



Faculteit Bio-ingenieurswetenschappen

Academiejaar 2010-2011

# Parametrisatie van stochastische neerslagmodellen met heuristische zoektechnieken

**Willem Jan Vanhaute**

Promotoren: Prof. dr. ir. N. Verhoest & Prof. dr. B. De Baets

Tutors: ir. S. Vandenberghe & ir. K. Scheerlinck

Masterproef voorgedragen tot het behalen van de graad van  
Master in de bio-ingenieurswetenschappen: Land- en Waterbeheer

De auteur en promotoren geven de toelating deze scriptie voor consultatie beschikbaar te stellen en delen ervan te kopiëren voor persoonlijk gebruik. Elk ander gebruik valt onder de beperkingen van het auteursrecht, in het bijzonder met betrekking tot de verplichting uitdrukkelijk de bron te vermelden bij het aanhalen van resultaten uit deze scriptie.

The author and promoters give the permission to use this thesis for consultation and to copy parts of it for personal use. Every other use is subject to the copyright laws, more specifically the source must be extensively specified when using from this thesis.

Gent, 10 Juni 2011

De promotoren

De begeleiders

De auteur

Prof. dr. ir. N. Verhoest

ir. S. Vandenberghe

Willem Jan Vanhaute

Prof. dr. B. De Baets

ir. K. Scheerlinck

# Woord vooraf

*Zoals gebruikelijk, wordt in het voorwoord van de gelegenheid gebruik gemaakt om enkele personen, die een onontbeerlijke bijdrage leverden aan deze masterproef, te bedanken.*

*Vooreerst wil ik mijn promotoren, Prof. Dr. ir. Verhoest en Prof. Dr. De Baets, bedanken, zij maakten het mogelijk dat ik me kon verdiepen in dit onderwerp. Niet alleen door me doorheen het proces te voorzien van de nodige begeleiding, maar vooral door mijn interesse in waterbeheer en computationele technieken doorheen de jaren aan te wakkeren.*

*Veel dank gaat uit naar mijn tutors, ir. Scheerlinck en ir. Vandenberghe. Zonder jullie begeleiding, de hulp met het zoeken van literatuur, het vele nalezen, de Matlab/L<sup>A</sup>T<sub>E</sub>X tips & tricks, en nog veel meer, zou het resultaat wellicht héél wat minder fraai geweest zijn.*

*Verder wil ik Rudi bedanken, voor de voorziening en de installatie van mijn voornaamste werk-instrument, de computer. De mensen van het Labo voor Bosbeheer en Ruimtelijke Informatietechnieken wil ik graag bedanken voor het dulden van mijn aanslag op hun koffievoorraad en het gebruik van het zogenaamde thesislokaal.*

*Heel veel dank ook aan mijn ouders, voor hun nooit aflatende steun en begrip. Niet alleen tijdens het maken van deze masterproef, maar ook in de daaraan voorafgaande jaren.*

*Tot slot wil ik mijn steun en toeverlaat Jolien bedanken, om op eender welk moment voor me klaar te staan en mij onvoorwaardelijk te steunen.*

*Willem Jan Vanhaute  
Gent 10 Juni 2011*



# Inhoudsopgave

<b>Lijst van afkortingen</b>	<b>1</b>
<b>Samenvatting</b>	<b>3</b>
<b>1 Inleiding</b>	<b>5</b>
1.1 Algemene omkadering . . . . .	5
1.2 Doelstellingen . . . . .	7
1.3 Indeling . . . . .	8
<b>2 Neerslagmodellen</b>	<b>11</b>
2.1 Inleiding . . . . .	11
2.2 Modelleren van puntneerslag . . . . .	12
2.2.1 Manifestatie van neerslagevents . . . . .	12
2.2.2 Duur en intensiteit . . . . .	14
2.2.3 Clustergebaseerde modellen . . . . .	14
<b>3 Beschrijving van de neerslag te Ukkel</b>	<b>19</b>
3.1 Inleiding . . . . .	19
3.2 Statistische analyse . . . . .	19
3.2.1 Eerste, tweede en derde orde momenten . . . . .	19
3.2.2 Voorwaardelijke distributiekarakteristieken . . . . .	21
3.2.3 Voorwaardelijke probabiliteiten inzake neerslagvolumes . . . . .	22
3.2.4 Extreme waarden analyse . . . . .	23
<b>4 Optimalisatie</b>	<b>27</b>
4.1 Inleiding . . . . .	27
4.2 Doelfunctie . . . . .	28
4.3 Selectie van gebruikte statistieken in de doelfunctie . . . . .	29
4.4 Parameterruimte . . . . .	30
4.5 Simplex-Simulated Annealing . . . . .	31
4.5.1 Simulated Annealing . . . . .	31
4.5.2 Simplex methode . . . . .	32
4.5.3 SIMPSA . . . . .	33
4.6 Particle Swarm Optimization . . . . .	37
4.6.1 Inleiding . . . . .	37
4.6.2 Theoretische aspecten . . . . .	38
4.6.3 Praktische aspecten . . . . .	39

<b>5</b>	<b>Vergelijking PSO en SIMPSA</b>	<b>47</b>
5.1	Inleiding . . . . .	47
5.2	Performantiecriteriën . . . . .	47
5.3	Resultaten . . . . .	48
5.3.1	Benadering 1 . . . . .	49
5.3.2	Benadering 2 . . . . .	53
5.4	Besluit . . . . .	56
<b>6</b>	<b>Evaluatie gekalibreerde modellen</b>	<b>59</b>
6.1	Inleiding . . . . .	59
6.2	Parameters . . . . .	59
6.3	Statistische analyse . . . . .	60
6.3.1	Eerste, tweede en derde orde momenten . . . . .	61
6.3.2	Voorwaardelijke distributiekarakteristieken . . . . .	63
6.3.3	Voorwaardelijke probabiliteiten inzake neerslagintensiteit . . . . .	64
6.3.4	Extreme waarden analyse . . . . .	66
6.4	Besluit . . . . .	66
<b>7</b>	<b>Introductie derde orde moment</b>	<b>71</b>
7.1	Inleiding . . . . .	71
7.2	Analytische uitdrukking . . . . .	72
7.3	Doelfunctie . . . . .	73
7.4	Resultaten . . . . .	74
7.4.1	Parameters . . . . .	74
7.4.2	Statistische analyse . . . . .	74
7.5	Besluit . . . . .	82
<b>8</b>	<b>Invloed doelfunctie</b>	<b>83</b>
8.1	Inleiding . . . . .	83
8.2	Beschrijving en visualisatie van het probleem . . . . .	83
8.3	Alternatieve configuraties van de doelfunctie . . . . .	85
8.3.1	Impliciete weging . . . . .	85
8.3.2	Relatieve bijdrage aan de fitness . . . . .	85
8.4	Invloed op analytische momenten . . . . .	86
8.4.1	Doelfunctie zonder derde orde moment . . . . .	86
8.4.2	Doelfunctie met derde orde moment . . . . .	88
8.5	Invloed op de gesimuleerde momenten . . . . .	89
8.5.1	Statistische analyse . . . . .	89
8.6	Besluit . . . . .	96
<b>9</b>	<b>Het Truncated Bartlett-Lewis model</b>	<b>97</b>
9.1	Inleiding . . . . .	97
9.2	Optimalisatie . . . . .	97
9.2.1	Selectie van de gebruikte karakteristieken in de doelfunctie . . . . .	98
9.2.2	Doelfunctie . . . . .	100
9.2.3	Parameterruimte . . . . .	100
9.2.4	Implementatie SIMPSA . . . . .	100
9.2.5	Geschatte parameters . . . . .	101
9.3	Statistische analyse . . . . .	102
9.3.1	Eerste, tweede en derde orde momenten . . . . .	104

9.3.2	Voorwaardelijke distributiekarakteristieken . . . . .	105
9.3.3	Voorwaardelijke probabiliteiten inzake neerslagintensiteit . . . . .	106
9.3.4	Extreme waarden analyse . . . . .	106
9.4	Invloed truncatieparameter $\epsilon$ . . . . .	109
9.5	Besluit . . . . .	111
<b>10</b>	<b>Besluit</b>	<b>115</b>
10.1	Algemene bevindingen . . . . .	115
10.2	Toekomstperspectieven . . . . .	116
	<b>Bibliografie</b>	<b>119</b>
<b>A</b>	<b>Analytische modelbeschrijvingen</b>	<b>125</b>
A.1	Het originele Bartlett-Lewis model (OBL) . . . . .	125
A.1.1	OBL eerste en tweede orde eigenschappen . . . . .	125
A.1.2	OBL Zero Depth Probability . . . . .	126
A.1.3	OBL derde orde moment . . . . .	126
A.1.4	Het gewijzigde Bartlett-Lewis model (MBL) . . . . .	127
A.1.5	MBL eerste en tweede orde eigenschappen . . . . .	127
A.1.6	MBL Zero Depth Probability . . . . .	128
A.1.7	MBL derde orde moment . . . . .	129
A.2	Het gewijzigde Bartlett-Lewis Gamma model (MBLG) . . . . .	129
A.2.1	MBLG eerste en tweede orde eigenschappen . . . . .	130
A.2.2	MBLG Zero Depth Probability . . . . .	130
A.3	Het Truncated Bartlett-Lewis model (TBL) . . . . .	131
A.3.1	TBL eerste en tweede orde eigenschappen . . . . .	131
A.3.2	TBL Zero Depth Probability . . . . .	132
A.3.3	TBL derde orde moment . . . . .	132
<b>B</b>	<b>Tabellen</b>	<b>133</b>
<b>C</b>	<b>Figuren</b>	<b>145</b>





# Lijst van afkortingen

## F

FES                      Functie-Evaluaties

## G

GRV                      Gumbel Reduced Variate

## M

MBL                      Modified Bartlett-Lewis  
MBLG                     Modified Bartlett-Lewis Gamma

## O

OBL                      Original Bartlett-Lewis

## P

POT                      Peak-Over-Threshold  
PSO                      Particle Swarm Optimization

## R

RMSE                      Root Mean Square Error

**S**

SIMPISA                  Simplex-Simulated Annealing

**T**

TBL                        Truncated Bartlett-Lewis

**Z**

ZDP                        Zero Depth Probability

# Samenvatting

In het kader van de Europese richtlijn over beoordeling en beheer van overstromingsrisico's, dienen door de lidstaten overstromingsrisicobeheerplannen ingediend te worden. De modellering van neerslag is een kritieke input in de daartoe gebruikte neerslag-afvoermodellen. Gezien in het verleden gebleken is dat de parametrisatie van deze neerslagmodellen een heikel punt is, wordt in het huidige werk onderzocht op welke manier de parametrisatie van verschillende puntneerslagmodellen (van het type Bartlett-Lewis), met behulp van heuristische zoektechnieken, efficiënter en gericht kan verlopen. Hiertoe wordt de vergelijking gemaakt tussen Particle Swarm Optimization en Simplex-Simulated Annealing. Hieruit blijkt dat, rekening houdend met nauwkeurigheid, efficiëntie, betrouwbaarheid en gebruiksvriendelijkheid, het gebruik van Simplex-Simulated Annealing leidt tot de beste resultaten. Naast het aanduiden van een geschikte zoektechniek, worden verschillende mogelijkheden om het parametrisatieproces van de Bartlett-Lewis modellen te verbeteren onderzocht. Uit de resultaten kan geconcludeerd worden dat het gebruik van het derde orde moment in de doelfunctie leidt tot een verbeterde representatie van de extremen bij simulatie, maar er tegelijk voor zorgt dat enkele lagere orde momenten minder goed benaderd worden. Om aan dit probleem tegemoet te komen, wordt de invloed van het gebruik van twee alternatieve doelfuncties onderzocht. De bekomen resultaten tonen aan dat een impliciete weging in de doelfunctie leidt tot een algemeen betere prestatie van de modellen, een weging volgens relatieve bijdrage aan de fitness leidt daarentegen niet tot betere resultaten. Tot slot worden de verworven inzichten gebruikt om een nieuw Bartlett-Lewis model, namelijk het Truncated Bartlett-Lewis model, te kalibreren, met goede resultaten tot gevolg.



# Hoofdstuk 1

## Inleiding

### 1.1 Algemene omkadering

Het modelleren van neerslag kent verscheidene toepassingen in diverse takken van de hydrologische wetenschappen. Typische hydrologische toepassingen van neerslagdata zijn het ontwerp van rioolssystemen, het dimensioneren van reservoirs en het inschatten van overstromingsgevaar (Onof *et al.*, 2000). In verband met deze laatste toepassing trad op 26 november 2007 de Europese richtlijn over beoordeling en beheer van overstromingsrisico's in werking. Deze richtlijn stelt voorop dat iedere lidstaat overstromingsgevoelige gebieden identificeert, overstromingsgevaar- en overstromingsrisicokaarten opstelt en daarvoor de respectievelijke overstromingsrisicobeheerplannen indient. Hierbij verwijst de term overstromingsrisico naar '*de combinatie van de probabiliteit op voorkomen van een overstroming en de potentiële averechtse effecten op volksgezondheid, het milieu, cultureel erfgoed en economische activiteit geassocieerd met een overstroming*' (Directive 2007/60/EC, 2007). De correcte inschatting van de terugkeerperiode van piekdebieten is hierbij een belangrijk aspect waaraan voldoende aandacht dient geschonken te worden (Verhoest *et al.*, 2010).

Ter voorspelling van deze piekdebieten wordt doorgaans gebruik gemaakt van zogenaamde neerslag-afvoermodellen, waarvoor neerslag een kritische input is. In vele gevallen worden deze piekdebieten bepaald in functie van één specifieke storm of bui, zijnde een significante historische neerslaggebeurtenis of een gesimuleerde bui met bepaalde statistische eigenschappen (Wheater, 2002). Het nadeel van deze benadering is dat geen rekening wordt gehouden met de vochttoestand van de bodem op het moment dat de storm zich voordoet. Deze initiële vochttoestand bepaalt niettemin de hoeveelheid oppervlakkige afvoer die plaatsvindt en dus grotendeels de respons van het rivierenstelsel. Een initieel vochtig stroomgebied zal immers een grotere neiging tot overstroming vertonen dan een initieel droog stroomgebied (Verhoest *et al.*, 2010). Hieruit volgt dat de conventionele aanpak niet gebruikt kan worden bij het opstellen van overstromingsfrequentiekaarten die dienen als input voor overstromingsrisicokaarten. Wanneer continue tijdreeksen gebruikt worden, kunnen voornoemde problemen vermeden worden en kan de frequentie-analyse van de bekomen afvoerdata gebruikt worden bij de berekening van de benodigde kans op overstroming. Deze methode wordt echter beperkt door de lengte van de beschikbare geobserveerde neerslag. Tijdreeksen van lange duur zijn nodig om overstromingen met lage frequentie nauwkeurig te kunnen voorspellen. Indien de beschikbare gegevens ontoereikend zijn, is het gebruik van neerslagmodellen een valabel alternatief (Boughton & Droop, 2003).

Neerslagmodellen kunnen over het algemeen ingedeeld worden in vier categorieën (Onof *et al.*, 2000): (1) meteorologische modellen die opgebouwd zijn uit complexe sets differentiaalvergelijkingen die verschillende fysische processen die betrekkingen hebben tot de vorming van neerslag beschrijven; (2) stochastische modellen die de ruimtelijke evolutie van het regenvalproces onafhankelijk van de schaal beschrijven; (3) statistische modellen die bepaalde trends modelleren; en (4) stochastische procesmodellen die, ondanks de eenvoudige veronderstellingen die ze maken omtrent de fysische processen die aan de grondslag van het regenvalproces liggen, in staat zijn de hiërarchische structuur van het regenvalproces goed te benaderen met een beperkt aantal parameters. Het huidige werk beperkt zich tot de laatste categorie modellen, meer bepaald de Bartlett-Lewis modellen welke reeds gebruikt worden door de Vlaamse overheid (VMM) voor het gebruik als input in neerslag-afvoermodellen.

Opdat de gesimuleerde neerslagreeksen representatief zouden zijn voor een bepaald studiegebied, dienen de gebruikte modellen gekalibreerd te worden. Het gebruik van formele statistische technieken, zoals *maximum likelihood estimators*, vergt een extensieve computationele inspanning en zijn niet praktisch, waardoor ze zelden gebruikt worden (Obeysekera *et al.*, 1987). Doorgaans wordt gebruik gemaakt van de momentenmethode, waarbij een stel primaire eigenschappen (momenten) van de data gelijkgesteld worden aan de overeenkomstige modelwaarden (Rodriguez-Iturbe *et al.*, 1988). De kalibratie van de Bartlett-Lewis modellen herleidt zich dus tot een optimalisatieprobleem. De keuze van de gebruikte momenten en hun gewicht zorgen er voor dat de kalibratie van de Bartlett-Lewis modellen geen voor de hand liggende opdracht is. De kwaliteit van het resultaat van de optimalisatie varieert bovendien in functie van de beginwaarden waardoor het vinden van een behoorlijk resultaat niet gewaarborgd is (Khaliq & Cunnane, 1996). Daardoor was de kalibratie in het verleden veeleer het resultaat van *trial-and-error* in combinatie met eenvoudige gradiëntgebaseerde technieken (Onof & Wheeler, 1993; Velghe *et al.*, 1994; Verhoest *et al.*, 1997). De nood aan een efficiënte en gerichte kalibratiemethode en een verbetering van het proces op zich dringen zich op.

Verschillende empirische studies, waarbij de Bartlett-Lewis modellen gefit werden aan empirische data, hebben aangetoond dat de modellen de eerste en tweede orde statistieken van de neerslaghoeveelheid goed inschatten op verschillende aggregatieniveaus (Rodriguez-Iturbe *et al.*, 1987b; Entekhabi *et al.*, 1989; Cowpertwait, 1991; Onof & Wheeler, 1993, 1994a; Velghe *et al.*, 1994; Cowpertwait, 1994; Khaliq & Cunnane, 1996; Calenda & Napolitano, 1999; Cameron *et al.*, 2000, 2001; Heneker *et al.*, 2001; Smithers *et al.*, 2002; Favre *et al.*, 2004; Cowpertwait *et al.*, 2007). De reproductie van extreme neerslaggebeurtenissen is echter nog voor verbetering vatbaar. Cowpertwait (1998) wijst erop dat de incorporatie van hogere orde momenten in de doelfunctie hiervoor een oplossing kunnen bieden.

Uit het bovenstaande blijkt dat verschillende aspecten van het kalibratieproces dienen fijngesteld te worden. In het huidige werk wordt hiertoe gebruik gemaakt van verschillende types Bartlett-Lewis modellen welke gefit zullen worden aan 105 jaar lange observaties, gemeten te Ukkel. Er dient vermeld te worden dat deze dataset uniek is in zijn soort omwille van de lengte en de meetfrequentie (10 minuten) van de gemeten neerslagwaarden en de kwaliteit ervan (Ntegeka & Willems, 2008).

## 1.2 Doelstellingen

In het huidige werk wordt gezocht naar een efficiënte en intelligente manier om de parametrisatie van enkele types Bartlett-Lewis modellen te voltrekken. Daarnaast wordt ook onderzocht op welke wijze het kalibratieproces op zich kan worden verbeterd, door de keuze van het type doelfunctie en de gebruikte momenten. Concreet kunnen deze doelstellingen opgedeeld worden in verschillende specifieke onderzoeksvragen:

- *Welke methoden lenen zich het best tot de kalibratie van de Bartlett-Lewis modellen?* De kalibratie van de Bartlett-Lewis modellen omvat het minimaliseren van bepaalde doelfuncties. Omwille van de aanwezigheid van talrijke lokale minima wordt ervoor gekozen gebruik te maken van heuristische zoektechnieken. Deze hanteren een vooraf bepaalde strategie die niet afhangt van de vorm van de doelfunctie, dit in tegenstelling tot lokale technieken die zich veelal laten leiden door informatie over de doelfunctie zoals de gradiënt ervan. Heuristische zoektechnieken zouden daarom beter in staat zijn lokale optima te vermijden.
- *Welke heuristische zoektechniek geniet de voorkeur en waarom?* Uit het uitgebreide gamma heuristische zoektechnieken wordt gekozen voor een veelvuldig gebruikte populatiegebaseerde methode, meer bepaald Particle Swarm Optimization (Kennedy & Eberhart, 1995). Particle Swarm Optimization laat zijn zoektocht leiden door de informatie die verworven wordt door de verschillende leden van de populatie, waardoor een meer globaal beeld van de doelfunctie wordt bekomen en lokale minima kunnen worden vermeden. In de hydrologie werd de methode voorsnog weinig gebruikt, hoewel de aanwezige toepassingen veelbelovend zijn (Scheerlinck *et al.*, 2009). Om evenwel na te gaan wat de meerwaarde van het gebruik van een populatiegebaseerd algoritme is, wordt ze vergeleken met een niet-populatiegebaseerd algoritme, namelijk Simplex-Simulated Annealing (SIMPISA) (Cardoso *et al.*, 1996). SIMPISA combineert de Simplex-methode van Nelder & Mead (1965) en Simulated Annealing (Kirkpatrick *et al.*, 1983; Kirkpatrick, 1984) met het oog op een verbeterde ontwijking van lokale minima. Ook deze methode is een onbekende in hydrologische toepassingen.
- *Hoe kan de performantie van beide methoden het best beoordeeld worden?* Gezien de methoden gebruikt worden om een bepaalde doelfunctie te minimaliseren kunnen de bekomen resultaten vergeleken worden op basis van de *fitness* van de doelfunctie. Daarnaast dient uiteraard ook rekening gehouden te worden met andere eigenschappen zoals benodigde rekentijd, betrouwbaarheid en gebruiksvriendelijkheid.
- *Wat is de invloed en de meerwaarde van de incorporatie van het derde orde moment in de doelfunctie?* Cowpertwait (1998) oppert dat de incorporatie van het derde orde moment in de doelfunctie leidt tot een betere reproductie van extreme neerslaggebeurtenissen door de modellen. In het huidige werk wordt dan ook nagegaan wat de invloed van deze incorporatie is op de algemene statistische eigenschappen en de extreme waarden van de gesimuleerde neerslagreeksen.
- *Wat is de invloed van de doelfunctie op de performantie van de gekalibreerde modellen?* De kalibratie van de modellen omvat enkel het minimaliseren van een vooropgestelde doel-

functie. Hierbij wordt enkel de fit tussen de gemodelleerde en geobserveerde eigenschappen die expliciet vervat zijn in de doelfunctie geoptimaliseerd. Logischerwijs kunnen niet alle eigenschappen waarvoor de gesimuleerde reeks in overeenstemming moet zijn met de observaties worden geïncorporeerd in de doelfunctie. Aangezien bepaalde alternatieve configuraties van de doelfunctie voorhanden zijn, wordt de invloed van deze alternatieven op de algemene statistische eigenschappen en de extremen onderzocht.

- *Kunnen de bevindingen uit vorige onderzoeksvragen succesvol worden vertaald naar de kalibratie van een nieuw model en wat is de meerwaarde van dit model?* Recentelijk werd een nieuw type Bartlett-Lewis model ontwikkeld met het oog op een verbeterde reproductie van extreme neerslaggebeurtenissen. Dit model zal worden gekalibreerd aan de hand van de meest geschikt bevonden zoekmethode, type doelfunctie en al dan niet met de incorporatie van het derde orde moment in de doelfunctie. Er wordt nagegaan of deze bevindingen hier tot goede resultaten leiden, waarna de performantie en de meerwaarde van dit nieuwe model zal worden beoordeeld.

### 1.3 Indeling

Om meer inzicht te verwerven in de opbouw van de Bartlett-Lewis modellen wordt in Hoofdstuk 2 een blik geworpen op de verschillende concepten die aan de basis liggen van stochastische neerslagmodellen. Vertrekkende van deze basisconcepten wordt een kort overzicht gegeven van de bestaande literatuur over de beschouwde modellen en aanverwanten. Hierbij wordt bijzondere aandacht besteed aan de evolutie die de Bartlett-Lewis modellen reeds gekend hebben en de redenen daarvoor.

De beschouwde modellen worden gekalibreerd aan de hand van observaties, geregistreerd te Ukkel. In Hoofdstuk 3 worden de statistische eigenschappen van de dataset toegelicht. De manier waarop de geregistreerde neerslag wordt geanalyseerd vormt ook meteen de grondslag voor de analyse van de gesimuleerde neerslag.

Een belangrijk onderdeel van deze thesis bestaat uit het zoeken naar een heuristische zoekmethode die de gebruiker in staat stelt om op efficiënte en gerichte wijze een parametrisatie uit te voeren. Er worden twee zoekmethoden (SIMPSA en PSO) voorgesteld die vooralsnog geen of weinig toepassing kenden binnen het gebied van de neerslagmodellering. Hoofdstuk 4 gaat dieper in op de verschillende aspecten van het parametrisatie- of optimalisatieproces van de Bartlett-Lewis modellen. Daarnaast wordt de theorie die schuilgaat achter de twee voorgestelde zoekmethoden uiteengezet, alsook wordt stilgestaan bij de praktische implementatie ervan.

In Hoofdstuk 5 wordt de nadruk gelegd op de prestaties van de voorgestelde zoekmethoden. Het doel van dit hoofdstuk is het bepalen welk van beide methoden ons als beste in staat stelt om de besproken modellen te kalibreren. Hierbij wordt niet alleen rekening gehouden met de kwaliteit van de parametrisatie, maar ook met meer praktische zaken als snelheid, gebruiksvriendelijkheid, algemene toepasbaarheid.

In Hoofdstuk 6 worden de prestaties van de gekalibreerde modellen beoordeeld, in vergelijking met de observaties en met elkaar. De vergelijking is hierbij gebaseerd op de verschillende eigenschappen van de gesimuleerde neerslagreeksen, zoals dit wordt voorgesteld in Hoofdstuk 3.

Met de in Hoofdstuk 6 gekalibreerde modellen als referentie, wordt in Hoofdstuk 7 een eerste



poging gedaan om het optimalisatieproces te verbeteren. Tot dan toe werden de beschouwde modellen gekalibreerd zonder rekening te houden met hogere-orde eigenschappen van de neerslag. In Hoofdstuk 7 wordt de invloed van de introductie van dergelijk hogere-orde moment (namelijk het derde centrale moment) op de prestaties onderzocht.

Naast het sleutelen aan de gebruikte eigenschappen van de neerslag tijdens de parametrisatie, kan ook gesleuteld worden aan het belang dat tijdens de parametrisatie aan elk van deze onderlinge eigenschappen gegeven wordt. Hoofdstuk 8 onderzoekt hoe op een objectieve wijze aan bepaalde eigenschappen een groter, dan wel kleiner, gewicht kan worden toegekend tijdens de parametrisatie.

Tot slot wordt in Hoofdstuk 9 een eerste parametrisatie van een nieuw Bartlett-Lewis model (het Truncated Bartlett-Lewis model) uitgevoerd. Hierbij wordt rekening gehouden met de opgedane kennis tijdens de voorgaande hoofdstukken. Hierbij wordt opnieuw stilgestaan bij het parametrisatieproces en wordt de gesimuleerde neerslag onderworpen aan een statistische analyse.

Hoofdstuk 10 sluit af met een samenvatting van de bekomen resultaten en de formulering van een algemeen besluit. Daarnaast wordt nog een blik geworpen op enkele belangrijke aandachtspunten bij toekomstig onderzoek.



## Hoofdstuk 2

# Neerslagmodellen

### 2.1 Inleiding

Hydrologie betekent letterlijk de studie van het water. Deze tak van de wetenschap onderzoekt de circulatie van water in de natuurlijke omgeving en de processen die zich hierbij afspelen (Brutsaert, 2005). Een belangrijk onderdeel van de hydrologische cyclus is neerslag. Niet alleen in hydrologie, maar ook in meteorologie, atmosferische fysica en klimatologie wordt neerslag uitvoerig bestudeerd (Onof *et al.*, 2000). De studie van neerslag en de achterliggende mechanismen zijn dus een interdisciplinair gegeven.

Neerslag is een belangrijke schakel in veel hydrologische toepassingen. Een accurate representatie van neerslag is bijvoorbeeld van belang bij het ontwikkelen van afvoer- en grondwaterstromingsmodellen, het voorspellen van overstromingen en droogte en het ontwerp van rioleringsnetwerken (Onof *et al.*, 1994). De hiervoor beschikbare observaties van historische neerslagreeksen zijn echter in veel gevallen te kort of werden opgemeten op een tijdschaal die niet geschikt is voor een bepaalde toepassing. In dat geval is het van belang dat historische reeksen kunnen worden gesynthetiseerd en gesimuleerd (Bras & Rodriguez-Iturbe, 1976).

De simulatie van neerslag gebeurt aan de hand van wiskundige modellen. Het ontwikkelen van dergelijke modellen is geen sinecure. De structuur van neerslag is danig divers en variabel dat het ontwerpen van een alomvattend model alvast onmogelijk blijkt te zijn (Waymire & Gupta, 1981). Er werden niettemin reeds vele pogingen ondernomen om zowel de temporele als de ruimtelijke structuur van neerslag in modelvorm te gieten. Deze inspanningen hebben geleid tot een groot aanbod aan neerslagmodellen met elk hun specifieke eigenschappen en conceptuele benadering, veelal onderling verschillend in functie van de toepassing. Deze verschillende types neerslagmodellen kunnen ruwweg in vier categorieën ingedeeld worden (Onof *et al.*, 2000).

- Meteorologische modellen : opgebouwd uit grote complexe sets differentiaalvergelijkingen die de fysische processen, die leiden tot regenval, exact trachten te beschrijven (Mason, 1986).
- Stochastische modellen : beschrijven de ruimtelijke evolutie van het regenvalproces op een schaal-onafhankelijke wijze. Hiertoe worden de *Multi-scaling* modellen gerekend, bijvoorbeeld multifractale cascades (Lovejoy & Schertzer, 1990; Gupta & Waymire, 1994).
- Statistische modellen (*single of multisite temporal*) zoals ontwikkeld door Stern & Coe (1984); Chandler & Wheeler (1998a,b), etc., die in staat zijn om ruimtelijke niet-stationariteit

en temporele trends weer te geven.

- Temporele en spatio-temporele stochastische procesmodellen. Hierbij worden eenvoudige veronderstellingen omtrent de fysische processen, die betrokken zijn bij de vorming van neerslag, aangenomen. In tegenstelling tot de stochastische modellen worden deze modellen beter in staat geacht de hiërarchische structuur van het neerslagproces weer te geven. In vergelijking met de statistische modellen laten ze toe om vele karakteristieken van de neerslag te bundelen met slechts een aantal parameters en om interseizoenale en intersite vergelijkingen te maken.

Dit werk beperkt zich tot temporele stochastische procesmodellen. Meer bepaald modellen die neerslag modelleren in één enkel punt. Aangezien in de meeste gevallen neerslag wordt gemeten met behulp van pluviometers, leunt de conceptuele benadering van de puntneerslagmodellen het dichtst aan bij de observaties (Haan *et al.*, 1976).

In aanloop naar de beschrijving van de in dit werk gebruikte modellen wordt in Sectie 2.2 dieper ingegaan op de temporele structuur van regenval en de wiskundige concepten die hiermee gepaard gaan.

## 2.2 Modelling van puntneerslag

De vorming van neerslag kan beschreven worden aan de hand van twee basisprincipes. Indien de neerslag in een bepaald punt, in functie van de tijd, wordt bekeken, treden intuïtief twee variabelen naar de voorgrond. Een eerste variabele is het aantal buien of *events*. Op een bepaald tijdstip regent het immers, of regent het niet. Wat daarbij van belang is, is hoe hard het dan precies regent. De tweede variabele associeert dus bepaalde intensiteiten met de events (Waymire & Gupta, 1981).

De hieronder besproken stochastische puntneerslagmodellen zijn allen gebaseerd op deze twee basisprincipes. De manier waarop de manifestatie van een bui, en de intensiteit ervan in functie van de tijd, wordt gemodelleerd zal echter verschillen van model tot model.

De verschillende manieren waarop deze eigenschappen van neerslag kunnen worden gemodelleerd worden in onderstaande secties besproken. Hierbij wordt in Sectie 2.2.1 dieper ingegaan op de manifestatie van neerslagevents in functie van de tijd. Sectie 2.2.2 tracht vervolgens een overzicht te geven van de verschillende manieren waarop de duur en intensiteit van een dergelijke bui wordt gemodelleerd. Sectie 2.2.3 sluit daarbij aan met een bespreking van de meest courant gebruikte puntneerslagmodellen.

### 2.2.1 Manifestatie van neerslagevents

Laten we aannemen dat een discrete tijdschaal gehanteerd wordt. En stel dat  $Y_1, Y_2, Y_3, \dots, Y_i, \dots$  een opeenvolging van willekeurige variabelen is die de neerslaghoeveelheid gedurende een bepaalde periode  $i$  weergeven. De duur van  $Y_i$  wordt het aggregatieniveau genoemd. Stel dat bijvoorbeeld gewerkt wordt op het dagniveau. In dat geval stelt  $Y_i$  de gevallen neerslag gedurende dag  $i$  voor. Definieer vervolgens een nieuwe reeks  $Z_i$  als volgt :

$$\begin{aligned} Z_i &= 1, & \text{als } Y_i > 0 \\ Z_i &= 0, & \text{als } Y_i = 0, \end{aligned} \tag{2.1}$$

en  $N_n = \sum_{i=0}^n Z_i$ . De variabele  $N_n$  telt met andere woorden het aantal dagen waarop regen valt in een vast tijdsinterval  $[0, n]$ . Om nu de probabilistische structuur van  $N_n$  te bepalen wordt uitgegaan van een eenvoudige veronderstelling omtrent de reeks  $Z_i$ . De variabelen  $Z_i$  wordt verondersteld onafhankelijk en identiek verdeeld te zijn met  $P(Z_i = 1) = p$  en  $P(Z_i = 0) = q = 1 - p$  voor  $i = 1, 2, \dots, n$ . Dan volgt hieruit dat  $N_n$  een binomiale verdeling volgt:

$$P(N_n = k) = \binom{n}{k} p^k q^{n-k} \quad , 0 \leq k \leq n \quad (2.2)$$

Niettegenstaande het feit dat deze methode weinig bijval kent in praktische toepassingen wordt zijn analogon in een continue tijdschaal wel veelvuldig toegepast. Wanneer een tijdsinterval  $[0, t]$  wordt gediscretiseerd in  $n$  subintervallen van lengte  $\Delta t$  en over deze discrete tijdsintervallen het bovenstaand binomiaal proces wordt gedefinieerd, kan worden gezien dat wanneer  $n \rightarrow \infty$ ,  $p \rightarrow 0$  zodat  $np \rightarrow \lambda t$ . Vergelijking 2.2 convergeert naar een Poisson verdeling (Feller, 1968). Indien  $N(t) \equiv N(0, t)$  het aantal regenevents voorstelt die voorkomen in het interval  $[0, t]$ , dan kan de probabiliteitsdichtheid van  $N(t)$  voorgesteld worden als

$$P(N(t) = k) = e^{-\lambda t} (\lambda t)^k / k! \quad , k = 0, 1, \dots \quad (2.3)$$

Deze vergelijking komt overeen met een homogeen Poisson proces. De parameter  $\lambda$  kan gezien worden als de snelheid waarmee events zich voordoen in de tijd. Een variant op dit homogeen Poisson proces wordt bekomen wanneer  $\lambda$  in functie van de tijd varieert. Vergelijking 2.3 wordt dan als volgt aangepast.

$$P(N(t) = k) = e^{-\Lambda(t)} (\Lambda(t))^k / k! \quad , k = 0, 1, \dots$$

met

$$\Lambda(t) = \int_0^t \lambda(s) ds \quad , 0 < s < \infty \quad (2.4)$$

Dit proces wordt het niet-homogeen Poisson proces genoemd (Waymire & Gupta, 1981). Voor meer uitleg hierover wordt verwezen naar Cox & Isham (1980).

Het voornaamste probleem van het gebruik van een Poisson proces voor de beschrijving van de manifestatie van een event is dat verondersteld wordt dat opeenvolgende gebeurtenissen zich onafhankelijk van elkaar voordoen (Waymire & Gupta, 1981). Onderzoek naar de temporele structuur van neerslag wijst uit dat niet het geval is. Het feit dat de dagelijkse weers- en neerslagpatronen zowel persistent als periodiek zijn (Wiser, 1964) en dat elk neerslagevent kan gezien worden als een cluster in de tijd (Kavvas & Delleur, 1981) doet vermoeden dat in werkelijkheid geen sprake is van onafhankelijkheid.

Een mogelijke tegemoetkoming hieraan is het aannemen dat de beschouwde incrementen de wetten van een discrete Markov-keten volgen. De probabilistische structuur van een Markov-keten wordt volledig bepaald door een transitieprobabiliteitsmatrix en een initiële probabiliteitsdistributie. Indien deze transitieprobabiliteit tijdsafhankelijk is, spreekt men van een homogene Markov-keten (Waymire & Gupta, 1981). Een natuurlijke variant bestaat erin die kans seizoensgebonden te laten variëren (Stern & Coe, 1984). Gezien deze modellen sterk verschillend zijn van de in dit werk gebruikte modellen, wordt hiervoor verwezen naar de literatuur (Gabriel & Neumann, 1962; Haan *et al.*, 1976; Todorovich & Woolhiser, 1975; Stern & Coe, 1984; Fofoula-Georgiou & Lettenmaier, 1987). Bovendien kan gesteld worden dat deze modellen over

het algemeen niet in staat zijn om neerslagfenomenen op verschillende tijdschalen weer te geven (Rodriguez-Iturbe *et al.*, 1987a).

Hoewel het gebruik van het Poisson proces een redelijke flexibiliteit bezit om neerslag te modelleren op een bepaalde tijdschaal is het vrij beperkt wanneer verschillende tijdschalen worden beschouwd. De prestatie van dergelijk model kan echter aanzienlijk verbeterd worden door de introductie van een geclusterd puntproces (Kavvas & Delleur, 1981). Twee zulke processen komen in aanmerking voor het opstellen van een clustergebaseerd puntproces, het Neyman-Scott proces en het Bartlett-Lewis proces (Rodriguez-Iturbe *et al.*, 1987a).

### 2.2.2 Duur en intensiteit

Gekoppeld aan de manifestatie van de neerslagevents dient een methode voorzien te worden die neerslaghoeveelheid en de duur van het event beschrijft. Hiervoor wordt doorgaans een rechthoekige puls gekoppeld aan elk punt dat via het Poissonproces wordt gegenereerd. Deze puls wordt gekenmerkt door een willekeurige lengte en een willekeurige intensiteit. De totale intensiteit op een bepaald tijdstip is de som van alle actieve cellen op dat tijdstip (Rodriguez-Iturbe *et al.*, 1987a). Naast het gebruik van rechthoekige pulsen werden ook reeds methoden ontwikkeld die gebruik maken van bijvoorbeeld driehoekige pulsen (Marien & Vandewiele, 1986) of pulsen die afwijken van de rechthoekige vorm door de introductie van een willekeurige component of *jitter* in het model (Rodriguez-Iturbe *et al.*, 1987a). Het gebruik van rechthoekige pulsen is gerechtvaardigd omdat er geen duidelijke patronen waar te nemen zijn in de vorm van neerslagpulsen. Bovendien minimaliseert deze aanpak het aantal benodigde parameters en bijgevolg de rekenintensiteit (Rodriguez-Iturbe *et al.*, 1987a).

### 2.2.3 Clustergebaseerde modellen

Zoals reeds vermeld in voorgaande secties wordt bij de modellering van puntneerslag veelal gebruik gemaakt van een geclusterd puntproces, dat de manifestatie van events in functie van de tijd beschrijft, waaraan rechthoekige pulsen, die de intensiteit van de neerslag beschrijven, worden gekoppeld. In dit werk wordt dan ook gebruik gemaakt van dergelijke clustergebaseerde modellen. In deze sectie wordt een overzicht gegeven van de specifieke eigenschappen van de verschillende varianten clustergebaseerde puntneerslagmodellen.

Twee belangrijke voorbeelden van geclusterde puntprocessen zijn het Bartlett-Lewis en het Neyman-Scott proces. Het verschil tussen beiden is vrij subtiel. Beide processen zijn gebaseerd op Poisson processen voor het genereren van stormmoorsprongen en met elk van deze stormmoorsprongen wordt een willekeurig aantal rechthoekige cellen geassocieerd. In het Neyman-Scott proces worden de posities van deze cellen bepaald door een reeks onafhankelijke en identiek verdeelde variabelen, waarbij de locatie van deze distributie wordt gegeven door de stormmoorsprong. In het Bartlett-Lewis proces zijn de intervallen tussen opeenvolgende cellen onafhankelijk en identiek verdeeld (Rodriguez-Iturbe *et al.*, 1987a). Het subtiel verschil tussen beide methoden maakt dat ze allebei vergelijkbare resultaten opleveren. Er kan dan ook enkel op basis van empirische analyse besloten worden welk van beide modellen de voorkeur geniet voor een bepaalde praktische toepassing in een bepaald gebied (Velghe *et al.*, 1994).

Het Neyman-Scott proces werd oorspronkelijk ontwikkeld om de ruimtelijke distributie van ster-

rebeelden te beschrijven (Neyman & Scott, 1958) en werd later door Vere-Jones (1970) gebruikt in de tijdsdimensie om het voorkomen van aardbevingen te modelleren. Het oorspronkelijke Neyman-Scott Rectangular Pulses model (NSRP) veronderstelt dat stormoorsprongen voorkomen als een Poisson proces met parameter  $\lambda$ . Een willekeurig aantal celoorsprongen  $C$  ( $C \geq 1$ ) wordt geassocieerd met de storm. De celoorsprongen worden onafhankelijk van elkaar verplaatst ten opzichte van de celoorsprong met afstanden die exponentieel verdeeld zijn met parameter  $\beta$ , waarbij geen enkele oorsprong samenvalt met de stormoorsprong. Met elk van deze celoorsprongen wordt dan een cel geassocieerd die een willekeurige duur en diepte heeft en waarbij de duur een exponentiële verdeling volgt met parameter  $\eta$ . De celclusters geassocieerd met bepaalde stormen zijn onafhankelijk. Voor de verdeling van  $C$  kan gebruik gemaakt worden van een geometrische verdeling, een Poisson verdeling of een Truncated Poisson verdeling (opdat  $C$  positief zou zijn).

In het oorspronkelijk Bartlett-Lewis Rectangular Pulses model (OBL) worden stormoorsprongen gegenereerd aan de hand van een Poisson proces met parameter  $\lambda$ . Elke oorsprong wordt gevolgd door een aantal celoorsprongen die ontstaan als gevolg van een Poisson proces met parameter  $\beta$ . Na een bepaalde tijd, exponentieel verdeeld met parameter  $\gamma$ , stopt het aanmaken van celoorsprongen. Rodriguez-Iturbe *et al.* (1987a) introduceert hierbij de dimensieloze parameters  $\kappa = \beta/\eta$  en  $\phi = \gamma/\eta$ . Het aantal celoorsprongen  $C$  per storm volgt een geometrische verdeling met gemiddelde  $\mu_c = 1 + \kappa/\phi$ . Elke celoorsprong wordt gekoppeld aan een cel met willekeurige duur en hoogte, waarbij de duur exponentieel verdeeld is. Ook de celhoogte of intensiteit volgt een exponentiële verdeling. Het OBL model kan met andere woorden beschreven worden aan de hand van vijf parameters. Deze worden weergegeven in Tabel 6.2.

Rodriguez-Iturbe *et al.* (1987a,b) bestudeerden uitvoerig beide modellen (NSRP en OBL) en kwamen daarbij tot de conclusie dat deze de algemene karakteristieken van de neerslag behoorlijk benaderden. Het aantal droge intervallen wordt echter voor grotere aggregatieniveaus door beide overschat en de keuze van een geometrische of Poisson verdeling voor het aantal cellen per storm bij NSRP blijkt te leiden tot een bias in de resultaten. Deze kans op droge intervallen of *zero depth probability* (ZDP) is een belangrijke eigenschap bij het modelleren van neerslag, zeker wanneer het neerslagmodel gekoppeld wordt aan een neerslag-afvoer model. Overeenkomsten van de modeloutput van een infiltratiemodel met gemeten runoff waren dan ook niet bevredigend.

Rekening houdend met deze bevindingen werd het model aangepast (Rodriguez-Iturbe *et al.*, 1988), en kreeg het de naam Modified Bartlett-Lewis Rectangular Pulse model (MBL). De gemiddelde celduur wordt toegestaan te variëren van storm tot storm. Zo houdt het model rekening met het voorkomen van structureel verschillende stormtypes. De parameter van de exponentieel verdeelde celduur  $\eta$  volgt nu een gamma verdeling met index  $\alpha$  en schaalparameter  $\nu$ . Hieruit volgt dat  $E[\eta] = \alpha/\nu$  en  $\text{VAR}[\eta] = \alpha/\nu^2$ . Opdat de verwachte levensduur van 1 cel eindig zou zijn wordt aangenomen dat  $\alpha > 1$ . Door  $\kappa$  en  $\phi$  constant te houden en  $\eta$  random wordt een gelijkaardige structuur per storm bekomen, maar hebben de verschillende stormen cellen waarvan het gemiddelde en de variantie variëren van storm tot storm. Het MBL model krijgt ten opzichte van het OBL een extra parameter (zie Tabel 6.2).

Geïnspireerd door het MBL model werd ook het NSRP model aangepast (Entekhabi *et al.*, 1989) tot het Modified Neyman-Scott Rectangular Pulse model. Ook hier wordt zoals bij MBL  $\eta$  variabel per storm volgens een gamma distributie. Het verschil met MBL is echter dat  $\kappa$  en

$\phi$  niet constant worden gehouden. Zoals reeds aangehaald leidde bij het NSRP de keuze tussen een geometrisch verdeeld of Poisson verdeeld aantal cellen per storm tot verschillende resultaten (Rodriguez-Iturbe *et al.*, 1987b). Het aantal cellen per storm wordt daarom door Entekhabi *et al.* (1989) als Poisson verdeeld verondersteld.

In eerste instantie blijken de aangepaste modellen te leiden tot betere resultaten (Rodriguez-Iturbe *et al.*, 1988; Entekhabi *et al.*, 1989). De modellen zijn in staat de ZDP beter te reproduceren, ook de correlatiestructuur leunt dichter aan bij de geobserveerde en er wordt een verbetering opgetekend bij de frequentie-analyse (Velghe *et al.*, 1994). Toch werden door Onof & Wheeler (1993) excessieve waarden voor de autocorrelaties voor  $lags^1$  groter dan 12 uur vastgesteld. Bovendien werden de extreme waarden voor terugkeerperiodes groter dan de lengte van de geobserveerde dataset volgens hen slecht gereproduceerd. Als antwoord hierop werd het Modified Bartlett-Lewis Gamma model (MBLG) ontworpen (Onof & Wheeler, 1994b). Hierbij wordt de celintensiteit, in plaats van door een exponentiële distributie, voorgesteld door een gamma verdeling met twee parameters. Hierdoor werd een betere schatting bekomen. De zeven parameters (zie Tabel 6.2), in plaats van zes bij MBL, bemoeilijken echter wel het kalibratieproces (Onof *et al.*, 1994; Onof & Wheeler, 1994b).

Een probleem dat opduikt bij het gebruik van de voornoemde modellen is dat wanneer lange simulaties worden geproduceerd, zich occasioneel onrealistische events voordoen. Het gaat hierbij om events die worden gekarakteriseerd door lange cellen in lange stormen. In vele gevallen leidt dit tot onverwacht grote extremen op grove tijdschalen zoals een dag (Verhoest *et al.*, 2010).

De reden hiervoor is dat de parameter  $\eta$ , welke uit een Gamma distributie getrokken wordt, een extreem lage waarde wordt toegekend (Wheeler *et al.*, 2006). Met als gevolg dat de kans op het trekken van een cel met heel lange duur heel groot is (zijn verwachtingswaarde is  $\eta^{-1}$ ). De verwachtingswaarde van de duur van de storm  $(\eta\phi)^{-1}$  wordt bijgevolg ook groot. Tijdens deze storm is de verwachtingswaarde van de manifestatie van regencellen echter laag ( $\kappa\eta$ ). Indien nu een gemiddeld tot hoge celintensiteit, onafhankelijk bekomen uit een andere exponentiële distributie, wordt toegekend aan deze lange cellen ontstaat een onrealistische regencel. Gezien deze overdreven cellen niet per se een grote intensiteit hebben worden ze niet ontmaskerd door een klassieke statistische analyse. Ze zijn niettemin verantwoordelijk voor de overschatting van extreme overstromingen (Verhoest *et al.*, 2010).

Om aan dit probleem tegemoet te komen werd het Truncated Bartlett-Lewis (TBL) model ontwikkeld (Onof, 2011), waarbij de extra parameter  $\epsilon$ , ten opzichte van het MBL model (zie Tabel 6.2), instaat voor de truncatie van de Gamma verdeling waaruit de  $\eta$  getrokken wordt. Hierdoor wordt verhinderd dat  $\eta$  een waarde aanneemt die lager is dan de drempelwaarde  $\epsilon$ . Zodoende kan het voornoemde probleem met onrealistische neerslagcellen verholpen worden. De truncatie van de gamma verdeling werd reeds toegepast door Wheeler *et al.* (2006) en Verhoest *et al.* (2010). Hun resultaten toonden aan dat de methode de verwachtingen inlost. Verder onderzoek is echter aan de orde. Meer bepaald de kalibratie van het model werd tot op heden nog niet nauwgezet uitgevoerd.

