) - G2 GEBREEKTE KLIP KROONLAAG KOMMERSIEEL





---

## AANHANGSEL Q

---

### BEDUIDENDSTE INVLOEDFAKTORE

---

Die beduidendste invloedfaktore kan bepaal word met die metode soos in *Afdeling 6.8.2* bespreek word. Hierdie metode is in die sigblad *Invloedrykste Invloedfaktore.xls* geïmplementeer en kan in *Aanhangsel Q* op die KS gevind word. Die stappe wat in die sigblad getoon word stem presies ooreen met die in die bespreking. Die eerste sewe stappe en *Stap 9* is alreeds volledig bespreek, maar in hierdie aanhangsel word *Stap 8* – Kostebydrae en Invloedfaktorsterkte aanpassing meer omvattend bespreek.

Die geskatte kostebydrae (indien daar nie inligting beskikbaar was om dit op te baseer nie) (*Stap 3*) en die invloedfaktore sterkte van die spesialiste (*Stap 4*) is nie noodwendig korrek nie. Dit veroorsaak dat die berekende korrelasies wat op hierdie geskatte waardes gebaseer is, sal verskil indien dit vergelyk word met die werklike korrelasies wat tussen die eenheidtariewe bestaan. In die sigblad (*Stap 8*) word 'n metode voorgestel hoe die geskatte waardes aangepas kan word, sodat die berekende korrelasies ooreenstem met die werklike korrelasie of dan ten minste die verskil minimeer.

Die metode is gebaseer op die veronderstelling dat die geskatte waardes van die spesialiste (*Stap 3 & 4*) korrek is en dat enige afwyking daarvan 'n fout is. Indien die geskatte waardes aangepas moet word, word dit gesien as 'n verlaging in die korrektheid van die waardes wat gebruik word om die korrelasies mee te bereken. Die korrektheid word gevind deur die verskil tussen die aangepaste en geskatte waardes te bereken. Hoe groter die verskil is hoe laer is die korrektheid. In die sigblad is dit geïmplementeer deur 'n verskil van 0, 1 en 2 as onderskeidelik 100%, 66% en 33% akkuraat te definieer.

---



Twee *Visual Basic*-makros met verskillende doelfunksie is geskryf om die geskatte waardes aan te pas sodat die berekende korrelasie en die werklike korrelasie geminimeer kon word. Dit is sodanig gedoen dat die aangepaste waardes so min as moontlik van die geskatte waardes verskil, m.a.w. dat die akkuraatheid so hoog as moontlik bly. Hiervoor is van *MS Excel* se “Solver”-funktionaliteit gebruik gemaak.

Die een optimeringmodel is so opgestel dat die verskil tussen die berekende en die werklike korrelasies geminimeer moet word, maar dat daar slegs 'n gemiddelde van 'n 10% fout gemaak moet word met die aanpassing van die geskatte waardes, *Vb Q-1*.

#### Vb Q-1: MINIMEER KORRELASIE VERSKIL

```
Sub Macro4()
'
' Macro4 Macro
' Macro recorded 2004/02/20 by André Bester
'
' Keyboard Shortcut: Ctrl+d
'

    SolverReset
    SolverAdd CellRef:="$BE$24", Relation:=3, FormulaText:="$BE$28"
    SolverAdd CellRef:="$AB$5:$AP$14", Relation:=1, FormulaText:="3"
    SolverAdd CellRef:="$AB$5:$AP$14", Relation:=3, FormulaText:="-3"
    SolverOptions MaxTime:=100, Iterations:=1000, Precision:=0.001, AssumeLinear:= _
        False, StepThru:=False, Estimates:=1, Derivatives:=1, SearchOption:=1, _
        IntTolerance:=1, Scaling:=False, Convergence:=0.001, AssumeNonNeg:=False
    SolverOk SetCell:="$BE$23", MaxMinVal:=2, ValueOf:="0", ByChange:="$AB$5:$AP$14"
    SolverSolve
End Sub
```

Die ander model is so opgestel dat die akkuraatheid gemaksimeer word van die aangepaste waardes, maar dat die verskil tussen die twee berekende en werklike korrelasie nie meer as gemiddeld 10% van mekaar verskil nie, *Vb Q-2*.

**Vb Q–2: MAKSIMEER AKKURAATHEID**

```

Sub Macro2()
'
' Macro2 Macro
' Macro recorded 2004/02/20 by André Bester
'
' Keyboard Shortcut: Ctrl+s
'

SolverReset
SolverAdd CellRef:="$D$23:$D$33", Relation:=1, FormulaText:="$BE$27"
SolverAdd CellRef:="$E$24:$E$33", Relation:=1, FormulaText:="$BE$27"
SolverAdd CellRef:="$f$25:$f$33", Relation:=1, FormulaText:="$BE$27"
SolverAdd CellRef:="$g$26:$g$33", Relation:=1, FormulaText:="$BE$27"
SolverAdd CellRef:="$h$27:$h$33", Relation:=1, FormulaText:="$BE$27"
SolverAdd CellRef:="$i$28:$i$33", Relation:=1, FormulaText:="$BE$27"
SolverAdd CellRef:="$j$29:$j$33", Relation:=1, FormulaText:="$BE$27"
SolverAdd CellRef:="$k$30:$k$33", Relation:=1, FormulaText:="$BE$27"
SolverAdd CellRef:="$l$31:$l$33", Relation:=1, FormulaText:="$BE$27"
SolverAdd CellRef:="$m$32:$m$33", Relation:=1, FormulaText:="$BE$27"
SolverAdd CellRef:="$n$33", Relation:=1, FormulaText:="$BE$27"
SolverAdd CellRef:="$AB$5:$AP$14", Relation:=1, FormulaText:="3"
SolverAdd CellRef:="$AB$5:$AP$14", Relation:=3, FormulaText:="-3"
SolverOptions MaxTime:=100, Iterations:=1000, Precision:=0.001, AssumeLinear:= _
    False, StepThru:=False, Estimates:=1, Derivatives:=1, SearchOption:=1, _
    IntTolerance:=1, Scaling:=False, Convergence:=0.001, AssumeNonNeg:=False
SolverOk SetCell:="$BE$24", MaxMinVal:=1, ValueOf:="0", ByChange:="$AB$5:$AP$14"
SolverSolve
End Sub

```





AANHANGSEL R

---

**KOMPAKTE SKYF**

---