De parametrisatie van deze modellen blijkt een complexe onderneming te zijn (Verhoest *et al.*, 1997). Het objectief van de huidige thesis is om na te gaan of met geavanceerde kalibratiemethoden sneller een optimale parameterset kan gevonden worden. Daartoe neemt het huidige

---

<sup>1</sup>De autocorrelatie met lag  $i$ , van een bepaalde tijdreeks, is de correlatie tussen de oorspronkelijke tijdreeks en diezelfde tijdreeks, verschoven over een interval  $i$ .



werk vier van de besproken modellen in beschouwing. Voor een analytische beschrijving van de gebruikte modellen wordt verwezen naar Bijlage A.

**Tabel 2.1:** Samenvattende tabel van de te optimaliseren parameters van de besproken Bartlett-Lewis modellen.

model	parameters					
OBL	$\lambda$	$\beta$	$\gamma$	$\mu_x$	$\eta$	
MBL	$\lambda$	$\kappa$	$\phi$	$\mu_x$	$\alpha$	$\nu$
MBLG	$\lambda$	$\kappa$	$\phi$	$\alpha$	$\nu$	$p$ $\delta$
TBL	$\lambda$	$\kappa$	$\phi$	$\mu_x$	$\alpha$	$\nu$ $\epsilon$



## Hoofdstuk 3

# Beschrijving van de neerslag te Ukkel

### 3.1 Inleiding

Voor de kalibratie van de in hoofdstuk 2 beschreven neerslagmodellen wordt gebruik gemaakt van een historische neerslagreeks, bekomen te Ukkel. De dataset heeft een resolutie van 10 minuten en bestaat uit neerslagregistraties (tot op 0.1 mm nauwkeurig) geaccumuleerd over intervallen van 10 minuten. Deze gegevens werden opgemeten, gedurende een periode van 105 jaar, van 1 januari 1898 tot 31 december 2002, aan de hand van een Fellman-Feuss pluviograaf in het klimatologische park van het Koninklijk Meteorologisch Instituut te Ukkel, nabij Brussel (Demaree, 2003). Er dient vermeld te worden dat dit een uitzonderlijke tijdreeks is. Ten eerste omdat de reeks uitzonderlijk lang is, en gemeten werd met een unieke meetfrequentie van 10 minuten. Ten tweede, de data is over de hele lijn van hoge kwaliteit. Dit komt doordat de neerslag sinds 1898 door hetzelfde toestel werd gemeten en op dezelfde manier verwerkt werd (Ntegeka & Willems, 2008; Vandenberghe *et al.*, 2010). Een onderdeel van deze reeks, meer bepaald 27 jaar, werd reeds gebruikt voor de kalibratie van de Bartlett-Lewis modellen door Verhoest *et al.* (1997). Voor een uitgebreidere analyse van de dataset met betrekking tot de aanwezigheid van trends en cycli in de 105 jaar lange neerslagreeks van Ukkel wordt verwezen naar De Jongh *et al.* (2006).

### 3.2 Statistische analyse

De statistische analyse van de neerslag is gebaseerd op eerste, tweede en derde-orde eigenschappen van de neerslagintensiteit (Sectie 3.2.1), in combinatie met enkele voorwaardelijke distributiekarakteristieken omtrent diens verdeling (Sectie 3.2.2 en 3.2.3) en het uitvoeren van een frequentie-analyse (Sectie 3.2.4).

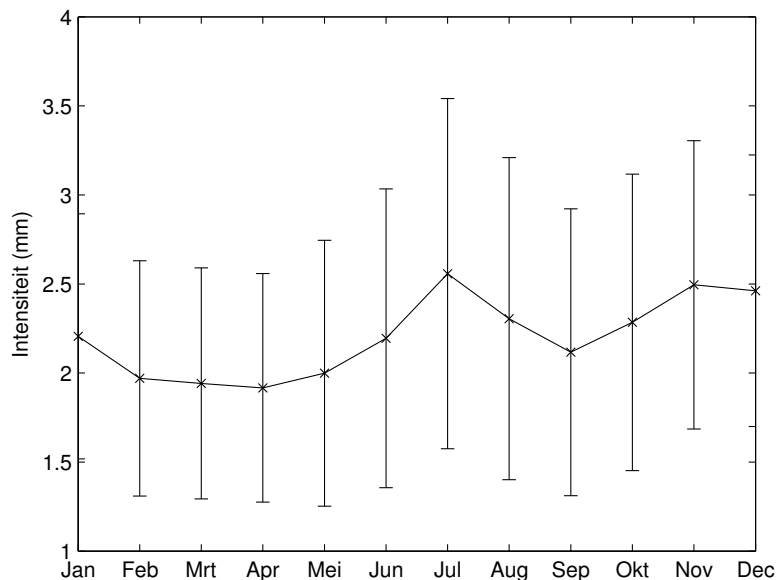
#### 3.2.1 Eerste, tweede en derde orde momenten

Een eerste stap in de beschrijving van neerslag in functie van de tijd, is de beschrijving van de eerste, tweede en derde orde momenten van de neerslagintensiteit. Hiermee wordt het gemiddelde, de variantie, de autocovariantie en de autocorrelaties, de ZDP en de scheefheid of *skewness*

van de neerslagintensiteit bedoeld. Bij de modellering van neerslag is het immers van belang dat de synthetische neerslagreeks hiervoor een goede overeenkomst vertoont met de historische reeks. Gezien de puntneerslagmodellen stationariteit van de neerslagmodellen veronderstellen en ons klimaat gekenmerkt wordt door seizoensale schommelingen (convectieve neerslag in de zomer en cyclonale neerslag in de winter), kan er niet worden voldaan aan deze veronderstellingen. Daarom wordt de tijdreeks opgesplitst volgens de verschillende maanden. Op die manier kan temporele homogeniteit gewaarborgd worden (Obeysekera *et al.*, 1987; Verhoest *et al.*, 1997). De historische reeks stelt metingen op het aggregatieniveau van 10 minuten ter beschikking. Bijgevolg beschikken we ook over informatie omtrent alle aggregatieniveau's die een veelvoud zijn van 10 minuten. Het lijkt geen twijfel dat er een zekere correlatie bestaat tussen deze verschillende aggregatieniveau's. Toch zijn er uitgesproken verschillen. Opdat het model voor verscheidene toepassingen kan gebruikt worden, wordt er naar gestreefd dat de synthetische reeksen op de verschillende aggregatieniveau's goede fits leveren (Rodriguez-Iturbe *et al.*, 1987a). Om deze doelstelling te bereiken wordt tijdens de kalibratie, op verschillende aggregatieniveau's, met de eigenschappen van de dataset rekening gehouden.

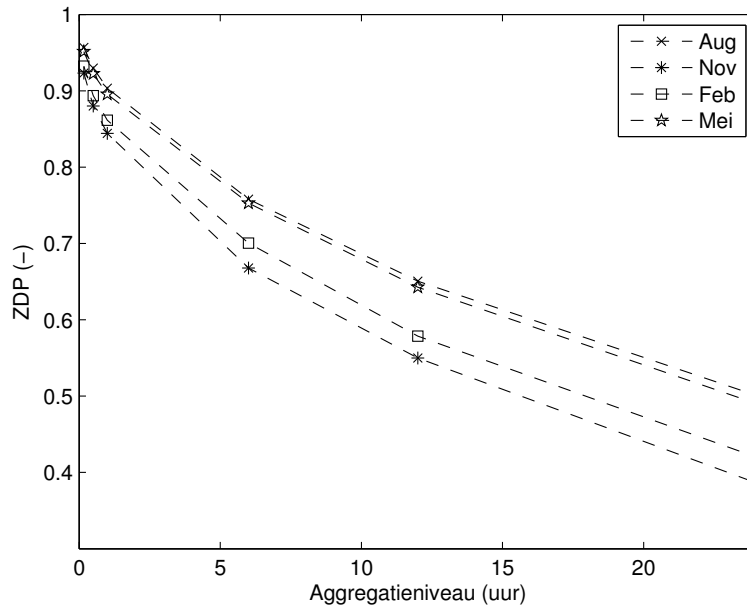
Tabel B.1 geeft de karakteristieken weer voor de historische tijdreeks, op verschillende aggregatieniveau's (AN). Deze tabellen omvatten het gemiddelde (Avg), de variantie (Var), de autocovariantie (Cov), de autocorrelatie met lag  $i$  (Cor( $i$ )) voor  $i = 1, \dots, 3$ , de kans op droge intervallen of Zero Depth Probability (ZDP) en de skewness van de verdeling (Skew).

Figuur 3.1 geeft de gemiddelde dagelijkse neerslag in Ukkel weer in functie van de maand. Hierin is geen duidelijke seizoensale trend te ontwaren. Figuur 3.2 geeft het verloop van de ZDP voor



**Figuur 3.1:** Maandelijks variatie van de dagelijkse gemiddelde neerslag met standaardfout

enkele maanden in functie van het aggregatieniveau weer. Hieruit kan worden afgeleid dat tijdens de lente en zomer de kans op droogte telkens iets hoger is dan in de herfst en winter. In samenspraak met Figuur 3.1 kan hieruit besloten worden dat er in de zomer meer droge perioden zijn maar niet noemenswaardig minder neerslag. De stormen of neerslaggebeurtenissen zijn met andere woorden intenser in de zomer.



**Figuur 3.2:** Verloop van de ZDP voor de maanden augustus, november, februari en mei

### 3.2.2 Voorwaardelijke distributiekarakteristieken

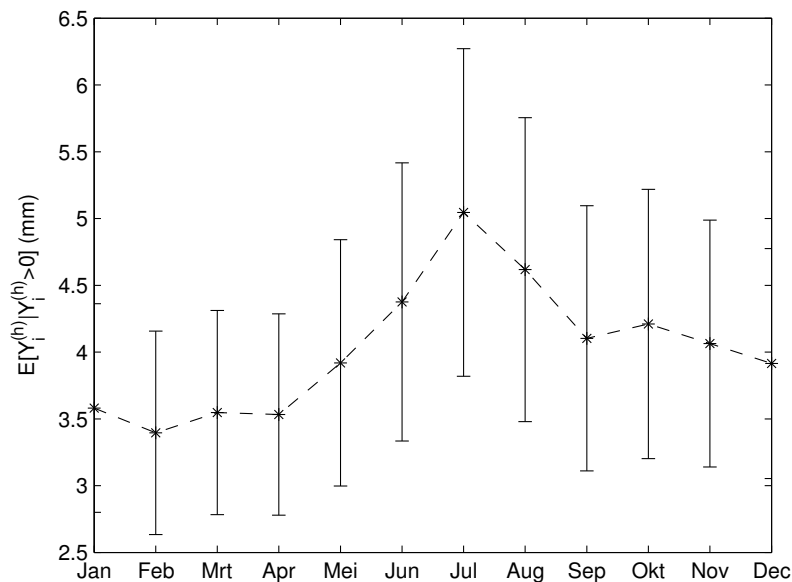
In eerste instantie dienen de gebruikte puntneerslagmodellen een correcte weergave te bieden van de geobserveerde momenten en de ZDP voor verschillende aggregatieniveau's. Om na te gaan of ook de opbouw van de gesimuleerde tijdreeks strookt met de observaties worden enkele distributiekarakteristieken gedefinieerd. In totaal worden negen voorwaardelijke karakteristieken gedefinieerd (Verhoest *et al.*, 1995). Onderstaande uitdrukkingen drukken deze voorwaardelijke distributiekarakteristieken uit in formulevorm. Hierbij is  $Y_i^{(h)}$  de regenintensiteit voor een aggregatieniveau  $h$  in het  $i$ -de interval, in mm.

$$\begin{aligned}
 \text{vwd1} &: P(Y_i^{(h)} = 0 | Y_{i-1}^{(h)} = 0) \\
 \text{vwd2} &: E[Y_i^{(h)} | Y_i^{(h)} > 0] \\
 \text{vwd3} &: \text{Var}[Y_i^{(h)} | Y_i^{(h)} > 0] \\
 \text{vwd4} &: E[Y_i^{(h)} | Y_i^{(h)} > 0, Y_{i-1}^{(h)} = 0] \\
 \text{vwd5} &: \text{Var}[Y_i^{(h)} | Y_i^{(h)} > 0, Y_{i-1}^{(h)} = 0] \\
 \text{vwd6} &: E[Y_i^{(h)} | Y_i^{(h)} > 0, Y_{i+1}^{(h)} = 0] \\
 \text{vwd7} &: \text{Var}[Y_i^{(h)} | Y_i^{(h)} > 0, Y_{i+1}^{(h)} = 0] \\
 \text{vwd8} &: E[Y_i^{(h)} | Y_i^{(h)} > 0, Y_{i-1}^{(h)} > 0] \\
 \text{vwd9} &: \text{Var}[Y_i^{(h)} | Y_i^{(h)} > 0, Y_{i-1}^{(h)} > 0]
 \end{aligned}$$

De eerste voorwaardelijke distributiekarakteristiek (vwd1) drukt de kans uit dat er geen regenval optreedt op voorwaarde dat er in het vorige tijdsinterval ook geen regen viel. Een tweede en derde karakteristiek drukken respectievelijk de verwachtingswaarden en variantie uit van de neerslagintensiteit in een tijdsinterval, op voorwaarde dat deze intensiteit niet nul is. Voorwaardelijke distributiekarakteristieken 4 en 5 (vwd4 en vwd5) drukken de verwachtingswaarde en variantie uit van de neerslagintensiteit in een tijdsinterval dat niet nul is en volgt op een droog interval. Voorwaardelijke distributiekarakteristieken 6 en 7 (vwd6 en vwd7) zijn gelijk aan 4 en 5 respectievelijk, met dat verschil dat de voorwaarde nu stelt dat de beschouwde cellen gevolgd

worden door een droog interval in de plaats van er door worden voorafgegaan. Ten slotte drukken voorwaardelijke distributiekarakteristieken 8 en 9 (vwd8 en vwd9) de verwachtingswaarde en variantie uit van de neerslagintensiteit in een interval dat niet droog is en volgt op een evenmin droog interval.

In Tabel B.2 worden deze voorwaardelijke distributiekarakteristieken voor de verschillende maanden en aggregatieniveau's voor Ukkel weergegeven. Een eerste blik op de tabel leert dat bij grotere aggregatieniveau's de variabiliteit toeneemt. Figuur 3.3 toont het verloop van de gemiddelde bui-intensiteit (vwd2) voor een aggregatieniveau van 24 uur. Hieruit blijkt dat de buien in de zomer intenser zijn dan in de winter. In tegenstelling tot de gemiddelde hoeveelheid neerslag in Figuur 3.1 is hier wel een meer uitgesproken seizoenaliteit waar te nemen. De winterse cyclonale buien zijn minder intens, duren langer (zie Figuur 3.2) en komen vaker voor (ze leiden namelijk tot een gelijkaardige totale hoeveelheid neerslag, Figuur 3.1). Dit in tegenstelling tot de intensievere en kortere zomerse buien.



**Figuur 3.3:** Verloop van de gemiddelde bui-intensiteit bij een aggregatieniveau van 24 uur, met standaardfout

### 3.2.3 Voorwaardelijke probabiliteiten inzake neerslagvolumes

Naast het behouden van de ZDP is het belangrijk dat de verdeling van de neerslag volgens intensiteit<sup>1</sup> behouden blijft bij de gesimuleerde reeks. Vooral bij het bepalen van representatieve buien voor het ontwerp van rioleringsystemen is dit van groot belang. Een onder- of overschatting van bepaalde buien is namelijk niet gewenst in dat opzicht. Hiertoe wordt de kans dat de neerslagintensiteit kleiner is dan een bepaalde drempelwaarde  $i$  ( $DP_i$ ) berekend. Aldus wordt

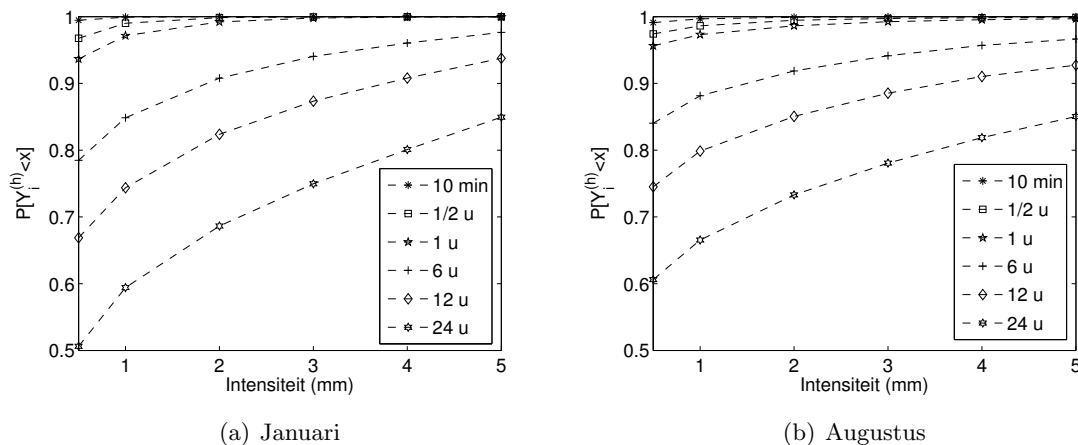
<sup>1</sup>Merk op dat het gebruik van het woord intensiteit eigenlijk niet volledig correct is. De voorwaardelijke probabiliteiten hebben eigenlijk betrekking tot gevallen volumes gedurende bepaalde tijdsperiodes. Strikt gezien zou dus de term volume moeten gebruikt worden. Aangezien deze volumes echter telkens besproken worden voor een bepaald aggregatieniveau, kunnen de besproken waarden gezien worden als volumes/aggregatieniveau, i.e. een intensiteit.

de verdeling van de neerslag volgens intensiteit in kaart gebracht. Onderstaande uitdrukkingen geven een overzicht van deze voorwaardelijke probabiliteiten, waarbij  $Y_i^{(h)}$  de neerslagintensiteit voor een aggregatieniveau  $h$  in het  $i$ -de interval voorstelt, in mm (Verhoest *et al.*, 1995).

$$\begin{aligned} DP_{0.5} &: P(Y_i^{(h)} < 0.5) \\ DP_1 &: P(Y_i^{(h)} < 1) \\ DP_2 &: P(Y_i^{(h)} < 2) \\ DP_3 &: P(Y_i^{(h)} < 3) \\ DP_4 &: P(Y_i^{(h)} < 4) \\ DP_5 &: P(Y_i^{(h)} < 5) \end{aligned}$$

Tabel B.3 geeft deze voorwaardelijke probabiliteiten weer voor de historische tijdreeks. Hieruit blijkt dat voor elk aggregatieniveau de kans op neerslag kleiner dan een bepaalde intensiteit stijgt, indien de voorwaardelijke intensiteiten stijgen. Zo is bijvoorbeeld te zien dat op een aggregatieniveau van 10 minuten, de neerslagintensiteit slechts heel uitzonderlijk groter is dan 2 mm ( $DP_2=0.9999$ ) en nooit groter dan 3 mm ( $DP_3=1.0000$ ). Deze voorwaardelijke probabiliteiten dalen logischerwijs met toenemend aggregatieniveau.

In Figuur 3.4 worden deze voorwaardelijke probabiliteiten grafisch weergegeven voor de maand januari en augustus. Hieruit blijkt dat in augustus de kans op het voorkomen van kleine neerslaghoeveelheden over het algemeen groter is dan in januari. Dit kan mede verklaard worden door het feit dat de ZDP in augustus groter is dan in januari.



**Figuur 3.4:** Kans op voorkomen van kleine neerslaghoeveelheden voor verschillende aggregatieniveau's in Ukkel (1898-2002). De ordinaat stelt de kans voor dat een neerslaghoeveelheid kleiner of gelijk aan de waarde in de abscis zich voordoet op een bepaald aggregatieniveau.

### 3.2.4 Extreme waarden analyse

Er zijn verschillende manieren waarop de extreme waarden van historische tijdreeks kunnen geanalyseerd worden. De meest eenvoudige en meest gebruikte manier is om telkens de jaarlijkse of maandelijkse maximale waarde te selecteren. Dit resulteert in een reeks van jaarlijkse of maandelijkse maxima (Verhoest *et al.*, 1997).

Een tweede wijdverspreide methode is de *Peak-Over-Threshold* (POT) methode. Hierbij worden neerslaggebeurtenissen die een bepaalde intensiteit overschrijden geselecteerd. De bekomen reeks

is uiteraard sterk afhankelijk van de subjectief gekozen drempelwaarde, en omdat de methode aanneemt dat de observaties zich onafhankelijk van elkaar voordoen is deze methode minder populair bij hydrologen (Verhoest *et al.*, 1997). De methode werd desondanks reeds meermaals gebruikt. Ntegeka & Willems (2008), bijvoorbeeld, bestuderen aan de hand van een POT-analyse de aanwezigheid van bepaalde trends in de extreme waarden te Ukkel. Hierbij stelde men een clustering van hoge extreme waarden vast, wat het opsporen van de invloed op de extreme waarden tengevolge van een klimaatsverandering bemoeilijkt. Het integreren van dergelijke clustering en al dan niet anthropogene invloeden in de analyse van de extreme liggen buiten het bestek van deze thesis. Daarom wordt voor de analyse van de historische reeks van Ukkel de analyse van de extreme waarden beperkt tot de eerstgenoemde methode. Deze wordt in onderstaande secties eerst uitgeklaard, waarna de resultaten worden weergegeven.

### 3.2.4.1 Retourperiode

Veronderstel dat een extreme gebeurtenis zich voordoet als een toevalsveranderlijke  $I$  groter is dan een zeker niveau  $I_T$ . Stel vervolgens dat  $\tau$  de tijd tussen twee opeenvolgende extreme gebeurtenissen is. Bijgevolg is  $\tau$  ook een toevalsveranderlijke waarvan de verwachtingswaarde  $E[\tau]$  de retourperiode  $T$  voor  $I \geq I_T$  genoemd wordt:

$$E[\tau] = T \quad (3.1)$$

De kans  $p$  dat de toevalsveranderlijke  $I$  groter of gelijk is aan  $I_T$  kan als volgt geschreven worden.

$$P(I \geq I_T) = \frac{1}{T} = p \quad (3.2)$$

### 3.2.4.2 Extreme waarden verdelingsfuncties

Zoals reeds vermeld wordt bij de analyse van de extremen hier enkel rekening gehouden met het maximum gedurende een bepaalde periode. De keuze van de periode waarin het maximum wordt beschouwd is uiteraard vrij. Zo kan met jaarlijkse of maandelijkse maxima gewerkt worden.

De verdeling van de bekomen reeks maxima is uiteraard verschillend van de populatie van waaruit ze afkomstig zijn. Verhoest *et al.* (1997) vond dat de data van Ukkel een Gumbel distributie volgt. De visualisatie van de extreme waarden gebeurt aan de hand van een semi-logaritmische schaal met in de abscis de *Gumbel reduced variate* (GRV) en in de ordinaat de neerslagintensiteit. De GRV wordt als volgt gedefinieerd:

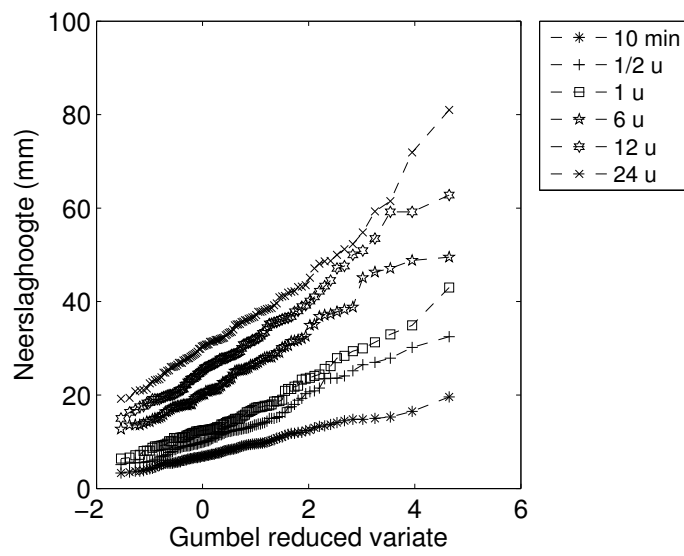
$$\text{GVR} = -\ln \left( \ln \left[ \frac{T}{T-1} \right] \right) \quad (3.3)$$

Indien een reeks extreme waarden de Gumbel verdeling volgt, dan vormt deze een rechte in de Gumbel plot.

### 3.2.4.3 Extreme waarden Ukkel

In Figuur 3.5 worden de jaarlijkse maxima voor verschillende aggregatieniveau's afgebeeld. Figuur C.1 geeft vervolgens de maandelijkse maxima.





**Figuur 3.5:** Extreme waarden diagram voor jaarlijkse maxima in Ukkel



## Hoofdstuk 4

# Optimalisatie

### 4.1 Inleiding

Kalibratie of optimalisatie duidt op het zoeken naar de meest geschikte parameters voor een bepaald model. Een parameterset wordt als meest geschikt bevonden als de resultaten van het model het best overeenkomen met de geobserveerde werkelijkheid. De schatting van de parameters gebeurt bijgevolg aan de hand van historische data. De geschatte parameters worden vervolgens gebruikt om een gesimuleerde neerslagreeks van gewenste lengte te produceren, waarbij het gewenst is dat de eigenschappen van deze reeks zo min mogelijk, of ten minste niet significant, verschillen van de geobserveerde reeks (Haan *et al.*, 1976).

Dat de kalibratie van de Bartlett-Lewis modellen geen voor de hand liggende taak is, valt af te leiden uit de uitgebreide aandacht die reeds aan het onderwerp werd besteed. Het toenemend aantal parameters bij de vernieuwde modellen dragen er toe bij dat het kalibratieproces steeds meer aandacht geniet (Verhoest *et al.*, 1997). Verscheidene methoden werden dan ook reeds aangewend voor de kalibratie van de parameters.

Het gebruik van formele statistische technieken voor de kalibratie van de Bartlett-Lewis modellen, zoals *maximum likelihood estimators*, vergt een extensieve computationele inspanning en is bijgevolg niet praktisch (Obeysekera *et al.*, 1987). Bovendien is het moeilijk om een bruikbare vorm van de likelihood functie te vinden en is niet meteen duidelijk of dit wel de beste benadering is (Rodriguez-Iturbe *et al.*, 1988). Rodriguez-Iturbe *et al.* (1988) stelde voor om een stelsel van primaire eigenschappen van de data gelijk te stellen aan bij voorkeur op theoretische wijze berekende modelwaarden. Indien dit onmogelijk blijkt te zijn, kan ook gebruik gemaakt worden van diezelfde eigenschappen, berekend via simulatie. De *Generalised Method of Moments*, die in de meeste toepassingen wordt aangewend voor het fitten van puntneerslagmodellen aan historische data (Rodriguez-Iturbe *et al.*, 1987b; Entekhabi *et al.*, 1989; Cowpertwait, 1991; Onof & Wheeler, 1993, 1994a; Velghe *et al.*, 1994; Cowpertwait, 1994; Khaliq & Cunnane, 1996; Calenda & Napolitano, 1999; Cameron *et al.*, 2000, 2001; Heneker *et al.*, 2001; Smithers *et al.*, 2002; Favre *et al.*, 2004; Cowpertwait *et al.*, 2007), lost simultaan een stelsel vergelijkingen op die de modelparameters in verband brengen met de geobserveerde momenten.

De resulterende set niet-lineaire vergelijkingen kan simultaan worden opgelost om de modelparameters af te leiden. Verschillende methoden kunnen hiertoe aangewend worden. Indien mogelijk, kunnen unieke oplossingen van de vergelijkingen bekomen worden (Rodriguez-Iturbe *et al.*, 1988; Khaliq & Cunnane, 1996). In de meeste gevallen wordt echter gebruik gemaakt

van een optimalisatieprocedure die de gebruiker in staat stelt een vooropgestelde doelfunctie te optimaliseren. In Sectie 4.2 wordt de in dit werk gebruikte doelfunctie nader toegelicht. De doelfunctie is een *least squares* doelfunctie die reeds meerdere malen werd gebruikt voor de kalibratie van Bartlett-Lewis en Neyman-Scott modellen (Entekhabi *et al.*, 1989; Cowpertwait, 1991; Velghe *et al.*, 1994; Verhoest *et al.*, 1997).

De optimalisatie van de doelfunctie werd door Velghe *et al.* (1994) en Verhoest *et al.* (1997) uitgevoerd aan de hand van *Powell's quadratically convergent algorithm*, terwijl Onof & Wheeler (1993) gebruik maken van de *Powell hybrid method*. Door de complexiteit van de doelfunctie, welke meerdere lokale minima bezit, slagen deze algoritmen er niet altijd in om tot een bevredigende oplossing te komen. Bovendien dient de gebruiker een startpunt te definiëren wat kan leiden tot een bias in de resultaten (Khaliq & Cunnane, 1996).

In het huidige werk wordt geopteerd voor meer recente algoritmen waarvan verwacht wordt dat zij meer kans bieden op het vinden van het globale optimum van de doelfunctie. De twee beschouwde algoritmen, namelijk het Simplex-Simulated Annealing (SIMPISA) algoritme (Cardoso *et al.*, 1996) en Particle Swarm Optimization (PSO) (Kennedy & Eberhart, 1995), worden in Sectie 4.5 en 4.6 respectievelijk beschreven. Alvorens over te gaan tot de beschrijving van de gekozen optimalisatietechnieken wordt eerst stilgestaan bij de algemene kenmerken van de optimalisatieprocedure. In Sectie 4.2 wordt de gebruikte doelfunctie toegelicht. Vervolgens wordt in Sectie 4.3 de selectie van de gebruikte eigenschappen in de doelfunctie besproken. Daaropvolgend wordt in Sectie 4.4 de afbakening van de parameterruimte nader bekeken, om dan over te gaan naar een uitgebreide beschrijving van de optimalisatietechnieken.

## 4.2 Doelfunctie

Zoals reeds vermeld worden de in Hoofdstuk 2 uiteengezette modellen gekalibreerd aan de hand van de *Generalised Method of Moments*. De modelparameters worden zo gekozen dat de met het model gesimuleerde neerslagreeks qua statistieken zo min mogelijk verschilt van de historische reeks. Dit wordt bewerkstelligd door de statistische eigenschappen, berekend voor de historische reeks, gelijk te stellen aan diezelfde eigenschappen, berekend aan de hand van analytische uitdrukkingen die een functie zijn van de modelparameters (Rodriguez-Iturbe *et al.*, 1987b; Entekhabi *et al.*, 1989; Cowpertwait, 1991; Onof & Wheeler, 1993, 1994a; Velghe *et al.*, 1994; Verhoest *et al.*, 1997). De algemene vorm van de doelfunctie  $f$  die door verschillende auteurs (Entekhabi *et al.*, 1989; Cowpertwait, 1991; Velghe *et al.*, 1994; Verhoest *et al.*, 1997) wordt gebruikt is de volgende :

$$f(\vec{x}) = \sum_{i=1}^N W_i \left( \frac{M_i(\vec{x})}{M'_i} - 1 \right)^2 \quad (4.1)$$

waarbij  $M_i(\vec{x})$  de analytische waarde is voor statistiek  $i$ , berekend aan de hand van parameter-vector  $\vec{x}$ ,  $M'_i$  de waarde van statistiek  $i$ , berekend aan de hand van de historische observaties,  $N$  het aantal statistieken, op bepaalde aggregatieniveau's, gebruikt bij de kalibratie en  $W_i$  het gewicht dat wordt toegekend aan statistiek  $i$ .

De uitkomst van de doelfunctie is met andere woorden een maat voor de overeenkomst tussen de parameterset  $\vec{x}$  en de geobserveerde momenten in de doelfunctie. Deze uitkomst of *fitness* wordt kleiner naarmate de analytische momenten beter overeenkomen met de geobserveerde momenten. PSO en SIMPISA (Secties 4.5 en 4.6) zullen dus ingezet worden om de doelfunctie

te minimaliseren. Indien in dit werk dus gesproken wordt over optimalisatie van de doelfunctie wordt hiermee telkens minimalisatie bedoeld. Dit neemt niet weg dat de besproken algoritmen even goed kunnen aangewend worden voor een maximalisatieprocedure.

De keuze van de gewichten is een subjectieve zaak. Velghe *et al.* (1994) en Verhoest *et al.* (1997) gebruikten bijvoorbeeld  $W_i = 1$  voor alle statistieken in de doelfunctie terwijl Cowpertwait (1991)  $W_i=100$  voor het gemiddelde gebruikten en  $W_i = 1$  voor de overige gebruikte eigenschappen.

In navolging van Verhoest *et al.* (1997) wordt  $W_i$  gelijkgesteld aan 1 voor alle gebruikte statistieken in de doelfunctie. Dit laat toe om resultaten verkregen met SIMPSA en PSO te vergelijken met wat zij hebben verkregen.

De doelfunctie wordt dus zoals vermeld geminimaliseerd aan de hand van twee optimalisatie-algoritmen, welke later in dit hoofdstuk in detail besproken worden. Deze algoritmen kalibreren alle parameters tegelijkertijd. Omwille van de complexiteit van de modellen en bijgevolg de doelfunctie wordt dit niet altijd zo uitgevoerd. Onof & Wheeler (1994b) stellen bijvoorbeeld voor het MBLG model, dat zeven parameters bezit, een parametrisatie in twee stappen voor. In een eerste stap wordt een slecht identificeerbare parameter gefixeerd terwijl de overige parameters aan de hand van de bovenstaande methode geschat worden. Dit proces wordt herhaald voor verschillende waarden van de gefixeerde parameter. Uit de bekomen set parametervectoren wordt dan aan de hand van een tweede doelfunctie de beste parameterset uitgekozen.

In het huidige werk wordt echter enkel gebruik gemaakt van de rechtstreekse methode. Van de gebruikte optimalisatie-algoritmen wordt immers verondersteld dat zij, zelfs wanneer ze geconfronteerd worden met complexe doelfuncties, het globaal optimum weten te vinden.

### 4.3 Selectie van gebruikte statistieken in de doelfunctie

De gebruikte statistieken in de doelfunctie moeten bij voorkeur zo weinig mogelijk met elkaar gecorreleerd zijn. De meeste eigenschappen dienen gevoelig te zijn aan temporele schaafeffecten op het niveau van een enkele cel en ten minste één eigenschap dient gekozen te worden op een aggregatieniveau dat overeenkomt met de tijd tussen stormen (Rodriguez-Iturbe *et al.*, 1988). De gekozen set variabelen moet dus vertegenwoordigd zijn door eigenschappen die betrekking hebben tot de intensiteit van de neerslag en tot de proportie aan droge intervallen (Onof & Wheeler, 1993). Hoe beter de analytische eigenschappen, bevat in de doelfunctie, overeenkomen met de historische waarden, des te meer kans dat de overeenkomst tussen de eigenschappen die niet werden opgenomen in de doelfunctie en de geobserveerde eigenschappen navenant is (Velghe *et al.*, 1994).

Khaliq & Cunnane (1996) vonden dat het gebruik van een groter aantal eigenschappen dan het aantal parameters in het model leidde tot een betere representatie van de historische reeks door het model. Volgens de auteurs is het beter om meerdere eigenschappen bij benadering te optimaliseren dan slechts enkele zeer goed.

Bij de keuze van de gebruikte eigenschappen in de doelfunctie en hun aantal dient rekening gehouden te worden met de gevoeligheid van de parameters. Rodriguez-Iturbe *et al.* (1987b) vonden dat de variatie in de parametersets die ontstaat door het gebruik van momenten op verschillende aggregatieniveau's gelijkaardig is aan de variatie die ontstaat door het uitvoeren van verschillende kalibraties met verschillende startpunten. Onof & Wheeler (1994b) stuitten

eveneens op aanzienlijk verschillende schattingen van de parameters bij het gebruik van verschillende momenten in de doelfunctie.

Rodriguez-Iturbe *et al.* (1988) rapporteren dat de uitkomst van de kalibratie van het MBL model sterk afhankelijk is van de gebruikte momenten in de doelfunctie. Hierbij zijn vooral de parameters van de Gamma verdeelde gemiddelde celduur,  $\alpha$  en  $\nu$ , moeilijk te identificeren. Onof & Wheeler (1993) tonen voor datzelfde model aan dat, op  $\mu_x$  en  $\lambda$  na, alle parameters bij verandering van de doelfunctie sterk variabel zijn, maar dat deze parameters niettemin bevredigende resultaten leverden. Ook Velghe *et al.* (1994) besluiten dat het MBL model gevoelig is aan de gebruikte momenten in de doelfunctie.

Er blijkt dus geen brede consensus te bestaan over de keuze van de momenten in de doelfunctie. Wel kan, op basis van de literatuur, gesteld worden dat zowel uitdrukkingen die de intensiteit van de neerslag beschrijven als uitdrukkingen die de proportie aan droge intervallen in rekening brengen vertegenwoordigd moeten zijn (Onof & Wheeler, 1993). Om de variabiliteit van de geschatte parameters te verkleinen is het bovendien aangewezen om de verschillende eigenschappen op verschillende aggregatieniveau's op te nemen. Hierbij is het van belang dat deze aggregatieniveau's relatief ver uiteen liggen (Calenda & Napolitano, 1999).

Voor het opstellen van de doelfuncties voor OBL, MBL en MBLG wordt beroep gedaan op het werk van Verhoest *et al.* (1997), die aan de hand van de gebruikte doelfuncties goede resultaten wisten te bekomen. De kalibratie van het TBL model werd nog niet beschreven in de literatuur. De invloed van de doelfunctie op de geschatte parameters is daarom nog onbekend. De keuze van de gebruikte momenten zal in dezelfde lijn liggen en wordt later besproken (Hoofdstuk 9). Tabel 4.1 geeft de gebruikte momenten voor de doelfuncties van de respectievelijke modellen die worden gebruikt bij de kalibratie weer (Verhoest *et al.*, 1997).

**Tabel 4.1:** Gebruikte momenten en aggregatieniveau voor de doelfunctie van de verschillende Bartlett-Lewis modellen.

	Gemiddelde	Variantie	Covariantie	ZDP
OBL	1 u	10 min en 24 u	10 min en 24 u	-
MBL	10 min	10 min	10 min en 24 u	10 min en 24 u
MBLG	10 min	10 min en 24 u	10 min en 24 u	10 min en 24 u

## 4.4 Parameterruimte

Bij de optimalisatie van de besproken modellen aan de hand van zowel SIMPSA als PSO moet aan het algoritme door de gebruiker duidelijk gemaakt worden waar de zoektocht naar de meest geschikte parameters mag plaatsvinden. Met andere woorden dienen de onder- en bovengrenzen van de parameters meegegeven te worden. Indien geen voorkennis voorhanden is omtrent het globale minimum van de doelfunctie kunnen de grenzen initieel best voldoende groot gekozen worden. Op die manier is de kans dat het optimum zich buiten de gekozen grenzen bevindt klein. Is er wel *a priori* kennis over het optimum, bijvoorbeeld resultaten van eerder uitgevoerde kalibraties, dan kunnen deze waarden als richtwaarden gebruikt worden om de grenzen af te bakenen, veronderstellend dat deze resultaten dicht bij het globaal optimum liggen.

De ondergrens van de parameters is omwille van de aard van de parameters gelijk aan nul.

De parameters kunnen enkel positieve waarden aannemen. De keuze van de bovengrenzen kan beredeneerd worden aan de hand van de fysische interpretatie van de parameters. Het doel van de Bartlett-Lewis modellen is om de structuur van het regenproces zodanig te reduceren dat het kan worden beschreven aan de hand van slechts enkele parameters. Zoals bij elk model zijn vereenvoudiging en idealisatie uiteraard onvermijdelijk. Daarom beschrijven de parameters van de beschouwde Bartlett-Lewis modellen geen gedetailleerde fysische processen, maar kan er niettemin een fysische interpretatie aan gegeven worden (Rodriguez-Iturbe *et al.*, 1987a). De keuze van de gebruikte grenzen en de interpretatie van de bekomen parameters dienen dus vanuit dit opzicht benaderd te worden. In Hoofdstuk 5 wordt dieper ingegaan op de keuze van de parametergrenzen.

## 4.5 Simplex-Simulated Annealing

Het Simplex-Simulated Annealing (SIMPSA) algoritme is een hybride tussen de Simplex methode van Nelder & Mead (1965) en Simulated Annealing (Kirkpatrick *et al.*, 1983). De algoritmen die aan de basis liggen van SIMPSA worden bijgevolg eerst apart verduidelijkt, om vervolgens de stap te maken naar de beschrijving van het algoritme dat het beste van beide algoritmen combineert.

In Sectie 4.5.1 wordt de werking van Simulated Annealing toegelicht, terwijl in Sectie 4.5.2 dieper wordt ingegaan op de opbouw van de Simplex methode. Daaropvolgend wordt in Sectie 4.5.3 SIMPSA eerst theoretisch besproken, gevolgd door een praktische bespreking van de implementatie voor de Bartlett-Lewis modellen.

### 4.5.1 Simulated Annealing

*Annealing* is een thermodynamisch proces waarbij een vaste stof door verhitting gesmolten wordt en vervolgens traag wordt afgekoeld om het te laten kristalliseren in een toestand met minimale vrije energie. Indien de snelheid waarmee de stof wordt afgekoeld niet nauwlettend in de gaten wordt gehouden of wanneer de initiële temperatuur te laag is, kan niet bij elke temperatuur een evenwicht bereikt worden, wat resulteert in een suboptimale schikking van de atomen. Dergelijke lokale imperfecties in de structuur van de stof beïnvloeden de sterkte en hardheid in negatieve zin. Thermisch evenwicht bij een bepaalde temperatuur wordt gekarakteriseerd door een Boltzmann verdelingsfunctie van de energietoestanden. Onder deze voorwaarden kan bij eender welke temperatuur een transitie plaatsvinden van een lage naar een hoge energietoestand. Zulke transities worden verondersteld verantwoordelijk te zijn voor het bereiken van een minimale energietoestand (Aarts & Van Laarhoven, 1987).

Metropolis *et al.* (1953) introduceerde een eenvoudig algoritme dat kan gebruikt worden om een efficiënte simulatie van een verzameling atomen in evenwicht bij een bepaalde temperatuur uit te voeren. Bij elke stap wordt een atoom willekeurig over een kleine afstand verplaatst en de resulterende energieverandering in het systeem  $\Delta E$  wordt berekend. Als  $\Delta E \leq 0$ , dan wordt de verplaatsing aanvaard en wordt de configuratie met het verplaatste atoom als uitgangspunt voor de volgende stap gebruikt. Indien  $\Delta E > 0$ , dan is de kans dat de verplaatsing plaatsgrijpt  $P(\Delta E) = \exp(-\Delta E/k_b T)$ , met  $T$  de huidige temperatuur van het systeem en  $k_b$  de Boltzmann constante. De kans dat een hogere energietoestand wordt aanvaard als een verbeterde ener-

gietoestand is hoog bij hoge temperatuur en nadert nul bij lage temperaturen. Hierbij wordt aangenomen dat het thermisch evenwicht bij elke temperatuur wordt bereikt.

Het algoritme dat door Metropolis *et al.* (1953) werd ontwikkeld, kan worden gebruikt als optimalisatietechniek (Kirkpatrick *et al.*, 1983; Kirkpatrick, 1984). De energietoestand van het systeem wordt vervangen door de uitkomst van de doelfunctie en de positionering van de atomen door de positie van de parameters in de parameterruimte. Net zoals de meeste iteratieve schema's beweegt het algoritme met kleine stappen doorheen de parameterruimte, maar de temperatuur zorgt ervoor dat het niet komt vast te zitten door beweging weg van een minimum mogelijk te maken. In het merendeel van de gevallen wordt Simulated Annealing toegepast op combinatorische minimalisatieproblemen. Een overzicht van dergelijke toepassingen wordt gegeven in Cardoso *et al.* (1996). Simulated annealing werd ook reeds toegepast bij de optimalisatie van continue parameterruimtes. Een bijkomend probleem dat zich hierbij stelt, in vergelijking met de optimalisatie van combinatorische problemen, is het genereren van alternatieve configuraties, uitgaande van de huidige configuratie. In het algemeen wordt gebruik gemaakt van willekeurige stappen rondom de huidige positie in de parameterruimte van de parametervector (Cardoso *et al.*, 1996). Het willekeurige karakter van deze werkwijze zorgt ervoor dat het gebruik van Simulated Annealing voor de optimalisatie van continue problemen in de meeste gevallen ondermaats is in vergelijking met andere methoden. Deze willekeurige zoektocht kan echter vermeden worden door het stochastische Simulated Annealing te combineren met de Simplex methode van Nelder & Mead (1965) (Press & Teukolsky, 1991). Deze aanpak wordt verder besproken in Sectie 4.5.3.

#### 4.5.2 Simplex methode

De Simplex methode is gebaseerd op een idee van Spendley *et al.* (1962) voor het opsporen van optimale operatie-omstandigheden, door het evalueren van de output van een systeem op een stel van punten die een simplex vormen in de parameterruimte, en het continu vormen van nieuwe simplices door een punt te reflecteren in het hypervlak van de overige punten. Dit idee is duidelijk toepasbaar voor het minimaliseren van wiskundige functies (Nelder & Mead, 1965). Het simplex beweegt zich hierbij zelfstandig doorheen de parameterruimte door een opeenvolging van *reflectie*, *contractie* en *expansie*.

Veronderstel dat een functie met  $D$  variabelen geoptimaliseerd dient te worden, zonder dat er restricties worden opgelegd.  $\vec{x}_0, \vec{x}_1, \dots, \vec{x}_D$  zijn dan de  $(D+1)$  punten in de  $D$ -dimensionale ruimte die het huidige simplex definiëren. De uitkomst van de doelfunctie in punt  $\vec{x}_i$  in de parameterruimte kan geschreven worden als  $y_i$ . Het punt met de laagste functiewaarde krijgt het suffix  $l$  ( $y_l = \min_i(y_i)$ ) en dat met de hoogste functiewaarde het suffix  $h$  ( $y_h = \max_i(y_i)$ ). Verder wordt  $\vec{\bar{x}}$  gedefinieerd als het zwaartepunt van de punten waarvoor geldt dat  $i \neq h$ . De afstand tussen  $\vec{x}_i$  en  $\vec{x}_j$  wordt voorgesteld aan de hand van  $d(\vec{x}_i, \vec{x}_j)$ .

In elke stap van het proces wordt  $\vec{x}_h$  vervangen door een nieuw punt. Hierbij wordt gebruik gemaakt van de drie vermelde operaties. Deze kunnen als volgt gedefinieerd worden. De reflectie van  $\vec{x}_h$  wordt voorgesteld door  $\vec{x}^*$  en diens coördinaten worden gevormd door de relatie

$$\vec{x}^* = (1 + \alpha)\vec{\bar{x}} - \alpha\vec{x}_h \quad (4.2)$$

waarbij  $\alpha$  een positieve constante is, de *reflectiecoëfficiënt*.  $\vec{x}^*$  ligt dus op de rechte die  $\vec{x}_h$  en  $\vec{\bar{x}}$  verbindt, aan de tegenovergestelde kant van  $\vec{\bar{x}}$  ten opzichte van  $\vec{x}_h$ , met  $d(\vec{x}^*, \vec{\bar{x}}) = \alpha d(\vec{x}_h, \vec{\bar{x}})$ .



Indien  $y_h > y^* > y_l$ , dan wordt  $\vec{x}_h$  vervangen door  $\vec{x}^*$  en begint het proces opnieuw met het nieuwe simplex.

Als  $y^* < y_l$ , met andere woorden de reflectie heeft een nieuw minimum gecreëerd, dan vindt een expansie van  $\vec{x}^*$  naar  $\vec{x}^{**}$  plaats aan de hand van

$$\vec{x}^{**} = \gamma \vec{x}^* + (1 - \gamma) \vec{\bar{x}} \quad (4.3)$$

De *expansiecoëfficiënt*  $\gamma$ , groter dan 1, is de verhouding van de afstand  $d(\vec{x}^{**}, \vec{\bar{x}})$  tot  $d(\vec{x}^*, \vec{\bar{x}})$ . Als  $y^{**} < y_l$ , wordt  $\vec{x}_h$  vervangen door  $\vec{x}^{**}$  en wordt het proces opnieuw gestart. Indien echter  $y^{**} > y_l$ , dan wordt de expansie als mislukt beschouwd en wordt  $\vec{x}_h$  vervangen door  $\vec{x}^*$  alvorens te herstarten.

Als bij de reflectie van  $\vec{x}$  naar  $\vec{x}^*$  blijkt dat  $y^* > y_i$  voor alle  $i \neq h$ , i.e. het vervangen van  $\vec{x}$  door  $\vec{x}^*$  creëert een nieuw maximum  $y^*$ , dan wordt een nieuwe  $\vec{x}_h$  gelijkgesteld aan  $\vec{x}_h$  of  $\vec{x}^*$ , degene met de laagste functiewaarde, en wordt  $\vec{x}^{**}$  als volgt berekend.

$$\vec{x}^{**} = \beta \vec{x}_h + (1 - \beta) \vec{\bar{x}}. \quad (4.4)$$

De *contractiecoëfficiënt*  $\beta$  ligt tussen 0 en 1 en is de ratio van de afstand  $[\vec{x}^{**}, \vec{\bar{x}}]$  tot  $[\vec{x}, \vec{\bar{x}}]$ .  $\vec{x}^{**}$  vervangt  $\vec{x}_h$  en het proces begint opnieuw, tenzij  $y^{**} > \min(y_h, y^*)$ . In dat geval is het resultaat van de contractie slechter dan de beste van  $\vec{x}_h$  en  $\vec{x}^*$ . In het geval van dergelijke mislukte contractie worden alle  $\vec{x}_i$  vervangen door  $(\vec{x}_i + \vec{x}_l)/2$  en wordt het proces herstart.

Het Nelder & Mead (1965) Simplex algoritme verschilt van eerder gepubliceerde algoritmen omdat het geen gebruik maakt van gradiënten (eerste orde afgeleiden) of van kwadratische vormen (tweede orde afgeleiden). De eenvoudige methodiek en beperkte veronderstellingen omtrent de doelfunctie maken dat het algoritme minder snel convergeert naar een lokaal minimum. De Simplex methode werd dan ook reeds veelvuldig gebruikt en uitgebreid bestudeerd.

### 4.5.3 SIMPSA

Het optimalisatie-algoritme dat in dit werk wordt gebruikt is een combinatie van Simulated Annealing en de Simplex methode. Simulated annealing maakt gebruik van willekeurige nieuwe parametercombinaties in zijn zoektocht naar een optimum. Om de zoektocht naar optimale parameters meer gestructureerd en gericht te laten verlopen wordt gebruik gemaakt van de Simplex methode. Het algoritme kan dus gezien worden als een Simplex methode waarbij het simplex niet steeds in de richting van een minimum evolueert. Er bestaat bij elke operatie van het simplex een kans dat een 'slechtere' zet wordt aanvaard, zoals dat ook het geval is bij Simulated Annealing. Dit zou ervoor zorgen dat het algoritme nog beter in staat is om lokale optima te vermijden. Het gebruik van zulke combinatie, Simplex-Simulated Annealing, werd het eerst voorgesteld door Press & Teukolsky (1991). Het werd later door Cardoso *et al.* (1996) aangepast en wordt hieronder aldus beschreven.

Zoals eerder vermeld is bij het Metropolis algoritme de kans dat een nieuwe configuratie of oplossing aanvaard wordt afhankelijk van de energietoestand of fitness van de nieuwe configuratie ten opzichte van de huidige. Indien de nieuwe configuratie een betere oplossing is ( $\Delta E < 0$ ), dan wordt deze altijd aanvaard. Als echter de nieuwe configuratie slechter is dan de oude ( $\Delta E \geq 0$ ), bestaat toch een kans  $P(\Delta E) = \exp(-\Delta E/k_b T)$  dat deze aanvaard wordt. In het originele Metropolis algoritme wordt verondersteld dat bij elke temperatuur thermisch evenwicht

bereikt wordt. Omdat deze veronderstelling leidt tot een te grote computationele last wordt ze door Cardoso *et al.* (1996) niet in acht genomen. De zogenaamde non-equilibrium variant leidt eveneens tot goede resultaten, zij het op een efficiëntere en dus snellere manier. In onderstaande tekst worden de belangrijkste aspecten van deze door Cardoso *et al.* (1996) beschreven non-equilibrium variant van het SIMPSA-algoritme besproken.

#### 4.5.3.1 Initiële annealing temperatuur

De initiële annealing temperatuur bepaalt de kans dat initieel minder goede oplossingen worden aanvaard. De initiële temperatuur moet dus zo gekozen worden dat bij het begin van de optimalisatie de kans op het aanvaarden van een verkeerde beweging groot is, om zo stelselmatig te dalen als het proces vordert. Om dit te bekomen wordt de initiële temperatuur berekend aan de hand van Vergelijking 4.5 (Cardoso *et al.*, 1996). Hiertoe wordt eerst een preliminaire temperatuurslus doorlopen. Tijdens deze preliminaire lus is het gewenst dat alle bewegingen aanvaard worden. De preliminaire temperatuur wordt daarom heel hoog genomen door de absolute waarde van de doelfunctie, bekomen in het startpunt, te vermenigvuldigen met een grote positieve waarde. Daaropvolgend wordt de cyclus van  $m_0$  bewegingen doorlopen, corresponderend met een volledige Metropolis cyclus. Tijdens deze cyclus wordt het aantal succesvolle bewegingen  $m_1$  ( $E^k < E^{k-1}$ , voor  $k = 2$  tot  $m_0$ ) en het aantal niet-succesvolle bewegingen  $m_2$  ( $E^k \geq E^{k-1}$ ) opgeslagen. Daarnaast wordt de gemiddelde verhoging van de fitness voor de  $m_2$  niet-succesvolle bewegingen ( $\Delta f^+$ ) berekend. Aan de hand hiervan wordt vervolgens de initiële temperatuur  $T_i$  voor het verdere verloop van de optimalisatie berekend:

$$X = \frac{m_1 + m_2 \exp\left(\frac{-\Delta f^+}{T_i}\right)}{m_1 + m_2} \quad (4.5)$$

waarbij  $X$  de aanvaardingsratio is. In het verdere verloop van de optimalisatie is dus de initiële kans op een verkeerde beweging gelijk aan  $X$ . Het koelschema vangt dan aan met de hieruit berekende temperatuurswaarde.

#### 4.5.3.2 Koelschema

Het koelschema dat gehanteerd wordt, is gebaseerd op dat van Aarts & Van Laarhoven (1985).

$$T^{j+1} = T^j \cdot c = \frac{T^j}{1 + \frac{T^j \cdot \ln(1+\delta)}{3\sigma}} \quad (4.6)$$

hierbij zijn  $\delta$  en  $\sigma$  trajectorie parameters,  $j$  de huidige iteratie en  $c$  de *koelfactor*. De parameter  $\delta$  controleert de koelsnelheid. Kleine waarden ( $< 1$ ) resulteren in een trage convergentie terwijl grotere waarden ( $> 1$ ) zorgen voor convergentie naar slechte lokale optima. Deze parameter is dus van cruciaal belang en wordt in het kader van dit werk dan ook afgesteld in functie van het gebruikte model. De tweede parameter  $\sigma$  is de standaardafwijking van de fitness van alle configuraties bij de huidige temperatuur  $T^j$ .

Zoals reeds vermeld, wordt gebruik gemaakt van de non-equilibrium variant van het algoritme. Volgens Press & Teukolsky (1991) dient bij elke temperatuur een vast aantal iteraties doorlopen worden vooraleer de temperatuur kan worden verlaagd. Bij de non-equilibrium variant gebeurt dit alleen tijdens de eerste cyclus, omdat dan een geschikte initiële temperatuur moet worden

berekend. In de daaropvolgende stappen wordt de temperatuur verlaagd zodra een betere oplossing gevonden wordt. Gezien er dus nood is aan een representatief punt om deze verbetering te beschrijven wordt gebruik gemaakt van het globale zwaartepunt, zonder uitsluiting van het slechtste punt (Cardoso *et al.*, 1996).

#### 4.5.3.3 Het initiële simplex

Voor een  $D$ -dimensioneel probleem wordt een simplex ontworpen van  $D + 1$  punten (Nelder & Mead, 1965). Aangezien bij aanvang van het zoekproces slechts een enkel startpunt ( $\vec{x}_0$ ) wordt opgegeven door de gebruiker dienen de overige punten nog berekend te worden. De resterende  $D$  punten worden berekend door telkens het startpunt in een bepaalde dimensie aan te passen. In het te berekenen punt  $\vec{x}_i$ , met  $i = 1, \dots, D$ , wordt telkens de  $a$ -de dimensie ( $a = 1, \dots, D$ ) aangepast ten opzichte van  $\vec{x}_0$  aan de hand van de onderstaande vergelijking.

$$x_{i,a} = x_{0,a} + (0.5 - \text{rand}) \times K^j \times (UB_a - LB_a), \quad (4.7)$$

waarbij  $\vec{x}_{i,a}$  de  $a$ -de dimensie van punt  $\vec{x}_i$ ,  $LB_a$  de ondergrens en  $UB_a$  de bovengrens van de  $a$ -de dimensie,  $\text{rand}$  een willekeurig getal tussen 0 en 1,  $K^j$  is een variabele factor (zie Sectie 4.5.3.5) en  $j$  is de huidige iteratie. De keuze van het startpunt kan berusten op eerder gepubliceerde kalibratieresultaten. Indien dit niet mogelijk is, kan het startpunt willekeurig bepaald worden.

#### 4.5.3.4 Alternatieve configuraties

De alternatieve configuraties zijn nieuwe oplossingen voor het optimalisatieprobleem. Deze worden gegenereerd op basis van de Simplex methode van Nelder & Mead (1965). Press & Teukolsky (1991) voegen een positieve, logaritmisches verdeelde variabele, proportioneel tot de controletemperatuur  $T$ , toe aan de functiewaarde geassocieerd met elk punt van het simplex. Op diezelfde manier wordt een gelijkaardige willekeurige variabele van het mogelijke nieuwe punt afgetrokken. Deze schema's worden voorgesteld door de volgende relaties :

$$f(\vec{x}_k)_{\text{perturbed}} = f(\vec{x}_k) - T \times \ln(\text{rand}); \quad k = 1, \dots, D + 1 \quad (4.8)$$

$$f(\vec{x}_{\text{new}})_{\text{perturbed}} = f(\vec{x}_{\text{new}}) + T \times \ln(\text{rand}) \quad (4.9)$$

met  $f(\vec{x}_k)$  de functiewaarde van punt  $k$ ,  $f(\vec{x}_{\text{new}})$  de functiewaarde op het nieuwe punt en  $f(\vec{x})_{\text{perturbed}}$  de geperturbeerde functiewaarde. Voor een minimalisatieprobleem worden de  $D + 1$  punten geperturbeerd richting hogere functiewaarden volgens Vergelijking 4.8 en het volgens de Simplex methode nieuwe punt wordt geperturbeerd naar een lagere functiewaarde, volgens Vergelijking 4.9. Indien het nieuwe punt correspondeert met een lagere functiewaarde wordt deze nieuwe configuratie aanvaard. Indien het nieuwe punt echter overeenkomt met een hogere functiewaarde bestaat de kans dat deze nieuwe configuratie aanvaard wordt, afhankelijk van de relatieve functiewaarden van de geperturbeerde waarden (Press & Teukolsky, 1991). Dit is het principe achter het Metropolis algoritme. Op deze manier wordt het mogelijk dat een minder goede oplossing toch aanvaard wordt. Deze benadering heeft bovendien het voordeel dat als  $T \rightarrow 0$ , het algoritme gereduceerd wordt tot de Simplex methode (Cardoso *et al.*, 1996).

#### 4.5.3.5 Grensoverschrijdend gedrag

Voor het omgaan met simplices die één of meerdere punten bevatten die buiten de parametergrenzen vallen bestaan verschillende methoden. Een mogelijke manier is om de parameters die buiten de gedefinieerde grenzen vallen te vervangen door de grenswaarde (Box, 1965). Deze methode leidt echter tot instabiele simplices die de optimalisatie belemmeren (Cardoso *et al.*, 1996). Anderzijds kunnen de punten van het simplex die niet voldoen aan de voorwaarden vervangen worden door willekeurig gekozen nieuwe punten, vertrekkend van het beste punt van het simplex en gebruik makende van Vergelijking 4.7. Om de convergentie te verzekeren varieert de factor  $K^j$  in functie van de iteratie  $j$ , volgens het Aarts & Van Laarhoven (1985) schema (Vergelijking 4.6). In dit werk wordt een punt dat zich buiten de parametergrenzen bevindt bestraft door er een grote waarde aan toe te kennen, waardoor  $k^j$  niet gebruikt wordt en telkens gelijk is aan 1, Vergelijking 4.7 wordt immers enkel aangewend voor het aanmaken van de initiële simplices. Hierdoor zal elke overtreding van de grenzen gevolgd worden door een contractiestap die het simplex uiteindelijk binnen de grenzen zal houden (Nelder & Mead, 1965).

#### 4.5.3.6 Stopcriteria

Voor het stopzetten van de optimalisatieprocedure wordt gebruik gemaakt van een combinatie van drie stopcriteria. Het algoritme wordt stopgezet indien een maximaal aantal functie-evaluaties (FEs) overschreden wordt. Daarnaast wordt het algoritme stopgezet indien het verschil tussen de coördinaten van de punten van het simplex kleiner wordt dan een gespecificeerde tolerantie (TOLX). Tenslotte wordt ook een tolerantie opgelegd voor het verschil in functiewaarden tussen de verschillende punten van het simplex. Is dit verschil kleiner dan de tolerantie (TOLFUN), dan wordt het algoritme beëindigd.

#### 4.5.3.7 Implementatie

Alvorens SIMPSA kan worden ingezet voor de succesvolle kalibratie van een bepaalde doelfunctie, dient voldoende aandacht gespendeerd te worden aan de instellingen van het algoritme. Elke doelfunctie is immers anders en het zonder meer gebruik maken van de standaardinstellingen zou leiden tot een ondermaats resultaat.

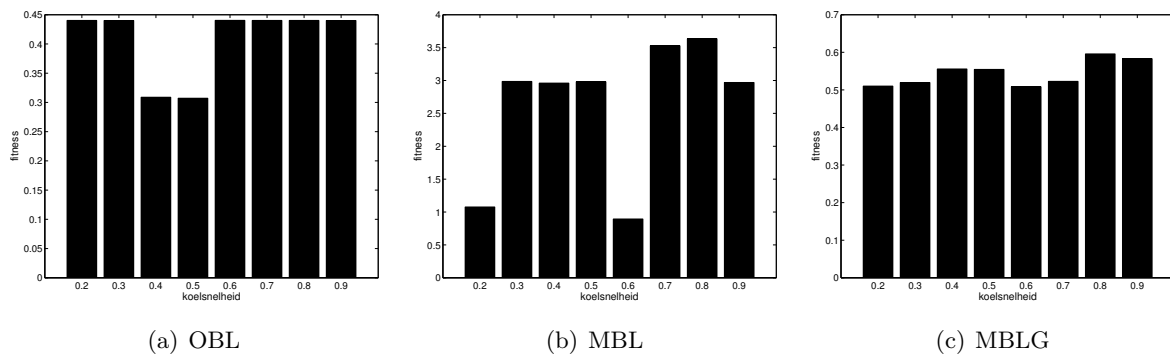
Zoals reeds vermeld wordt de initiële temperatuur bepaald door het doorlopen van een preliminaire temperatuurslus. De initiële temperatuur wordt dan zo berekend dat de kans op een beweging in de 'verkeerde' richting gelijk is aan de aanvaardingsratio  $X$  (Vergelijking 4.5), welke meestal gelijkgesteld wordt aan 95%. Hierbij werd ook vermeld dat bij aanvang van deze preliminaire temperatuurslus een grote temperatuur wordt ingesteld waardoor alle bewegingen mogelijk zijn. Deze grote waarde wordt bekomen door het product van de functiewaarde van het startpunt te vermenigvuldigen met  $10^5$  (Cardoso *et al.*, 1996). Indien dit product kleiner is dan  $10^5$  dan wordt de preliminaire temperatuur gelijkgesteld aan  $10^5$ . Het aantal iteraties  $m_0$  dat tijdens deze eerste lus wordt doorlopen wordt gelijkgesteld aan 50 (Donckels, 2009).

In het koelschema dat na de preliminaire lus doorlopen wordt, dienen ook enkele parameters ingesteld te worden. Vooreerst moet vermeld worden dat bij de overige temperatuurslussen telkens maximaal 10 iteraties doorlopen worden. Dit kan minder zijn aangezien overgegaan wordt naar een nieuwe temperatuur van zodra een betere oplossing gevonden wordt. In het koelschema zijn twee parameters van belang (zie Vergelijking 4.6). De koelsnelheid  $\delta$  en de koelfactor  $c$ .

Deze laatste wordt aan banden gelegd door het instellen van een maximale koelfactor welke standaard gelijkgesteld wordt aan 0.9. De temperatuur zal met andere woorden ten opzichte van de vorige iteratie telkens met minimaal 10% dalen. Voor de keuze van koelsnelheid worden verschillende mogelijkheden uitgetest. Door telkens een beperkt aantal iteraties uit te voeren voor verschillende koelsnelheden kan op objectieve wijze nagegaan worden welke koelsnelheid het meest geschikt is voor een bepaalde doelfunctie. De koelsnelheid  $\delta$  wordt getest in het interval  $[0.2, 0.9]$  met stappen van 0.1. Bij elke koelsnelheid worden vijf herhalingen uitgevoerd met per herhaling 100 iteraties. Vervolgens wordt de mediaan van de herhalingen berekend en het resultaat uitgezet in een grafiek. De koelsnelheid die na 100 iteraties de beste functiewaarde oplevert wordt weerhouden en gebruikt voor verdere kalibraties.

Figuur 4.1 toont deze resultaten voor de verschillende modellen. In Figuur 4.1(a) is te zien dat voor het OBL model de beste resultaten bekomen worden bij een koelsnelheid gelijk aan 0.5. Voor MBL wordt  $\delta$  vastgelegd op 0.6 (Figuur 4.1(b)). Figuur 4.1(c) toont dat de beste keuze van  $\delta$  voor het MBLG model ook hier 0.6 is.

Voorts dienen drie stopcriteria, die bepalen wanneer het algoritme zijn zoekproces naar een minimum afbreekt, bepaald te worden. De bepaling van deze criteria zijn cruciaal voor het uiteindelijke resultaat. Indien de keuze van zowel TOLX, TOLFUN en het maximale aantal iteraties té klein wordt gekozen, wordt het algoritme te vroeg stopgezet waardoor de kans groot is dat de optimale oplossing niet bereikt wordt. Door een methode van *trial-and-error* werden TOLFUN en TOLX vastgelegd op respectievelijk  $10^{-7}$  en  $10^{-10}$ , voor alle modellen.



**Figuur 4.1:** Invloed van de koelsnelheid op de *fitness* voor de verschillende Bartlett-Lewis modellen.

## 4.6 Particle Swarm Optimization

### 4.6.1 Inleiding

Het bereiken van complexe doelstellingen met behulp van eenvoudige individuen is een inspiratiebron voor sommige artificiële intelligentietechnieken, gegroepeerd onder de noemer *Swarm Intelligence*. Eén van deze technieken is Particle Swarm Optimization (PSO), geïntroduceerd voor de simulatie van sociaal gedrag en later voorgesteld als optimalisatiemethode (Kennedy & Eberhart, 1995). Deze optimalisatiemethode is gebaseerd op het gedrag van kuddedieren zonder een duidelijke leidende figuur. Ondanks de afwezigheid van een leider zijn bijvoorbeeld vogelzwermen, visscholen, schapen en dergelijke in staat zich te organiseren en zich in groep te verplaatsen aan de hand van lokale interacties tussen de individuen. Deze lokale interacties

leiden tot een complex collectief gedrag dat de kudde of school in staat stelt een bepaald doel te bereiken, bijvoorbeeld het zoeken van voedsel. De simulatie van dit gedrag en het aanpassen van de doelstelling aan het gestelde probleem zorgt er voor dat deze methode uiterst geschikt is voor het oplossen van een optimalisatievraagstuk (Engelbrecht, 2006).

In volgende secties wordt eerst het algoritme beschreven naar analogie met de beschrijving van Engelbrecht (2006) (Sectie 4.6.2), waarna verschillende algemene en meer specifieke aspecten van de implementatie besproken worden (Sectie 4.6.3).

#### 4.6.2 Theoretische aspecten

Een PSO algoritme is opgebouwd uit een zwerm van  $N$  partikels, waarbij elk partikel een mogelijke oplossing van het probleem voorstelt. Het partikel beweegt doorheen de multidimensionele zoekruimte, op zoek naar het globale optimum. Hierbij wordt de positie van het partikel telkens bijgestuurd aan de hand van zijn reeds afgelegde traject en dat van de volledige zwerm. In een  $D$ -dimensionele zoekruimte worden de positie en de snelheid van een partikel  $i$ , met  $i = 1, \dots, N$ , voorgesteld door een  $D$ -dimensionele vector  $\vec{x}_i = (x_{i1}, x_{i2}, \dots, x_{iD})$  en  $\vec{v}_i = (v_{i1}, v_{i2}, \dots, v_{iD})$ , respectievelijk. De positie  $\vec{x}_i$  van een partikel wordt aangepast door de snelheid  $\vec{v}_i$  op te tellen bij de huidige positie :

$$\vec{x}_i(t+1) = \vec{x}_i(t) + \vec{v}_i(t+1) \quad (4.10)$$

waarbij  $t$  en  $t+1$  de iteratiestap uitdrukken.

De snelheidsvector drijft het optimalisatieproces en drukt zowel de persoonlijk verworven kennis als de sociaal uitgewisselde informatie tussen de individuen uit. De persoonlijke ervaring wordt doorgaans de *cognitieve component* genoemd, welke proportioneel is aan de afstand van het partikel tot de beste positie die dat partikel tot dan toe bezocht heeft (het partikel's *persoonlijk beste* positie). De sociale kennis wordt de *sociale component* van de snelheidsvergelijking genoemd.

Op basis van de manier waarop informatie wordt uitgewisseld wordt een onderscheid gemaakt tussen *global best* (*gbest*) PSO en *local best* (*lbest*) PSO. Bij *gbest* PSO wordt informatie uitgewisseld tussen alle partikels in de zwerm. Bij *lbest* PSO wordt enkel informatie uitgewisseld tussen naburige partikels. De definitie van de *neighbourhood* van de partikels weerspiegelt zich in de formulering van de snelheidsvergelijking. Aangezien in dit werk het *gbest* algoritme wordt gebruikt, wordt in de volgende paragrafen enkel *gbest* PSO beschreven.

Voor het *gbest* PSO algoritme is de *neighbourhood* voor elk partikel de volledige zwerm. De sociale component van de snelheidsvergelijking beschrijft de informatie die door de gehele zwerm werd verworven. In dat geval is de sociale informatie de beste positie gevonden door de zwerm,  $\vec{p}_g$ . De cognitieve component houdt daarentegen rekening met de persoonlijke ervaring van een partikel. Deze persoonlijke ervaring, en meer bepaald de tot dan toe best bezochte positie  $\vec{p}_i$  van partikel  $i$ , wordt samen met de sociale informatie gebruikt om de snelheid van het partikel tijdens de volgende iteratie te bepalen. De snelheid van partikel  $i$  wordt dan als volgt berekend:

$$\begin{aligned} \vec{v}_i(t+1) = & w \cdot \vec{v}_i(t) + c_1 \cdot r_1(t) \cdot [\vec{p}_i(t) - \vec{x}_i(t)] \\ & + c_2 \cdot r_2(t) \cdot [\vec{p}_g(t) - \vec{x}_i(t)], \end{aligned} \quad (4.11)$$

met  $\vec{v}_i(t)$  de snelheidsvector van partikel  $i$  op tijdstap  $t$ ,  $\vec{x}_i(t)$  de positie van het partikel  $i$  op tijdstap  $t$ ,  $c_1$  en  $c_2$  zijn positieve acceleratieconstanten die gebruikt worden om de invloed van de cognitieve en sociale componenten respectievelijk te schalen,  $r_1(t)$  en  $r_2(t)$  zijn uniform verdeelde

willekeurige parameters tussen 0 en 1 en worden bij elke iteratiestap gegenereerd en  $w$  is het inertiegewicht. Het inertiegewicht  $w$  kan een positieve constante zijn of een positieve lineaire of niet-lineaire functie van de tijd. Het inertiegewicht speelt een rol in het balanceren tussen een globale en lokale zoektocht (Shi & Eberhart, 1998). In Sectie 4.6.3 wordt dieper ingegaan op de betekenis van deze parameters.

De persoonlijk beste positie  $\vec{p}_i$  van partikel  $i$  is, zoals reeds vermeld, de beste positie die het partikel bezocht heeft sinds de eerste tijdstap. De persoonlijk beste positie op de volgende tijdstap,  $t + 1$ , wordt als volgt berekend.

$$\vec{p}_i(t+1) = \begin{cases} \vec{p}_i(t) & \text{als } f(\vec{x}_i(t+1)) \geq f(\vec{p}_i(t)) \\ \vec{x}_i(t+1) & \text{als } f(\vec{x}_i(t+1)) < f(\vec{p}_i(t)) \end{cases} \quad (4.12)$$

waarbij  $f$  de doelfunctie is. De doelfunctie drukt de performantie of kwaliteit van de positie van het partikel uit.

De globaal beste positie waarvoor de laagste waarde van de doelfunctie  $f$  bereikt werd, wordt gevonden als :

$$g = \arg \min_{i=1}^N [f(\vec{p}_i(t+1))], \quad (4.13)$$

met  $g$  de positie waarvoor  $f(\vec{p}_i(t+1))$  minimaal is. De globaal beste positie wordt dus gekozen als de beste van de persoonlijk beste posities  $\vec{p}$ .

### 4.6.3 Praktische aspecten

Vooraleer PSO kan gebruikt worden moet de gebruiker de parameterruimte afbakenen (Sectie 4.6.3.1), PSO-parameters kiezen (Sectie 4.6.3.2) en het stopcriterium vastleggen (Sectie 4.6.3.3). Hieronder worden eerst de algemene aspecten toegelicht terwijl vervolgens de specifieke implementie besproken wordt.

#### 4.6.3.1 Grensoverschrijdend gedrag

Omwille van het willekeurige karakter van PSO dient zoals reeds vermeld in Sectie 4.4 de parameterruimte waarin de partikels zich kunnen bewegen afgebakend te worden. De implementatie van PSO moet ook voorzien worden van enkele regels waaraan een partikel dat zich buiten de opgelegde grenzen van de parameterruimte bevindt wordt onderworpen. Er zijn verschillende manieren mogelijk om deze restricties in acht te nemen (Engelbrecht, 2006).

Een eerste optie is het negeren van de grensvoorwaarden. Hierbij bestaat de kans dat alle partikels buiten de zoekruimte worden getrokken. In bepaalde gevallen is dit niet problematisch maar in vele gevallen dient een (fysisch) aanvaardbare oplossing gezocht te worden binnen bepaalde grenzen van de parameterruimte. Om dit te bewerkstelligen wordt in bepaalde gevallen de snelheid van een partikel dat zich buiten de parameterruimte begeeft op nul gezet. Op die manier kan het partikel terug in de parameterruimte getrokken worden onder invloed van de persoonlijk beste en de globaal beste positie (Engelbrecht, 2006).

Partikels die zich buiten de zoekruimte begeven kunnen anderzijds toegestaan worden, zonder dat het gevaar op convergentie buiten de zoekruimte optreedt. Er wordt meer bepaald toegestaan dat partikels zich buiten de grenzen begeven, evenwel kan deze positie niet de persoonlijk beste positie worden (Engelbrecht, 2006).

Indien men echter niet wil dat de partikels hoe dan ook buiten de parametergrenzen terecht komen, kan geopteerd worden om elk partikel dat de grensvoorwaarden schendt te herinitialiseren binnen de grenzen. Dit kan door ze een willekeurige positie binnen de parameterruimte toe te kennen of door ze op de grens te plaatsen. Bij de willekeurige herinitialisatie kan eventueel de snelheid behouden worden. Bij het plaatsen op de grens zijn wederom verschillende mogelijkheden voorhanden zoals de snelheid op nul zetten of de snelheid te reflecteren. Voor een meer gedetailleerd overzicht van de verschillende mogelijkheden met betrekking tot de huidige en de volgende paragrafen wordt verwezen naar Engelbrecht (2006).

#### 4.6.3.2 PSO parameters

De resultaten van PSO zijn ook afhankelijk van de keuze van verschillende parameters. Meer bepaald de populatiegrootte  $N$ , de cognitieve parameter  $c_1$ , de sociale parameter  $c_2$  en het inertiegewicht  $w$ . Engelbrecht (2006) vond dat een goede waarde voor de populatiegrootte  $N=30$  is. De keuze van  $c_1$ ,  $c_2$  en  $w$  heeft een belangrijke invloed op de snelheid, convergentie en efficiëntie van het algoritme.

De acceleratiecoëfficiënten  $c_1$  en  $c_2$ , samen met de willekeurige parameters  $r_1$  en  $r_2$ , controleren de stochastische invloed van de cognitieve en de sociale component op de snelheid van een partikel (Engelbrecht, 2006). Indien  $c_1 = c_2 = 0$ , blijven partikels vliegen aan hun initiële snelheid totdat ze op de rand van de parameterruimte botsen (aangenomen dat  $w = 0$ ). In het geval dat  $c_1 > 0$  en  $c_2 = 0$  dan onderneemt elk partikel een onafhankelijke lokale zoektocht waarbij het de omgeving van zijn persoonlijk beste positie doorzoekt en deze vervangt indien een nieuwe positie beter bevonden wordt. Als anderzijds  $c_1 = 0$  en  $c_2 > 0$ , dan wordt de ganse zwerm aangetrokken door één punt,  $p_g$ . De zwerm gedraagt zich als een stochastische *hill-climber*.

De sterkte van PSO ligt in de exploratie/exploitatie-verhouding. Met andere woorden de verhouding tussen de invloed van de sociale component die instaat voor de exploratie van de zoekruimte, en de cognitieve component die leidt tot een meer specifieke exploitatie. De zoektocht is dan ook het meest gebalanceerd wanneer  $c_1 = c_2$ . In dat geval worden de partikels aangetrokken tot het gemiddelde van  $p_g$  en  $p_i$ . In de meeste gevallen wordt dan ook  $c_1 = c_2$  genomen, hoewel de verhouding tussen deze twee parameters probleemafhankelijk is.

Als  $c_1 \gg c_2$  wordt elk partikel veel meer door zijn persoonlijk beste positie aangetrokken dan door de globaal beste positie wat resulteert in lokale zoektochten met weinig kans op convergentie van de zwerm naar een globaal optimum.

Anderzijds zal wanneer  $c_2 \gg c_1$  zullen de partikels te veel aangetrokken worden door de globaal beste positie waardoor ze te snel convergeren naar een lokaal minimum. Voor unimodale problemen met een zwak reliëf in de doelfunctie zal een grote sociale component het meest efficiënt zijn terwijl voor doelfuncties met een ruw reliëf de optimalisatie meer baat zal hebben aan een grotere cognitieve component.

Ook de grootte van  $c_1$  en  $c_2$  is van belang. Lage waarden voor  $c_1$  en  $c_2$  vertalen zich in een trage vlucht waardoor partikels ruim de tijd hebben om gebieden ver weg van een optimum te verkennen alvorens ze uiteindelijk naar dat optimum toegetrokken worden. Grote waarden voor  $c_1$  en  $c_2$  veroorzaken meer snelheid, meer abrupte bewegingen naar of weg van goede regionen. Het inertiegewicht  $w$  werd door Shi & Eberhart (1998) geïntroduceerd als een mechanisme om de exploratie- en exploitatie-eigenschappen van de zwerm te controleren. Het inertiegewicht



bepaalt in welke mate de snelheid van een partikel beïnvloed wordt door diens voorafgaande snelheid. De waarde van  $w$  is zeer belangrijk voor het convergent gedrag van de zwerm. Als  $w \geq 1$  dan nemen bij elke tijdstap de snelheden toe waardoor de zwerm uiteindelijk divergeert. Partikels slagen er niet in van richting te veranderen om zo te convergeren naar veelbelovende gebieden. Indien  $w < 1$ , dan neemt de snelheid stelselmatig af totdat de snelheid nul bedraagt (afhankelijk van de cognitieve en sociale componenten). Met andere woorden zorgt een grote  $w$  voor meer exploratie met meer diversiteit terwijl een kleine  $w$  lokale exploitatie bevoordeelt. Hoe kleiner de waarde van  $w$ , des te groter de invloed van de cognitieve en de sociale component. De keuze van  $w$  is wederom probleemafhankelijk en gebeurt best in combinatie met de keuze van  $c_1$  en  $c_2$  omdat deze elkaar sterk beïnvloeden. Meestal worden  $c_1$ ,  $c_2$  en  $w$  op empirische wijze bepaald in functie van het gestelde probleem. Als leidraad in deze procedure kan beroep gedaan worden op de literatuur. De invloed van de keuze van deze parameters op de convergentie wordt immers in de literatuur uitvoerig beschreven voor zowel vereenvoudigde (Trelea, 2003) als voor het zoals hierboven beschreven PSO algoritme (Jiang *et al.*, 2007). Aan de hand daarvan kunnen parameters gezocht worden in het theoretisch convergentiedomein. De keuze van de parameters in dit domein zou er toe moeten leiden dat de zwerm uiteindelijk convergeert naar één bepaald punt.

#### 4.6.3.3 Stopcriterium

Het stopcriterium bepaalt wanneer het iteratieve proces wordt afgebroken. Verschillende benaderingen zijn mogelijk. In vele gevallen wordt een combinatie van verschillende stopcriteria gehanteerd. Vooraf worden dan bijvoorbeeld een maximaal aantal iteraties, een maximale kalibratieduur, een drempelwaarde voor de uitkomst van de doelfunctie, etc. gespecificeerd. Voor zwermgebaseerde optimalisatietechnieken zoals PSO kan ook gebruik gemaakt worden van een convergentiecriterium. Het algoritme wordt dan gestopt als een bepaald deel van de populatie reeds geconvergeerd is naar eenzelfde oplossing.

#### 4.6.3.4 Implementatie

In bovenstaande secties werden enkele algemene aspecten van de implementatie van het PSO-algoritme toegelicht. De specifieke toepassing van deze principes voor de kalibratie van de Bartlett-Lewis modellen wordt hieronder besproken.

**Grensoverschrijdend gedrag** Bij de kalibratie van de Bartlett-Lewis modellen wordt ervoor gekozen om een partikel dat zich buiten de grenzen van de parameterruimte begeeft terug te plaatsen op de grens en zijn snelheid te reflecteren.

**PSO parameters** De bepaling van de PSO-parameters gebeurt exhaustief (naar analogie met Scheerlinck *et al.* (2009)). De populatiegrootte  $N$  wordt hier niet bij betrokken. De populatie wordt steeds opgebouwd uit 30 partikels. Verder wordt geopteerd voor een constante inertiefactor  $w$ . Om parameters  $c_1$ ,  $c_2$  en  $w$  te bepalen wordt een exhaustieve zoektocht uitgevoerd voor de maand januari. Hierbij wordt voor verschillende combinaties van  $c_1$ ,  $c_2$  en  $w$  een kalibratie uitgevoerd, waarbij telkens 100 iteraties worden doorlopen. De parameter  $c_1$  varieert tussen 0.5 en 2.5,  $c_2$  tussen 0.5 en 2.5 en  $w$  tussen 0.2 en 1, met stappen van 0.2. Deze waarden werden

gekozen in het convergentiedomein van PSO (Jiang *et al.*, 2007).

Figuren 4.2, 4.3 en 4.4 geven het resultaat van deze exhaustieve parameterselectie weer voor OBL, MBL en MBLG, respectievelijk. Telkens wordt de invloed van de verschillende combinaties van  $c_1$ ,  $c_2$  en  $w$  op de fitness na 100 iteraties weergegeven. Op basis van deze figuren wordt voor elk model een set PSO parameters vastgelegd.

**OBL** Figuur 4.2(a) toont de invloed op de fitness na honderd iteraties voor verschillende combinaties van  $c_1$  en  $c_2$  bij een vaste inertie ( $w=0.4$ ). Uit Figuur 4.2(a) blijkt dat de laagste fitnesswaarden bekomen worden voor  $c_2 \in [0.9, 2.5]$  en  $c_1 \in [0.5, 2.5]$ . Een combinatie van een lage  $c_1$  en hoge  $c_2$  en vice versa leveren lage fitnesswaarden op. Om echter de kans op voortijdige convergentie in een lokaal minimum te beperken, wordt geopteerd de waarde  $c_2$  niet te hoog te kiezen. Figuur 4.2(b) toont het verloop van de fitness bij een vaste waarde voor  $c_2$  ( $c_2=1.5$ ), waaruit blijkt dat voornamelijk bij  $w \in [0.4, 0.6]$  en  $c_1 \in [0.5, 2.5]$  goede resultaten bekomen worden. Figuur 4.2(c) toont de evolutie van de fitness voor verschillende waarden van  $w$  en  $c_2$  bij vaste  $c_1$  ( $c_1=1.9$ ). De betere resultaten worden bekomen voor  $w \in [0.2, 0.4]$  en  $c_2 \in [1.1, 1.9]$ . Hoewel er geen absoluut beste parametercombinatie kon gevonden worden liggen alle combinaties in het convergentiedomein en kan aan de hand van voorgaande resultaten een weloverwogen keuze gemaakt worden. Voor OBL worden de parameters voor PSO vastgelegd op:  $c_1=1.9$ ,  $c_2=1.5$  en  $w=0.3$ .

**MBL** Figuur 4.3(a) toont de invloed op de fitness na honderd iteraties voor verschillende combinaties van  $c_1$  en  $c_2$  bij een vaste inertie ( $w = 0.4$ ). Uit Figuur 4.3(a) blijkt dat de laagste fitnesswaarden bekomen worden voor  $c_2 \in [0.9, 2.5]$  met  $c_1 \in [0.9, 2.5]$ . Ook hier wordt gekozen om de waarde van  $c_2$  niet te hoog te kiezen in combinatie met een lage  $c_1$ , om convergentie in lokale minima te vermijden. Figuur 4.3(b) toont het verloop van de fitness bij een vaste waarde voor  $c_2$  ( $c_2=1.7$ ), waaruit blijkt dat de beste fitness bekomen wordt bij  $w=0.4$  en  $c_1 \in [0.9, 2.1]$ . Figuur 4.3(c) toont de evolutie van de fitness voor verschillende waarden van  $w$  en  $c_2$  bij vaste  $c_1$  ( $c_1=1.8$ ). Hierbij worden de beste resultaten worden bekomen indien  $w \in [0.2, 0.6]$  en  $c_2 \in [1.5, 2.3]$ . Gezien de resultaten ook hier niet eenduidig zijn worden de volgende parameters gekozen:  $c_1=1.9$ ,  $c_2=1.8$  en  $w=0.3$ .

**MBLG** Figuur 4.4(a) toont de invloed op de fitness na honderd iteraties voor verschillende combinaties van  $c_1$  en  $c_2$  bij een vaste inertie ( $w=0.4$ ). Uit Figuur 4.4(a) blijkt dat de laagste fitnesswaarden bekomen worden voor  $c_2 \in [1.3, 2.3]$  met  $c_1 \in [1.5, 2.3]$ . Figuur 4.4(b) toont het verloop van de fitness bij een vaste waarde voor  $c_2$  ( $c_2=1.8$ ), waaruit blijkt dat de beste fitness bekomen wordt bij  $w=0.4$  en  $c_1 \in [1.5, 2.1]$ . Figuur 4.4(c) toont de evolutie van de fitness voor verschillende waarden van  $w$  en  $c_2$  bij vaste  $c_1$  ( $c_1=1.8$ ). Hierbij worden de beste resultaten worden bekomen indien  $w=0.4$  en  $c_2 \in [1.5, 2.5]$ . Volgende parameters worden derhalve gekozen:  $c_1=1.9$ ,  $c_2=1.4$  en  $w=0.4$ . Tabel 4.2 geeft een overzicht van de gekozen parameters voor de verschillende modellen.

**Stopcriterium** Voor de keuze van het meest geschikte stopcriterium voor PSO wordt een onderzoek uitgevoerd naar de impact van het stopcriterium op het uiteindelijke resultaat. Het eerste criterium houdt in dat het algoritme wordt stopgezet als minstens de helft van de popu-

**Tabel 4.2:** Parameters voor kalibratie met PSO.

model	$N$	$c_1$	$c_2$	$w$
OBL	30	1.9	1.5	0.3
MBL	30	1.9	1.8	0.3
MBLG	30	1.9	1.4	0.4

latie  $N$  samenvalt met de globaal beste oplossing. Indien met andere woorden de Euclidische afstand tussen de positie van de  $N/2$  beste partikels en  $\vec{p}_g$  kleiner is dan een gekozen tolerantie stopt het algoritme met zoeken. In dit geval wordt gekozen voor een tolerantie gelijk aan  $10^{-6}$ . Het tweede stopcriterium houdt bij hoelang de globaal beste positie onveranderd blijft. Indien  $\vec{p}_g$  gedurende 10 iteraties ter plekke blijft, wordt de optimalisatieprocedure een halt toegeroepen. Bij de analyse van deze twee criteria is het van belang om met verschillende zaken rekening te houden. In eerste instantie wordt nagegaan hoe goed het gemiddeld resultaat is bij een aantal herhalingen en hoe sterk de resultaten, bekomen door een bepaald stopcriterium, gespreid liggen. Indien deze gemiddelde prestaties gelijkaardig zijn, kan vervolgens rekening gehouden worden met het gemiddeld aantal benodigde FEs. Een stopcriterium dat sneller tot hetzelfde resultaat leidt is uiteraard te verkiezen boven zijn alternatief. Tenslotte kan beargumenteerd worden dat niet de gemiddelde resultaten van belang zijn maar eerder de gevonden minima. Want daar zijn we immers naar op zoek. Ook al leiden de twee criteria tot eenzelfde gemiddeld resultaat, toch zal de benadering die de parameterset met de laagste functiewaarde oplevert onze voorkeur genieten.

Om na te gaan welk van beide criteria in dit geval tot de meest wenselijke resultaten leidt, worden voor de drie modellen telkens 30 herhalingen van de kalibratie voor de maand januari uitgevoerd met elk van beide stopcriteria. De onder- en bovengrenzen voor de verschillende parameters worden hier ingesteld als  $[0,100]$ . Het resultaat van deze 30 herhalingen kan worden samengevat door het minimum en de mediaan van de bekomen functiewaarden en de standaardafwijking. Deze statistieken worden weergegeven in Tabel 4.3.

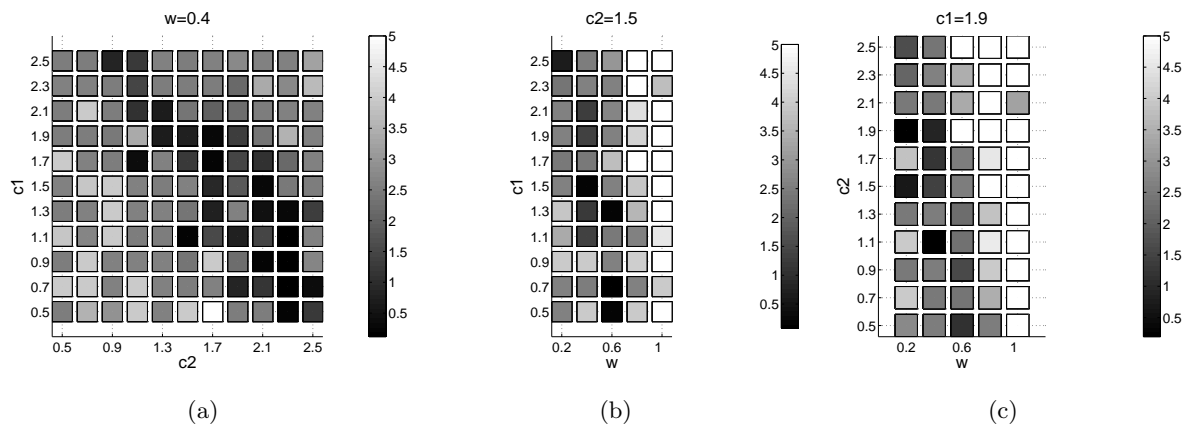
Gebaseerd op de fitnesswaarden die bekomen worden na 30 herhalingen kan een eenduidig besluit gevormd worden voor alle modellen. De mediaan van de resultaten van de twee criteria is voor alledrie de modellen gelijkaardig. Het eerste stopcriterium vertoont echter telkens een kleinere standaardafwijking. Het beste resultaat dat wordt voorgesteld door de parameterset met de minimale fitness wijst erop dat deze grotere spreiding bij het tweede criterium te wijten is aan een groter aandeel aan hogere functiewaarden. Deze minima zijn namelijk gelijkaardig voor OBL en MBLG en zelfs beter bij gebruik van stopcriterium 1 voor MBL. Het gebruik van het eerste stopcriterium leidt dus voor de drie modellen tot een robuuster resultaat, dit wil zeggen dat er minder kans bestaat dat het algoritme te vroeg stopt in een lokaal minimum. Dit is van belang bij verdere kalibratie omdat dus de kans op een beter resultaat groter is dan bij het gebruik van het tweede criterium. Tenslotte kan gezien worden dat het gemiddeld aantal benodigde FEs kleiner is als gebruik gemaakt wordt van het eerste criterium, behalve voor MBL. Het nadeel van het hoger aantal benodigde FEs bij MBL weegt echter niet op tegen de betere resultaten.

Er kan besloten worden dat bij kalibratie met PSO de voorkeur uitgaat naar het eerste stopcriterium. Tijdens het verdere verloop van dit werk wordt dan ook uitsluitend gebruik gemaakt

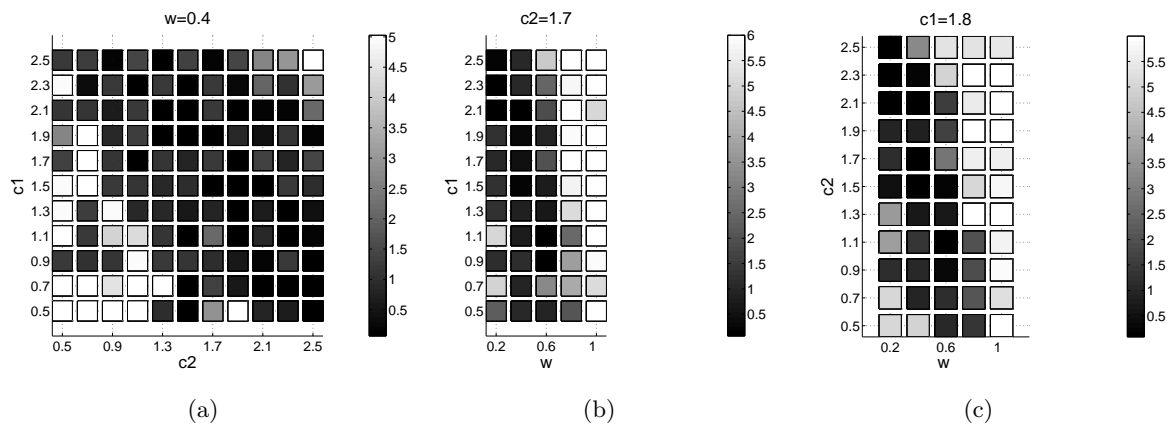
van dit criterium.

**Tabel 4.3:** Vergelijking tussen de resultaten van twee verschillende stopcriteria bij herhaaldelijke kalibratie met PSO voor de maand januari.

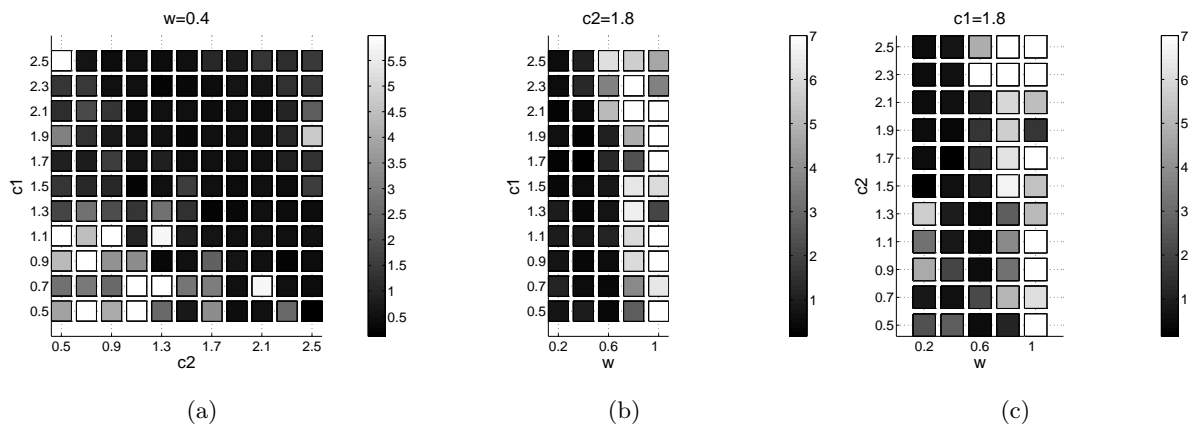
model	criterium	<i>fitness</i>			FEs
		minimum	mediaan	StDev	gemiddelde
OBL	1	2,00E-01	4,40E-01	1,21E-01	18744
	2	2,04E-01	4,40E-01	1,55E+00	21109
MBL	1	2,15E-03	6,43E-02	1,54E-02	114937
	2	1,91E-02	8,44E-02	1,32E+00	27533
MBLG	1	7,75E-02	1,24E-01	2,11E-01	25780
	2	7,75E-02	1,63E-01	1,89E+00	44179



**Figuur 4.2:** Invloed van de PSO parameters op de *fitness* voor OBL.



**Figuur 4.3:** Invloed van de PSO parameters op de *fitness* voor MBL.



**Figuur 4.4:** Invloed van de PSO parameters op de *fitness* voor MBLG.



# Hoofdstuk 5

## Vergelijking PSO en SIMPSA

### 5.1 Inleiding

In dit hoofdstuk wordt getracht een vergelijking te maken tussen de in Hoofdstuk 4 beschreven optimalisatietechnieken. Hierbij wordt gekeken naar de mate waarin de verschillende optimalisatietechnieken de doelfunctie minimaliseren en de efficiëntie waarmee dit gebeurt, alsook de computationele tijd die daarvoor vereist is. Hierbij wordt, om het overzicht te behouden en vanuit de veronderstelling dat de resultaten extrapoleerbaar zijn naar de overige maanden, enkel rekening gehouden met de maanden januari en juli. Voor beide methoden worden telkens 30 herhalingen van de kalibratie uitgevoerd. De bekomen resultaten kunnen daarop vergeleken worden aan de hand van een niet-parametrische statistische toets, zijnde de Kruskal-Wallis toets.

### 5.2 Performantiecriteriën

De verschillende optimalisatietechnieken worden beoordeeld aan de hand van de hieronder beschreven performantiecriteriën, namelijk accuraatheid, robuustheid en efficiëntie.

**Accuraatheid** Accuraatheid refereert naar de kwaliteit van een oplossing bekomen door een bepaalde optimalisatie. In het geval van PSO wordt dit op elk tijdstip  $t$  voorgesteld door de globaal beste positie, terwijl dit bij SIMPSA het zwaartepunt van het simplex op tijdstip  $t$  is. Indien geen *a priori* kennis voorhanden is omtrent het globale optimum van de doelfunctie  $f$  kan de accuraatheid van een optimalisatietechniek beoordeeld worden aan de hand van de *fitness* op tijdstip  $t$  (Engelbrecht, 2006), waarbij

$$\text{accuraatheid}(t) = f(\vec{x}(t)) = \sum_{i=1}^N W_i \left( \frac{M_i(\vec{x}(t))}{M'_i} - 1 \right)^2 \quad (5.1)$$

met  $\vec{x}(t)$  de parametervector met beste fitness op tijdstip  $t$ . Om verschillende algoritmen met elkaar te kunnen vergelijken wordt de accuraatheid van de oplossing na een bepaald aantal functie-evaluaties (FEs) beschouwd in tegenstelling tot iteraties. PSO is immers een populatiegebaseerd algoritme en SIMPSA niet. Door dus de bekomen fitness te vergelijken in samenspraak met het benodigd aantal FEs kan de accuraatheid van de methoden beoordeeld worden. Gebaseerd op deze definitie van Engelbrecht (2006) wordt hieronder de accuraatheid ingeschat als de mediaan van de fitness bij een bepaald aantal herhalingen van de kalibratieprocedure.

**Robuustheid** Hoe kleiner de variantie van een performantiecriterium gedurende een bepaald aantal FEs, des te robuuster of stabielere de optimalisatiemethode. Bij het kiezen van de meest geschikte methode is het namelijk van belang dat de methode betrouwbaar is en het resultaat dus niet veel afwijkt bij verschillende herhalingen. Op deze manier kan bij kalibratie met een beperkt aantal herhalingen met meer zekerheid gesteld worden dat een goede oplossing bereikt werd (Engelbrecht, 2006). De standaardafwijkingen van de *fitness* gedurende een bepaald aantal herhalingen kunnen dus gebruikt worden om hierover een uitspraak te doen.

**Efficiëntie** De efficiëntie van een algoritme kan uitgedrukt worden als het aantal FEs dat nodig is om een oplossing met een bepaalde accuraatheid  $\epsilon$  te bereiken (Engelbrecht, 2006). Als ook de tijd per functie-evaluatie berekend wordt kan de gemiddelde computationele efficiëntie van een bepaald aantal herhalingen vergeleken worden. De efficiëntie van de optimalisatietechnieken zal dus worden beoordeeld aan de hand van het gemiddeld aantal benodigde FEs om tot een oplossing te komen en de tijd die daarbij nodig is.

**Kruskal-Wallis toets** Naast het gebruik van beschrijvende statistieken om de prestatie van de twee algoritmen te beoordelen is er nood aan een meer objectieve maatstaf die uitsluitsel kan geven omtrent de prestatie van de onderlinge methoden. De Kruskal-Wallis toets, welke de niet-parametrische variant is van de *one way*-ANOVA gaat na of verschillende steekproeven kunnen beschouwd worden als zijnde afkomstig van eenzelfde populatie. Deze steekproeven zullen uiteraard onderlinge verschillen vertonen. De vraag is of deze verschillen toevallig ontstaan of dat ze veroorzaakt worden door het feit dat de steekproeven getrokken werden uit verschillende populaties (Kruskal & Wallis, 1952). De nulhypothese stelt dan ook dat de verschillende steekproeven afkomstig zijn uit eenzelfde populatie. Indien deze hypothese verworpen wordt bij een bepaald significantieniveau, kan gesteld worden dat minstens één van de steekproeven getrokken werd uit een verschillende populatie. Gezien hier een vergelijking wordt gemaakt tussen slechts twee methoden kan thans besloten worden dat de resultaten van de kalibratieprocedure van een bepaalde techniek significant verschilt van de andere. Hoewel deze test dus niet expliciet nagaat of ook de mediaan van de populaties verschillend is, wordt dit in de praktijk veelal zo geïnterpreteerd (Kruskal & Wallis, 1952). De methode is dus niet helemaal geschikt, maar geeft toch de beste basis om in samenspraak met visuele hulpmiddelen zoals een boxplot van de onderlinge resultaten een uitspraak te doen over de prestaties van de optimalisatietechnieken.

### 5.3 Resultaten

De vergelijking aan de hand van de hierboven aangegeven performantiecriteriën wordt in twee situaties opgedeeld. De twee situaties hebben met betrekking tot de afbakening van de parametergrenzen een ander uitgangspunt. In eerste instantie worden de grenzen (voor alle modellen en beide optimalisatiemethoden) voldoende groot genomen. Deze werkwijze zou ook gehanteerd worden bij de kalibratie van een model waarbij geen enkele kennis aanwezig is omtrent mogelijke optima. Op deze manier kan dus worden nagegaan welk van beide optimalisatietechnieken de andere overtreft in het aanduiden van een optimale regio voor de parameters. Aan de hand van de resultaten en gebruik makend van de literatuur worden vervolgens de grenzen verkleind om na te gaan welk van beide technieken het meest geschikt is voor het vinden van een optimum



wanneer wel *a priori* kennis voorhanden is. Hieronder worden beide benaderingen beschouwd als aparte gevalstudies (Benaderingen 1 en 2) waarbij telkens een besluit wordt geformuleerd aan de hand van de beschreven performantiecriteria en de finale parametersets.

### 5.3.1 Benadering 1

Omdat verondersteld wordt dat er geen *a priori* kennis bestaat over het optimum worden de parametergrenzen tijdens de kalibratie vastgelegd als  $[0,100]$ . Dit geldt voor alle parameters. Vervolgens worden 30 herhalingen van de kalibratieprocedure uitgevoerd. Hieronder wordt bijgevolg voor de verschillende performantiecriteria de mediaan van deze 30 herhalingen weergegeven. De gekozen parameters van de optimalisatietechnieken werden reeds beschreven in Hoofdstuk 4. De stopcriteria voor SIMPSA zijn voor alle modellen dezelfde (zie Hoofdstuk 4). Hetzelfde geldt voor PSO.

Tabellen 5.1 en 5.2 vatten de resultaten voor de kalibratie met zowel SIMPSA als PSO, voor de drie beschouwde modellen, samen. Hierbij werd bij beide algoritmes de kalibratieprocedure stopgezet omdat convergentie bereikt werd. Voor PSO betekent dit dat de helft van de populatie convergeert naar eenzelfde punt en voor SIMPSA dat het verschil in de functiewaarden van de verschillende punten van het simplex kleiner werd dan de gespecificeerde tolerantie (TOLFUN) (zie Hoofdstuk 4). In Tabellen 5.1 en 5.2 worden de resultaten van de kalibratie van de Bartlett-Lewis modellen door SIMPSA en PSO weergegeven voor respectievelijk januari en juli. Hierin worden de mediaan en standaardafwijking van de fitness gedurende 30 herhalingen weergegeven, alsook het gemiddeld aantal benodigde FEs. Figuur 5.1 geeft de boxplots van de fitness bekomen bij 30 herhalingen door de twee zoektechnieken voor de verschillende modellen en maanden. De bespreking van de resultaten voor de eerste benadering wordt hieronder gegeven voor elk model afzonderlijk.

**Tabel 5.1:** Benadering 1: minimum, mediaan en standaardafwijking (StDev) van de fitness en het gemiddeld aantal benodigde FEs bij 30 herhalingen van de kalibratie van de maand januari met SIMPSA en PSO van de verschillende Bartlett-Lewis modellen.

model	methode	<i>fitness</i>			FEs
		minimum	mediaan	StDev	gemiddelde
OBL	SIMPSA	0.177	0,440	0,100	13349
	PSO	0.200	0,440	0,121	18744
MBL	SIMPSA	0.067	0,112	1,287	17346
	PSO	0.002	0,064	0,015	114937
MBLG	SIMPSA	0.078	0,459	0,672	19622
	PSO	0.077	0,124	0,211	25780

**OBL** De resultaten van de kalibratie van het OBL model laten zich opmerken omdat hierbij door beide optimalisatietechnieken voor de maand januari (Tabel 5.1) en juli (Tabel 5.2) eenzelfde mediaan bekomen wordt. Dit komt door de aanwezigheid van een lokaal minimum waarnaar beide methoden veelvuldig convergeren. Deze mediaan kan dus als een aanwijzing gezien worden voor het feit dat beide methoden even accuraat zijn. Anderzijds kan gesteld worden dat beide methoden bij de kalibratie van OBL niet genoeg in staat zijn om het aanwezige lokale

**Tabel 5.2:** Benadering 1: minimum, mediaan en standaardafwijking (StDev) van de fitness en het gemiddeld aantal benodigde FEs bij 30 herhalingen van de kalibratie van de maand juli met SIMPSA en PSO van de verschillende Bartlett-Lewis modellen.

model	methode	<i>fitness</i>			FEs
		minimum	mediaan	StDev	gemiddelde
OBL	SIMPSA	0.4004	0,816	0,222	14015
	PSO	0.0097	0,816	0,423	20033
MBL	SIMPSA	0.1093	0,249	1,448	16446
	PSO	0.0003	0,099	0,062	64275
MBLG	SIMPSA	0.0972	0,856	1,003	19132
	PSO	0.0965	0,425	0,408	21939

minimum consequent te vermijden.

Rekening houdend met de standaardafwijking blijkt dat SIMPSA, zowel voor de maand januari als voor juli, net iets meer robuust is dan zijn tegenhanger PSO. Ook het gemiddeld aantal benodigde FEs is lager bij SIMPSA. SIMPSA blijkt dus op het eerste zicht robuuster en efficiënter. Om na te gaan hoe de resultaten van beide zoektechnieken gespreid zijn, wordt voor de maanden januari en juli een boxplot gemaakt ter vergelijking van de beide methoden. Figuren 5.1(a) en 5.1(b) geven deze boxplots weer voor de maanden januari en juli. Op basis van de boxplots is te zien dat het verschil in standaardafwijking voor de maand januari hoogstwaarschijnlijk te wijten is aan de uitschieter die aanwezig is bij PSO. Zonder deze zullen ook de standaardafwijkingen voor januari vermoedelijk in dezelfde lijn liggen wat impliceert dat het verschil tussen beide methoden niet significant zal worden bevonden. De p-waarde van de Kruskal-Wallis toets bevestigt dit vermoeden, deze is gelijk aan 0.82 (zie ook Tabel 5.3). Voor de maand juli kan eenzelfde besluit getrokken worden. Het verschil in spreiding wordt grotendeels veroorzaakt door de grotere aanwezigheid van uitschieters bij PSO. Hierdoor lijkt PSO minder robuust te zijn, maar dit kan aan het toeval worden toegeschreven. De p-waarde is gelijk aan 0,43 en het verschil wordt op het 5% significantieniveau als niet significant bevonden.

**Tabel 5.3:** Benadering 1: vergelijking tussen bekomen *fitness*-waarden bij 30 herhalingen van kalibratie met SIMPSA en PSO aan de hand van de Kruskal-Wallis toets.

model	p-waarde	
	januari	juli
OBL	0,8245	0,4333
MBL	1,48E-09	4,40E-10
MBLG	0,0459	0,0002

**MBL** Voor MBL kan gezien worden dat voor zowel januari als voor juli de kalibratie met PSO aanzienlijk accurater en robuuster is. De mediaan en standaardafwijking van de fitness bij 30 herhalingen zijn een grootte-orde kleiner dan die bekomen door kalibratie met SIMPSA. Een nadeel is dat PSO minder efficiënt is bij het bekomen van deze resultaten. Het gemiddeld aantal FEs ligt namelijk een stuk hoger bij de kalibratie met PSO. In Figuur 5.1(c) en Figuur 5.1(d) is te zien dat SIMPSA voor de maand januari tweemaal en voor juli eenmaal convergeert naar

een slecht lokaal minimum terwijl PSO deze ondermaatse minima weet te vermijden wat neerkomt op een robuuster resultaat zoals ook wordt weerspiegeld door de standaardafwijking. De Kruskal-Wallis toets wijst uit dat het verschil in resultaten voor het MBL model op het 5% significantieniveau weldegelijk als significant kan worden bevonden. p-waarden zijn respectievelijk voor januari en juli gelijk aan  $1.48 \cdot 10^{-9}$  en  $4.40 \cdot 10^{-10}$  (Tabel 5.3). Er kan dus besloten worden dat het gebruik van PSO tot betere resultaten leidt voor de kalibratie van het MBL model in een groot parameterbereik (cf. benadering 1). Hierbij dient uiteraard in het achterhoofd gehouden worden dat de kalibratieduur groter is.

**MBLG** De kalibratie van het MBLG model lijkt, afgaande op de mediaan en standaardafwijking, ook hier accurater en robuuster uitgevoerd door PSO. Hoewel het verschil minder groot is dan het geval was bij MBL kunnen toch duidelijke verschillen opgemerkt worden. Dit vermoeden wordt bevestigd door Figuur 5.1(e) en 5.1(f). Het verschil in efficiëntie is verwaarloosbaar wat erop wijst dat de keuze van de optimalisatietechniek in dit geval zou uitgaan naar PSO. In de boxplots is bovendien te zien dat ook hier SIMPSA geplaagd wordt door de aanwezigheid van lokale minima die door het algoritme niet altijd kunnen ontweken worden. PSO ondervindt ook hinder van enkele uitschieters voor de kalibratie van de maand juli hoewel deze minder extreem zijn dan bij SIMPSA. De Kruskal-Wallis toets levert uitsluitel met p-waarden van respectievelijk 0.046 en 0.0002 voor januari en juli. Op het 5% significantieniveau kan dus besloten worden dat ook hier het gebruik van PSO betere resultaten oplevert dan SIMPSA en dat bij een vergelijkbare kalibratieduur.

Zoals beschreven in de bovenstaande paragrafen kan op basis van de bekomen fitnesswaarden besloten worden dat bij de kalibratie van de Bartlett-Lewis modellen, waarbij de parametergrenzen ruim gekozen worden, PSO voor MBL en MBLG significant betere resultaten oplevert terwijl voor OBL geen verschil wordt opgetekend. Hieruit zou kunnen afgeleid worden dat PSO beter dan SIMPSA in staat is om een optimale parameterregio aan te duiden ingeval er geen of weinig kennis is hieromtrent. Wanneer echter de beste parametersets die bekomen worden bij deze 30 herhalingen nader bekeken worden is te zien dat deze redenering niet volledig correct is. Tabel 5.4 toont de beste parametersets voor de respectievelijke modellen en twee afzonderlijke maanden die werden bekomen met de beide optimalisatietechnieken.

Voor alledrie de modellen wordt  $\lambda$  telkens vrij goed geschat als men de bekomen waarden vergelijkt met de literatuur (Verhoest *et al.*, 1997).  $\lambda$  is namelijk de frequentie waarmee buien zich voordoen, uitgedrukt in aantal buien per uur. Een waarde van 0,02 komt dan bijvoorbeeld overeen met ongeveer 15 buien per maand. Voor OBL is te zien dat bij de rest van de parameters vrij grote schommelingen waar te nemen zijn. Dit duidt op een veelvuldige aanwezigheid van lokale minima daar de fitnesswaarden niet significant verschillen tussen beide methoden doch de bekomen parameterwaarden ver uiteen liggen. Voor OBL kan dus op basis van deze resultaten de parameter ruimte niet met zekerheid verkleind worden, daarvoor zal beroep moeten gedaan worden op de beschikbare literatuur.

Voor MBL geldt een gelijkaardige conclusie.  $\lambda$  wordt vrij goed benaderd terwijl bij de overige parameters geen duidelijke lijn waar te nemen is in de resultaten. De fitnesswaarden van de parametersets voor januari en juli, bekomen door PSO, zijn aanzienlijk lager, hoewel dit geen garantie blijkt te zijn voor het effectief aanduiden van een veelbelovend parametergebied. Voor

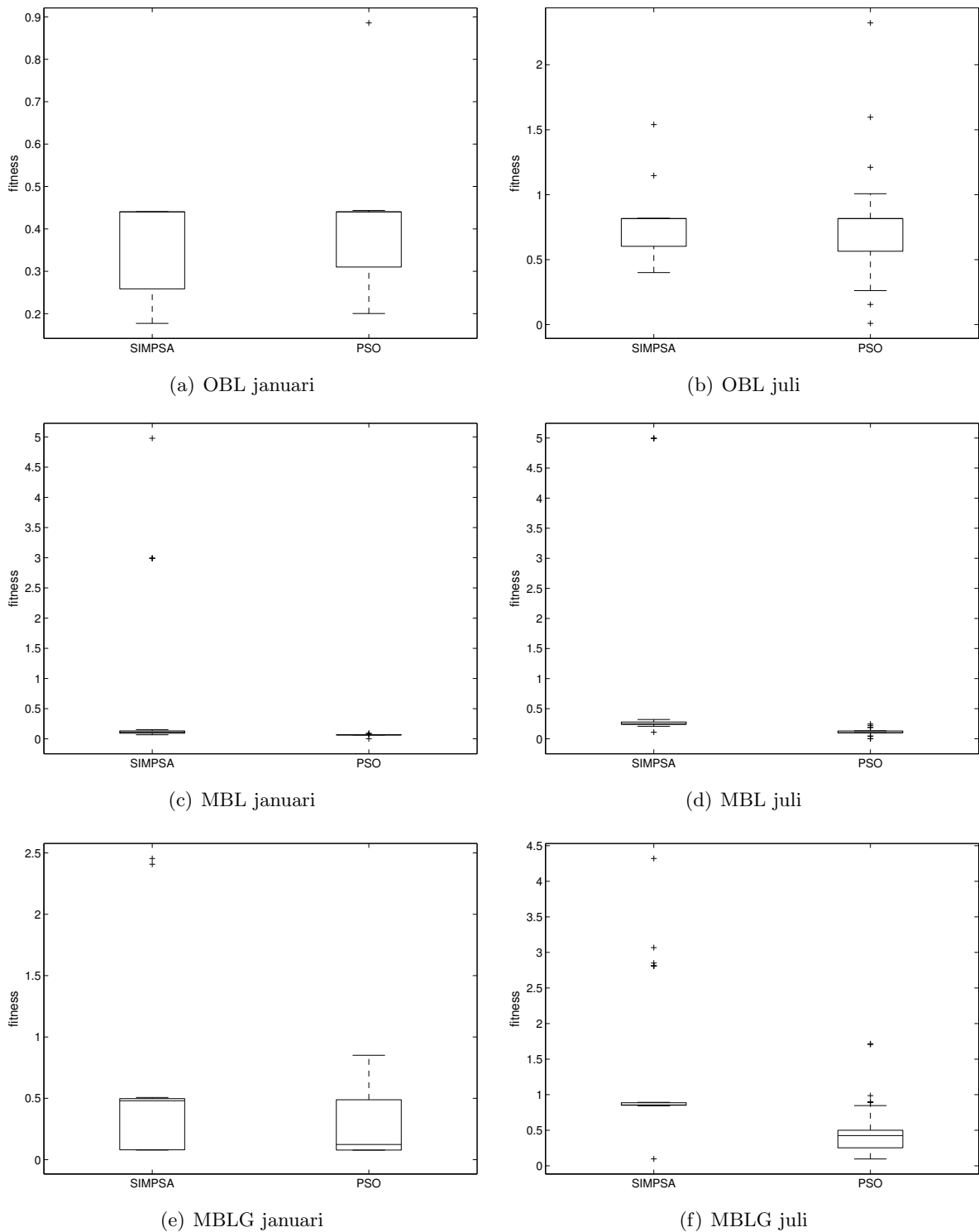
de maand juli is namelijk te zien dat PSO voor  $\kappa$  tegen de grens aanleunt, dit terwijl de overige parameters behalve  $\phi$  wel realistisch ingeschat worden. Dit wijst misschien op een interactie tussen de parameters  $\kappa$  en  $\phi$ .

Voor MBLG kunnen wederom gelijkaardige trends waargenomen worden.  $\lambda$  wordt goed ingeschat terwijl de overige parameters variëren en, gebaseerd op resultaten van Verhoest *et al.* (1997), onwaarschijnlijk hoge waarden aannemen. Beschouwen we bijvoorbeeld de waarden voor  $\alpha$  en  $\nu$  voor juli, bekomen door SIMPSA ( $\alpha = 37,3$ ;  $\nu = 42,1$ ). De parameters  $\alpha$  en  $\nu$  zijn respectievelijk de index en vormparameter van de gamma verdeling waaruit  $\eta$  getrokken wordt. De verwachtingswaarde van  $\eta$  is gelijk aan  $\alpha/\nu = 0,89$ , bijgevolg is de verwachte gemiddelde celduur ( $1/\eta$ ) gelijk aan 1,1 uur. Deze onrealistisch lange celduur zorgt er ook voor dat andere parameters onrealistische waarden aannemen. Er is dus overduidelijke interactie tussen de parameters.  $\phi$  is bijvoorbeeld ook opmerkelijk hoog. Dit is niet verwonderlijk gezien  $(\eta\phi)^{-1}$  de gemiddelde duur van een bui voorstelt. Omdat de cellen in de buien onrealistisch lang zijn moeten de buien kort zijn opdat de eigenschappen in de doelfunctie overeen zouden komen met de observaties. Dit fenomeen van equifinaliteit wordt ook weerspiegeld in de waarden van  $\kappa$ ,  $p$  en  $\delta$ . Het is dus duidelijk dat de structuur van de neerslag bij de validatie van het gekalibreerde model bijzondere aandacht verdient. De momenten in de doelfunctie kunnen namelijk vrij goed benaderd worden terwijl de structuur van de neerslag en de opbouw van de buien geheel onrealistisch is.

Over het algemeen kan dus gesteld worden dat geen van beide methoden echt in staat is om in een slecht afgebakende parameter ruimte de optimale regio te vinden. De keuze van de grenzen van de parameters weegt dus zwaar door op het uiteindelijke resultaat. In de volgende sectie zal worden uitgemaakt of een meer gegronde afbakening leidt tot een meer robuust resultaat voor de gekalibreerde parameters.

**Tabel 5.4:** Benadering 1: beste parametersets bij 30 herhalingen kalibratie SIMPSA en PSO.

model	methode	maand	<i>fitness</i>	parameters							
OBL	SIMPSA	Jan	0,1771	$\lambda$	$\beta$	$\gamma$	$\mu_x$	$\eta$			
		Jul	0,4004	0,0171	15,3417	0,1795	2,0514	32,8721			
	PSO	Jan	0,2004	0,0199	49,7219	56,2963	2,2776	0,8020			
		Jul	0,0097	0,0171	14,8632	0,1747	3,7466	59,8388			
	Verhoest <i>et al.</i> (1997)	Jan	9,32E-12	0,0133	8,4787	0,3580	29,5965	88,0422			
		Jul	2,54E-12	0,01462	0,53596	0,09194	0,80850	1,54149			
	MBL	SIMPSA	Jan	0,0674	$\lambda$	$\kappa$	$\phi$	$\mu_x$	$\alpha$	$\nu$	
			Jul	0,1093	0,0375	37,0093	36,4543	0,6617	2,9089	3,2924	
PSO		Jan	0,0022	0,0206	0,0001	79,6311	2,9642	23,2186	46,1913		
		Jul	0,0003	0,0299	0,3090	0,0227	1,1877	4,5136	0,6337		
Verhoest <i>et al.</i> (1997)		Jan	1,02E-07	0,0310	100,0000	43,3276	2,5806	2,4672	0,5276		
		Jul	2,30E-04	0,03792	0,58884	0,03426	1,21522	3,18989	0,23938		
MBLG		SIMPSA	Jan	0,0784	$\lambda$	$\kappa$	$\phi$	$\alpha$	$\nu$	$p$	$\delta$
			Jul	0,0972	0,0402	48,5396	66,7902	2,2944	1,2317	67,4511	54,4739
	PSO	Jan	0,0775	0,0165	77,0734	29,2503	37,2744	42,1286	74,3892	44,4946	
		Jul	0,0965	0,0405	42,5049	100,0000	2,2733	1,1690	100,0000	65,3404	
	Verhoest <i>et al.</i> (1997)	Jan	9,73E-05	0,0171	79,2540	30,2152	20,3826	21,9480	99,9819	59,8786	
		Jul	5,87E-04	0,03715	0,29167	0,02182	3,34956	0,19954	2,02340	1,00289	
			Jul	0,02216	0,0410	0,01224	3,06739	0,20324	3,32887	0,326499	



**Figuur 5.1:** Benadering 1: Boxplots kalibratie Bartlett-Lewis modellen.

### 5.3.2 Benadering 2

Aan de hand van de resultaten in Sectie 5.3.1 kon geen duidelijk besluit getrokken worden over de optimale parameterregio's voor de verschillende modellen en maanden. Daarom wordt op basis van de resultaten van Verhoest *et al.* (1997) hier de parameterruimte verkleind om vervolgens het proces uit de vorige sectie te herhalen. In dit tweede scenario (benadering 2) wordt getracht een preciezere schatting van de parameters te bekomen. Tabel 5.5 geeft een overzicht van deze

verkleinde grenzen voor de verschillende modellen. Met deze nieuwe grenzen worden wederom 30 kalibraties uitgevoerd met zowel PSO als SIMPSA. Hierbij blijven de instellingen onveranderd ten opzichte van benadering 1.

**Tabel 5.5:** Onder- en bovengrenzen voor de kalibratie van de respectievelijke modellen in benadering 2 (geldig voor alle maanden).

model	grens	parameters						
OBL		$\lambda$	$\beta$	$\gamma$	$\mu_x$	$\eta$		
	Ondergrens	0	0	0	0	0		
	Bovengrens	0,1	10	1	10	10		
MBL		$\lambda$	$\kappa$	$\phi$	$\mu_x$	$\alpha$	$\nu$	
	Ondergrens	0	0	0	0	0	0	
	Bovengrens	0,1	10	1	15	20	5	
MBLG		$\lambda$	$\kappa$	$\phi$	$\alpha$	$\nu$	$p$	$\delta$
	Ondergrens	0	0	0	0	0	0	0
	Bovengrens	0,1	10	1	5	5	20	10

**Tabel 5.6:** Benadering 2: minimum, mediaan en standaardafwijking (StDev) van de fitness en het gemiddeld aantal benodigde FEs bij 30 herhalingen van de kalibratie van de maand januari met SIMPSA en PSO van de verschillende Bartlett-Lewis modellen.

model	methode	<i>fitness</i>			FEs
		minimum	mediaan	StDev	gemiddelde
OBL	SIMPSA	8.39E-09	1,28E-07	1,12E-01	13236
	PSO	2.99E-11	1,05E-04	1,08E-01	12125
MBL	SIMPSA	2.04E-06	7,98E-05	3,77E-04	21481
	PSO	4.89E-07	7,25E-06	3,70E-04	357862
MBLG	SIMPSA	4.27E-06	3,81E-05	6,66E-04	25617
	PSO	2.95E-04	1,44E-02	2,99E-02	73854

In Tabel 5.6 wordt een overzicht gegeven van de mediaan en standaardafwijking van de fitness bij 30 kalibraties door PSO en SIMPSA en het gemiddeld aantal benodigde FEs voor de maand januari. Tabel 5.7 geeft op haar beurt het overzicht voor de maand juli. Net zoals bij de eerste benadering kon door beide algoritmen telkens convergentie bereikt worden.

**OBL** Voor het OBL model is te zien in zowel Tabel 5.6 als Tabel 5.7 dat de mediaan van de fitness lager ligt voor de kalibraties met SIMPSA dan deze bij PSO. Dit wijst er dus op dat SIMPSA voor OBL meer accuraat is en dit bij een gelijkaardig aantal FEs. Dit is ook duidelijk zichtbaar in Figuur 5.2(a). Verder is te zien dat de standaardafwijking voor de maand januari gelijkaardig is als bij PSO, beide blijken ongeveer even robuust, zij het dat SIMPSA meer accuraat is. Tabel 5.6 toont aan dat PSO een beter minimum vindt. Desalniettemin kan de voorkeur gegeven worden aan het gebruik van SIMPSA. De mediaan geeft namelijk duidelijk aan dat de kans op het vinden van een aanvaardbaar minimum groter is dan bij het gebruik van PSO, waar het gevonden minimum eerder kan gezien worden als een toevalstreffer. Voor de maand juli

**Tabel 5.7:** Benadering 2: mediaan en standaardafwijking (StDev) van de fitness en het gemiddeld aantal benodigde FEs bij 30 herhalingen van de kalibratie van de maand juli met SIMPSA en PSO van de verschillende Bartlett-Lewis modellen.

model	methode	<i>fitness</i>		FEs
		mediaan	StDev	gemiddelde
OBL	SIMPSA	3,94E-08	9,85E-08	9371
	PSO	5,24E-05	4,65E-03	10278
MBL	SIMPSA	4,38E-04	5,10E-02	20005
	PSO	3,90E-04	4,19E-03	229879
MBLG	SIMPSA	1,01E-03	3,53E-02	36005
	PSO	8,85E-02	3,98E-02	50811

is SIMPSA niet alleen accurater maar ook meer robuust, wat blijkt uit de standaardafwijking van de fitness na 30 herhalingen. In Figuur 5.2(a) zien we dat SIMPSA op één uitschieter na zeer lage fitnesswaarden bekommt. Indien deze uitschieter niet zou worden meegenomen in de berekening van de standaardafwijking zou deze aanzienlijk kleiner zijn. Voor PSO is in diezelfde figuur te zien dat de methode net iets meer uitschieters heeft en dus een lichtjes hogere spreiding rond het gemiddelde heeft. Tabel 5.8 geeft aan dat het verschil tussen beide methoden op het 5% significantieniveau als significant kan bevonden worden ( $p = 0.0003$ ). Figuur 5.2(b) geeft aan dat SIMPSA voor de maand juli overduidelijk beter presteert. De spreiding is miniem ten opzichte van deze van PSO, wat reeds bleek uit Tabel 5.7. Dit wordt tevens bevestigd door de Kruskal Wallis toets ( $p = 1.66 \cdot 10^{-7}$ ).

**MBL** De resultaten voor het MBL model, eveneens weergegeven in Tabel 5.6 en Tabel 5.7 en Figuren 5.2(c) en 5.2(d), wijzen erop dat PSO ditmaal als beste uit de strijd komt wanneer enkel rekening gehouden wordt met de bekomen fitnesswaarden. Voor de maand januari is de mediaan van de bekomen fitnesswaarden voor PSO lager dan die bekomen door SIMPSA terwijl de robuustheid, gegeven door de standaardafwijking, voor beide methoden zo goed als gelijk is. Hierbij dient wel opgemerkt te worden dat PSO er een stuk langer over doet dan SIMPSA, wat ook het geval is bij de maand juli. De fitness voor juli is voor beide methoden gelijkaardig met het verschil dat PSO zich iets meer robuust toont. In Figuur 5.2(d) is te zien dat dit verschil in standaardafwijking te wijten is aan enkele lokale minima waarin SIMPSA terecht komt, terwijl PSO daar duidelijk minder moeite mee blijkt te hebben. Voor januari is in Figuur 5.2(c) duidelijk te zien dat PSO de betere van de twee is wat ook wordt bevestigd door de statistische toets ( $p = 0.0012$ ). Voor de maand juli is het verschil minder groot door het toedoen van de voornoemde problemen die SIMPSA blijkt te hebben met enkele lokale minima, hier wordt een  $p$ -waarde van 0.0008 bekomen. Voor beide maanden is PSO met andere woorden significant beter dan SIMPSA als rekening gehouden wordt met het 5% significantieniveau.

**MBLG** De resultaten voor het MBLG model blijken, afgaand op figuren 5.2(e) en 5.2(f), uit te draaien in het voordeel van SIMPSA. Dit vermoeden wordt kracht bijgezet door de resultaten zichtbaar in Tabel 5.6 en Tabel 5.7. Daarin is te zien dat SIMPSA voor beide maanden accurater, robuuster en efficiënter is. Tabel 5.8 bevestigd dat op het 5% significantieniveau de kalibraties,

bekomen door SIMPSA, significant beter zijn dan deze bekomen door PSO. De keuze voor de optimalisatietechniek voor het MBLG model gaat dus uit naar SIMPSA.

**Tabel 5.8:** Benadering 2: vergelijking tussen bekomen *fitness*-waarden bij 30 herhalingen van kalibratie met SIMPSA en PSO aan de hand van de Kruskal-Wallis toets.

model	p-waarde	
	januari	juli
OBL	0,0003	1,66E-07
MBL	0,0012	0,0008
MBLG	1,2658E-10	3,13E-07

Op basis van bovenstaande resultaten kan besloten worden voor OBL en MBLG dat SIMPSA als beste van beide optimalisatietechnieken uit de hoek komt. Voor MBL blijkt het omgekeerde waar te zijn hoewel de benodigde rekentijd voor PSO in dat geval aanzienlijk hoger ligt. Net zoals bij benadering 1 worden in Tabel 5.9 de parametersets met laagste fitness, bekomen door PSO en SIMPSA, voor beide maanden en de verschillende Bartlett-Lewis modellen weergegeven. In deze tabel is te zien dat deze ditmaal niet ver uiteen liggen zoals het geval was bij de eerste benadering. Beide methoden slagen er dus in om de parameters goed te schatten. Voor OBL is te zien dat, ondanks het feit dat SIMPSA gemiddeld beter scoort, PSO toch een parameterset bekomt met een lagere fitnesswaarde. De geschatte parameters voor januari verschillen op zich weinig. Het minieme verschil doet vermoeden dat geen van beide methoden het absolute minimum heeft kunnen bereiken, hoogstwaarschijnlijk kunnen nog betere fitnesswaarden bereikt worden door de zoekruimte stelselmatig te verkleinen. Aangezien we te maken hebben met een stochastisch verschijnsel kan beargumenteerd worden dat deze minieme verschillen bij simulatie waarschijnlijk geen significant verschil zullen opleveren. Voor de maand juli is het verschil voor de parameters  $\beta$ ,  $\mu_x$  en  $\eta$  wel noemenswaardig. Voor het MBL model kan een gelijkaardig betoog gevoerd worden. De parameters bekomen door SIMPSA en PSO voor januari verschillen amper terwijl voor juli hier en daar de verschillen opmerkelijk zijn (zie  $\kappa$  en  $\alpha$ ) ondanks de vrijwel gelijke fitness. Dit kan erop wijzen dat MBL niet voldoende flexibel is om zowel de winterse als de zomerse neerslag te simuleren. Voor het MBLG model zijn wederom verschillen in de bekomen parametersets waar te nemen voor PSO en SIMPSA, ditmaal voor zowel januari als juli. Dit kan te wijten zijn aan de toegenomen complexiteit van het model, wat de kans op equifinaliteit vergroot. Anderzijds kan het zijn dat PSO hier naast het optimum grijpt, terwijl SIMPSA dit niet doet.

## 5.4 Besluit

In dit hoofdstuk werd gezocht naar de meest geschikte optimalisatiemethode voor de OBL, MBL en MBLG modellen. Hiervoor werd gebruik gemaakt van twee benaderingen. De eerste benadering ging uit van een situatie waarin geen kennis bestaat omtrent de optimale regio van de parameters voor de verschillende modellen. In de tweede benadering werd een beroep gedaan op de literatuur om de parametergrenzen af te bakenen. Om op basis daarvan te trachten uit te maken welk van beide methoden (SIMPSA of PSO) als beste de parameter ruimte exploiteert.

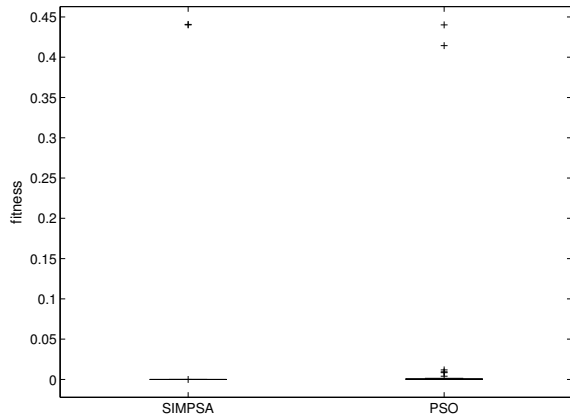


**Tabel 5.9:** Benadering 2: beste parametersets bij 30 herhalingen kalibratie SIMPSA en PSO.

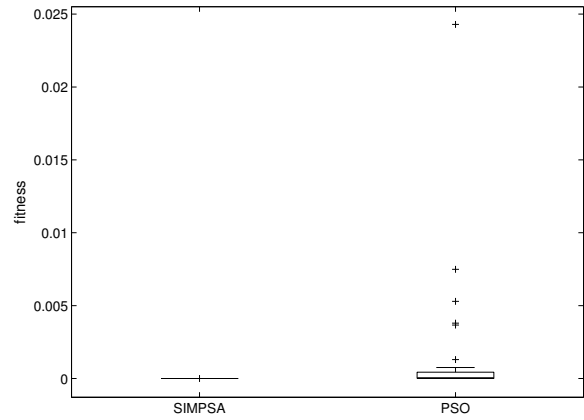
model	methode	maand	<i>fitness</i>	parameters						
OBL	SIMPSA	Jan	8,386E-09	0,01977	2,42439	0,11333	1,16896	5,63091		
		Jul	9,409E-09	0,01653	0,73205	0,13658	7,41204	7,30627		
	PSO	Jan	2,985E-11	0,01977	2,42442	0,11333	1,16885	5,63027		
		Jul	1,526E-13	0,01572	0,90428	0,14679	8,41465	8,96080		
	Verhoest <i>et al.</i> (1997)	Jan	9,32E-12	0,02558	0,53596	0,09194	0,80850	1,54149		
		Jul	2,54E-12	0,01462	0,15243	0,04648	4,01946	2,60690		
MBL	SIMPSA	Jan	2,039E-06	0,03664	0,95948	0,01965	1,21280	2,66352	0,06905	
		Jul	3,467E-04	0,02285	0,05555	0,01460	7,72147	6,91783	0,74491	
	PSO	Jan	4,887E-07	0,03665	0,96766	0,01949	1,21084	2,66153	0,06796	
		Jul	3,404E-04	0,02677	0,14666	0,03039	8,00960	3,20008	0,18784	
	Verhoest <i>et al.</i> (1997)	Jan	1,02E-07	0,03792	0,58884	0,03426	1,21522	3,18989	0,23938	
		Jul	2,30E-04	0,01802	0,04686	0,00866	6,28524	7,75433	0,89866	
MBLG	SIMPSA	Jan	4,273E-06	0,03519	0,36746	0,02884	2,81116	0,17500	4,73050	2,40662
		Jul	3,672E-04	0,02657	0,05365	0,02493	3,03546	0,18288	7,20407	0,50654
	PSO	Jan	2,945E-04	0,03431	0,40584	0,02817	2,98274	0,21179	2,34134	1,42209
		Jul	1,093E-03	0,03377	0,29234	0,08251	2,36316	0,06286	19,97579	1,42829
	Verhoest <i>et al.</i> (1997)	Jan	9,73E-05	0,03715	0,29167	0,02182	3,34956	0,19954	2,02340	1,00289
		Jul	5,87E-04	0,02216	0,0410	0,01224	3,06739	0,20324	3,32887	0,326499

Uit de resultaten is gebleken dat geen van beide methoden echt in staat is om bij de eerste benadering de optimale parameterregio aan te duiden. Er kan besloten worden dat deze methoden niet genoeg in staat zijn de zoekruimte te exploreren en de lokale minima te ontwijken.

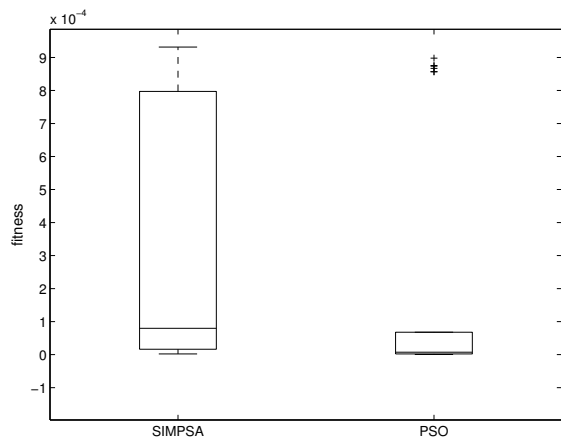
In de tweede benadering slaagden beide methoden er wel in om tot goede resultaten te komen. Voor het OBL en MBLG model presteert SIMPSA significant beter dan PSO. Voor het MBL model levert PSO dan weer betere resultaten op. Hierbij is de benodigde rekentijd voor de kalibratie echter wel aanzienlijk hoger bij PSO. Wanneer de door SIMPSA best bekomen parameterset vergeleken wordt met diegene die bekomen werd met PSO, blijken deze slechts minieme verschillen te vertonen. Er kan dus gesteld worden dat uit praktische overwegingen het gebruik van SIMPSA voor de kalibratie van alledrie de onderzochte modellen gerechtvaardigd is.



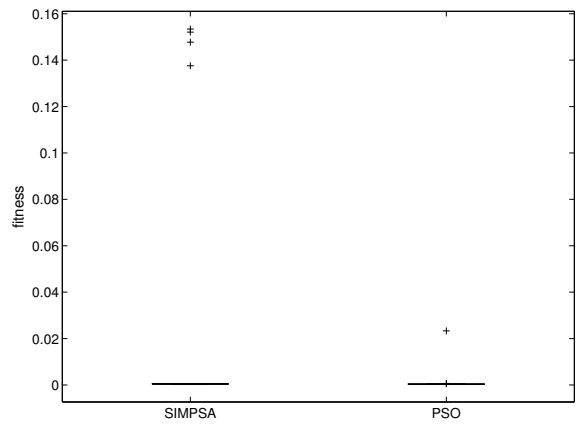
(a) OBL januari



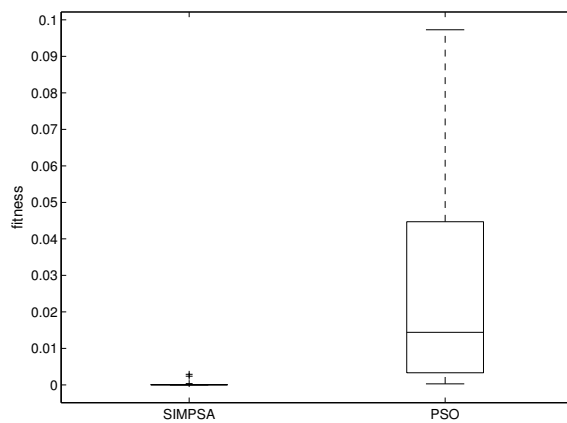
(b) OBL juli



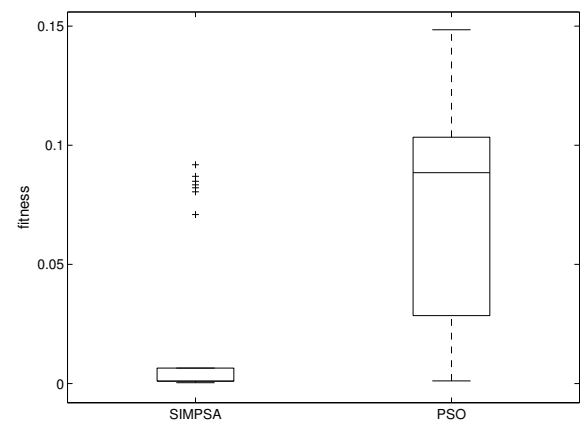
(c) MBL januari



(d) MBL juli



(e) MBLG januari



(f) MBLG juli

**Figuur 5.2:** Benadering 2: boxplots kalibratie Bartlett-Lewis modellen.

## Hoofdstuk 6

# Evaluatie gekalibreerde modellen

### 6.1 Inleiding

In Hoofdstuk 5 werd besloten dat de kalibratie van de OBL, MBL en MBLG modellen bij voorkeur gebeurt aan de hand van SIMPSA. Bij de kalibratie werd gebruik gemaakt van de in Hoofdstuk 4 voorgestelde doelfunctie. Om na te gaan hoe de eigenschappen van de neerslagreeksen, gevormd aan de hand van de gekalibreerde modellen, zich verhouden tot de observaties, dienen ook de overige maanden gekalibreerd te worden. De analyse van de resulterende synthetische neerslagreeksen houdt (cf. Hoofdstuk 3) rekening met de eerste, tweede en derde orde eigenschappen van de neerslagintensiteit, enkele voorwaardelijke distributiekarakteristieken omtrent diens verdeling, en de verdeling van de extreme waarden. Hiertoe worden voor de verschillende modellen 50 tijdreeksen gesimuleerd met een lengte van 105 jaar. Vervolgens wordt de mediaan van de berekende eigenschap in kwestie over 50 herhalingen berekend, teneinde deze te kunnen vergelijken met de geobserveerde waarden. In Sectie 6.2 worden de geschatte parameters besproken. Sectie 6.3 vervolgt met de analyse van de gesynthetiseerde neerslagreeksen.

### 6.2 Parameters

Tabellen B.4, B.5 en B.6 geven de met behulp van SIMPSA geschatte parameters voor de respectievelijke Bartlett-Lewis modellen weer. Daarbij wordt de uitkomst van de doelfunctie of *fitness* meegegeven. De bekomen fitnesswaarden kunnen vergeleken worden met diegene die door Verhoest *et al.* (1997) werden bekomen. Op basis daarvan kan besloten worden dat ook hier goede resultaten worden bekomen. Bovendien worden met SIMPSA sneller resultaten geboekt dan het geval was bij Verhoest *et al.* (1997), gezien de kalibratie toen zeer veel tijd in beslag nam (Verhoest, persoonlijke communicatie, 2010). Het gebruik van SIMPSA heeft er dus toe geleid dat het kalibratieproces sneller en doeltreffender verloopt, wat bij aanvang ook de opzet was. Zoals eerder vermeld kan aan de parameters van de Bartlett-Lewis modellen een fysische interpretatie gegeven worden (Rodriguez-Iturbe *et al.*, 1987a). Door deze fysisch interpreteerbare variabelen uit te zetten in een grafiek kunnen bepaalde trends in de geschatte parameters blootgelegd worden. Daarenboven kan nagegaan worden in welke mate deze variabelen door de verschillende modellen anders benaderd worden.

Figuur 6.1 geeft deze grootheden in functie van de maand en het model weer. Hierbij worden  $\lambda$ ,  $\mu_c$ ,  $\mu_x$  en  $\eta^{-1}$  weergegeven, welke gezien kunnen worden als :

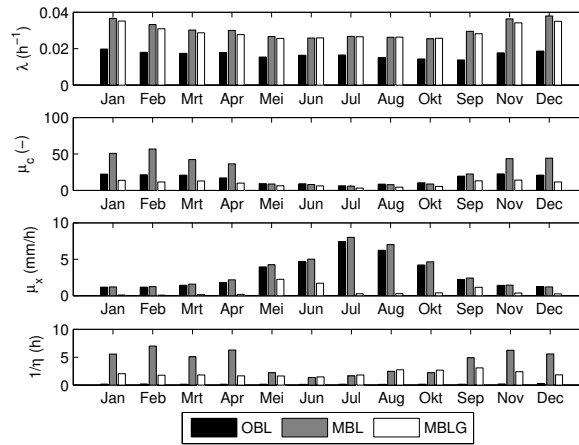
- $\lambda$  ( $\text{h}^{-1}$ ) : de snelheid waarmee buien zich manifesteren,
- $\mu_c$  (-) : het gemiddeld aantal cellen per bui of storm,
- $\mu_x$  (mm) : de gemiddelde celhoogte en
- $1/\eta$  (h) : de gemiddelde celduur.

Het gebruik van deze notatie kan leiden tot verwarring. Deze parameters worden immers niet bij alle modellen op deze manier behandeld bij de optimalisatie. Toch kunnen ze voor elk model afgeleid worden uit de verschillende parameters (zie Tabel 6.2 en Bijlage A, voor meer uitleg).  $\mu_c$  is bijvoorbeeld gelijk aan  $1 + \kappa/\phi$ . Bij het MBL model worden  $\kappa$  en  $\phi$  expliciet behandeld als modelparameters, terwijl bij het OBL model deze afgeleid worden uit  $\beta$ ,  $\eta$  en  $\gamma$  ( $\kappa = \beta/\eta$  en  $\phi = \gamma/\eta$ ). Voor het MBL model en het MBLG model is de verwachtingswaarde van  $\eta$  op zijn beurt gelijk aan het gemiddelde van de gamma verdeling met vormparameter  $\alpha$  en schaalparameter  $\nu$ . De gemiddelde celduur kan dan geschreven worden als  $\eta^{-1} = (\alpha\nu)^{-1}$ . Daarbij komt nog dat, voor het MBLG model,  $\mu_x = (p\delta)^{-1}$ , de inverse van het gemiddelde van de gamma verdeling met vormparameter  $p$  en schaalparameter  $\delta$ . Voor meer uitleg omtrent de analytische beschrijving van de beschouwde modellen wordt verwezen naar Bijlage A.

Voor alledrie de modellen is het gemiddeld aantal cellen 's winters aanzienlijk hoger dan in de zomer (Figuur 6.1). Het gemiddeld aantal cellen per storm blijkt duidelijk afhankelijk van het modeltype. Het MBL model bijvoorbeeld, genereert per storm gemiddeld 2 tot 3 keer meer cellen dan de andere twee modellen. De gemiddelde celhoogte is over het algemeen hoger in de zomer dan in de winter, met opmerkelijke verschillen tussen de modellen. De gemiddelde celduur geeft op het eerste zicht geen blijk van seizoenaliteit. De snelheid van buimanifestatie is dermate klein dat ze niet op de weergegeven figuur kan worden geïnterpreteerd. Raadpleging van Tabellen B.4, B.5 en B.6 toont aan dat er in de winter meer buien plaatsvinden dan in de zomer. Dit leidt tot de conclusie dat de geschatte parameters de geobserveerde seizoenaliteit in de historische data weerspiegelen (zie Hoofdstuk 3). De zomermaanden worden gekenmerkt door minder buien (lagere  $\lambda$ ), met hogere intensiteiten (hogere  $\mu_x$ ). De parameters  $\mu_c$  en  $\eta^{-1}$  zijn iets moeilijker te interpreteren. Het aantal cellen en de gemiddelde celduur zijn immers weinig betekenisvol vanuit het standpunt van de waarnemer. Bovendien kunnen we vermoeden dat deze parameters onderling wel eens sterk met elkaar verbonden zouden kunnen zijn. Een bui met veel, maar korte cellen, zou inderdaad dezelfde statistische eigenschappen kunnen hebben als een korte bui met lange cellen. De grote verschillen tussen deze gemiddelde celduur en gemiddeld aantal cellen in functie van de modellen (zie Figuur 6.1) zijn daar misschien het gevolg van.

### 6.3 Statistische analyse

Met behulp van de geschatte parameters (Sectie 6.2) worden per model 50 neerslagreeksen, met een lengte van 105 jaar, gesimuleerd. Voor elk van de herhalingen kunnen verschillende statistische eigenschappen worden berekend (cf. Hoofdstuk 3), waarvan vervolgens de mediaan genomen worden om de vergelijking met de observaties te vergemakkelijken. Hiervoor wordt gebruik gemaakt van de procentuele afwijking ten opzichte van de observaties. Deze procentuele



**Figuur 6.1:** Vergelijking van de seizoensale trend, vervat in de geschatte parameters voor de verschillende Bartlett-Lewis modellen. Weergave van enkele fysisch interpreteerbare eigenschappen van de neerslag in functie van de parameters.

afwijking kan als volgt worden berekend:

$$\left( \frac{M_i'' - M_i'}{M_i'} \right) \times 100 \quad (6.1)$$

met  $M_i''$  de mediaan van de eigenschap  $i$ , berekend overheen het ensemble van 50 herhalingen en  $M_i'$  de waarde van statistiek  $i$ , berekend aan de hand van de historische reeks. Hieruit volgt dat een positieve procentuele afwijking overeenkomt met een overschatting van de modelresultaten, terwijl een negatieve afwijking het gevolg is van een onderschatting ten opzichte van de observaties. Merk op dat Vergelijking 6.1 analoog is aan Vergelijking 4.1 uit Hoofdstuk 4, met dit verschil dat hier geen gebruik gemaakt wordt van analytische maar gesimuleerde eigenschappen. Niettemin is de link tussen beide voor de hand liggend. Een grote procentuele afwijking van de geobserveerde waarden zal immers overeenkomen met een proportionele afwijking in de doelfunctie (indien deze statistiek bij het bepaalde aggregatieniveau werd opgenomen in de doelfunctie uiteraard). De hieronder beschreven procentuele afwijkingen geven dus een idee van hoe de doelfunctie door de verschillende momenten wordt beïnvloed, wat interessant is indien overwogen zou worden om de gebruikte gewichten in de doelfunctie te herzien (zie Hoofdstuk 4). In Sectie 6.3.1 worden de eerste, tweede en derde orde momenten nader bekeken. Sectie 6.3.2 voorziet in tekst en uitleg omtrent de in Hoofdstuk 3 gedefinieerde voorwaardelijke distributiekarakteristieken, terwijl Sectie 6.3.3 dieper ingaat op de voorwaardelijke probabiliteiten. Ten slotte worden de resultaten van de frequentie-analyse besproken in Sectie 6.3.4.

### 6.3.1 Eerste, tweede en derde orde momenten

In Figuur C.2 worden voor de verschillende maanden de procentuele afwijking van de gemiddelde neerslaghoogte, de ZDP, de skewness, de variantie en de covariantie weergegeven. Een blik op de resultaten leert dat deze niet direct als consistent kunnen bestempeld worden. De gemiddelde neerslaghoogte wordt niet telkens door hetzelfde model als beste benaderd. Het OBL model levert de kleinste afwijking op voor de maanden maart (Figuur C.2(e)), augustus (Figuur C.2(o)) en oktober (Figuur C.2(s)). Het MBL model benadert de gemiddelde neerslaghoogte beter voor de maanden april (Figuur C.2(g)), mei (Figuur C.2(i)), september (Figuur C.2(q)) en december

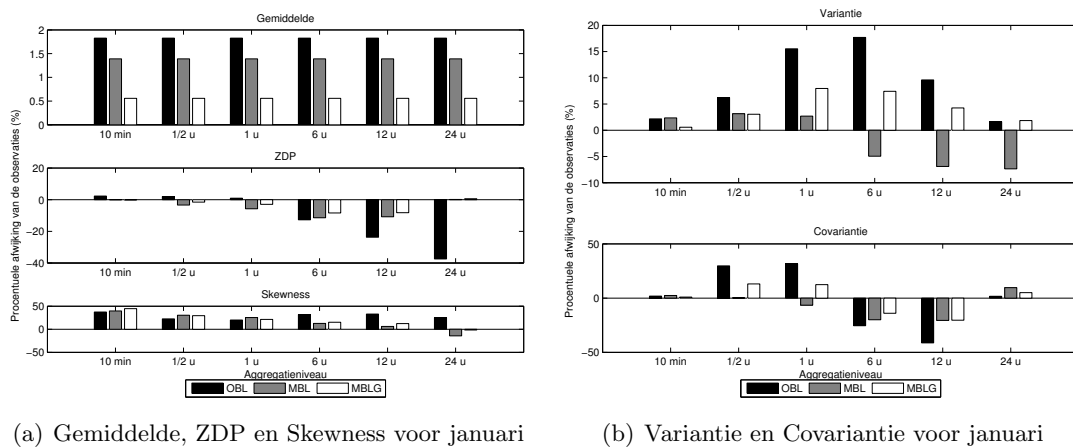
(Figuur C.2(w)), terwijl het MBLG model de beste benadering voor januari (Figuur C.2(a)) en februari (Figuur C.2(c)) oplevert. De procentuele afwijking is in vrijwel alle gevallen relatief klein. De afwijking bedraagt maximaal 3%, wat hoogstwaarschijnlijk te verwaarlozen valt. De modellen leveren met andere woorden allemaal goede benaderingen van de gemiddelde neerslaghoogte.

De procentuele afwijking van de variantie van de neerslaghoogte is daarentegen wel uitgesproken. Het is duidelijk dat het OBL model de variabiliteit van de neerslaghoogte het slechtst inschat. Dit is bijvoorbeeld te zien in Figuur 6.2(b). De resultaten voor de overige maanden worden weergegeven in Figuur C.2, daar kan doorgaans eenzelfde tendens opgemerkt worden. Vooral op een aggregatieniveau van 1, 6 en 12 uur is de afwijking ten opzichte van de observaties hoog, tot bijna 20 %. Ook de simulaties, uitgevoerd door het MBLG model, vertonen een gelijkaardig patroon met lage procentuele afwijkingen op aggregatieniveau's van 10 minuten en 24 uur en grotere afwijkingen daar tussenin. Dit heeft wellicht te maken met de keuze van de momenten in de doelfunctie. De afwijkingen zijn wel aanzienlijk kleiner dan bij het OBL model. Voor zowel het OBL model als het MBLG model wordt in de doelfunctie de variantie bij een aggregatieniveau van 10 minuten en 24 uur gebruikt (zie ook Tabel 4.1). Deze worden bijgevolg beter benaderd dan de tussenliggende. Voor de doelfunctie van het MBL model werd enkel gebruik gemaakt van de variantie bij een aggregatieniveau van 10 minuten. De procentuele afwijking van de variantie van de neerslaghoogte ten opzichte van de observatie is daarom niet minder goed, zoals blijkt uit de figuren. Over het algemeen is de afwijking voor aggregatieniveau's van 1/2 u tot 12 u in bepaalde gevallen zelfs kleiner dan bij de andere modellen, wat echter niet voor alle maanden het geval is. De afwijking op dagniveau is wel meestal groter, gezien deze statistiek niet wordt opgenomen in de doelfunctie. Het MBL model is ook de enige van de drie die de variantie van de neerslaghoogte onderschat. Deze vergelijking tussen de opgenomen momenten in de doelfunctie laat opnieuw blijken dat het opnemen van meerdere aggregatieniveau's in de doelfunctie leidt tot een betere algemene representatie van de statistieken op verschillende tijdsniveau's, zoals wordt geopperd door Khaliq & Cunnane (1996) (zie ook Sectie 4.3).

Een gelijkaardige conclusie kan gevormd worden voor de ZDP. Zoals vermeld in Hoofdstuk 4, Tabel 4.1, wordt bij de kalibratie van het OBL model de ZDP niet betrokken in de doelfunctie. De resulterende onderschatting van de ZDP is bijgevolg aanzienlijk, tot wel 40% voor januari (zie Figuur 6.2(a)). Bij het MBL model en het MBLG model zit de ZDP wel vervat in de doelfunctie, wat duidelijk kleinere afwijkingen van de observaties oplevert. Op basis van Figuur C.2 kunnen geen opmerkelijke verschillen tussen het MBL model en het MBLG model voor de afwijking van de ZDP opgemerkt worden.

Wanneer de procentuele afwijkingen van de covariantie van de neerslaghoogte van de synthetische reeksen, bekomen door de respectievelijke modellen, ten opzichte van de historische reeks, onder de loep genomen worden, kan een duidelijk patroon blootgelegd worden. Voor aggregatieniveau's tot en met 1 uur leveren alledrie de modellen (met uitzondering van het MBL model voor januari-april) een overschatting van de covariantie. Hierbij kan opgemerkt worden dat het MBL model en het MBLG model aanzienlijk kleinere afwijkingen opleveren, in tegenstelling tot het OBL model. De covariante wordt onderschat door de verschillende modellen op aggregatieniveau's tussen 1 en 24 uur (zie Figuur 6.2(b)). Op het dagniveau is de afwijking voor elk van de drie modellen opnieuw vrij klein, gezien de covariantie op het niveau van 24 uur is opgenomen in de doelfunctie.

Tot slot wordt de afwijking van de skewness van de verdeling weergegeven (Figuur C.2 en Figuur 6.2(a)). Deze wordt, bij lagere aggregatieniveau's, aanzienlijk overschat door elk van de onderzochte modellen. De afwijking neemt af naarmate het aggregatieniveau groter wordt maar blijft niettemin groot. Over het algemeen presteert het OBL model het minst goed. Volgens Cowpertwait (1998) komt het opnemen van deze skewness de reproductie van extreme waarden ten goede. Het opnemen van een analytische uitdrukking voor het derde orde moment (analoog aan de skewness) in de doelfunctie leidt tot een betere representatie van extreme waarden bij de Neyman-Scott modellen. Sinds kort zijn analytische uitdrukkingen voor het centrale derde moment voor de Bartlett-Lewis modellen beschikbaar (Onof, 2011), waardoor de invloed op de extreme waarden in deze context kan worden onderzocht (zie Hoofdstuk 7)



(a) Gemiddelde, ZDP en Skewness voor januari

(b) Variantie en Covariantie voor januari

**Figuur 6.2:** Eerste, tweede en derde orde momenten voor januari

### 6.3.2 Voorwaardelijke distributiekarakteristieken

Zoals reeds vermeld in Hoofdstuk 3 zijn niet alleen de eerste, tweede en derde orde karakteristieken van belang bij de evaluatie van de verschillende gebruikte modellen. De gesimuleerde neerslag dient bij voorkeur een opbouw te hebben die gelijkaardig is aan die van de historische reeks. Net zoals in de voorgaande sectie wordt de evaluatie gebaseerd op procentuele afwijkingen van de observaties. De waarden van de voorwaardelijke distributiekarakteristieken (zoals gedefinieerd in Sectie 3.2.2) voor Ussel zijn terug te vinden in Tabel B.2 en Tabel B.2. Om de overzichtelijkheid te waarborgen, wordt de gemiddelde procentuele afwijking overheen de verschillende maanden weergegeven. Op die manier kan een algemeen beeld gevormd worden over hoe de verschillende modellen zich tegenover elkaar verhouden op vlak van de opbouw en verloop van de neerslag.

De eerste voorwaardelijke distributiekarakteristiek (vwd1) drukt de kans op twee opeenvolgende droge intervallen uit. Het spreekt voor zich dat deze kans afneemt naarmate het aggregatieniveau toeneemt. De kans dat op twee opeenvolgende dagen geen regen valt is immers veel kleiner dan de kans dat op twee opeenvolgende uren geen neerslag valt. Figuur 6.3(a) toont de procentuele afwijkingen van de gesimuleerde reeksen ten opzichte van de historische reeks voor vwd1. Op basis van deze figuur kan besloten worden dat deze afwijking klein is voor alle modellen voor aggregatieniveau's tot en met 1 uur. Het OBL model vertoont hierbij de kleinste afwijking. Bij hogere aggregatieniveau's is de afwijking die door het OBL model gemaakt wordt veel groter

dan bij de andere modellen. Op het dagniveau wordt vwd1 door het MBL model en het MBLG model overschat, terwijl het OBL model deze kans onderschat. De kans dat het de dag na een droge dag regent wordt met andere woorden veel te laag ingeschat door het OBL model, terwijl deze kans overschat wordt door de overige twee modellen.

De overige voorwaardelijke distributiekarakteristieken worden weergegeven in Figuur 6.3. Hierbij valt op dat er een grote discrepantie is tussen de gesynthetiseerde waarden en de geobserveerde waarden. De gemiddelde bui-intensiteit (vwd2) wordt door het MBL model en het MBLG model bij kleinere aggregatieniveau's (zie Figuur 6.3(b)) onderschat, terwijl de afwijking kleiner wordt bij toename van de lengte van het aggregatieniveau. Het omgekeerde is waar voor het OBL model. De afwijking van de variantie van de bui-intensiteit (vwd3) geeft geen blijk van bepaalde verschillen tussen de modellen. Ook de procentuele afwijkingen ten opzichte van de observaties van de gemiddelde intensiteit bij aanvang van een bui (vwd4), vertonen een gelijkaardig patroon (Figuur 6.3(c)). Het OBL model levert goede resultaten bij lagere aggregatieniveau's en minder goede bij hogere aggregatieniveau's, terwijl het MBL model en MBLG model een omgekeerde tendens vertonen. De bekomen procentuele afwijkingen zijn wederom hoog. In Figuur 6.3(d) zijn de afwijkingen te zien voor de gemiddelde bui-intensiteit net voor het einde van de bui (vwd6) en de variantie van die intensiteit of neerslaghoogte (vwd7). Deze afwijkingen zijn zeer hoog, voor alle modellen. Alle modellen onderschatten deze intensiteit en de variabiliteit ervan sterk, vooral bij lagere aggregatieniveau's. De variabiliteit wordt door alle modellen bij lagere aggregatieniveau's zwaar onderschat. Figuur 6.3(e) toont ten slotte dat ook de intensiteit tijdens de bui (vwd8) en de variantie van die intensiteit (vwd9) vooral bij grotere aggregatieniveau's slecht wordt weergegeven. Er kan dus besloten worden dat alledrie de modellen gebreken vertonen bij de simulatie van de interne stormstructuur. Er blijkt onder de modellen een algemene tendens te bestaan om de intensiteit tijdens het verloop van de bui te onderschatten. Het OBL model modelleert de opbouw van de buien doorgaans beter bij lagere aggregatieniveau's dan bij hogere aggregatieniveau's, terwijl dit omgekeerd is voor de aangepaste modellen.

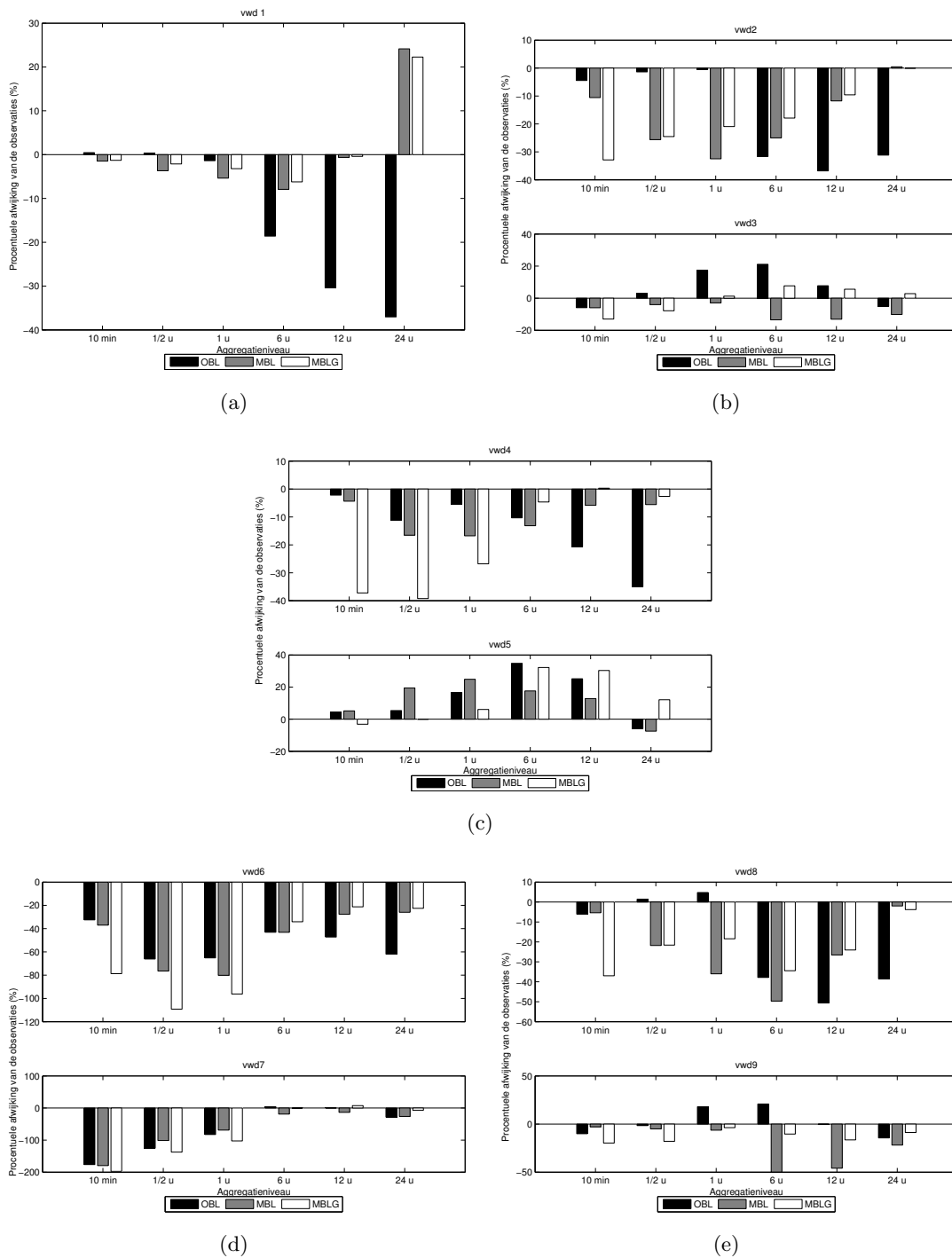
### 6.3.3 Voorwaardelijke probabiliteiten inzake neerslagintensiteit

Naast de algemene statistische eigenschappen (Sectie 6.3.1) en enkele voorwaardelijke distributiekarakteristieken met betrekking tot de opbouw en het verloop van de gesimuleerde neerslag (Sectie 6.3.2) is het van belang dat op verschillende aggregatieniveau's aan bepaalde voorwaardelijke probabiliteiten inzake de neerslagintensiteit voldaan wordt. Hiermee wordt de verdeling van de neerslag volgens intensiteit bedoeld (zie ook Sectie 3.2.3). Het is wenselijk dat de gesimuleerde neerslag, naast de ZDP, de verdeling van neerslag volgens intensiteit niet onder- of overschat.

Figuur 6.4 geeft de jaarlijkse gemiddelde procentuele afwijking voor de verschillende modellen weer. Figuur 6.4(a) geeft dit weer voor een aggregatieniveau van 10 minuten, Figuur 6.4(b) voor een aggregatieniveau van 1/2 uur, en zo verder. De figuren dienen als volgt te worden geïnterpreteerd: de ordinaat toont de procentuele afwijking van de kans dat de neerslaghoogte op het beschouwde aggregatieniveau kleiner is dan de neerslaghoogte in de abscis.

Op het eerste zicht is te zien dat de bekomen afwijkingen vrij klein zijn bij de kleinere aggregatieniveau's en groter worden bij grotere aggregatieniveau's. Tevens blijkt dat toenemend grotere afwijkingen bij afnemende neerslaghoogten worden bekomen. Op een aggregatieniveau van 10





**Figuur 6.3:** Gemiddelde procentuele afwijking van de voorwaardelijke distributiekarakteristieken overheen de verschillende maanden.

minuten is de kans dat er minder dan 0.5 mm neerslag valt groter bij de gesimuleerde reeksen dan bij de geobserveerde reeks, zij het dat de afwijking heel klein is ( $<0.15\%$ ). Op een aggregatieniveau van 24 uur zien we dat het OBL model  $DP_{0.5}$  en  $DP_2$  onderschat, dit wil zeggen dat de proportie aan intervallen waarin minder dan respectievelijk 0.5 mm en 2 mm neerslag valt kleiner is in de synthetische neerslagreeksen, geproduceerd door het OBL model. Het MBL model en het MBLG model overschatten dit, er zijn dus te veel intervallen waarin de neerslaghoogte kleiner is dan 0.5 mm en 2 mm. Deze resultaten liggen min of meer in lijn met deze bekomen in Sectie 6.3.2. Daaruit bleek ook al dat het OBL model de gemiddelde bui-intensiteit bleek te onderschatten bij hogere aggregatieniveau's, terwijl deze onderschatting bij de overige twee modellen kleiner was. Figuur 6.4(c) toont dat op een aggregatieniveau van 1 uur het OBL model beter presteert, vooral voor  $DP_{0.5}$  en  $DP_1$ . Dit heeft wellicht te maken met het feit dat voor de kalibratie van het OBL model het gemiddelde op een aggregatieniveau van 1 uur wordt gebruikt in de doelfunctie, wat niet het geval is voor de andere twee modellen (zie Hoofdstuk 4). Over het algemeen kan dus geen van de modellen als beste naar voor worden gebracht. Wel kan gesteld worden dat de afwijkingen voor het MBL model en het MBLG model over de hele lijn relatief klein zijn. Het OBL model is bij hogere aggregatieniveau's duidelijk minder accuraat, met aanzienlijke afwijkingen (tot wel 10%) tot gevolg.

### 6.3.4 Extreme waarden analyse

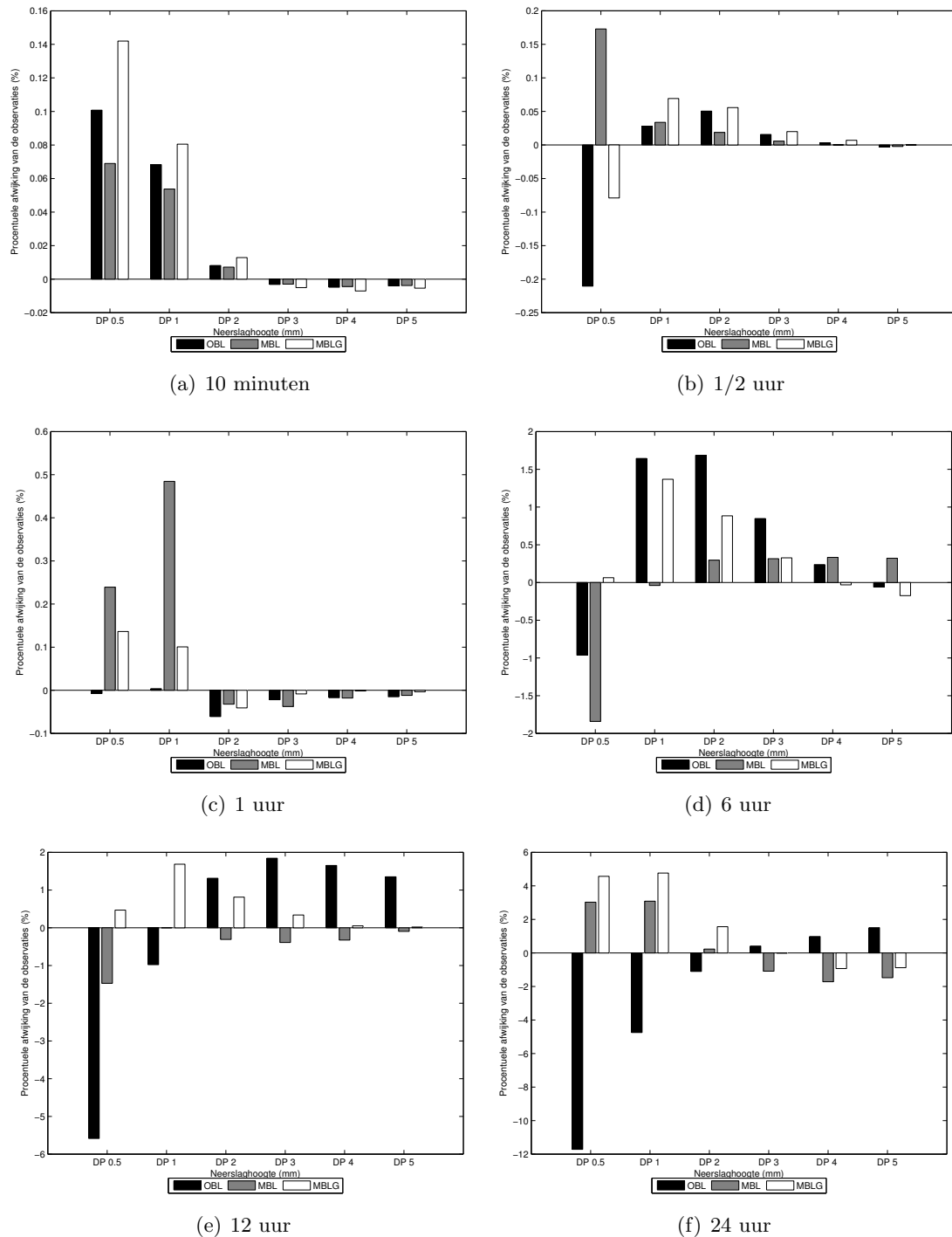
Tot slot worden de gesimuleerde neerslagreeksen onderworpen aan een extreme waarden analyse. De analyse van deze extremen gebeurt zoals dit reeds werd besproken in Sectie 3.2.4. De evaluatie van de resultaten is gebaseerd op de visuele interpretatie van de Gumbel plots waarop de jaarlijkse extremen vergeleken worden met de geobserveerde neerslagreeks.

Figuur 6.5 toont de vergelijkende plots voor de verschillende aggregatieniveau's. Hierop is te zien dat bij lagere aggregatieniveau's (10 tot 30 minuten, Figuur C.1(a) en C.1(b)) door alle modellen een onderschatting van de extreme waarden wordt gegeven. Deze onderschatting wordt groter bij grotere GRV's en dus bij grotere retourperiodes. Het dient vermeld te worden dat de door het OBL model en het MBL model gesimuleerde reeksen dichter aanleunen bij de historische waarden. Voor een aggregatieniveau van 1 uur (Figuur C.1(c)) is de fit met de historische waarden beter dan bij de lagere aggregatieniveau's. Bij hogere aggregatieniveau's presteert het OBL model steeds minder goed dan de aangepaste modellen (Figuur C.1(d) tot C.1(f)). Voor een aggregatieniveau van 6 uur (Figuur C.1(d)) volgen de Gumbel plots van het MBL model en MBLG model de historische waarden vrij goed. Voor gotere retourperiodes geven ze echter wel overschattingen. Dit laatste is ook het geval bij 12 en 24 uur (Figuur C.1(e) en C.1(f)). Bij lagere retourperiodes is de fit telkens goed, mits een kleine onderschatting, terwijl bij hogere retourperiodes telkens de neerslaghoogten worden overschat.

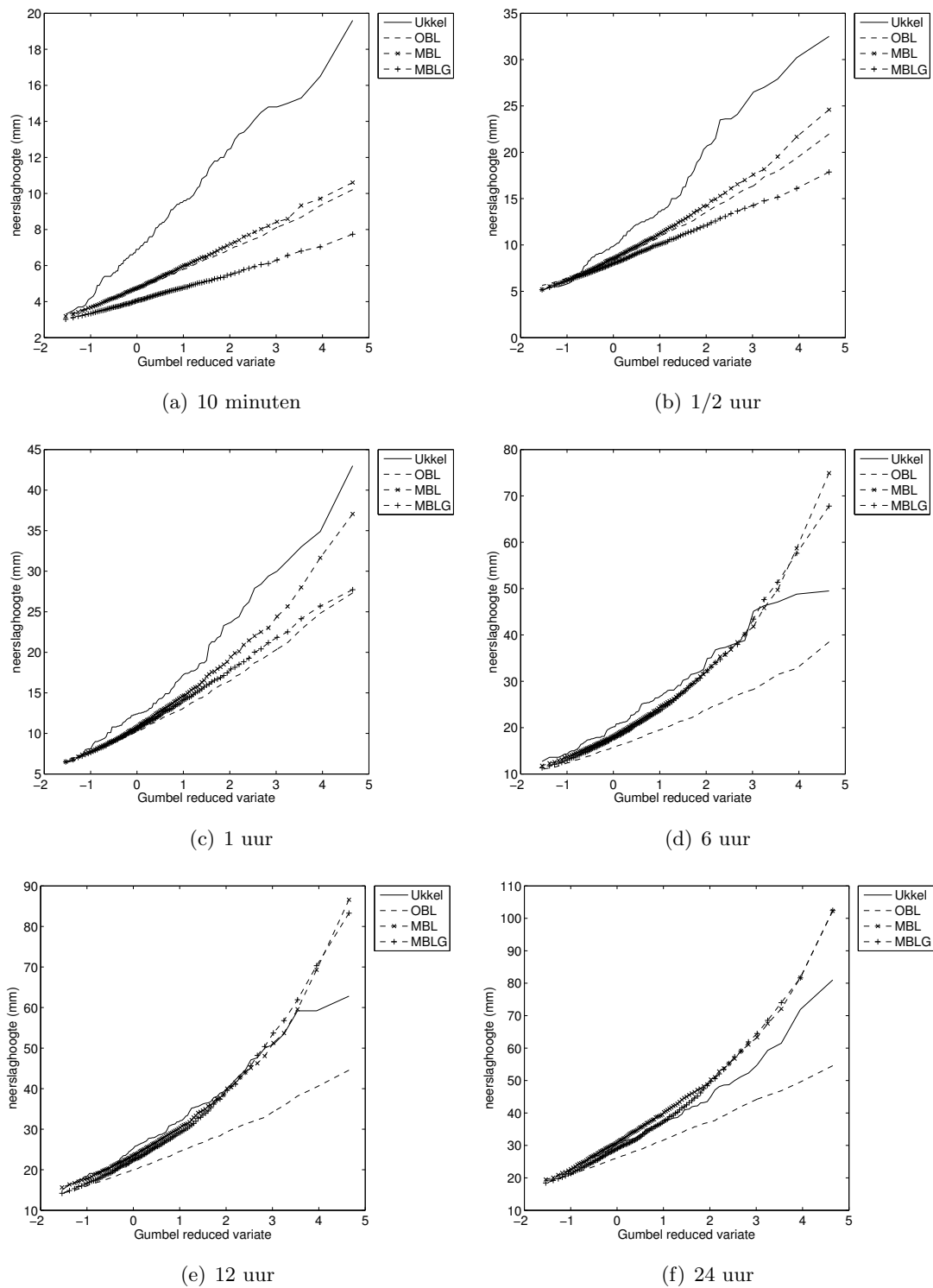
## 6.4 Besluit

In het huidige hoofdstuk werden de resultaten van de met SIMPSA gekalibreerde Bartlett-Lewis modellen geanalyseerd. Het gebruik van SIMPSA blijkt uitermate doeltreffend (lage bekomen fitness) en gebruiksvriendelijk te zijn.

Met het oog op een verdere fijnstelling van het optimalisatieproces worden de gesynthetiseerde



**Figuur 6.4:** Gemiddelde procentuele afwijking van de voorwaardelijke probabiliteiten inzake neerslagintensiteit, overheen de verschillende maanden.



**Figuur 6.5:** Vergelijkende jaarlijkse Gumbel plots in functie van het aggregatieniveau

tijdreeksen geanalyseerd. Hieruit blijkt dat alledrie de modellen de gemiddelde neerslaghoogte goed benaderen. De variantie van de neerslaghoogte wordt daarentegen, vooral door het OBL model, overschat. De covariantie wordt door alledrie de modellen bij lage aggregatieniveau's overschat en onderschat bij hogere aggregatieniveau's. Ook hier is het OBL model de zwakste van de drie. Het MBL model en het MBLG model vertonen relatief kleinere afwijkingen dan het OBL model voor de ZDP. Tot slot wordt de skewness van de verdeling van de neerslaghoogte door alle modellen, vooral bij lage aggregatieniveau's, slecht benaderd.

Uit de eerste voorwaardelijke karakteristiek (kans op twee opeenvolgend droge intervallen) blijkt dat het OBL model bij kleine aggregatieniveau's wel goede droog-nat verhoudingen oplevert, maar dat het omgekeerde waar is bij grotere aggregatieniveau's. De afwijking die daar door het OBL model wordt verkregen is heel wat hoger dan bij de aangepaste modellen. Uit de analyse van de overige voorwaardelijke distributiekarakteristieken blijkt dat alledrie de modellen het intensiteitsverloop van een regenbui, en de variabiliteit ervan sterk onderschatten. Het OBL model blijkt het verloop van de bui-intensiteit doorgaans beter te modelleren bij lagere aggregatieniveau's dan bij hogere aggregatieniveau's. Het omgekeerde geldt voor het MBL model en het MBLG model.

Naast de toch wel gebrekkige modellering van het intensiteitsverloop van de buien leert de analyse van de verdeling van de neerslagintensiteiten (de voorwaardelijke probabiliteiten inzake neerslagintensiteit) dat de verschillende modellen hiervoor vrij goede resultaten opleveren. Vooral voor het MBL model en het MBLG model zijn de afwijkingen ten opzichte van de geobserveerde waarden relatief klein. Het OBL model presteert ook hier minder goed.

Ten slotte werd ook een frequentie-analyse uitgevoerd op de bekomen synthetische neerslagreeksen. Uit de vergelijking van de bekomen jaarlijkse extremen met de geobserveerde jaarlijkse extremen blijkt dat de modellen vooral bij lage aggregatieniveau's de extreme waarden aanzienlijk onderschatten. Het MBLG model, dat werd aangepast met het oog op een betere representatie van de extreme waarden, presteert niet noemenswaardig beter dan het MBL model. Het levert gelijkaardige resultaten op bij grotere aggregatieniveau's en minder goede, in vergelijking met zowel het MBL model als het OBL model, bij de laagste aggregatieniveau's. Bij de hogere aggregatieniveau's onderschat het OBL model nog steeds de extreme waarden. Bij kleine retourperiodes wordt een vrij goede fit bekomen terwijl bij hogere retourperiodes een overschatting van de extremen wordt waargenomen.



# Hoofdstuk 7

## Introductie derde orde moment

### 7.1 Inleiding

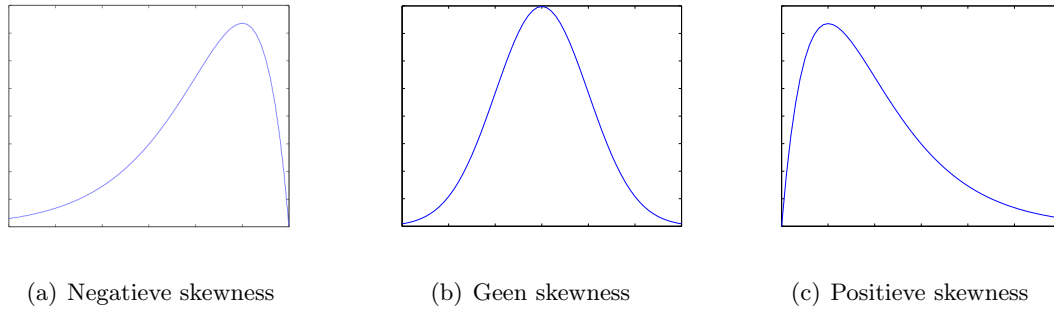
Zoals blijkt uit Hoofdstuk 6 presteren de beschouwde modellen op sommige vlakken vrij goed, terwijl voor andere criteria minder goede resultaten worden bekomen. De overeenkomst met de geobserveerde extreme waarden blijkt niet optimaal te zijn. Cowpertwait (1998) beargumenteert dat de staart van de empirische verdeling van de neerslaghoogte niet goed kan worden benaderd zonder dat bij de kalibratie wordt rekening gehouden met een of andere hogere orde statistiek. Hij voegt daarom het derde orde moment toe aan de doelfunctie voor de kalibratie van een Neyman-Scott model, wat leidt tot een betere reproductie van de extreme waarden.

Het derde orde moment is een gestandaardiseerde vorm van de *skewness* van een verdeling. De skewness of scheefheid van een statistische verdeling is een maat voor de asymmetrie van de verdeling. De normale verdeling, welke gekenmerkt wordt door zijn symmetrie, heeft bijgevolg een skewness die gelijk is aan nul, zoals getoond in Figuur 7.1(b). Naarmate de verdeling meer asymmetrisch wordt, zal de skewness toenemen. Deze kan zowel negatieve (Figuur 7.1(a)) als positieve waarden (Figuur 7.1(c)) aannemen. In dit werk wordt bij de kalibratie geen gebruik gemaakt van de skewness maar van de eerder vermelde gestandaardiseerde vorm. Het derde centrale moment  $E \left[ \left( Y_i^{(h)} - E[Y_i^{(h)}] \right)^3 \right]$ , met  $Y_i^{(h)}$  de cumulatieve neerslag over interval  $i$  met lengte  $h$ , verhoudt zich als volgt tot de skewness:

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}} = \frac{E \left[ \left( Y_i^{(h)} - E[Y_i^{(h)}] \right)^3 \right]}{E \left[ \left( Y_i^{(h)} - E[Y_i^{(h)}] \right)^2 \right]^{3/2}} \quad (7.1)$$

met  $g_1$  de skewness,  $m_3$  het derde gestandaardiseerde moment en  $m_2$  de variantie. Tijdens het verdere verloop van dit hoofdstuk en de volgende hoofdstukken zal naar het derde centrale moment verwezen worden als het derde orde moment (Weisstein, 2011).

Het loont de moeite om voor de Bartlett-Lewis modellen het effect van de introductie van het derde orde moment in de doelfunctie te onderzoeken. Dit was vooralsnog niet mogelijk, gezien hiervoor geen analytische uitdrukkingen voorhanden waren. Deze zijn sinds kort wel voor het OBL model, het MBL model en het TBL model (zie Hoofdstuk 9) beschikbaar (Onof, 2011) en kunnen bijgevolg geïmplementeerd en gebruikt worden bij de kalibratie van de respectievelijke modellen. Daartoe dienen de beschikbare analytische uitdrukkingen eerst gecontroleerd te worden op hun correctheid (Sectie 7.2). Sectie 7.3 vervolgt met de bespreking van de nieuwe



**Figuur 7.1:** Visuele voorstelling van het begrip skewness

doelfunctie, voorzien van het derde orde moment. De resultaten van de kalibratie en de evaluatie van de bekomen parameters worden besproken in Sectie 7.4.

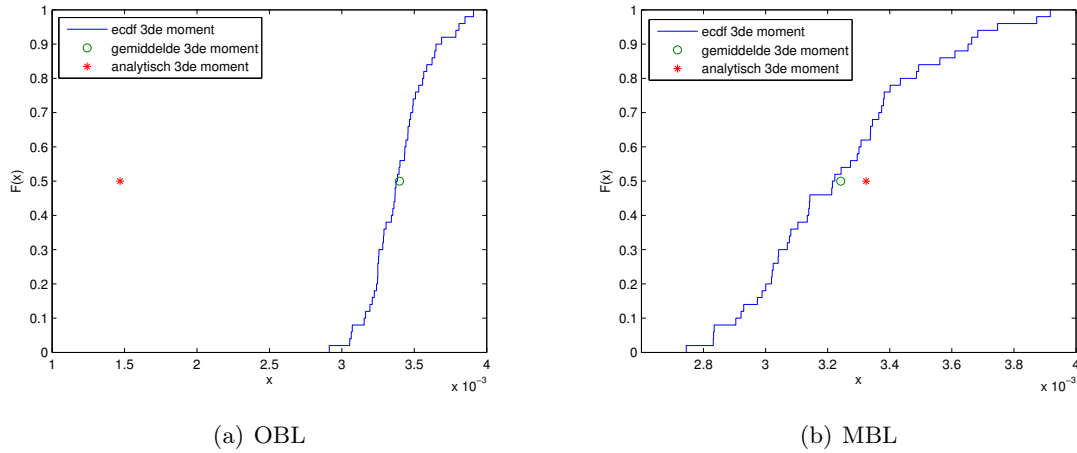
## 7.2 Analytische uitdrukking

Alvorens het derde orde moment te kunnen gebruiken in de doelfunctie, moet worden nagegaan of de analytische uitdrukkingen correct zijn. De analytische uitdrukkingen voor de lagere orde momenten werden reeds meermaals gebruikt zodat de controle ervan niet aan de orde is. De uitdrukkingen voor het derde orde moment daarentegen zijn nog niet gepubliceerd (Onof, 2011), waardoor niet met zekerheid kan gesteld worden dat ze correct zijn. Hiertoe worden met een willekeurige parameterset verscheidene simulaties uitgevoerd. Vervolgens kan de empirische distributiefunctie van de resultaten opgesteld worden. Net zoals bij de andere analytische uitdrukkingen (zie Bijlage A) wordt verwacht dat de uitdrukking van het derde orde moment bij benadering samenvalt met de mediaan van de bekomen distributie. Indien dit niet het geval is, kan besloten worden dat er iets schort met de uitdrukking en wordt deze om voor de hand liggende redenen beter niet gebruikt.

Gezien enkel voor het OBL model (zie Sectie A.1.3) en het MBL model (Sectie A.1.7) uitdrukkingen voor het derde orde moment beschikbaar zijn (het TBL model wordt in Hoofdstuk 9 besproken) wordt de analyse van de analytische uitdrukkingen hiertoe beperkt. Voor zowel het OBL model als het MBL model wordt een willekeurig parametercombinatie gekozen. Aan de hand van deze parameters wordt een ensemble van 50 simulaties van 105 jaar lengte uitgevoerd. Vervolgens wordt voor elk van de 50 herhalingen het derde orde moment op een aggregatieniveau van 10 minuten berekend. De experimentele cumulatieve distributiefunctie wordt in een grafiek uitgezet, samen met het gemiddelde van de 50 herhalingen. Het gemiddelde blijkt samen te vallen met de mediaan waardoor het gebruik ervan gerechtvaardigd is. Ten slotte wordt aan de hand van diezelfde parameterset het derde orde moment berekend en tevens uitgezet op grafiek. In Figuur 7.2 worden de resultaten voor de respectievelijke modellen weergegeven. In een oogopslag wordt duidelijk dat de uitdrukking voor het OBL model niet correct is (Figuur 7.2(a)). De analytische uitdrukking van het derde orde moment voor het MBL model blijkt daarentegen wel beter overeen te komen met de gesimuleerde waarden, hoewel ook hier de overeenkomst niet volledig correct is. De analytisch berekende waarden vallen niet samen met het gemiddelde van de gesimuleerde waarden voor het derde orde moment, maar correspondeert eerder met het 60% percentiel. Hoewel de overeenkomst dus niet perfect is, kan toch gebruik gemaakt worden van



de analytische uitdrukking voor het derde orde moment in de doelfunctie.



**Figuur 7.2:** Verificatie analytische uitdrukkingen derde orde moment

### 7.3 Doelfunctie

Net zoals in voorgaande hoofdstukken wordt bij de kalibratie gebruik gemaakt van de in Hoofdstuk 4 beschreven doelfunctie. Deze wordt geminimaliseerd met behulp van SIMPSA, gezien deze methode de meest efficiënte methode bleek te zijn in Hoofdstuk 5. Teneinde de invloed van het derde orde moment op de reproductie van de extreme waarden na te gaan wordt de analytische uitdrukking van het derde orde moment toegevoegd aan de doelfunctie. In Sectie 7.2 bleek dat de beschikbare uitdrukking voor het OBL model niet correct is. De invloed van het derde orde moment kan dus voorlopig enkel worden nagegaan aan de hand van het MBL model.

Gezien volgens Khaliq & Cunnane (1996) het gebruik van meerdere eigenschappen in de doelfunctie leidt tot een betere representatie van de historische reeks door het model, wordt ervoor geopteerd om de doelfunctie aan te vullen met twee termen voor het derde orde moment (zie ook Hoofdstuk 4). Hierbij wordt gekozen voor een aggregatieniveau van 10 minuten en 24 uur. In Tabel 7.1 worden nog eens alle gebruikte momenten weergegeven. De bekomen resultaten worden vergeleken met de in Hoofdstuk 6 bekomen resultaten voor het MBL model. Daarom wordt in Tabel 7.1 de samenstelling van de doelfunctie, gebruikt in Hoofdstuk 6, herhaald. Tijdens het verdere verloop van dit hoofdstuk wordt hiernaar gerefereerd als Fit1 en Fit2. Fit1 is de parameterset die wordt bekomen na kalibratie zonder het opnemen van het derde orde moment in de doelfunctie. Fit2 is de parameterset die bekomen wordt na kalibratie met de doelfunctie die wel het derde orde moment opneemt.

**Tabel 7.1:** Gebruikte momenten en aggregatieniveau voor doelfuncties ter verduidelijking van de invloed van het derde orde moment.

	Gemiddelde	Variantie	Covariantie	ZDP	Derde orde moment
Fit1	10 min	10 min	10 min en 24 u	10 min en 24 u	-
Fit2	10 min	10 min	10 min en 24 u	10 min en 24 u	10 min en 24 u

## 7.4 Resultaten

Voor de kalibratie van het MBL model met inbegrip van het derde orde moment, wordt opnieuw een beroep gedaan op SIMPSA. Hierbij worden dezelfde instellingen gebruikt als bij de kalibratie van het MBL model in Hoofdstukken 5 en 6, welke zijn terug te vinden in Sectie 4.5.3.7. De resulterende parameterset wordt hieronder eerst kort toegelicht (Sectie 7.4.1), waarna (cf. Hoofdstuk 6) simulatieresultaten worden vergeleken met observaties.

### 7.4.1 Parameters

In Tabel B.7 worden de geschatte parameters weergegeven. Figuur 7.3 geeft grafisch weer in welke mate de geschatte parameters van Fit1 en Fit2 verschillen. Hierop is te zien dat er duidelijke verschillen zijn opgetreden door het gebruik van het derde orde moment in de doelfunctie. Parameters  $\lambda$ ,  $\kappa$  en  $\phi$  zijn duidelijk lager bij Fit2. De overige parameters ( $\mu_x$ ,  $\alpha$  en  $\eta$ ) worden groter ingeschat tengevolge van de ingebruikname van het derde orde moment in de doelfunctie (Fit2). Figuur 7.4 geeft de verschillende fysisch interpreteerbare variabelen, afgeleid uit de parameters, weer (zie ook Sectie 6.2). Hieruit blijkt dat het model, gekalibreerd met inbegrip van het derde orde moment, minder buien creëert (lagere  $\lambda$ ) met in de wintermaanden minder neerslagcellen (lagere  $\mu_c$ ), die weliswaar gemiddeld een hogere intensiteit hebben (hogere  $\mu_x$ ). Ten slotte kan gezien worden dat de gemiddelde celduur ook kleiner wordt ingeschat (kleinere  $(\alpha\nu)^{-1}$ ). Uitgaande van de interpretatie van deze parameterwaarden lijkt het dus alsof er simpelweg minder neerslag valt. De volgende sectie zal moeten uitwijzen of deze bevinding zich ook laat voelen in de statistische eigenschappen van de synthetische neerslag.

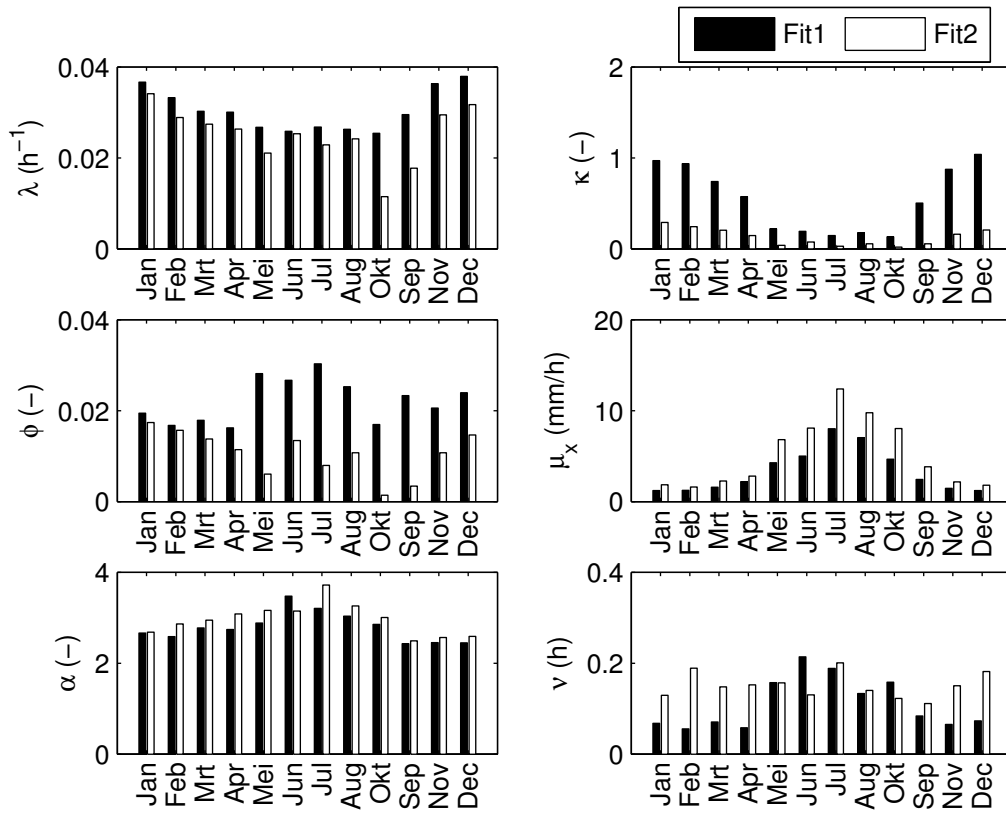
### 7.4.2 Statistische analyse

Net zoals in Hoofdstuk 6 worden voor beide parametersets 50 tijdreeksen met een lengte van 105 jaar gesimuleerd. Voor het ensemble van 50 herhalingen worden verschillende eigenschappen berekend, waarvan vervolgens de mediaan genomen wordt. Om een inschatting te kunnen maken van de prestatie van beide fits en om de interpretatie van de resultaten te vergemakkelijken wordt gebruik gemaakt van de procentuele afwijking van de beschouwde eigenschap ten opzichte van de observaties te Ukkel.

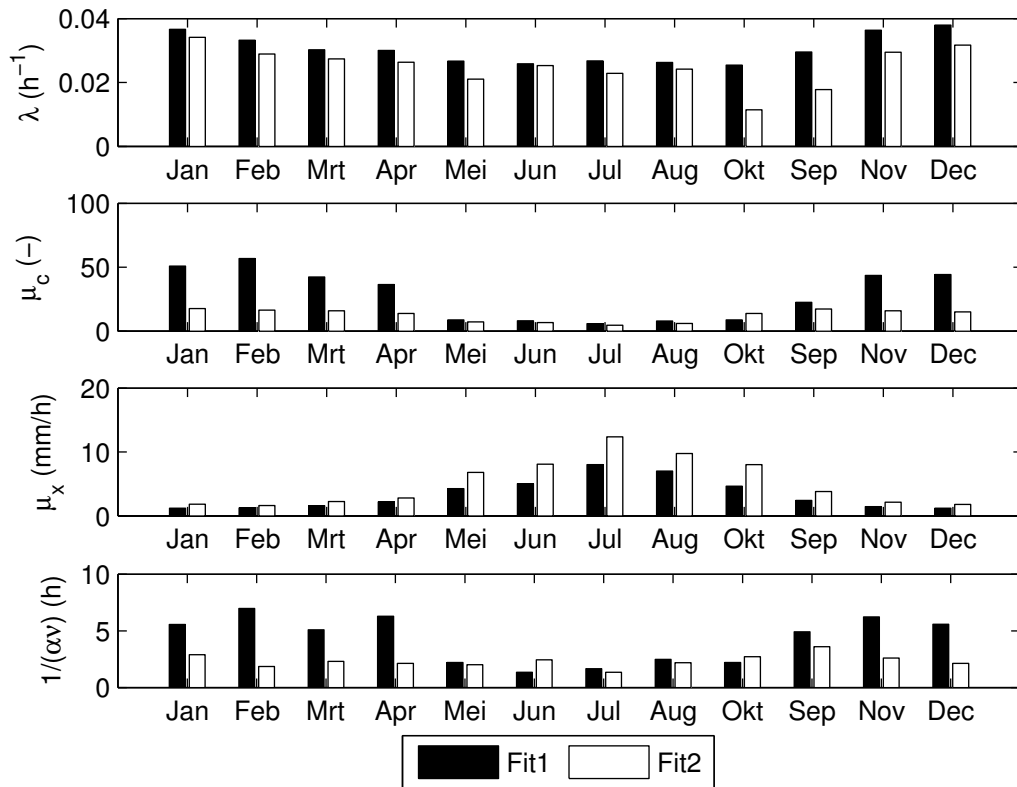
#### 7.4.2.1 Eerste, tweede en derde orde momenten

Figuur C.3 geeft deze procentuele afwijkingen voor de eerste, tweede en derde orde statistieken, voor beide parametersets, uitgezet in functie van het aggregatieniveau weer. Figuur 7.5 geeft deze eigenschappen voor de maand januari weer. De gemiddelde neerslaghoogte wordt door Fit2 systematisch overschat (Figuur 7.5(a)). Daar waar de afwijking, bekomen door Fit1, schommelt rond een absolute waarde van 1%, worden voor Fit2 waarden bekomen die tot 10% afwijken van de gemeten waarden te Ukkel. Eenzelfde besluit geldt voor het verloop van de variantie van de neerslaghoogte (Figuur 7.5(b)). Ook hier blijken de door Fit2 gesimuleerde neerslagreeksen een aanzienlijk hogere variantie te vertonen dan het geval is bij Fit1.

Voor de ZDP (Figuur 7.5(a)) kan gezien worden dat Fit2 in vrijwel alle gevallen een betere benadering oplevert, behalve voor een aggregatieniveau van 10 minuten. Ondanks de grote overschatting van de neerslaghoeveelheid en de overschatting van de variabiliteit daarvan, blijkt



**Figuur 7.3:** Vergelijking parameters Fit1 (MBL model zonder derde orde moment) en Fit2 (MBL model met derde orde moment).

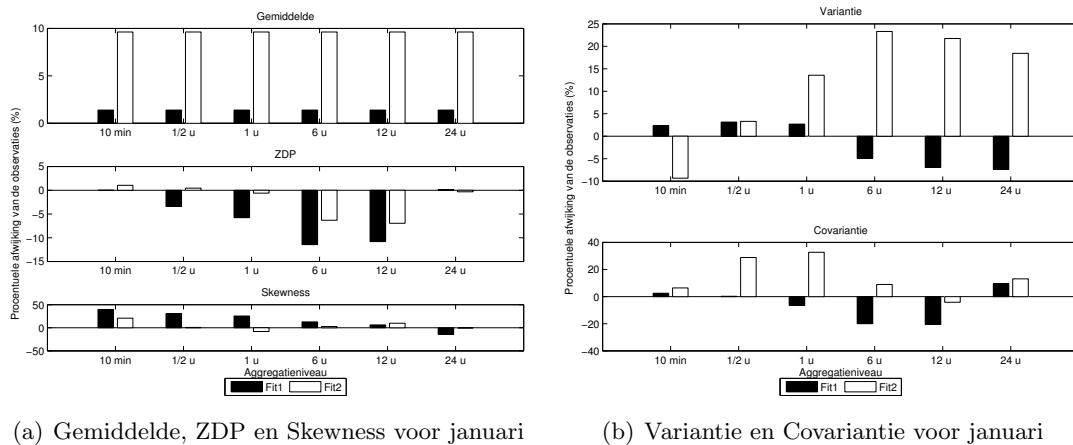


**Figuur 7.4:** Vergelijking fysieke variabelen Fit1 en Fit2.

de temporele schikking van de buien dichter aan te leunen bij de geobserveerde neerslag te Ukkel. Fit1 onderschat de proportie aan droge intervallen, wat betekent dat er te weinig droge intervallen gecreëerd worden. Er worden door Fit2 meer droge intervallen gecreëerd, terwijl de algemene neerslaghoeveelheid wordt overschat. Er kan dus verwacht worden dat de gemiddelde bui-intensiteit (vwd2) groter zal zijn bij simulatie met Fit2. In Sectie 6.3.2 werd besloten dat de OBL, MBL (Fit1) en MBLG modellen deze gemiddelde bui-intensiteit onderschatten. Het vermoeden rijst dus dat Fit2 neerslagreeksen oplevert waarbij deze voorwaarde beter benaderd wordt. Hier wordt echter dieper op ingegaan in Sectie 7.4.2.2.

De invloed van de betrekking van het derde orde moment in de doelfunctie op de covariantie van de neerslaghoeveelheid is wederom systematisch (Figuur 7.5(b)). Voor bijna alle maanden kan gezien worden dat voor aggregatieniveau's kleiner dan 6 uur een veel sterkere overschatting wordt bekomen met Fit2, dan het geval is met Fit1. Voor aggregatieniveau's groter dan 6 uur wordt dan weer een kleinere afwijking bekomen door Fit2. Hoewel deze trend voor het gros van de maanden geldt, kunnen uiteraard hier en daar kleine uitzonderingen op de regel ontdekt worden. De vraag is natuurlijk of dit dan niet eerder toe te schrijven is aan het toeval.

Als laatste wordt nagegaan hoe de skewness van de gesimuleerde data zich verhoudt tot de geobserveerde data (Figuur 7.5(a)). Er kan verwacht worden dat de skewness door Fit2 beter benaderd wordt dan door Fit1, aangezien Fit2 het resultaat is van een kalibratie met een doelfunctie waarin het derde orde moment is opgenomen. Deze veronderstelling wordt bevestigd door de figuren (Figuur 7.5(a) en C.3). Hoewel de procentuele afwijking voor een aggregatieniveau van 10 minuten nog steeds rond de 20% schommelt, kan gezien worden dat voor hogere aggregatieniveau's de skewness behoorlijk wordt benaderd door de met Fit2 gesynthetiseerde neerslagreeksen.



**Figuur 7.5:** Eerste, tweede en derde orde momenten voor januari

#### 7.4.2.2 Voorwaardelijke distributiekarakteristieken

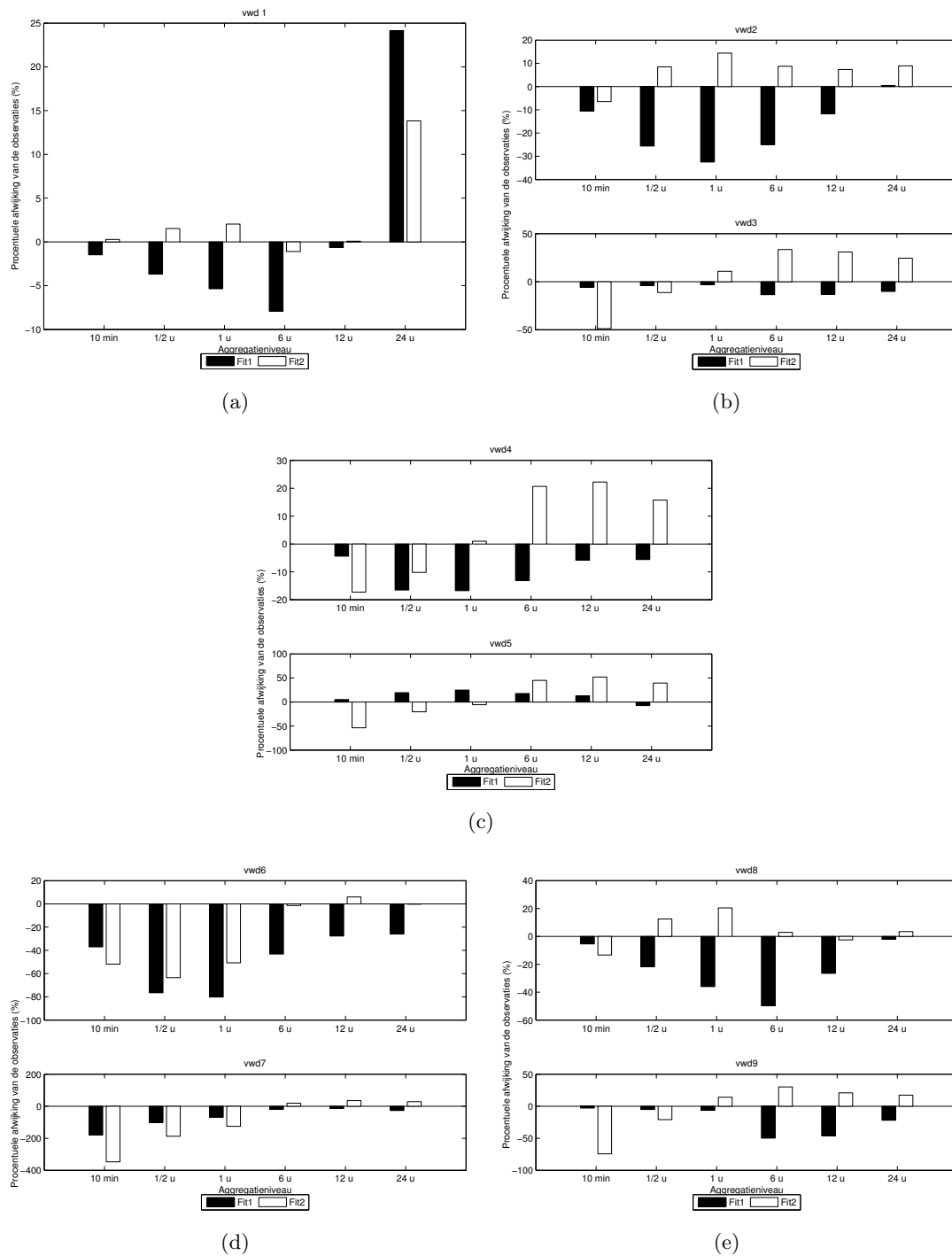
De analyse van enkele voorwaardelijke distributiekarakteristieken verschaft meer inzicht in de opbouw van de gesimuleerde neerslag. Het is de bedoeling dat de gesimuleerde neerslag niet alleen overeenstemt met de observaties op vlak van algemene statistische eigenschappen, maar ook dat de opbouw en het verloop van de gesimuleerde buien in zekere zin goede gelijknissen vertonen met de observaties. Voor meer informatie omtrent deze voorwaardelijke distributiekarakteristieken wordt verwezen naar de literatuur.

rakteristieken, hun berekeningswijze en de situatie te Ukkel wordt verwezen naar Sectie 3.2.2. Figuren 7.6(a) en 7.6(b) geven de procentuele afwijkingen van de observaties voor beide fits in functie van het aggregatieniveau weer. Voor de eerste voorwaarde, welke de kans uitdrukt dat twee opeenvolgende intervallen neerslagvrij zijn, is duidelijk te zien dat Fit2 een minder sterke afwijking ten opzichte van de geobserveerde kans vertoont. Uit sectie 7.4.2.1 bleek reeds dat Fit2 de neiging heeft om de gemiddelde neerslag te overschatten. Bovendien bleek de ZDP over het algemeen beter benaderd door Fit2. Het feit dat ook de eerste voorwaardelijke distributiekarakteristiek beter benaderd wordt door Fit2 toont aan dat de nat-droog verhoudingen beter gesimuleerd worden door Fit2. Het nadeel is dat als het regent, de neerslaghoeveelheid aanzienlijk overschat wordt. Dit blijkt des te meer uit Figuur 7.6(b). Terwijl Fit 1 de gemiddelde bui-intensiteit (vwd2) doorgaans onderschat, blijkt Fit2 deze beter te benaderen (kleinere procentuele afwijking ten opzichte van de observaties). Dit bevestigt de vermoedens die hieromtrent gerezen waren. De variantie van deze voorwaardelijke karakteristiek vertoont overigens sterkere afwijkingen ten opzichte van de geobserveerde waarden dan het geval is bij Fit1. Uit Figuren 7.6(c), 7.6(d) en 7.6(e) blijkt dat dit ook het geval is voor de andere voorwaardelijke karakteristieken hun variantie.

Door Fit1 wordt de ZDP en vwd1 onderschat, terwijl de gemiddelde neerslag toch beter benaderd wordt dan bij Fit2. Dit komt dus doordat er bij Fit 1 te veel intervallen zijn waarin er neerslag valt aan lage intensiteit. Hierdoor wordt de ZDP onderschat, want er zijn te weinig droge intervallen, en wordt de bui-intensiteit onderschat, want de gevallen neerslag wordt te veel gespreid over verschillende intervallen. Desalniettemin wordt de gemiddelde neerslaghoeveelheid wel goed benaderd, dit komt doordat bij de berekening hiervan de droge intervallen ook worden meegerekend. Deze eigenschap zegt dus enkel iets over de totale cumulatieve neerslag en niets over de verdeling ervan. Hierbij wordt meteen ook het belang van het gebruik van verschillende maatstaven bij de beoordeling van de modellen duidelijk. Voor Fit2 werd reeds duidelijk dat de verdeling van de neerslag in de tijd beter benaderd wordt (kleinere afwijking ZDP). Bovendien blijkt dat ook de gemiddelde bui-intensiteit beter benaderd wordt, er is sprake van een lichte overschatting. Ondanks een verbetering op vlak van ZDP en gemiddelde bui-intensiteit wordt het gemiddelde nu meer overschat.

Uit Figuur 7.6(c) blijkt Fit2 de gemiddelde intensiteit bij aanvang van een bui (vwd4) doorgaans slechter te benaderen dan Fit1. Net zoals bij de gemiddelde bui-intensiteit, wordt de gemiddelde intensiteit bij aanvang van een bui onderschat bij lagere aggregatieniveau's terwijl bij hogere aggregatieniveau's eerder een overschatting plaatsvindt. Deze trend is ook zichtbaar bij de variantie van deze voorwaardelijke karakteristiek (vwd5) en deze van de andere karakteristieken (vwd3, vwd7, vwd9). Voor vwd4 levert Fit2 bij intermediaire aggregatieniveau's (1/2 u en 1 u) betere benaderingen met een lagere variantie (vwd5) op.

De gemiddelde neerslag in een interval dat een droog interval voorafgaat (vwd6) wordt door Fit2 duidelijk beter benaderd dan Fit1. Vooral bij aggregatieniveau's van 6, 12 en 24 u worden slechts verwaarloosbaar kleine afwijkingen ten opzichte van de observaties bekomen, met een variantie (vwd7) die ook slechts enkele procentpunten afwijkt. Een gelijkaardig besluit geldt voor vwd8. Over het algemeen is de afwijking bekomen door Fit2 kleiner dan deze bekomen door Fit1.



**Figuur 7.6:** Gemiddelde procentuele afwijking van de voorwaardelijke distributiekarakteristieken overheen de verschillende maanden.

### 7.4.2.3 Voorwaardelijke probabiliteiten inzake neerslagintensiteit

Figuur 7.7 toont de procentuele afwijkingen ten opzichte van de observaties voor de verschillende voorwaardelijke probabiliteiten inzake neerslagintensiteit. Zoals aangetoond in Sectie 6.3.3 zijn deze afwijkingen voor Fit1 telkens relatief klein. Op Figuur 7.7 is te zien dat ook voor Fit2 deze afwijkingen relatief klein blijven, hoewel hier en daar opmerkelijke verschillen met Fit1 kunnen worden opgetekend. De procentuele afwijking van de observaties, bekomen met Fit2 blijkt over het algemeen groter te zijn dan deze bekomen door Fit1. Dit betekent dat de verdeling van de neerslagintensiteiten door Fit2 minder goed benaderd wordt.

### 7.4.2.4 Extreme waarden analyse

De statistische analyse van de simulaties, bekomen aan de hand van Fit2, wordt afgerond met een bespreking van de extreme waarden. Het initiële doel van de introductie van het derde orde moment in de doelfunctie was een betere representatie van de extreme waarden (Cowpertwait, 1998). Zoals weleer worden in Figuur 7.8 de Gumbel plots voor de verschillende aggregatieniveau's weergegeven.

Voor de aggregatieniveau's van 10 minuten tot en met een uur kan een duidelijke verbetering waargenomen worden. Voor de overige aggregatieniveau's is een visuele beoordeling moeilijker. Teneinde over een objectievere maatstaf te beschikken die een onderscheid tussen beide resultaten maakt wordt voor zowel Fit1 als Fit2 de RMSE berekend voor de bekomen extreme waarden. Zoals reeds vermeld zijn de weergegeven resultaten de mediaan van 50 simulaties van 105 jaar. Om de RMSE te berekenen wordt dan ook gebruik gemaakt van deze herhalingen. De RMSE wordt als volgt berekend:

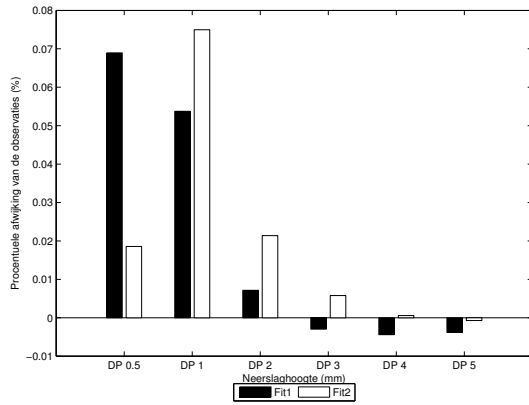
$$\text{RMSE} = \sqrt{\frac{1}{nry \cdot rep} \sum_{i=1}^{nry} \sum_{j=1}^{rep} (EV_{ij} - EV'_i)^2} \quad (7.2)$$

met  $nry$  het aantal gesimuleerde jaren,  $rep$  het aantal herhalingen,  $EV_{ij}$  de  $i$ -de grootste gesimuleerde extreme waarde voor herhaling  $j$  en  $EV'_i$  de  $i$ -de grootste geobserveerde extreme waarde. Vergelijking 7.2 is analoog aan de klassieke uitdrukking voor de RMSE, met dat verschil dat hier rekening wordt gehouden met de individuele herhalingen in plaats van bijvoorbeeld het gemiddelde van een aantal herhalingen.

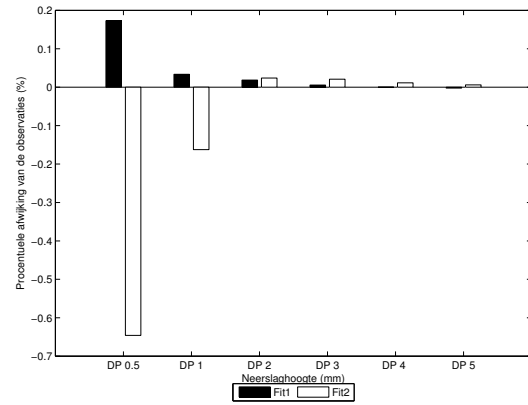
Tabel 7.2 geeft voor Fit1 en Fit2 de bekomen RMSE's weer in functie van het aggregatieniveau. Hieruit kan ondubbelzinnig besloten worden dat de introductie van het derde orde moment in de doelfunctie heeft geleid tot een betere representatie van extreme waarden voor de lagere aggregatieniveau's (10 min, 1/2 u en 1u). Voor de overige aggregatieniveau's kan gezien worden dat Fit1 steeds betere resultaten levert.

**Tabel 7.2:** RMSE extreme waarden voor Fit1 en Fit2

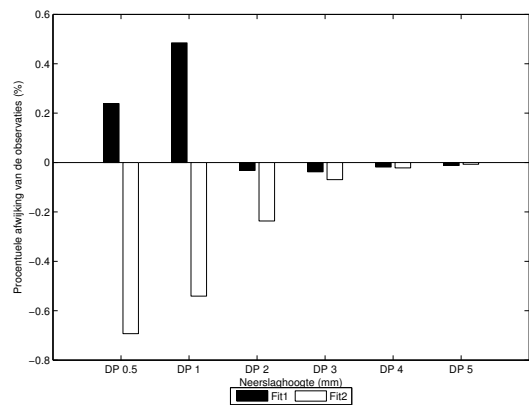
	10 min	1/2 u	1 u	6 u	12 u	24 u
Fit1	3.45	3.48	2.74	5.44	5.86	7.30
Fit2	1.91	1.78	1.90	6.34	7.93	7.94



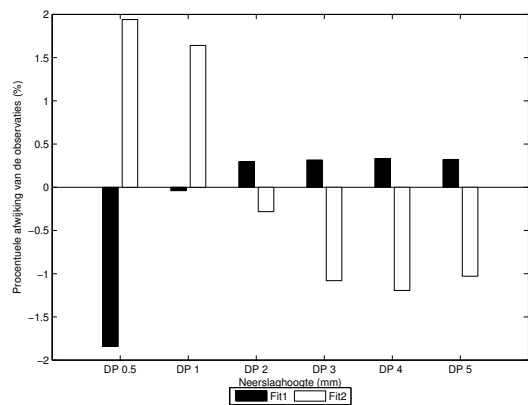
(a) 10 minuten



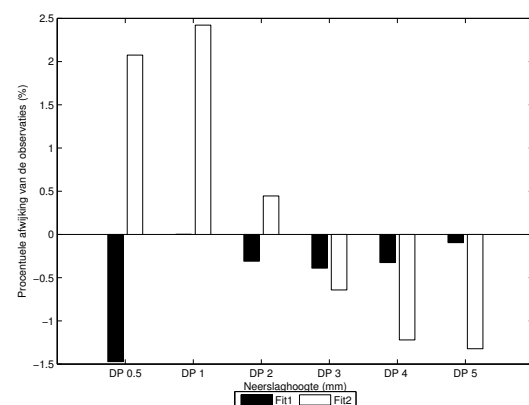
(b) 1/2 uur



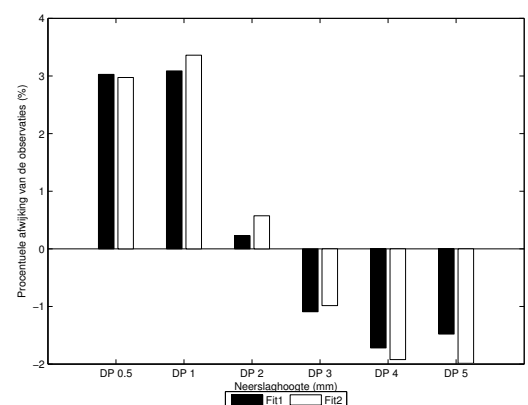
(c) 1 uur



(d) 6 uur



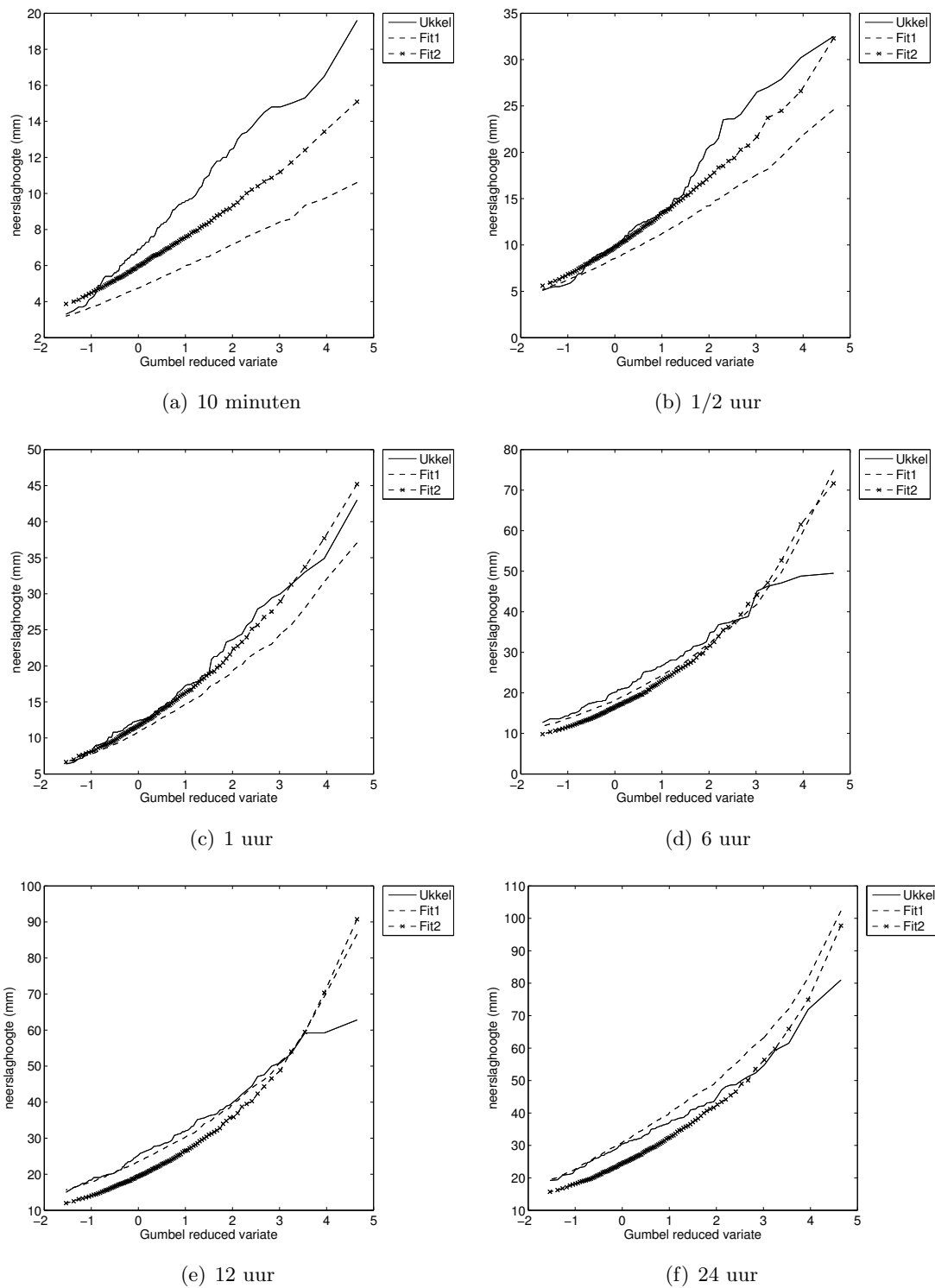
(e) 12 uur



(f) 24 uur

**Figuur 7.7:** Gemiddelde procentuele afwijking van de voorwaardelijke probabiliteiten inzake neerslagintensiteit, overheen de verschillende maanden.





**Figuur 7.8:** Vergelijkende jaarlijkse Gumbel plots in functie van het aggregatieniveau

## 7.5 Besluit

In dit hoofdstuk werd getracht om na te gaan wat de invloed van de introductie van het derde orde moment in de doelfunctie op de reproductie van de extreme waarden is. Gezien de analytische uitdrukkingen die voorhanden zijn voor het OBL model niet correct bleken te zijn, werd enkel rekening gehouden het MBL model. Er kan besloten worden dat de introductie van het derde orde moment in de doelfunctie een belangrijke invloed heeft op de prestatie van het model. Zo wordt de gemiddelde neerslaghoeveelheid sterk overschat. Bovendien varieert deze neerslaghoeveelheid meer dan het geval is bij de observaties en bij het MBL model, ook de covariantie vertoont opmerkelijk grotere afwijkingen. Eigenschappen zoals het derde orde moment en ZDP worden dan wel weer beter gerepresenteerd door het met het derde orde moment gekalibreerde model.

Uit de voorwaardelijke distributiekarakteristieken (meer bepaald  $v_{wd1}$ ) blijkt, in samenspraak met de ZDP, de introductie van het derde orde moment in de doelfunctie een positieve invloed te hebben op de nat-droog verhoudingen van de neerslag. De overige voorwaarden wijzen daarentegen op grotere fouten in vergelijking met de observaties. Hoewel de afwijking op de voorwaardelijke probabiliteiten inzake neerslagintensiteit beperkt blijven kan toch een mindere prestatie opgemerkt worden onder invloed van de bestudeerde introductie van het derde orde moment in de doelfunctie.

De extreme waarden worden voor lage aggregatieniveau's beter weergegeven na introductie van het derde orde moment. Voor hogere aggregatieniveau's is het resultaat echter omgekeerd.

Uit deze resultaten blijkt duidelijk dat de verbetering van de reproductie van de extreme waarden een duidelijke invloed heeft op andere eigenschappen van de synthetische neerslagreeks. Vooral de grove overschatting van de gemiddelde neerslag is onaanvaardbaar. De oorzaak hiervoor dient dus nader onderzocht te worden en indien mogelijk aangepast.

# Hoofdstuk 8

## Invloed doelfunctie

### 8.1 Inleiding

In Hoofdstuk 7 werd de invloed van de introductie van het derde orde moment in de doelfunctie van het MBL model nagegaan. Hieruit bleek dat dit een positieve invloed heeft op de verdeling van de extreme waarden. Een nadeel is dat de eerste en tweede orde momenten van de synthetische neerslagreeksen geen goede fit met de historische reeks meer vertonen. Het huidige hoofdstuk gaat na wat de oorzaak hiervan is en hoe dit kan verholpen worden, zonder de verbeterde representatie van de extreme waarden in het gedrang te brengen. In Sectie 8.2 wordt in de eerste plaats stilgestaan bij de oorzaak van het gestelde probleem. Op basis van de beschrijving en de visualisatie van het probleem worden in Sectie 8.3 alternatieve configuraties van de doelfunctie voorgesteld. De invloed van deze aanpassingen op de analytische momenten wordt besproken in Sectie 8.4, terwijl de invloed op de gesimuleerde momenten in Sectie 8.5 uit de doeken gedaan wordt.

### 8.2 Beschrijving en visualisatie van het probleem

Zoals reeds vermeld blijkt uit de resultaten in Hoofdstuk 7 dat de introductie van het derde orde moment in de doelfunctie ervoor zorgt dat de eerste en tweede orde momenten van de neerslagreeks minder goed benaderd worden. Om na te gaan wat hiervan de oorzaak is, wordt het aandeel van elk van de gebruikte momenten in de doelfunctie nader bekeken. In Tabel B.7 is immers te zien dat voor alle maanden de fitness voor de verschillende parametersets een stuk hoger ligt dan het geval was bij de kalibratie zonder derde orde moment (Tabel B.5). Deze fitness komt tot stand door de som te nemen van de verschillende kwadratische afwijkingen van de gebruikte momenten (zie ook Sectie 4.2). De verhoogde waarde van de doelfunctie kan het gevolg zijn van een grotere afwijking van al deze momenten in de doelfunctie, of van het feit dat bepaalde momenten wel goed, en andere minder goed benaderd worden. In dat laatste geval zou deze kennis interessant zijn om eventueel de gewichten in de doelfunctie daaraan aan te passen. Dit kan betracht worden door de minder goed benaderde termen in de doelfunctie meer gewicht te geven ten opzichte van de andere, waardoor de fouten van de in de doelfunctie opgenomen momenten ten opzichte van de geobserveerde waarden meer gespreid worden.

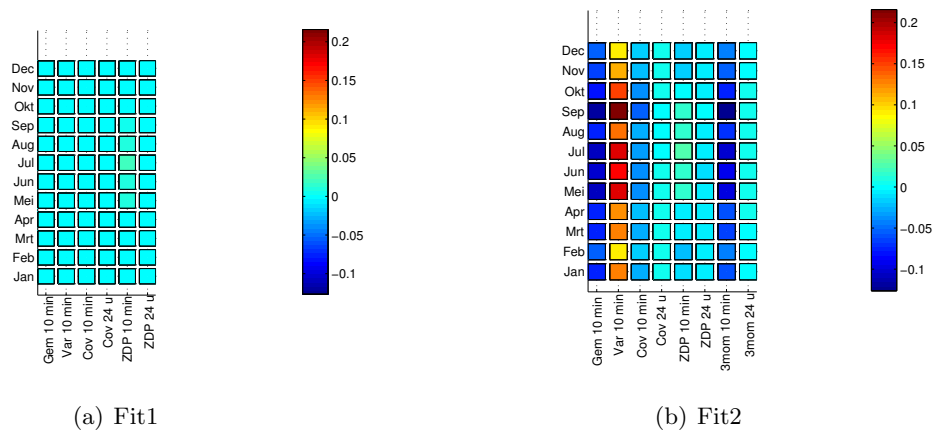
De verschillende termen in de beschouwde doelfuncties worden gevisualiseerd in Figuur 8.1. Merk op dat in de doelfunctie (Vergelijking 4.1, Sectie 4.2) gebruik gemaakt wordt van gekwadrateerde

termen. Om ook een beeld te krijgen van de gemaakte onder- of overschattingen wordt in plaats van de gekwadrateerde term hier de niet-gekwadrateerde term weergegeven. De resulterende afwijking zijn in feite analoog aan procentuele afwijkingen die gebruikt worden om de fit van de gesimuleerde momenten te beoordelen, zij het dan hier gebruik gemaakt wordt van analytische momenten (zie ook Bijlage A). Deze afgebeelde afwijking is dus telkens van de vorm:

$$\left( \frac{M_i(\vec{x})}{M'_i} - 1 \right) \tag{8.1}$$

met  $M_i(\vec{x})$  de analytische waarde voor statistiek  $i$  en  $M'_i$  de geobserveerde waarde. Figuur 8.1(a) geeft de zes termen van Fit1 (Tabel 7.1) in functie van de verschillende maanden weer. Figuur 8.1(b) geeft de acht termen van Fit2 (Tabel 7.1) weer. Figuur 8.1(a) toont aan dat de afwijkingen van de analytische momenten in de doelfunctie voor Fit1, de doelfunctie voor het MBL model zonder derde orde moment, vrij homogeen gespreid zijn overheen de verschillende gebruikte momenten en diens aggregatieniveau's. Enkel voor de maanden tussen april en oktober draagt de ZDP op 10 minuten aanzienlijk meer bij tot de fitness dan de andere termen. Op Figuur 8.1(b) is te zien dat voor Fit2 de afwijkingen minder goed gespreid zijn. Vooral de analytische waarde van de variantie op 10 minuten wijkt sterk af van de geobserveerde waarden. Daarnaast is te zien dat het gemiddelde en vooral de variantie op een aggregatieniveau van 10 minuten een hogere afwijking dan de overige termen vertonen. Deze bevinding liggen in lijn met de resultaten die in Hoofdstuk 7 werden bekomen. Daar werden voor Fit2 systematisch grotere afwijkingen voor het gemiddelde en de variantie opgetekend. Dit wordt rechtstreeks verklaard door Figuur 8.1.

Er kan dus gesteld worden dat we te kampen hebben met een onevenwicht in de doelfunctie. De analytische momenten voor het gemiddelde en de variantie worden in vergelijking met de covariantie, de ZDP en in zeker mate het derde orde moment, slecht benaderd. SIMPSA besluit de zoekprocedure stop te zetten, ook al is het duidelijk dat deze resultaten suboptimaal zijn. Om aan deze tekortkoming in de kalibratieprocedure tegemoet te komen worden enkele alternatieven met elkaar vergeleken.



**Figuur 8.1:** Visualisatie opbouwende termen van doelfuncties Fit1 en Fit2

### 8.3 Alternatieve configuraties van de doelfunctie

Totnogtoe werd bij de kalibratie van de reeds besproken modellen telkens gebruik gemaakt van de doelfunctie die in Sectie 4.2 werd voorgesteld. Om tegemoet te komen aan de problemen die zich stellen bij de introductie van het derde orde moment in de doelfunctie worden in het huidige hoofdstuk enkele alternatieve configuraties van de doelfunctie voorgesteld en geëvalueerd. Hierbij zal gebruik gemaakt worden van verschillende manieren om de gebruikte momenten in de doelfunctie te wegen. Vooralsnog werd in het huidige werk gebruik gemaakt van  $W=1$ , in navolging van Verhoest *et al.* (1997). Deze benadering leidt, zoals blijkt uit Hoofdstuk 7, niet altijd tot even goede resultaten. In een poging om de weging van de doelfunctie aan te passen en daarbij zo objectief mogelijk te werk te gaan, wordt gebruik gemaakt van twee alternatieven die onderstaande secties worden beschreven.

#### 8.3.1 Impliciete weging

Een eerste manier om de weging in de doelfunctie aan te passen, wordt voorgesteld door Cowpertwait *et al.* (2007). Deze stellen de volgende doelfunctie voor:

$$f(\vec{x}) = \sum_{i=1}^N \left[ \left( \frac{M_i(\vec{x})}{M'_i} - 1 \right)^2 + \left( \frac{M'_i}{M_i(\vec{x})} - 1 \right)^2 \right] \quad (8.2)$$

met  $M_i(\vec{x})$  de analytische waarde voor statistiek  $i$ , berekend aan de hand van parametervector  $\vec{x}$ ,  $M'_i$  de waarde van statistiek  $i$ , berekend aan de hand van de historische tijdreeks en  $N$  het aantal statistieken, op bepaalde aggregatieniveau's, gebruikt bij de kalibratie (zie ook Sectie 4.2).

In tegenstelling tot de reeds gebruikte doelfunctie (Vergelijking 4.1) heeft Vergelijking 8.2 per gebruikte eigenschap twee termen. In de ene term wordt de analytische waarde van de eigenschap gedeeld door de observatie, terwijl in de tweede term net het omgekeerde gebeurt. Dit helpt om er voor te zorgen dat het gevonden optimum geen bias boven of onder de observaties vertoont, wanneer de exacte oplossing niet gevonden wordt (Cowpertwait *et al.*, 2007).

#### 8.3.2 Relatieve bijdrage aan de fitness

Een alternatieve manier om op objectieve wijze een weging in de doelfunctie aan te brengen is door gebruik te maken van informatie uit voorgaande kalibraties. Zo werd in Hoofdstuk 7 en in Sectie 8.2 opgemerkt dat wanneer het derde orde moment wordt gebruikt bij de kalibratie van het MBL model, de fit van het gemiddelde en de variantie met de observaties niet meer optimaal is. Op basis van de relatieve bijdrage aan de fitness van de verschillende termen in de doelfunctie kunnen gewichten worden afgeleid die in een daaropvolgende kalibratie een groter gewicht zullen toekennen aan die termen die voorlopig slecht benaderd worden.

Er wordt dus als volgt te werk gegaan. Net zoals in Figuur 8.1 kunnen de verschillende termen van de doelfunctie apart berekend worden voor de verschillende maanden. Let wel dat bij de berekening van de relatieve bijdrage aan de fitness gebruik gemaakt wordt van de kwadratische termen, zoals die expliciet gebruikt worden in de doelfunctie (zie Vergelijking 4.1). Daaropvolgend kan voor elke maand de procentuele bijdrage van elke term aan de fitness worden berekend. Daaropvolgend kan de gemiddelde procentuele bijdrage aan de fitness, overheen de verschillende maanden, worden berekend. Deze gemiddelde procentuele bijdrage aan de fitness wordt dan bij kalibratie gebruikt als gewicht. De variantie op een aggregatieniveau van 10

minuten bijvoorbeeld, zal telkens een grote bijdrage leveren aan de doelfunctie en bijgevolg een groot gewicht krijgen. Als dit principe toegepast wordt op Fit1 en Fit2, bekomen in Hoofdstuk 6 en Hoofdstuk 7 (zie Tabel B.5 en Tabel B.7), worden de volgende gewichten bekomen:

**Tabel 8.1:** Relatieve bijdrage aan de fitness, gebruikt als gewichten (in %).

Fit1	Gem 10 min	Var 10 min	Cov 10 min	Cov 24 u	ZDP 10 min	ZDP 24 u		
$W_i$	2.5	0.54	0.47	4.2E-03	95.4	1.05		
Fit2	Gem 10 min	Var 10 min	Cov 10 min	Cov 24 u	ZDP 10 min	ZDP 24 u	3mom 10 min	3mom 24 u
$W_i$	20.3	59.1	2.8	0.1	0.02	16.9	0.7	0.03

## 8.4 Invloed op analytische momenten

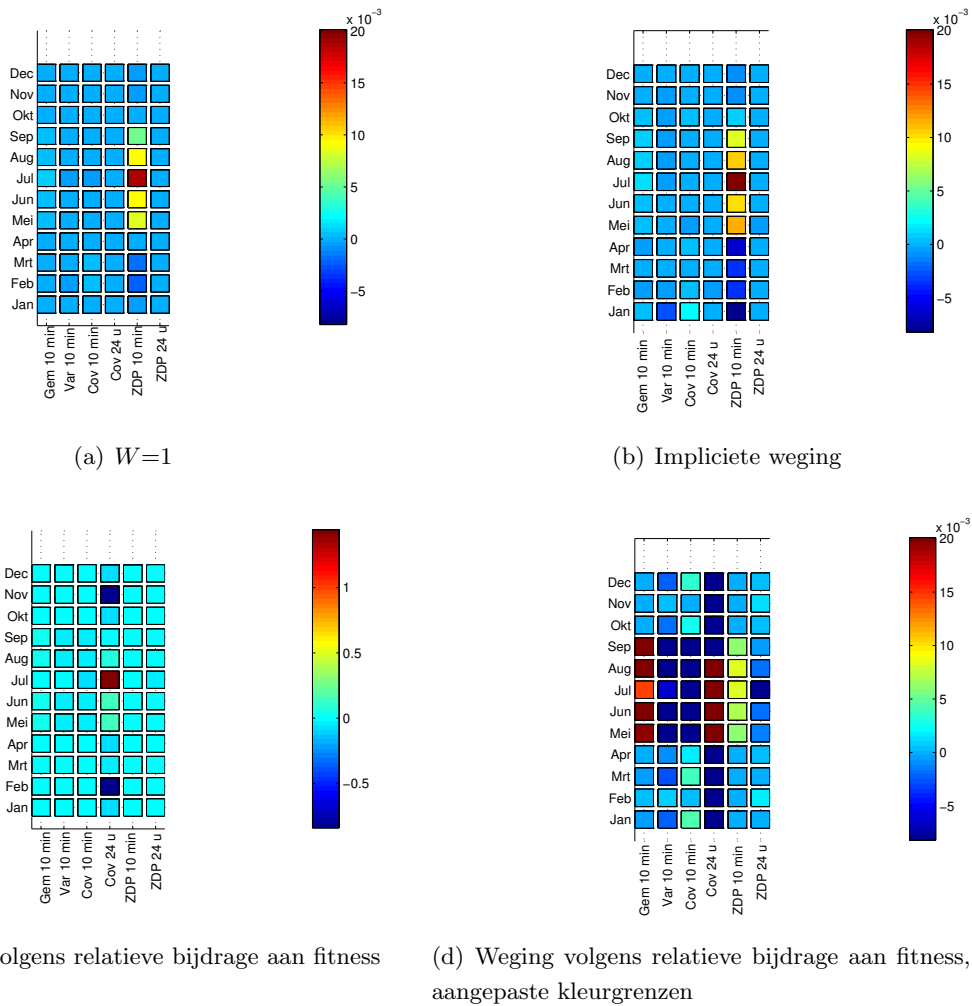
Om de invloed van de voorgestelde alternatieve doelfuncties op de fit van de verschillende momenten na te gaan, wordt in de eerste plaats onderzocht hoe de analytische momenten in de doelfunctie hierop reageren. Naar analogie met Sectie 8.2 worden de verschillende termen van de doelfunctie uitgezet in functie van de maand.

### 8.4.1 Doelfunctie zonder derde orde moment

In Figuur 8.2 worden de bekomen afwijkingen van de analytische momenten, berekend volgens Vergelijking 8.1, weergegeven. Figuur 8.2(a) geeft deze afwijkingen weer voor de reeds meermaals besproken parameters voor het MBL model, bekomen aan de hand van de klassieke doelfunctie (zie Sectie 4.2). Deze parameters zijn tevens terug te vinden in Tabel B.5. De resultaten, weergegeven in Figuur 8.2(b), zijn bekomen aan de hand van kalibratie met impliciete weging, zoals besproken in Sectie 8.3.1. Deze parameters zijn ook terug te vinden in Tabel B.8. Indien de vergelijking tussen beide figuren wordt gemaakt, kan opgemerkt worden dat bij het gebruik van de originele doelfunctie (Figuur 8.2(a)) er een overwegend positieve bias is in de termen van de doelfunctie. Dit is vooral te zien bij de ZDP op een aggregatieniveau van 10 minuten. Er kan dus verwacht worden dat dit ook het geval zal zijn bij simulatie. Figuur 8.2(b) toont de afwijkingen bekomen door kalibratie met een impliciete weging. Voor de ZDP op 10 minuten is te zien dat voor bijvoorbeeld de maand januari, de bekomen afwijking of bias meer negatief is dan bij de kalibratie zonder weging. Dit lijkt erop te wijzen dat de algemene positieve bias inderdaad verkleind wordt. Voor de maand juli blijft de afwijking van de ZDP op 10 minuten daarentegen ongeveer gelijk. De impliciete weging blijkt weinig invloed te hebben op de algemene verdeling van de afwijkingen, het is nog steeds de ZDP op 10 minuten waarvoor de analytische waarden de grootste afwijkingen vertonen ten opzichte van de observaties.

Figuur 8.2(c) toont op haar beurt de geschatte parameters die het gevolg zijn van een weging volgens relatieve bijdrage aan de fitness (zie Tabel B.10). Merk op dat op deze figuur een andere kleurenschaal gehanteerd wordt dan bij Figuur 8.2(a) en 8.2(b). Dit komt omdat we te maken hebben met enkele uitschieters voor de covariantie op 24 u, meer bepaald voor de maand februari, juli en november. Dit maakt het uiteraard moeilijk om de overige termen te beoordelen. In ieder geval kan gesteld worden dat wanneer gebruik gemaakt wordt van de relatieve weging volgens bijdrage aan de fitness de kalibratieprocedure moeizamer verloopt. Er zal meer aandacht moeten worden besteed aan het ontwijken van lokale minima, zoals blijkt

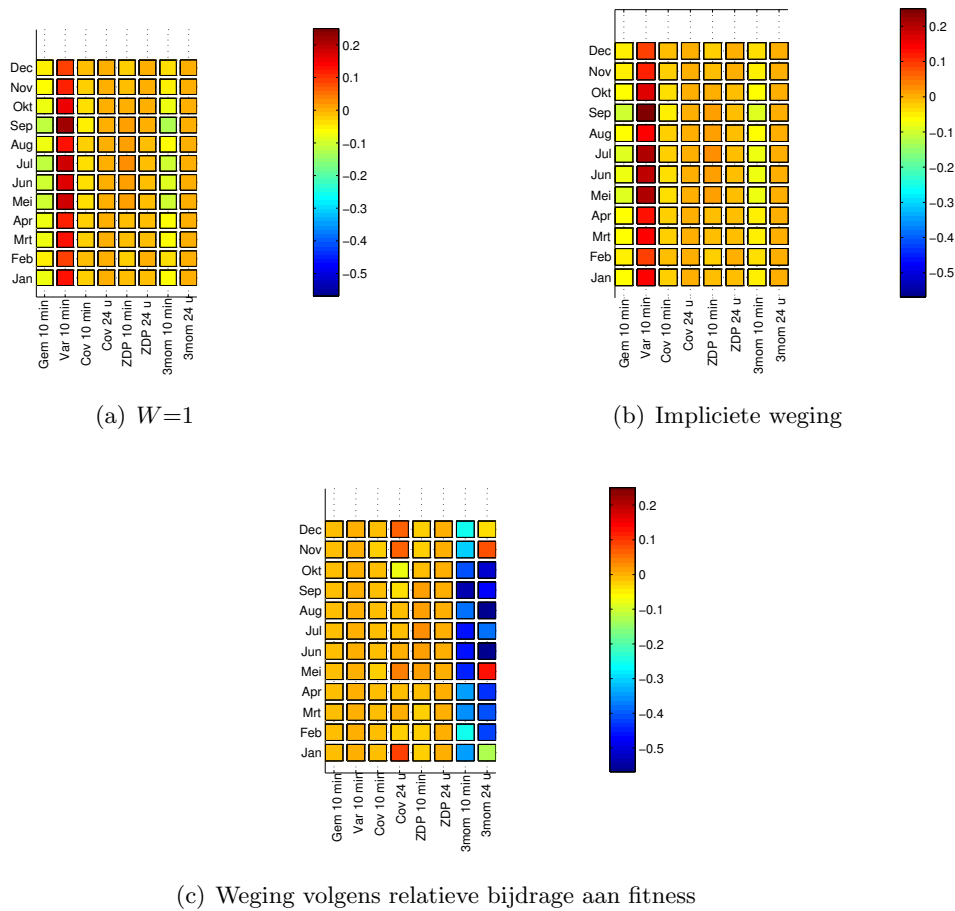
uit het voorkomen van de uitschieters in het resultaat. Figuur 8.2(d) geeft ter vergelijking met de oorspronkelijke doelfunctie en de impliciete weging een plot waarbij de kleurgrenzen zijn aangepast. Hierbij dient rekening gehouden te worden met het feit dat bepaalde waarden niet meer getrouw worden weergegeven door de kleurenschaal. Er dient dus enkel aandacht gegeven te worden aan de waarden die zich binnen de grenzen van de schaal bevinden. Die waarden die samenvallen met de grenzen overschrijden deze, maar op basis van de kleur kan niet gezien worden hoeveel deze grens overschreden wordt. De figuur toont aan dat vooral in de zomermaanden aanzienlijk slechtere resultaten worden bekomen dan het geval was bij de eerste twee methoden, dit vooral voor het gemiddelde, de variantie en de covariantie. De ZDP wordt zoals beoogd beter benaderd, weliswaar ten koste van de overige momenten. Algemeen kan dus gesteld worden dat de weging volgens relatieve bijdrage aan de fitness de kalibratie bemoeilijkt, waardoor het gros van de momenten minder goed benaderd wordt.



**Figuur 8.2:** Visualisatie afwijkingen niet-gekwadrateerde termen van doelfunctie zonder en met impliciete weging.

### 8.4.2 Doelfunctie met derde orde moment

Het gebruik van een alternatieve doelfunctie heeft vooral het verbeteren van de benadering van het gemiddelde en de variantie, wanneer het MBL model gekalibreerd wordt met inbegrip van het derde orde moment, tot doel. Hiertoe wordt de invloed van de impliciete weging, alsook de invloed van de weging naar de relatieve bijdrage aan de fitness nagegaan. Figuur 8.3 geeft de resultaten weer voor deze drie alternatieven. Net zoals in Sectie 8.4.1 kan worden besloten dat het gebruik van de impliciete weging (Figuur 8.3(b), Tabel B.9) niet tot een noemenswaardige verandering van de geschatte analytische momenten leidt. Figuur 8.3(c) toont aan dat het gebruik van de weging volgens de relatieve bijdrage aan de fitness wel heeft geleid tot een aanzienlijke verandering in de resultaten (zie ook Tabel B.11). De figuur toont aan dat zowel het gemiddelde als de variantie nu wel goed benaderd worden. Het gevolg hiervan is dat het derde orde moment minder goed geschat wordt. Over het algemeen wordt een onderschatting van het derde orde moment op beide aggregatieniveaus wordt bekomen. Analyse van de gesimuleerde neerslagreeksen zal moeten uitwijzen of het gebruik van het derde orde moment in dit geval nog steeds leidt tot een verbeterde representatie van de extreme waarden.



**Figuur 8.3:** Visualisatie afwijkingen niet-gekwadrateerde termen van doelfunctie zonder en met impliciete weging, met derde orde moment.



## 8.5 Invloed op de gesimuleerde momenten

Om na te gaan wat de invloed van de voorgestelde alternatieven van de doelfunctie op de gesimuleerde neerslagreeksen en diens eigenschappen is, wordt voor elke bekomen parameterset simulaties uitgevoerd. Net zoals in voorgaande hoofdstukken wordt een ensemble van 50 simulaties uitgevoerd. De gesynthetiseerde neerslagreeksen hebben een duur van 105 jaar. Vervolgens worden verschillende eigenschappen van de neerslagreeksen berekend en wordt de mediaan overheen het ensemble van simulaties onderling vergeleken. Om een inschatting van de overeenkomst met de observaties te kunnen maken, worden telkens de procentuele afwijkingen van de verschillende eigenschappen ten opzichte van de observaties gepresenteerd.

### 8.5.1 Statistische analyse

In totaal worden zes parametersets met elkaar vergeleken. Een parameterset die gekalibreerd is aan de hand van de doelfunctie uit Sectie 4.2 worden bestempeld als DF1. Parameters bekomen aan de hand van een impliciet gewogen doelfunctie (Sectie 8.3.1) krijgen het label DF2 en parameters die het gevolg zijn van kalibratie aan de hand van de doelfunctie, gewogen volgens de bijdrage aan de fitness, wordt bestempeld als DF3. Daarnaast wordt ook een onderscheid gemaakt naargelang de momenten die gebruikt worden in doelfunctie. Hierbij wordt een onderscheid gemaakt tussen Fit1 en Fit2, zoals weergegeven in Tabel 7.1. Fit1 duidt op een doelfunctie voor de kalibratie van het MBL model zonder derde orde moment, terwijl bij Fit2 het derde orde moment wel werd gebruikt bij de kalibratie.

#### 8.5.1.1 Eerste, tweede en derde orde momenten

In Figuren C.4 geven de procentuele afwijkingen van de verschillende berekende eerste, tweede en derde orde momenten ten opzichte van de observaties per maand weer. Om het overzicht te behouden worden hieronder de resultaten voor de maand januari weergegeven, voor de overige maanden kunnen gelijkaardige besluiten getrokken worden (Figuur C.4). Hierop kan gezien worden dat de gemiddelde neerslaghoogte (Figuur 8.4(a)) slecht benaderd wordt indien het derde orde moment wordt meegenomen in de kalibratieprocedure (Fit2), zowel wanneer DF1 als DF2 gebruikt worden. Wanneer echter de momenten in de doelfunctie gewogen worden volgens hun bijdrage aan de doelfunctie tijdens preliminaire kalibraties is de afwijking sterk verminderd voor (Fit2 DF3). Het gebruik van impliciete weging (DF2) leidt tot een lichte verbetering naar de afwijking van het gemiddelde ten opzichte van de observaties toe, zij het dat de verbetering te verwaarlozen is. Het gebruik van de alternatieve doelfunctie leidt niet tot noemenswaardige verschillen bij Fit1. De resultaten voor de gesimuleerde momenten ligt dus in dezelfde lijn als deze bekomen voor de analytische momenten (Sectie 8.4.2).

De analyse van de analytische momenten toonde ook aan dat het gebruik van het derde orde moment in de doelfunctie leidde tot een relatief grote afwijking op de variantie van de neerslaghoogte, ten opzichte van de observaties. Op Figuren 8.4(b) en C.4 is te zien dat Fit1 DF1 nog steeds de geobserveerde variantie als beste benadert. De impliciete weging van Fit1 (Fit1 DF2) leidt in de periode oktober-april tot iets betere resultaten dan Fit1 DF1. In de zomermaanden is de afwijking van Fit1 DF2 aanzienlijk groter dan deze van Fit1 DF1. Fit1 DF3 presteert slechter dan Fit1 DF1 en DF2 in de zomermaanden, terwijl 's winters gelijkaardige resultaten

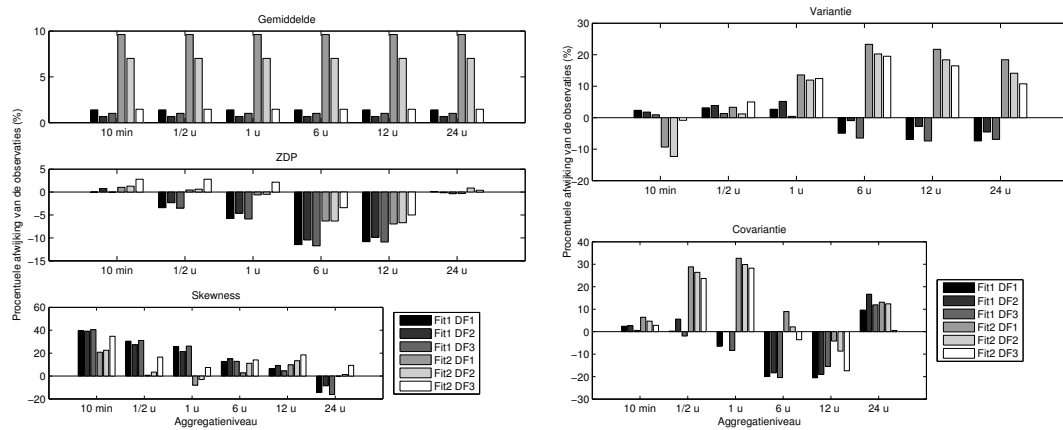
worden bekomen. Dit steekt het vermoeden dat rees bij Figuur 8.2(d) kracht bij. Voor Fit2 leidt het gebruik van DF2 ertoe dat de variantie beter wordt benaderd, zij het dat de afwijking in vergelijking met Fit1 DF1 nog steeds te groot is. Het gebruik van DF3 had vooral tot doel het verkleinen van de afwijking van het gemiddelde en de variantie ten opzichte van de observaties ingeval het derde orde moment gebruikt wordt in de doelfunctie. Er is te zien dat dit goed werkt voor een aggregatieniveau van 10 minuten. Daarvoor is de afwijking van Fit2 DF3 veel kleiner dan Fit2 DF1 en DF2. Voor grotere aggregatieniveau's wordt daarom beter gebruik gemaakt van DF1 of DF2, gezien de weging volgens bijdrage aan de fitness hier leidt tot een nog grotere afwijking van de variantie van de neerslaghoogte ten opzichte van de geobserveerde neerslagreeksen. Een mogelijke oplossing hiervoor zou het opnemen van meerdere aggregatieniveau's van de variantie in de doelfunctie kunnen zijn.

Voor de covariantie kunnen gelijkaardige besluiten getrokken worden (Figuur 8.4(b)). Het gebruik van de impliciete weging leidt in de winterse maanden tot een iets betere benadering van de covariantie in het geval van Fit1. In de zomermaanden is de afwijking die bekomen wordt door Fit1 DF2 aanzienlijk groter dan deze bekomen door Fit1 DF1. Zoals reeds bleek uit de analyse van de analytische waarden (Sectie 8.4.1) wordt de covariantie door Fit1 DF3 doorgaans minder goed ingeschat. Dit blijkt uit de analyse van de verschillende maanden. Figuur 8.4(b) zou hieromtrent een vertekend beeld kunnen geven, gezien voor januari wel een vrij goed resultaat bekomen wordt. Voor Fit2 is er geen verschil tussen zomer- en wintermaanden, zoals het geval was bij Fit1 DF2. Het gebruik van de impliciete weging leidt meestal tot een iets betere benadering van de covariantie van de neerslaghoogte ten opzichte van de observaties. DF3 presteert bij een aggregatieniveau van 10 minuten en 24 uur vrij goed, soms zelfs beter dan Fit1. In de tussenliggende aggregatieniveau's vertoont DF3 procentuele afwijkingen die gelijkaardig zijn aan deze bekomen door FIT2.

Aan de hand van de figuren kunnen voor de reproductie van de ZDP twee besluiten genomen worden (Figuur 8.4(a)). Ten eerste, dat het gebruik van impliciete weging over het algemeen leidt tot een iets beter resultaat dan wanneer dit niet gebeurt, en ten opzichte van het gebruik van DF3. Ten tweede, dat het gebruik van Fit2 DF1 en DF2 ervoor zorgt dat de ZDP overheen de verschillende aggregatieniveau's vrij goed wordt ingeschat, wat in Hoofdstuk 7 reeds duidelijk werd. Bij kleine aggregatieniveau's, zoals 10 minuten, is de overeenkomst met de observaties door de andere fits beter, wat ook het geval is voor 24 uur (de aggregatieniveau's opgenomen in de doelfunctie). De tussenliggende aggregatieniveau's worden echter door Fit2 DF1 en DF2 beter benaderd, wat ertoe leidt dat deze parametersets over de hele lijn het beste resultaat kunnen voorleggen. Voor Fit1 DF3 werd een lichte verbetering van de benadering van de analytische waarden opgetekend. Dit vertaalt zich echter niet tot een betere benadering van de observaties door de simulaties.

De afwijking van de skewness van de neerslaghoogte in de synthetische reeksen ten opzichte van deze waargenomen te Ukkel, is kleiner voor Fit2, aangezien het derde orde moment daar bij de kalibratie in de doelfunctie werd opgenomen (Figuur 8.4(a)). Het gebruik van DF3 geeft een kleiner gewicht aan het derde orde moment in de doelfunctie, om er voor te zorgen dat de andere momenten zoals het gemiddelde en de variantie in de eerste plaats goed benaderd zouden worden. Het is dan ook logisch dat in de synthetische reeksen, die aan de hand van Fit2 DF3 werden bekomen, de skewness meer afwijkt van de observatie. Deze afwijking is echter doorgaans wel kleiner dan deze bekomen door Fit1, wanneer het derde orde-moment niet meegerekend wordt

tijdens de kalibratie. Analyse van de extreme waarden zal moeten uitwijzen of een lichte verbetering van de fit van de skewness van de neerslaghoogte tot een betere representatie van de extreme waarden verdeling leidt of niet. Het gebruik van impliciete weging heeft geen duidelijk effect op de fit van de skewness, of het effect is althans te verwaarlozen.



(a) Gemiddelde, ZDP en Skewness voor januari

(b) Variantie en Covariantie voor januari

**Figuur 8.4:** Eerste, tweede en derde orde momenten voor januari

### 8.5.1.2 Voorwaardelijke distributiekarakteristieken

De procentuele afwijking van de gesimuleerde neerslagreeksen ten opzichte van de observaties voor de in Sectie 3.2.2 beschreven voorwaardelijke distributiekarakteristieken wordt weergegeven in Figuren 8.5(a) en 8.5. Figuur 8.5(a) toont de kans op twee opeenvolgend droge intervallen. Aangezien in Sectie 8.5.1.1 besloten werd dat de ZDP beter benaderd wordt door Fit2 DF1 en Fit2 DF2, wordt verwacht dat deze ook hier beter zullen zijn. Figuur 8.5(a) bevestigt dit vermoeden. vwd1 wordt doorgaans beter ingeschat wanneer het derde orde moment wordt opgenomen in de doelfunctie (Fit2). Daarnaast leidt het gebruik van de impliciete weging tot een betere benadering van vwd1, over de hele lijn voor Fit1 en voor Fit2 bij lagere aggregatieniveau's. De weging volgens relatieve bijdrage aan de fitness leidt tot beter resultaten voor Fit1. Voor Fit2 zijn de resultaten tengevolge van de relatieve weging beter voor grotere aggregatieniveau's en minder goed voor de lagere niveau's.

Voor de overige voorwaardelijke distributiekarakteristieken lijkt het gebruik van de impliciete weging in de doelfunctie en de weging volgens relatieve bijdrage aan de fitness telkens een positieve invloed te hebben, waarbij DF3 telkens iets beter dan DF2 presteert. De verwachtingswaarde van de gemiddelde bui-intensiteit wordt voor Fit2 bij een aggregatieniveau van 10 minuten, 6 uur, 12 uur en 24 uur, aanzienlijk beter benaderd wanneer de doelfunctie gewogen wordt naar de bijdragen aan de fitness (DF3), terwijl de invloed van DF3 op Fit1 eenduidig beter is. Dit is te zien op Figuur 8.5(b). Op een aggregatieniveau van een half uur en een uur presteert Fit2 DF3 minder goed. De variantie van deze gemiddelde bui-intensiteit (vwd3) wordt dan weer minder goed weergegeven door deze DF3.

Voorwaarden 4, 6 en 8 geven de gemiddelde neerslaghoogte in het begin van een bui, op het einde en ergens middenin, weer. Voorwaarden 5, 7 en 9 geven de respectievelijke varianties aan. Net zoals bij de algemene gemiddelde bui-intensiteit kan gezien worden dat de kleinste procentuele

afwijkingen het gevolg zijn van een kalibratie met gewichten volgens de relatieve bijdrage aan de fitness. Uiteraard zijn voor bepaalde aggregatieniveau's uitzonderingen waar te nemen maar er kan toch gesproken worden van een bepaalde tendens. Daarbij komt nog dat de varianties afwisselend beter en slechter benaderd worden door DF3, wat het vormen van een algemeen besluit weeral bemoeilijkt.

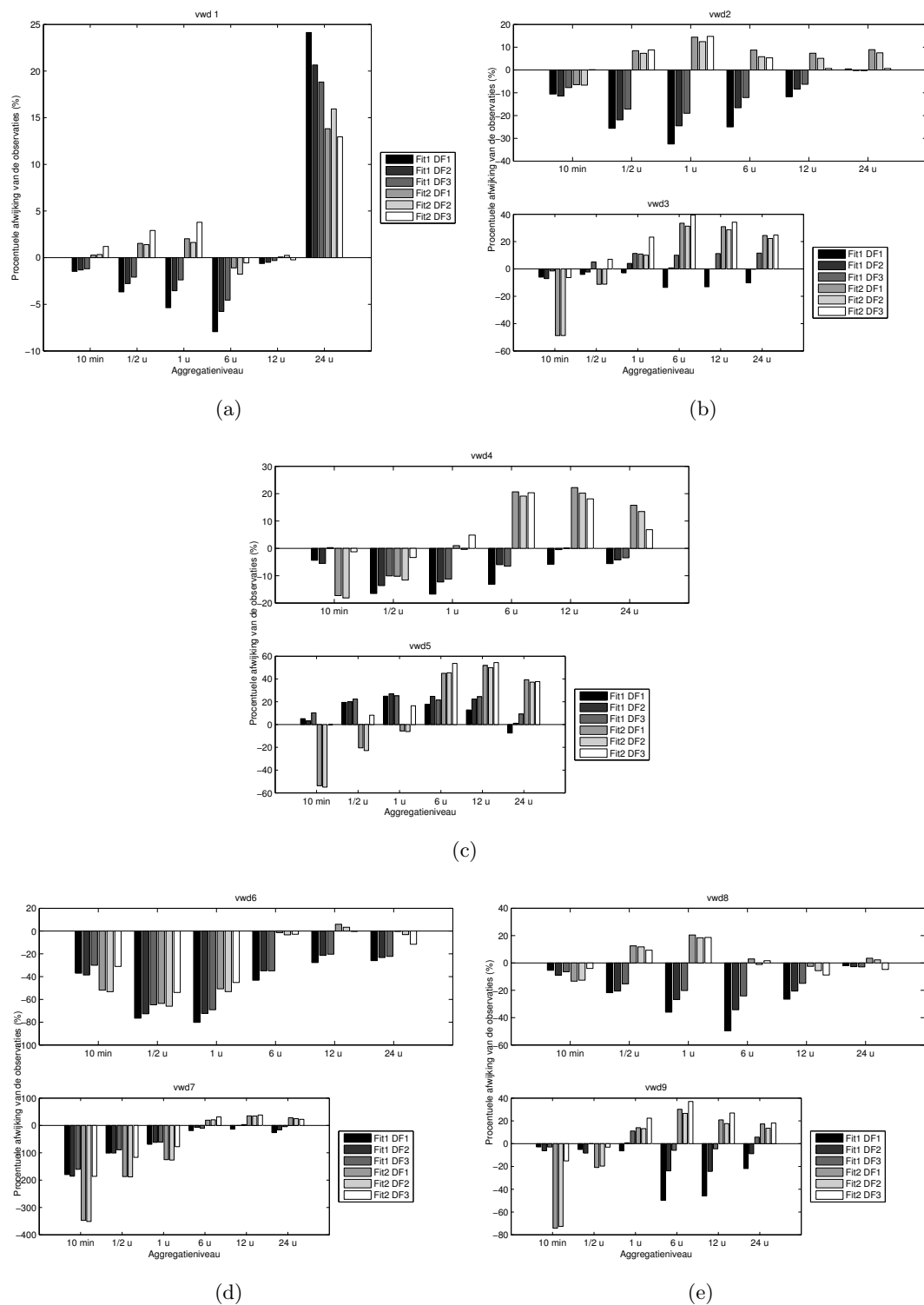
### 8.5.1.3 Voorwaardelijke probabiliteiten inzake neerslagintensiteit

De verdeling van de neerslagintensiteit wordt bestudeerd aan de hand van enkele voorwaardelijke karakteristieken inzake neerslagintensiteit. Figuur 8.6 geeft de procentuele afwijking van deze probabiliteiten ten opzichte van de observaties. Hierbij valt op dat DF3 leidt tot een foute verdeling van de neerslagintensiteit, voor zowel Fit1 als Fit2. Vooral bij grotere aggregatieniveau's wordt te veel neerslag met een lage intensiteit gegenereerd. Bijvoorbeeld bij een aggregatieniveau van 6 uur is te zien dat de kans dat de neerslaghoogte kleiner is dan 0.5 mm tot bijna 5% wordt overschat door Fit2 DF3, terwijl de andere parametersets hier kleinere afwijkingen tonen (maximaal ongeveer 2%). Bij grotere drempelwaarden presteert Fit2 DF3 dan weer beter. Dit is bijvoorbeeld te zien op Figuur 8.6(e), waar  $DP_4$  door Fit2 DF3 slechts heel licht overschat wordt. Omtrent het gebruik van de impliciete weging kan moeilijk een uitspraak gedaan worden gezien de resultaten niet opmerkelijk beter of slechter zijn dan zonder impliciete weging.

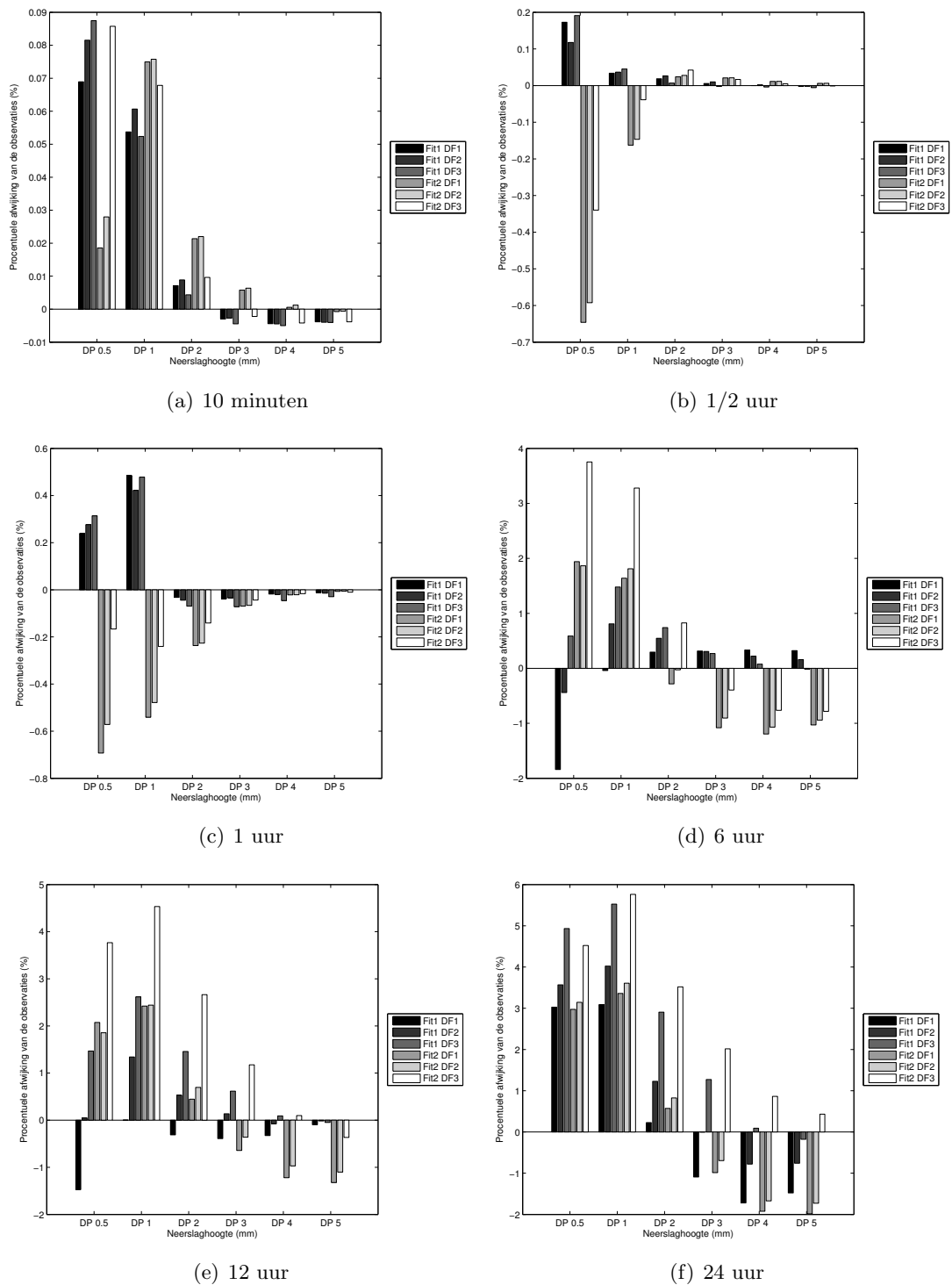
### 8.5.1.4 Extreme waarden analyse

In Hoofdstuk 7 bleek de introductie van het derde orde moment in de doelfunctie een positieve invloed te hebben op de reproductie van extreme waarden bij de lagere aggregatieniveau's. De belangrijkste vraag die bijgevolg in deze sectie kan gesteld worden is of de relatief kleine verbetering van de reproductie van de skewness (zie Sectie 8.5.1.1) door het invoeren van een weging volgens de relatieve bijdrage aan de fitness een bevorderend effect heeft op de extreme waarden. Figuur 8.7 toont de plots voor de jaarlijkse extremen in functie van het aggregatieniveau. Wanneer Fit1 DF1 vergeleken wordt met Fit2 DF3, kan gezien worden dat voor vrijwel alle aggregatieniveau's Fit1 DF1 beter aanleunt bij de historische extreme waarden. Fit1 DF3 lijkt daarentegen wel een verbetering teweeg te brengen. Dit kan vooral opgemerkt worden bij hogere GRV's. Voor Fit2 DF3 kan besloten worden dat de minimale verbetering van de fit van de skewness, geobserveerd in Sectie 8.5.1.1, niet leidt tot een verbetering van de reproductie van de extreme waarden.

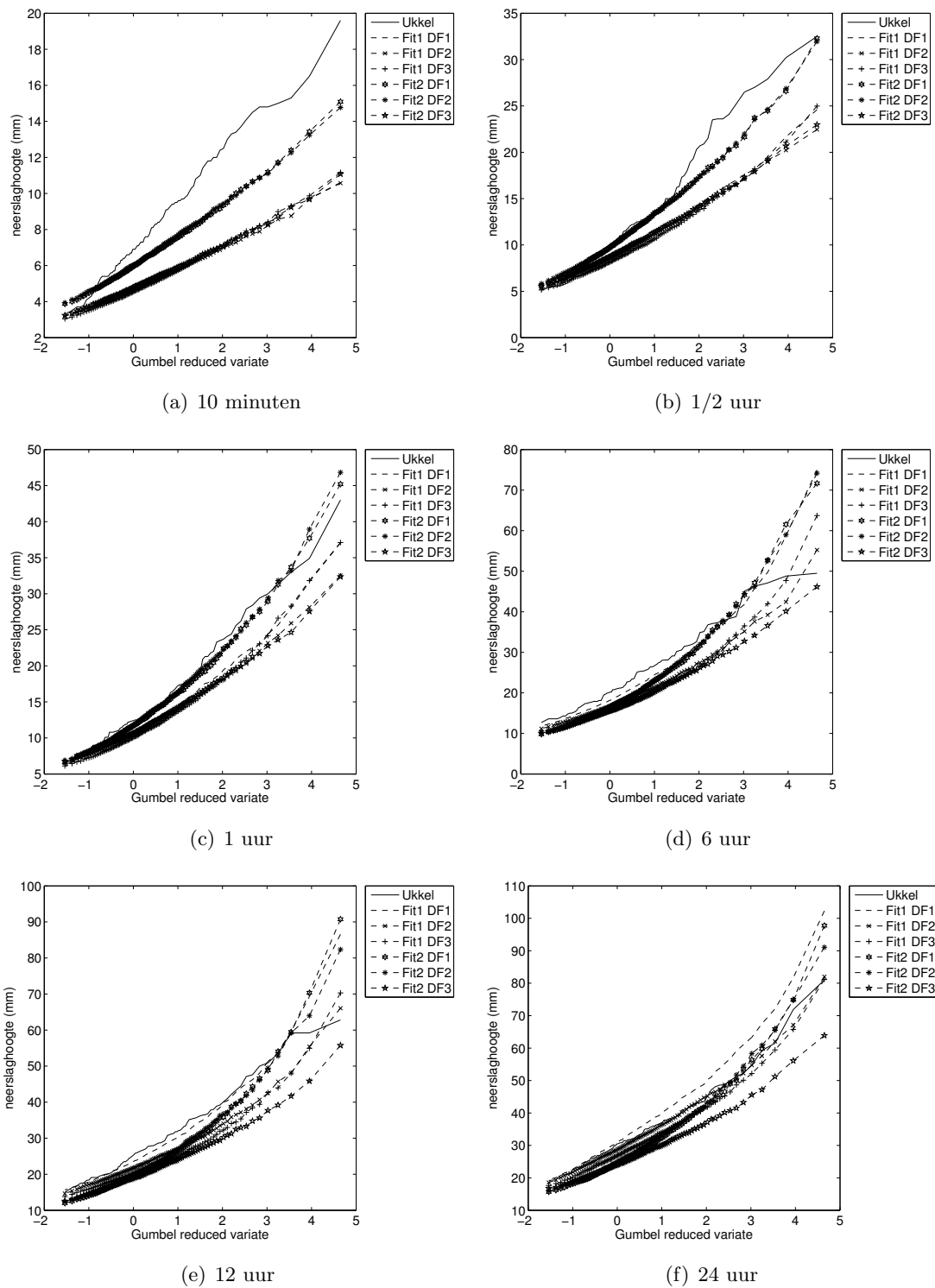
Om de interpretatie van de prestatie van de verschillende parametersets te vereenvoudigen wordt in Tabel 8.2 de RMSE (zie Vergelijking 7.2) voor de verschillende sets in functie van het aggregatieniveau weergegeven. Zoals Figuur 8.7 reeds suggereerde levert het gebruik van DF3 geen verbeteringen op voor Fit2. Voor Fit1 wordt ten opzichte van Fit1 DF1 enkel op het dagniveau een verbetering waargenomen. Voor de overige aggregatieniveau's lijkt de zichtbaar betere fit bij hogere GRV's niet geleid te hebben tot een algeheel beter fit bij alle aggregatieniveau's. Wat wel opmerkelijk is, is dat het gebruik van DF2 in het geval van Fit1 leidt tot een aanzienlijke verbetering van de extreme waarden bij een aggregatieniveau van 24 uur, ten opzichte van Fit1 DF1. Ook bij Fit2 leidt de impliciete weging tot een lichte verbetering, zij het over de hele lijn en niet enkel op dagniveau.



Figuur 8.5: Gemiddelde procentuele afwijking van de voorwaardelijke distributiekarakteristieken overheen de verschillende maanden.



**Figuur 8.6:** Gemiddelde procentuele afwijking van de voorwaardelijke probabiliteiten inzake neerslagintensiteit, overheen de verschillende maanden.



**Figuur 8.7:** Vergelijkende jaarlijkse Gumbel plots in functie van het aggregatieniveau

**Tabel 8.2:** RMSE extreme waarden Fit1, Fit2, DF1, DF2 en DF3

	10 min	1/2 u	1 u	6 u	12 u	24 u
Fit1 DF1	3.45	3.48	2.74	5.44	5.86	7.30
Fit1 DF2	3.58	3.60	3.30	5.08	5.67	4.04
Fit1 DF3	3.55	3.79	3.50	6.70	7.14	5.64
Fit2 DF1	1.91	1.78	1.90	6.34	7.93	7.94
Fit2 DF2	1.89	1.73	1.87	5.88	7.52	7.75
Fit2 DF3	3.46	3.51	3.38	5.89	7.68	7.21

## 8.6 Besluit

In het huidige hoofdstuk werd getracht de kalibratieprocedure te verbeteren door de vorm van de doelfunctie aan te passen. Verder onderzoek van de analytische momenten in dit hoofdstuk toont aan dat de gebrekkige fit het gevolg is van een onevenwicht in de doelfunctie. Daarom worden twee alternatieve configuraties van de doelfunctie voorgesteld. Een eerste voegt per term van de doelfunctie een extra term toe. Op die manier worden deze termen impliciet gewogen, wat zou leiden tot een kleinere bias van de geschatte parameters. Daarnaast wordt een methode voorgesteld om op objectieve manier gewichten aan de verschillende momenten in de doelfunctie toe te kennen. Hierbij worden de momenten gewogen volgens hun relatieve bijdrage aan de fitness tijdens preliminaire kalibraties.

Uit de analyse van de analytische momenten kan besloten worden dat het gebruik van de impliciete weging in de doelfunctie niet tot opmerkelijke verschillen in de fit van de analytische momenten met de observaties leidt, noch is er een zichtbare invloed op de verdeling van de fouten in de doelfunctie. Het gebruik van de weging naargelang de bijdrage aan de fitness brengt daarentegen wel opmerkelijke verschillen in de verdeling van de afwijkingen in de doelfunctie teweeg. Momenten zoals het gemiddelde en de variantie worden nu beter benaderd terwijl het derde orde moment nu een grotere afwijking vertoont.

Naast de fit van de analytische momenten worden opnieuw een reeks simulaties uitgevoerd en geanalyseerd. De belangrijkste conclusies hierbij zijn dat het gebruik van de impliciete weging over het algemeen een, weliswaar gering, positief effect heeft op de resultaten. Het gebruik ervan lijkt dus aangewezen. Toch is deze geringe verbetering onvoldoende om aan de problemen van de fit van het gemiddelde, de variantie en andere eigenschappen van de neerslagreeks, zoals opgemerkt in Hoofdstuk 7, tegemoet te komen. Het gebruik van de weging volgens de bijdrage aan de doelfunctie slaagt er wel in om deze fit te verbeteren, zij het dat deze verbetering ten koste is van meerdere andere eigenschappen, waarvan het meest belang gehecht wordt aan de extreme waarden, zeker in de aanloop naar Hoofdstuk 9. Daardoor kan het gebruik van de weging volgens de bijdrage aan de fitness vooralsnog niet als een verbetering worden aanschouwd. Verdere inspanningen zijn dus vereist om het derde orde moment op een succesvolle manier te integreren in het kalibratieproces. Onderzoek van de extreme waarden wees immers uit dat dit leidt tot een verbeterde prestatie op dat vlak.



## Hoofdstuk 9

# Het Truncated Bartlett-Lewis model

### 9.1 Inleiding

Het TBL model werd ontwikkeld door Onof (2011) om het voorkomen van onrealistisch lange neerslagcellen bij de eerder besproken modellen tegen te gaan. Deze doen zich sporadisch voor wanneer voor lange duur simulaties worden uitgevoerd. In vele gevallen leiden deze overdreven neerslagcellen tot onverwacht grote extremen op grove tijdschalen zoals een dag (Verhoest *et al.*, 2010).

De parameters van het TBL model zijn dezelfde als deze van het MBL model, met dit verschil dat er een zevende parameter bijkomt die een restrictie aan  $\eta$  oplegt, namelijk de truncatie parameter  $\epsilon$ . Op die manier wordt verhinderd dat extreem lage waarden voor  $\eta$  getrokken worden uit de gamma distributie, wat ervoor zorgt dat de gemiddelde celduur tijdens een storm ( $1/\eta$ ) en de gemiddelde duur van de storm  $(\eta\phi)^{-1}$  niet te lang zijn en het gemiddeld aantal celoorsprongen per storm  $(\eta\kappa)$  niet onrealistisch klein wordt. Er wordt dus vermeden dat er zich onrealistisch lange stormen met slechts enkele cellen van gemiddelde intensiteit en lange duur voordoen die door de statistische en extreme waarden analyse niet opgemerkt worden.

In de voorgaande hoofdstukken werd de bespreking van het TBL model steeds uitgesteld. Dit komt omdat er vooralsnog geen uitgebreide kalibratieresultaten zijn. Het onderzoek dat in de voorgaande hoofdstukken werd uitgevoerd stelt ons in staat op een meer onderbouwde manier de kalibratie van het model aan te vatten. De lessen die eruit getrokken werden, kunnen worden toegepast op de kalibratie van het TBL model. Zo wordt ervoor gekozen om de kalibratie met SIMPSA uit te voeren en om gebruik te maken van het derde orde moment, met een impliciete weging van de doelfunctie. Hierbij is geweten dat de introductie van het derde orde moment in de doelfunctie leidt tot een minder goede benadering van het gemiddelde en de variantie. Dit probleem verdient uiteraard bijzondere aandacht bij verder onderzoek, maar ligt buiten het bestek van deze thesis. In Sectie 9.2 wordt vooreerst aandacht besteed aan de optimalisatie van het TBL model, waarna in Sectie 9.3 de simulatieresultaten van het gekalibreerde model geanalyseerd worden.

### 9.2 Optimalisatie

Naar analogie met Hoofdstuk 4 worden in de huidige sectie de verschillende aspecten die betrekking hebben tot de kalibratie van het TBL model belicht. Hierbij wordt stilgestaan bij de

gebruikte doelfunctie (Sectie 9.2.2), de selectie van de gebruikte karakteristieken in de doelfunctie (Sectie 9.2.1), de afbakening van de parameterruimte (Sectie 9.2.3) en de specifieke implementatie van SIMPSA (Sectie 9.2.4), waarna de geschatte parameters weergegeven en besproken worden (Sectie 9.2.5).

### 9.2.1 Selectie van de gebruikte karakteristieken in de doelfunctie

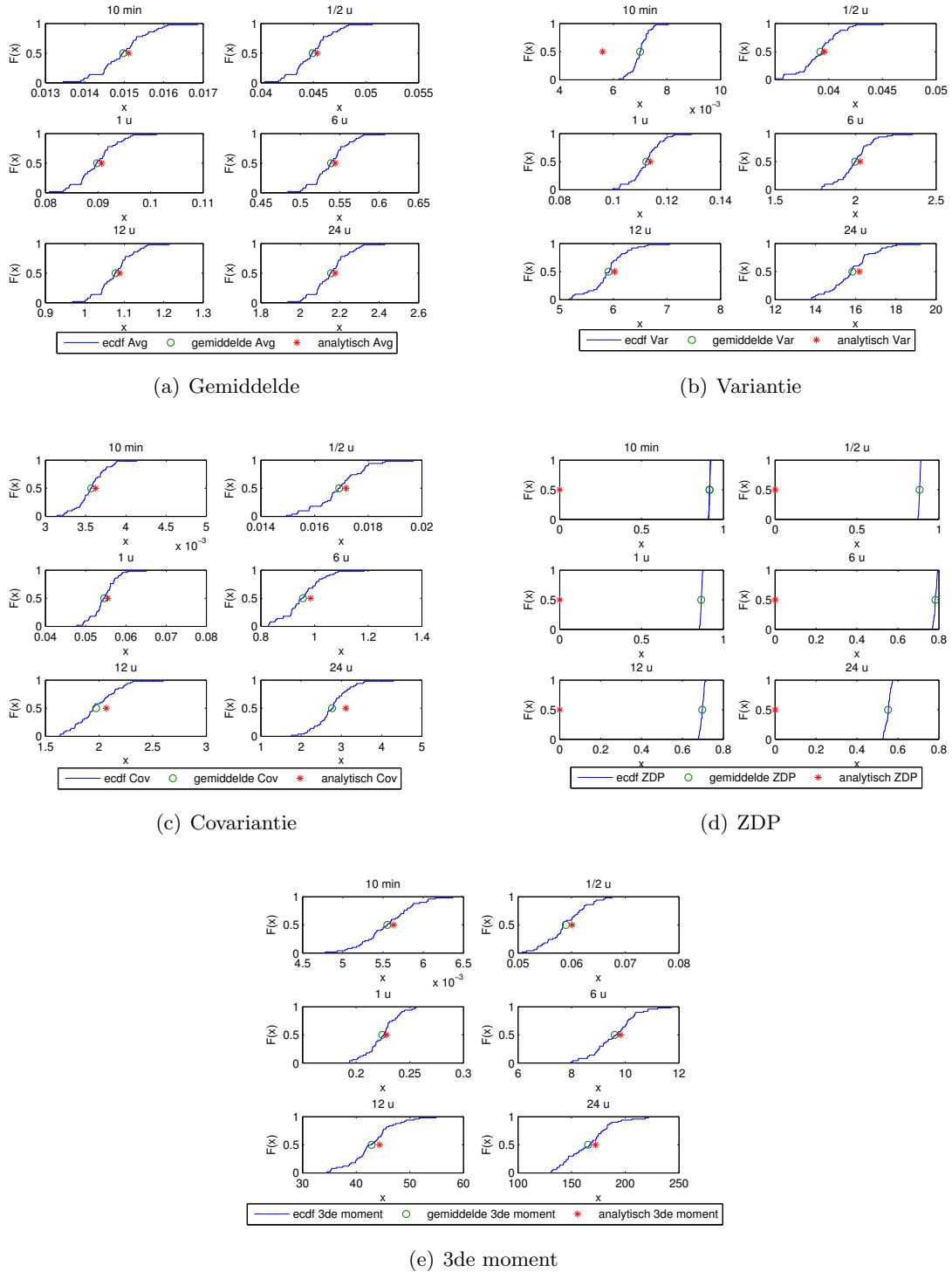
Voor de keuze van de gebruikte karakteristieken in de doelfunctie kan ditmaal geen beroep gedaan worden op de literatuur. In Hoofdstuk 6 blijkt dat de in Sectie 4.3 voorgestelde karakteristieken tot goede resultaten leiden, vooral voor het MBL model en het MBLG model. Hoofdstuk 7 wijst erop dat een introductie van het derde orde moment in de doelfunctie leidt tot een verbeterde representatie van de extreme waarden bij lagere aggregatieniveau's. Er wordt gekozen om de doelfunctie analoog aan deze van het MBL model en het MBLG model op te stellen. Hiertoe dienen eerst de beschikbare analytische uitdrukkingen voor het TBL model gecontroleerd te worden (zie Bijlage A.3), daar de uitdrukkingen immers nog niet eerder gebruikt werden voor het uitvoeren van een kalibratie.

Net zoals in Sectie 7.2 wordt voor een willekeurige parameterset een ensemble van 50 simulaties uitgevoerd. Vervolgens worden voor elk van de 50 herhalingen verschillende gesimuleerde momenten bij verschillende aggregatieniveau's berekend. Daaruit kunnen de experimentele cumulatieve distributiefuncties berekend worden en uitgezet worden in een grafiek, samen met het gemiddelde van de 50 herhalingen en de analytisch berekende waarde (Bijlage A.3). Indien de uitdrukkingen correct zijn, vallen de analytische waarden zo goed als samen met het gemiddelde van de 50 herhalingen. Figuur 9.1 toont de resultaten van dit experiment. De uitdrukkingen voor het gemiddelde, de covariantie en het derde orde moment (Figuur 9.1(a), 9.1(c) en 9.1(e)) lijken voor de gecontroleerde aggregatieniveau's te kloppen, gezien de analytisch berekende waarden zo goed als samenvallen met het gemiddelde van de 50 herhalingen. Voor de variantie van de neerslaghoogte op een aggregatieniveau van 10 minuten is duidelijk een afwijking waar te nemen (Figuur 9.1(b)). Deze wordt daarom best niet opgenomen in de doelfunctie. Figuur 9.1(d) toont aan dat de uitdrukking voor de ZDP (Sectie A.3.2) voor geen enkel van de onderzochte aggregatieniveau's correct blijkt te zijn.

De resulterende selectie van de te gebruiken karakteristieken in de doelfunctie wordt weergegeven in Tabel 9.1. Er wordt gebruik gemaakt van de variantie bij een aggregatieniveau van 1/2 u in plaats van 10 min, omwille van voornoemde redenen. De ZDP is logischerwijs niet opgenomen in de doelfunctie. Net zoals in Hoofdstuk 7 wordt gebruik gemaakt van het derde orde moment bij aggregatieniveau's van 10 minuten en 24 uur. In tegenstelling tot de doelfuncties van het MBL model (zie Tabel 7.1) wordt de variantie op een aggregatieniveau van 24 u wel meegerekend in de doelfunctie.

**Tabel 9.1:** Gebruikte momenten en hun aggregatieniveau voor het TBL model

Gemiddelde	Variantie	Covariantie	ZDP	Derde orde moment
10 min	1/2 u en 24 u	10 min en 24 u	-	10 min en 24 u



Figuur 9.1: Verificatie analytische uitdrukkingen TBL model.

### 9.2.2 Doelfunctie

In Hoofdstuk 8 is gebleken dat het gebruik van een impliciete weging in de doelfunctie geen opmerkelijke maar toch een algemeen positieve invloed heeft op de resulterende geschatte parameters. Deze positieve invloed was vooral op te merken bij de reproductie van de extreme waarden. De weging volgens de relatieve bijdrage aan de fitness slaagde in zijn opzet. De systematische afwijking van het gemiddelde en de variantie, tengevolge van de introductie van het derde orde moment, werd hiermee weggewerkt. Daartegenover stond dat bepaalde andere eigenschappen, waaronder de extreme waarden, minder goed benaderd werden. Bijgevolg wordt het TBL model gekalibreerd aan de hand van de doelfunctie met impliciete weging, die wordt voorgesteld in Sectie 8.3.1.

### 9.2.3 Parameterruimte

Ook voor de afbakening van de parametergrenzen kan geen beroep gedaan worden op reeds bestaande literatuur. Er zou gebruik kunnen gemaakt worden van een initieel grote parameter-ruimte, met de bedoeling dat het optimalistie-algoritme vervolgens de meest veelbelovende regio aanduidt. Uit Hoofdstuk 5 blijkt echter dat geen van beide voorgestelde optimalisatietechnieken (SIMPSA en PSO) daartoe in staat is. Daarom wordt voor het opstellen van de parametergrenzen een beroep gedaan op de grenzen van het MBL model. Het MBL model heeft, op  $\epsilon$  na, dezelfde parameters als het TBL model. Gebaseerd op Tabel 5.5 worden parametergrenzen opgesteld welke weergegeven worden in Tabel 9.2. Merk op dat de grenzen voor  $\kappa$  en  $\alpha$  niet gelijk zijn aan deze voor het MBL model. Proefondervindelijk kan namelijk worden besloten dat betere resultaten verkregen worden wanneer de in Tabel 9.2 voorgestelde grenzen gebruikt worden.

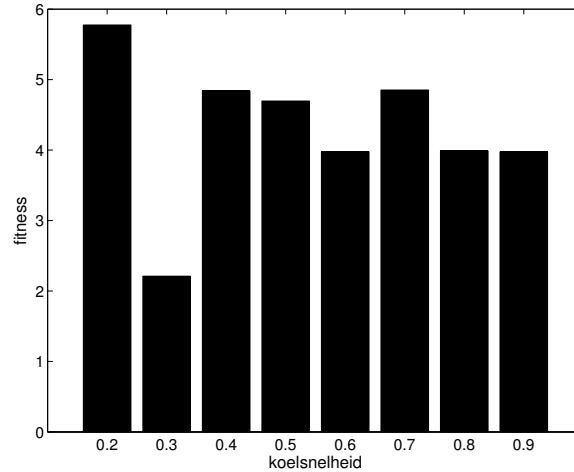
**Tabel 9.2:** Onder- en bovengrenzen voor de kalibratie van het TBL model (geldig voor alle maanden).

	$\lambda$	$\kappa$	$\phi$	$\mu_x$	$\alpha$	$\nu$	$\epsilon$
Ondergrens	0	0	0	0	0	0	0
Bovengrens	0,1	1	1	15	5	5	5

### 9.2.4 Implementatie SIMPSA

Vooraleer SIMPSA kan gebruikt worden om het TBL model te kalibreren dient (cf. Sectie 4.5.3.7) stilgestaan te worden bij enkele aandachtspunten. De instellingen die voor de overige Bartlett-Lewis modellen werden gebruikt zijn immers niet noodzakelijk de meest ideale voor het TBL model. Daarom wordt ook voor het TBL model een optimale koelsnelheid bepaald. Hiertoe wordt een reeks preliminaire kalibraties uitgevoerd waarbij de koelsnelheid gevarieerd wordt. Tijdens deze preliminaire kalibraties wordt telkens slechts een beperkt aantal iteraties uitgevoerd. Zodoende kan berekend worden welke koelsnelheid gemiddeld de beste resultaten oplevert. De koelsnelheid  $\delta$  wordt getest in het interval  $[0.2, 0.9]$  met stappen van 0.1. Bij elke koelsnelheid worden vijf herhalingen uitgevoerd met per herhaling 100 iteraties. Vervolgens wordt de mediaan van de herhalingen berekend en het resultaat uitgezet in een grafiek. De koelsnelheid die na 100 iteraties de beste fitness oplevert wordt weerhouden en gebruikt voor verdere kalibraties. Figuur 9.2 geeft de resultaten van de zoektocht naar de ideale koelsnelheid weer. Op basis van

deze figuur kan besloten worden dat met een koelsnelheid van 0.3 de beste resultaten bekomen worden. De kalibraties van het TBL model zullen dan ook uitgevoerd worden met een koelsnelheid van 0.3. De overige instellingen worden niet afgesteld in functie van het TBL model, zoals ook niet het geval was bij de overige gekalibreerde modellen. Voor meer details wordt verwezen naar Sectie 4.5.3.7 en Sectie 4.5 in het algemeen.



**Figuur 9.2:** Invloed van de koelsnelheid op de fitness voor het TBL model.

### 9.2.5 Geschatte parameters

Tabel 9.3 geeft het resultaat weer van de kalibratie met behulp van SIMPSA, gebruik makende van de doelfunctie met impliciete weging, voor het TBL model. Hieruit blijkt dat overwegend goede fitnesswaarden bekomen worden, voor zover daar voorlopig uitspraken over kunnen gedaan worden.

**Tabel 9.3:** Parameters met bijbehorende fitness voor de verschillende maanden voor het TBL model.

	<i>fitness</i>	$\lambda$	$\kappa$	$\phi$	$\mu_x$	$\alpha$	$\nu$	$\epsilon$
Jan	4.62E-03	0.02666	0.37619	0.01527	1.83873	3.91242	0.21116	0.83935
Feb	5.37E-03	0.02257	0.33243	0.01505	1.58109	4.99947	0.40060	0.84515
Mrt	2.07E-05	0.02310	0.25812	0.01301	2.24342	4.39776	0.25511	0.87170
Apr	1.74E-04	0.02407	0.18816	0.01205	2.78966	4.20008	0.22927	0.84903
May	6.31E-04	0.02175	0.08236	0.01125	6.90281	4.16241	0.21206	6.67E-06
Jun	5.90E-07	0.02646	0.10887	0.01561	8.40653	3.37045	0.12233	0.69176
Jul	1.79E-03	0.01381	0.03088	0.00385	12.67771	4.17139	0.21595	0.89062
Aug	1.11E-05	0.02017	0.07071	0.00983	9.93168	3.59908	0.15241	0.87837
Sep	8.16E-03	0.02123	0.05962	0.00733	8.20955	3.90010	0.16358	1.83E-08
Okt	8.73E-04	0.01871	0.20309	0.01094	3.75665	3.91489	0.20056	0.60839
Nov	3.01E-03	0.02204	0.31564	0.01342	2.07570	4.78780	0.34970	0.83955
Dec	1.68E-02	0.02243	0.37548	0.01575	1.72578	4.99735	0.42425	0.84346

Figuur 9.3 geeft aan in welke mate de intrede van de truncatieparameter  $\epsilon$  de schatting van de overige parameters heeft beïnvloed. Hiertoe wordt de vergelijking gemaakt met het MBL model dat gekalibreerd werd met inbegrip van het derde centrale moment en impliciete weging.

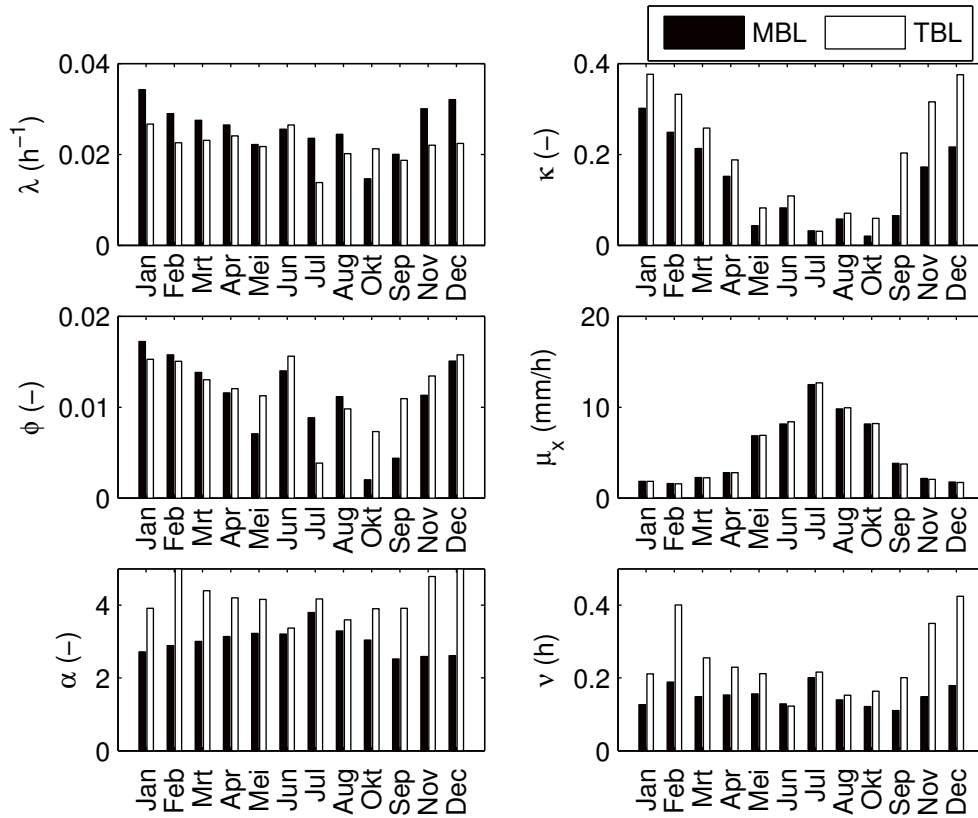
Er dient vermeld te worden dat laatstgenoemde kalibratie ook rekening hield met de ZDP, wat niet het geval was bij het TBL model. De veranderingen in de schatting van de parameters van het TBL model zijn dus niet ondubbelzinnig aan de intrede van de truncatieparameter toe te schrijven, ook het ontbreken van de ZDP zal zijn invloed hebben op de resultaten.

De schatting van  $\mu_x$  blijkt (zie Figuur 9.4) nagenoeg onveranderd te blijven. De waarden van  $\kappa$  worden doorgaans hoger ingeschat, terwijl  $\phi$  geen eenduidige stijging of daling ten opzichte van het MBL model vertoont. Dit resulteert in een hoger gemiddeld aantal cellen per bui ( $\mu_c = 1 + \kappa/\phi$ ), zoals te zien op Figuur 9.4. Voor het aantal stormmanifestaties kan gezien worden dat deze over het algemeen verminderd zijn (zie Figuren 9.3 en 9.4), terwijl de gemiddelde celduur ( $1/\eta = 1/(\alpha\nu)$ ) een algemene daling kent door de hoger geschatte  $\alpha$  en  $\nu$  (Figuur 9.4). Er worden met andere woorden minder buien aangemaakt, die op hun beurt meer cellen met gelijke gemiddelde celhoogte bevatten. Voorgaande bespreking van de geschatte parameters houdt geen rekening met  $\epsilon$ . Gezien de truncatie van de gamma verdeling een belangrijke invloed zal hebben op de verdeling van de celduur kunnen de voorlopige uitspraken omtrent de celduur enkel als indicatief beschouwd worden.

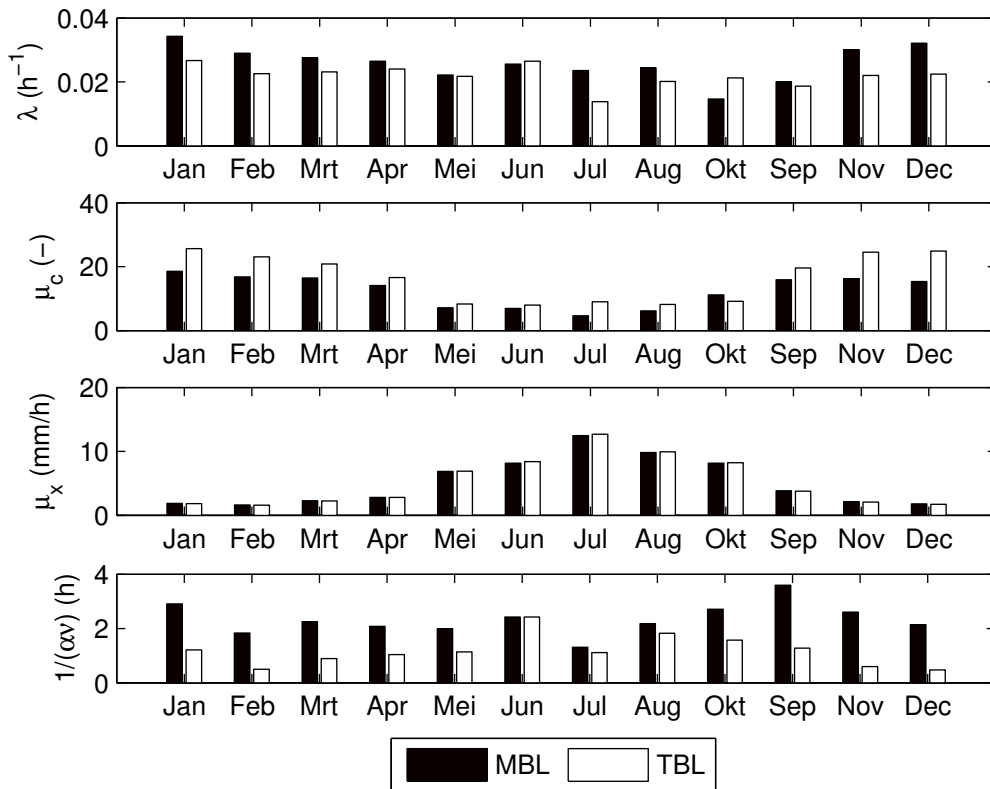
De verdeling van de gemiddelde celduur wordt onderworpen aan een truncatie met parameter  $\epsilon$ . Indien de  $\eta$ , getrokken uit de gamma verdeling met vorm- en schaalparameters  $\alpha$  en  $\nu$ , kleiner is dan  $\epsilon$  wordt deze verworpen tot een waarde groter dan  $\epsilon$  wordt bekomen. De werkelijke gemiddelde celduur zal bijgevolg nog kleiner zijn, gezien waarden voor  $\eta$  kleiner dan  $\epsilon$  niet meer mogelijk zijn. Er worden dus over het algemeen al kortere celduren bekomen dan het geval was bij MBL (cf. Fit2 DF2). Om de invloed van  $\epsilon$  concreet te illustreren worden in Figuur 9.5 de probabiliteitsdichtheidsfuncties (PDF) van  $\eta$  voor de maand januari weergegeven, voor het MBL model en voor het TBL model (voor vorm- en schaalparameters  $\alpha$  en  $\nu$ , zie Tabel B.9 en Tabel 9.3, voor MBL en TBL model respectievelijk). Hierop is inderdaad te zien dat de geschatte PDF bij het TBL naar grotere waarden toe is verschoven, met dus kortere verwachte celduren  $1/\eta$  tot gevolg. Daarnaast is ook  $\epsilon$  weergegeven, voor de maand januari is deze gelijk aan 0.84. Concreet betekent dit dat de celduur, welke exponentieel verdeeld is met parameter  $1/\eta$ , in 95% van de gevallen kleiner zal zijn 2 uur en 40 minuten, terwijl bij het MBL model de cel in 95% van de gevallen slechts kleiner is dan 16 uur. Het mag duidelijk zijn dat de truncatie van de gamma verdeling van  $\eta$  om onrealistisch lange cellen te voorkomen in zijn opzet slaagt. Tabel 9.3 toont aan dat voor de meeste maanden gelijkaardige resultaten bekomen worden, behalve voor mei en september. Voor laatstgenoemde maanden is  $\epsilon$  dermate klein dat de invloed verwaarloosbaar wordt. Voor deze waarden van  $\epsilon$  worden alsnog goede fitnesswaarden bekomen. In verdere kalibraties en tijdens de statistische analyse dient hieraan bijzondere aandacht besteed te worden.

### 9.3 Statistische analyse

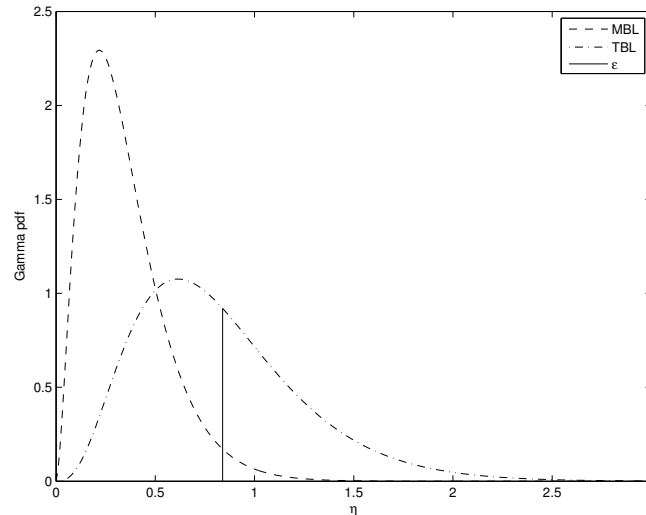
Net zoals in de voorgaande hoofdstukken wordt voor het TBL model een ensemble aan simulaties uitgevoerd. Hierbij worden 50 neerslagreeksen met een lengte van 105 jaar gemodelleerd. De bekomen neerslagreeksen worden vergeleken met de observaties aan de hand van procentuele afwijkingen, zoals besproken in Sectie 6.3. De bekomen resultaten worden vergeleken met de resultaten van het MBL model, meer bepaald Fit1 DF1 en Fit2 DF2. Er kan, zoals vermeld in Sectie 9.1 verwacht worden dat de invloed van de truncatieparameter niet zichtbaar is in de



Figuur 9.3: Vergelijking parameters MBL (cf. Fit2 DF2) en TBL.



Figuur 9.4: Vergelijking fysieke variabelen MBL (cf. Fit2 DF2) en TBL.



**Figuur 9.5:** PDF voor MBL en TBL en truncatie, januari.

onderstaande statistische analyse. Toch is het zinnig deze analyse uit te voeren, om na te gaan of de algemene statistische eigenschappen niet in het gedrang komen door de introductie van  $\epsilon$ .

### 9.3.1 Eerste, tweede en derde orde momenten

In Figuur C.5 worden per maand de procentuele afwijkingen van de eerste, tweede en derde orde momenten van de synthetische neerslagreeksen weergegeven. In onderstaande figuur worden de resultaten voor de maand januari weergegeven (Figuur 9.6).

Figuur 9.6(a) toont aan dat het probleem van de systematische overschatting van de gemiddelde neerslaghoogte (zie Hoofdstuk 7 en Hoofdstuk 8) zich niet voordoet. Niet alleen voor januari (Figuur 9.6(a)), maar ook voor de overige maanden is de procentuele afwijking van de gemiddelde neerslaghoogte, ten opzichte van de observaties, vergelijkbaar met de resultaten die bekomen worden door MBL Fit1 DF1.

Voor de afwijkingen van de variantie ten opzichte van de observaties te Ukkel kan worden verwacht dat het TBL model bij een aggregatieniveau van 1/2 en 24 u betere resultaten zal boeken dan de twee parametersets van het MBL model waarmee ze vergeleken worden. Dit heeft alles te maken met de samenstelling van de doelfunctie, waarin deze momenten zijn opgenomen. Bij het MBL model werd enkel de variantie bij een aggregatieniveau van 10 min opgenomen. Figuur 9.6(b) geeft de procentuele afwijkingen voor de verschillende beschouwde parametersets en modellen weer. De belangrijkste conclusie is dat het TBL model op een aggregatieniveau van 10 minuten een aanzienlijk grotere afwijking vertoont dan het MBL model, terwijl tussen 1/2 en 24 u opmerkelijk betere resultaten worden bekomen. Er kan uiteraard beargumenteerd worden dat dit vooral te wijten is aan de kalibratie op zich en dus weinig met de modellen te maken heeft.

De procentuele afwijkingen voor de ZDP voor januari worden weergegeven in Figuur 9.6(a). Zoals te verwachten, wordt voor hogere aggregatieniveau's een aanzienlijke afwijking ten opzichte van de observaties bekomen. Voor de lagere aggregatieniveau's (10 min, 1/2 u en 1 u) wijkt de ZDP, bekomen door het TBL model, niet noemenswaardig af van de resultaten die door het MBL model werden verkregen.

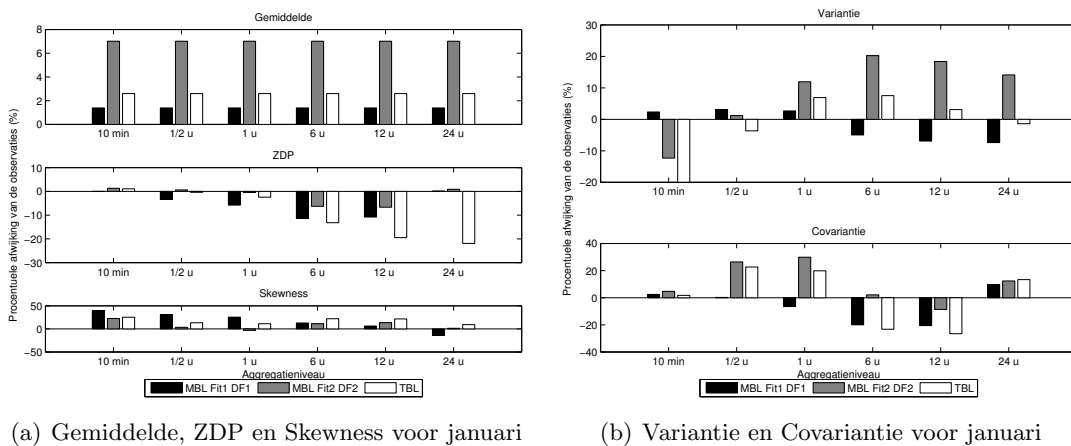
Figuur 9.6(b) geeft de procentuele afwijking van de covariantie weer. Hierop kan gezien wor-



den dat bij 10 min en 24 u (de momenten in de doelfunctie), de afwijking ten opzichte van de observaties klein is en nagenoeg gelijk aan deze bekomen door het MBL model. Voor de aggregatieniveau's 1/2 en 1 u is de afwijking, die door het TBL model wordt bekomen, ongeveer gelijk aan deze bekomen door MBL Fit2 DF2. Deze afwijking is tevens 2 tot 3 maal hoger dan bij MBL Fit1 DF1, wat zou kunnen toegeschreven worden aan het gebruik van het derde orde moment in de doelfunctie (zie ook Hoofdstukken 7 en 8). Bij aggregatieniveau 6 en 12 u komt de afwijking meer overeen met deze bekomen door MBL Fit1 DF1. Dit geldt echter niet als algemene conclusie, Figuur C.5 toont aan dat de bekomen afwijking van de covariantie verschilt van maand tot maand. Voor de maand juli (Figuur C.5(n)) worden zelfs afwijkingen tot -100 %, voor een aggregatieniveau van 24 u, bekomen.

De procentuele afwijking van de skewness wordt weergegeven in Figuur 9.6(a), voor de maand januari. Een algemene bevinding, die ook blijkt te gelden voor de andere maanden, is dat de skewness beter wordt benaderd in vergelijking met MBL Fit1 DF1. Toch wordt de beste benadering bekomen door MBL Fit2 DF2. In bepaalde maanden wordt de skewness beter benaderd wanneer deze niet is opgenomen in de doelfunctie (zie bijvoorbeeld augustus, Figuur C.5(o)). Dit laatste duidt nog maar eens aan dat de analyse van de gesimuleerde momenten nooit echt een eenduidig resultaat oplevert.

In Sectie 9.2.5 werd opgemerkt dat voor de maanden mei en september de truncatieparameter verwaarloosbaar klein was. Zoals reeds werd geopperd aan het begin van deze sectie heeft dit geen noemenswaardige invloed op de reproductie van de eerste, tweede en derde orde eigenschappen, zoals te zien is op Figuur C.5(i), C.5(j), C.5(q) en C.5(r). Dit bewijst nogmaals dat het probleem met de onrealistisch lange neerslagcellen onopgemerkt blijft bij de statistische analyse.



**Figuur 9.6:** Eerste, tweede en derde orde momenten voor januari

### 9.3.2 Voorwaardelijke distributiekarakteristieken

In Figuur 9.7 worden de voorwaardelijke distributiekarakteristieken, of beter de procentuele afwijkingen ten opzichte van de observaties, weergegeven. Voor meer informatie omtrent de definitie hiervan wordt verwezen naar Hoofdstuk 3, Sectie 3.2.2.

De kans op twee opeenvolgend droge intervallen (vwd1, Figuur 9.7(a)) wordt zichtbaar beter ingeschat door het TBL model voor aggregatieniveau's beneden de 6 u. Voor de grotere aggre-

gationiveau's is de afwijking van het TBL model groter. Daar wordt de betreffende kans tussen de 10 tot 20 % onderschat. Op een aggregatieniveau van 24 u is de afwijking ten opzichte van het MBL model, welke een grote overschatting levert, minder groot. Het aantal droge intervallen wordt door het TBL model op grote aggregatieniveau's dus aanzienlijk onderschat, dit resultaat werd reeds bekomen bij het bestuderen van de ZDP (zie Sectie 9.3.1). Het gebruik van de ZDP in de doelfunctie kan hiervoor een oplossing bieden.

Voor de overige voorwaardelijke distributiekarakteristieken, kan vooreerst opgemerkt worden dat de variabiliteit van de neerslagintensiteit (vwd3, vwd5, vwd9) in vele gevallen beter benaderd wordt door het TBL model. Daarnaast wordt de gemiddelde bui-intensiteit (vwd2) bij de lagere aggregatieniveau's (1/2 u en 1 u), het best door het TBL benaderd.

Voorop het dagniveau kunnen voor de overige distributiekarakteristieken grotere afwijkingen dan het MBL gevonden worden, bijvoorbeeld voor de neerslagintensiteit middenin een bui (vwd8, Figuur 9.7(e)). Voor de neerslagintensiteit bij aanvang van een bui (vwd4, Figuur 9.7(c)) en op het einde van een bui (vwd6, Figuur 9.7(d)) worden over het algemeen procentuele afwijkingen, ten opzichte van de observaties, bekomen die in lijn liggen met de resultaten die door het MBL werden bekomen.

### 9.3.3 Voorwaardelijke probabiliteiten inzake neerslagintensiteit

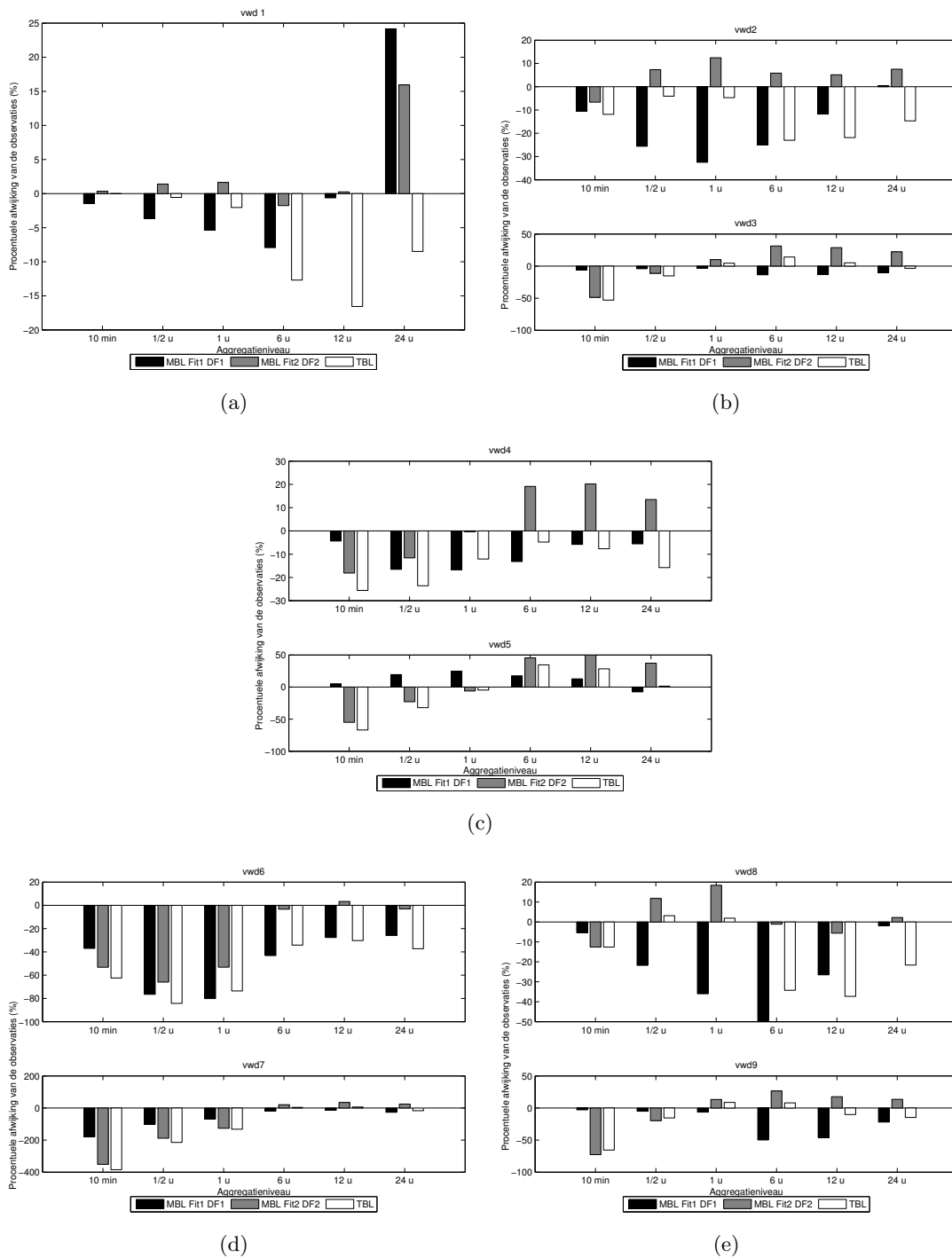
Figuur 9.8 geeft de procentuele afwijkingen van de voorwaardelijke probabiliteiten inzake neerslagintensiteit weer. Deze drukken de kans uit dat gedurende een bepaald tijdsinterval minder neerslag valt dan een bepaalde drempelwaarde (zie Sectie 3.2.3).

Op een aggregatieniveau van 10 min worden de verschillende voorwaardelijke probabiliteiten inzake neerslagintensiteit aanzienlijk overschat door het TBL model. Er zijn met andere woorden te veel 10 min intervallen waarin een te kleine neerslaghoogte valt. Op een aggregatieniveau van 12 u en 24 u wordt  $DP_{0.5}$  sterk onderschat door het TBL model (zie Figuur 9.8(e) en 9.8(f)). Dit hangt samen met het feit dat op die aggregatieniveau's de ZDP sterk onderschat wordt door het TBL model. De overige DP's worden dan wel weer vrij goed benaderd bij deze aggregatieniveau's. Voor 1/2 en 1 u kan gezien worden dat de benadering door het TBL model vrij goed is, beter dan het MBL model (zie Figuur 9.8(b) en 9.8(c)).

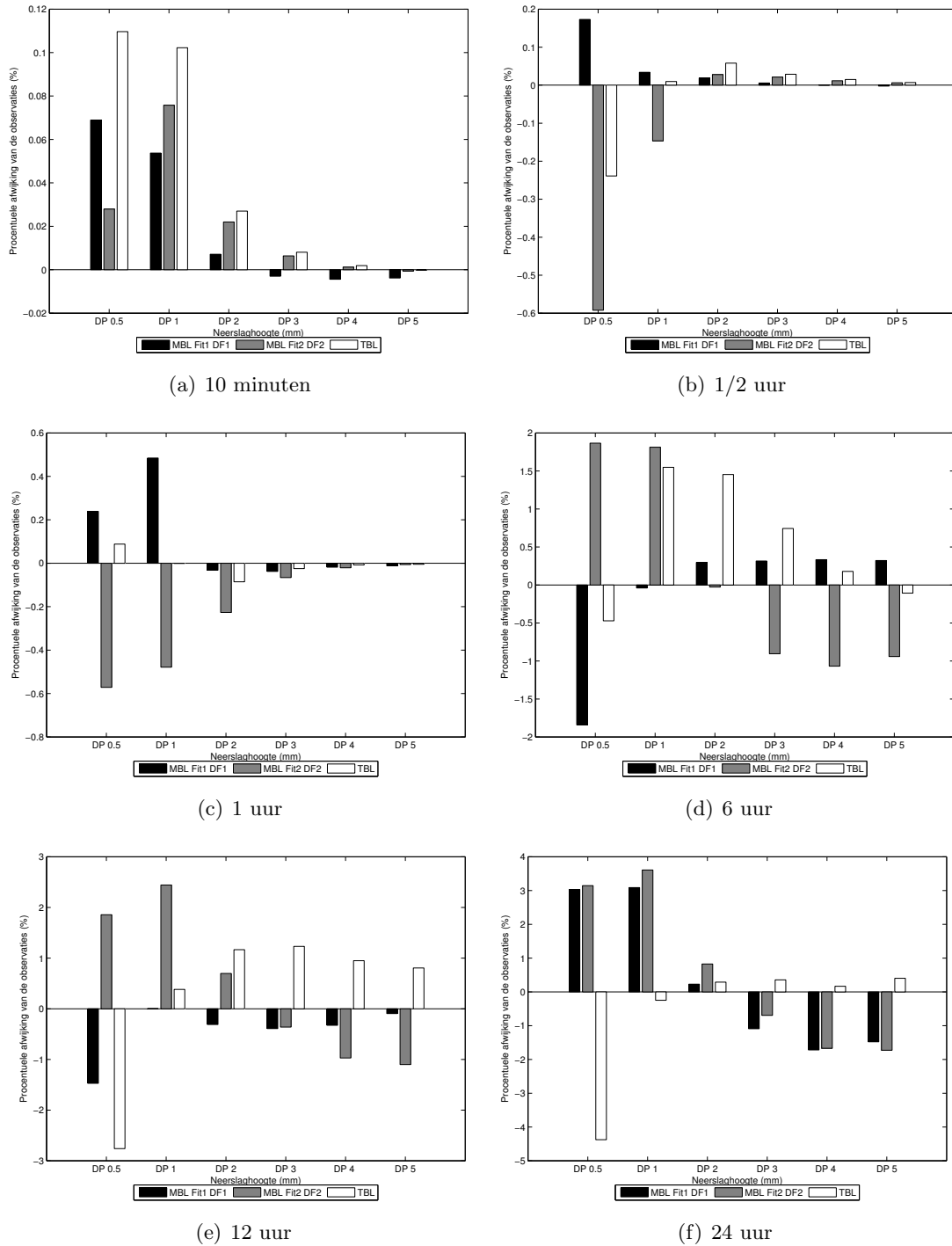
### 9.3.4 Extreme waarden analyse

De analyse van de extremen gebeurt opnieuw op twee manieren. Enerzijds wordt visueel nagegaan wat de invloed van de truncatie op de fit van de extreme waarden van de synthetische reeks met de observaties is (Figuur 9.9). Anderzijds wordt voor de verschillende aggregatieniveau's de RMSE berekend (cf. Hoofdstuk 7, Sectie 7.4.2.4, Vergelijking 7.2). Tabel 9.4 geeft een samenvatting voor de verschillende modellen die tijdens het huidige werk werden gekalibreerd, waarbij voor het MBL model een onderscheid wordt gemaakt tussen de verschillende gebruikte kalibratiemogelijkheden.

Op Figuren 9.9(a) en 9.9(b) kan duidelijk gezien worden dat op een aggregatieniveau van 10 min en 1/2 u een betere fit bekomen wordt met de observaties, in vergelijking met het MBL model. Voor de andere aggregatieniveau's is het moeilijk om hierover een uitspraak te doen gebaseerd op de figuren. Daarom wordt beter rekening gehouden met Tabel 9.4. In deze tabel wordt telkens ook het gemiddelde van de RMSE's overheen de verschillende aggregatieniveau's weergegeven.



**Figuur 9.7:** Gemiddelde procentuele afwijking van de voorwaardelijke distributiekarakteristieken overeen de verschillende maanden.



**Figuur 9.8:** Gemiddelde procentuele afwijking van de voorwaardelijke probabiliteiten inzake neerslagintensiteit, overheen de verschillende maanden.

Ook al kan het TBL model enkel voor een aggregatieniveau van 10 min de beste resultaten voorleggen, toch vertoont het gemiddeld gezien de beste fit met de observaties, aangezien de gemiddelde RMSE het kleinst is.

**Tabel 9.4:** RMSE extreme waarden, samenvattende tabel

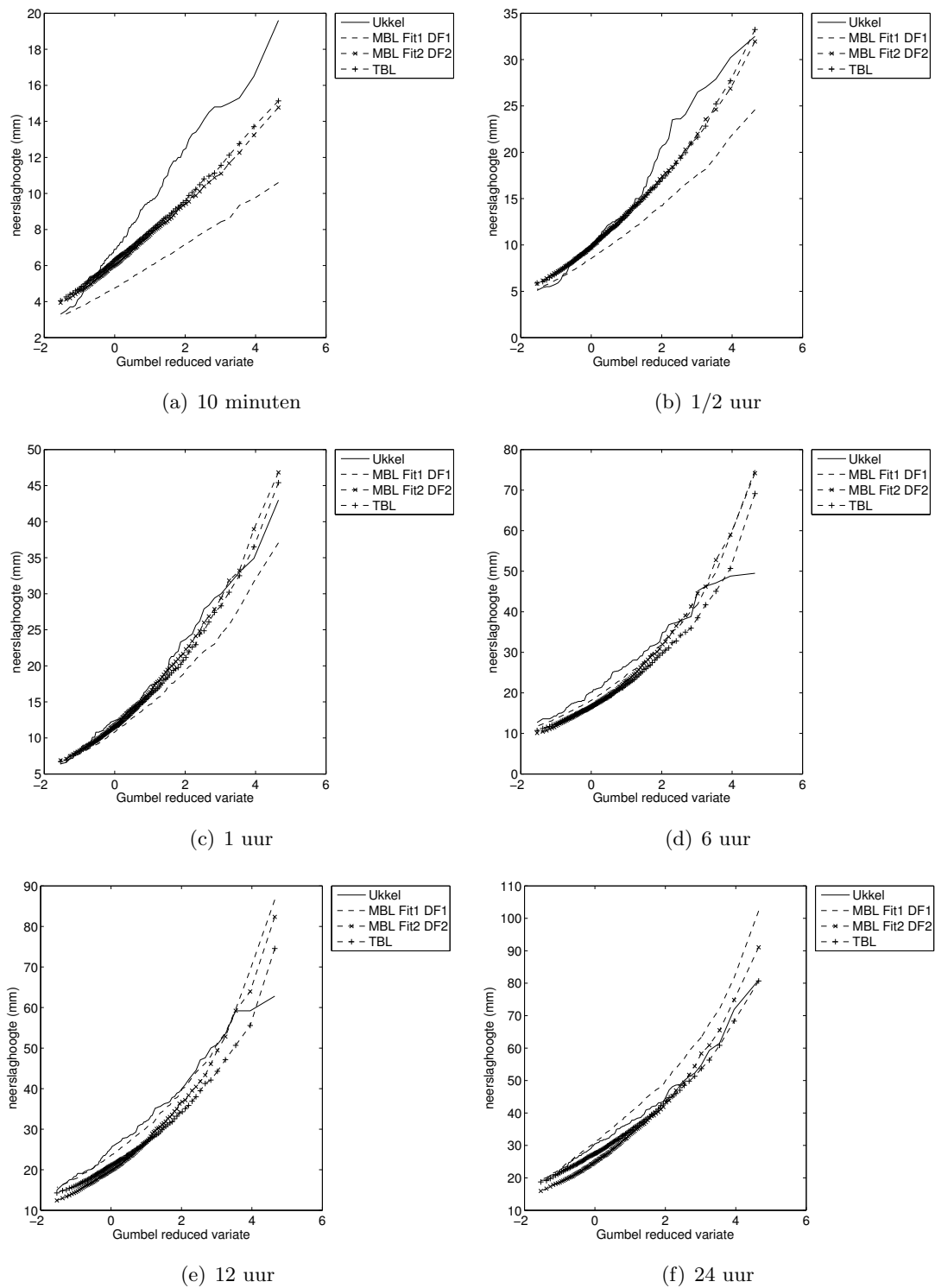
model	specificaties	10 min	1/2 u	1 u	6 u	12 u	24 u	Gemiddelde
OBL		3.59	3.96	4.37	6.96	7.88	6.21	5.50
MBLG		4.58	4.88	3.74	4.87	5.75	8.72	5.42
MBL	Fit1 DF1	3.45	3.48	2.74	5.44	5.86	7.30	4.71
	Fit1 DF2	3.58	3.60	3.30	5.08	5.67	4.04	4.21
	Fit1 DF3	3.55	3.79	3.50	6.70	7.14	5.64	5.05
	Fit2 DF1	1.91	1.78	1.90	6.34	7.93	7.94	4.63
	Fit2 DF2	1.89	1.73	1.87	5.88	7.52	7.75	4.44
	Fit2 DF3	3.46	3.51	3.38	5.89	7.68	7.21	5.19
	TBL		1.68	1.86	2.08	5.97	7.07	6.54

## 9.4 Invloed truncatieparameter $\epsilon$

Zoals reeds meermaals vermeld (Sectie 2.2.3, Sectie 9.1) geven de Bartlett-Lewis modellen in bepaalde gevallen aanleiding tot het ontstaan van onrealistisch lange regencellen. Gezien deze regencellen niet ontmaskerd worden door klassieke analyses (zie Sectie 9.3), werd het aandeel ervan in de overschatting van extreme overstromingen lange tijd genegeerd (Verhoest *et al.*, 2010). Verhoest *et al.* (2010) tonen aan dat het effect van de betreffende onrealistische cellen tot uiting komt wanneer lange tijdreeksen gesimuleerd worden (grootte orde 1000 jaar). Daarbij kan gezien worden dat de extremen bij terugkeerperiodes groter dan de geobserveerde dataset sterk overschat worden. Er wordt immers verondersteld dat de extreme waarden de Gumbel verdeling volgen, wat niet het geval is.

Om na te gaan of de introductie van de truncatieparameter een invloed heeft gehad, worden enkele tijdreeksen van 1000 jaar gesimuleerd. Om de prestatie van het TBL te vergelijken met de overige modellen en om eveneens de invloed van het derde orde moment na te gaan worden de resultaten vergeleken met de andere modellen. Voor het MBL model wordt met twee parametersets rekening gehouden, namelijk MBL Fit1 DF2 en MBL Fit2 DF2. De keuze gaat uit naar deze twee parametersets omdat zij van de zes MBL parametersets de beste gemiddelde RMSE vertonen (zie Tabel 9.4).

Figuur 9.10 geeft de vergelijking tussen de gesimuleerde waarden en de observaties voor de verschillende modellen weer. Op Figuur 9.10(a) en Figuur 9.10(b) worden respectievelijk de uurlijkse en dagelijkse jaarlijkse extremen weergegeven. Op basis van Figuur 9.10(a) kan besloten worden dat bij een simulatie van 1000 jaar het in gebruik nemen van het derde orde moment tot aanzienlijk betere resultaten leidt. MBL Fit2 DF2 en het TBL model (zie Sectie 9.2) vallen immers veel beter samen met de geobserveerde reeks dan de overige modellen. Bij hogere terugkeerperiodes (hogere Gumbel Reduced Variate) vertoont geen van beide de neiging om af te wijken van de Gumbel rechte (niet weergegeven). Dit laatste is wel merkbaar op Figuur 9.10(b).

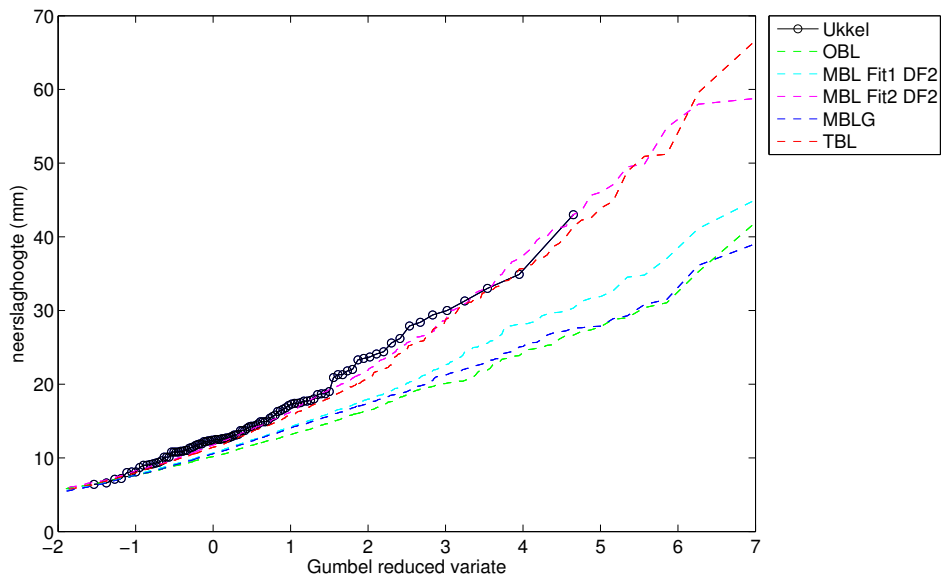


**Figuur 9.9:** Vergelijkende jaarlijkse Gumbel plots in functie van het aggregatieniveau

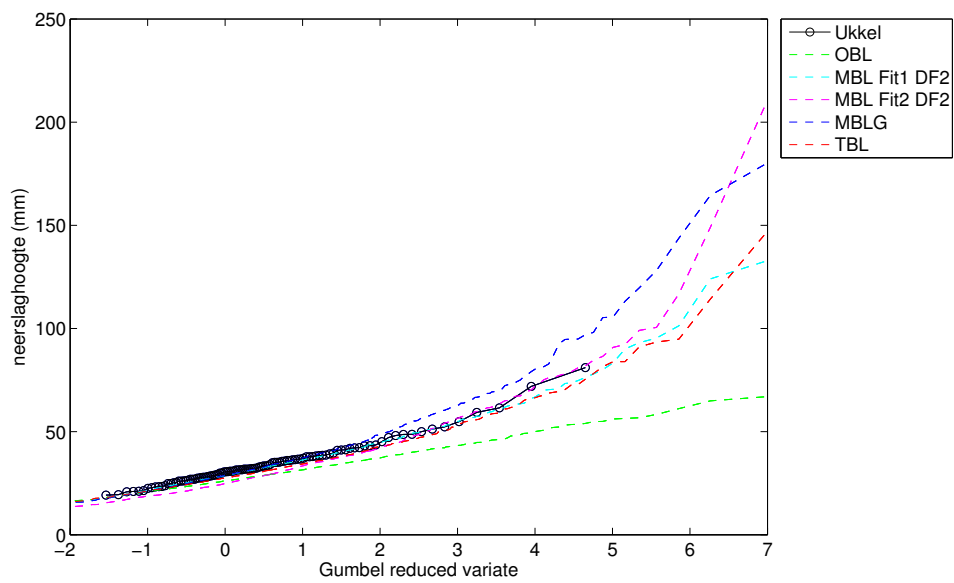
Op deze figuur vertonen het TBL model en het MBL model (zowel Fit1 als Fit2) de beste overeenkomst met de observaties. Bij hogere terugkeerperiodes hebben de modellen echter wel de neiging om af te wijken van een rechte lijn, wat overeenkomt met een overschatting van de extreme neerslaggebeurtenissen. Het TBL model levert een minder sterke overschatting dan MBL Fit2 DF2. Dit zou kunnen wijzen op een positieve invloed van de truncatie. Zij het niet dat ook MBL Fit1 DF2 slechts beperkte overschattingen oplevert. Het is dus niet met zekerheid te concluderen of de verkleinde overschatting door het TBL model te wijten is aan de introductie van de truncatieparameter. Implementatie van de synthetische neerslag in een hydrologisch model zal moeten uitwijzen of de overstromingen nu minder overschat worden of niet en of dit rechtstreeks kan worden toegeschreven aan de truncatie van de gamma verdeling van de gemiddelde celduur.

## 9.5 Besluit

In dit hoofdstuk werd voor een eerste maal getracht het Truncated Bartlett-Lewis model (Onof, 2011) naar behoren te kalibreren. Bij de kalibratie wordt gebruik gemaakt van de kennis die werd opgedaan omtrent de Bartlett-Lewis modellen in vorige hoofdstukken. Bijgevolg werd het TBL model gekalibreerd aan de hand van SIMPSA, gezien in Hoofdstuk 5 werd besloten dat de kalibratie van de Bartlett-Lewis modellen hiermee het meest opportuun is. SIMPSA wordt net zoals voor de andere Bartlett-Lewis modellen afgesteld in functie van het TBL model. Verder wordt ervoor gekozen om het derde orde moment op te nemen in de doelfunctie. Hoofdstuk 7 wees er immers op dat de representatie van de extreme waarden hiermee verbeterd werd in vergelijking met een doelfunctie zonder derde orde moment. De ZDP werd weggelaten uit de doelfunctie omdat na controle van de analytische uitdrukkingen deze niet correct bleek te zijn. De weging van de momenten die wel worden opgenomen is impliciet, gezien dit leidt tot een lichte verbetering van de algemene statistische eigenschappen en in zekere mate ook de reproductie van de extremen (Hoofdstuk 8). In de aldus bekomen parameters kan de invloed van de truncatieparameter duidelijk gezien worden. Het kan worden aangetoond dat de verwachtingswaarde van de celduur nu aanzienlijk verkort wordt tot meer realistische waarden. De analyse van de statistische eigenschappen van de gesimuleerde neerslagreeksen wijst erop dat de introductie van de truncatieparameter  $\epsilon$  geen negatieve invloed heeft op de reproductie van deze eigenschappen. Voor de eerste, tweede en derde orde momenten kan zelfs gezien worden dat de gemiddelde neerslaghoogte niet meer systematisch overschat wordt, zoals dat in Hoofdstuk 7 het geval was. Ook de variantie van de neerslaghoogte kent op bepaalde aggregatieniveau's minder grote afwijkingen ten opzichte van de observatie bij het TBL model. Dat de ZDP minder goed benaderd wordt is uiteraard te wijten aan het ontbreken van de ZDP in de doelfunctie. Bij lagere aggregatieniveau's wordt zelfs een verbetering van de droog-nat verhoudingen waargenomen. De algemene verdeling van de intensiteit tijdens de buien (gekenmerkt door de voorwaardelijke distributiekarakteristieken) wordt niet noemenswaardig gewijzigd in vergelijking met de andere modellen. Voor de voorwaardelijke probabiliteiten inzake neerslagintensiteit kan ook over het algemeen besloten worden dat nu eens betere, dan weer minder goede resultaten worden bekomen door het TBL model. Wanneer de extreme waarden geanalyseerd worden blijkt het TBL gemiddeld het best te presteren in vergelijking met alle andere modellen en fits die reeds werden voorgesteld in het huidige werk. Ten slotte werd een reeks simulaties met lengte 1000 jaar uitgevoerd, om



(a) Uurlijkse extremen



(b) Dagelijkse extremen

**Figuur 9.10:** Vergelijking van de jaarlijkse maximale extremen, gebaseerd 105 jaar observatie en simulaties met de verschillende Bartlett-Lewis modellen van telkens 1000 jaar.



na te gaan of het TBL model extreme neerslaggebeurtenissen met grote retourperiode minder overschat dan de andere Bartlett-Lewis modellen. Hierover kunnen voorlopig echter nog geen sluitende uitspraken gedaan worden, verder onderzoek in deze richting lijkt sterk aangewezen. Wel bevestigen de figuren nogmaals de positieve invloed van het derde orde moment op de reproductie van de extreme waarden.



# Hoofdstuk 10

## Besluit

### 10.1 Algemene bevindingen

In het huidige werk wordt getracht om de kalibratie van enkele Bartlett-Lewis modellen efficiënter en gericht te laten verlopen met behulp van een heuristische zoekmethode. Daarnaast wordt onderzocht op welke manier andere aspecten van het kalibratieproces, zoals het type doelfunctie en de gebruikte momenten in de doelfunctie, verder kunnen worden verfijnd.

Om de parametrisatie van de modellen te faciliteren en vooral om een optimalisatiemethode aan te duiden die consequent goede kalibratieresultaten levert worden twee heuristische zoekmethoden met elkaar vergeleken. Beide methoden worden aangewend om het OBL, MBL en het MBLG model te kalibreren. In eerste instantie wordt hierbij nagegaan of de methoden al dan niet in staat zijn om de parameterruimte op effectieve wijze te exploreren. Er wordt met andere woorden nagegaan of beide methoden in staat zijn om, wanneer geen of weinig kennis omtrent de optimale parameters bestaat, na te gaan waar zich de optimale parameterregio bevindt. Uit de resultaten in Hoofdstuk 5 kan besloten worden dat geen van beide methoden hiertoe in staat blijkt.

Wanneer gebruik gemaakt wordt van enige voorkennis omtrent de optimale parameterregio blijken beide methoden in staat behoorlijke resultaten te bekomen. Toch wordt tot het besluit gekomen dat bij toekomstige kalibraties van de Bartlett-Lewis modellen het best geopteerd wordt voor SIMPSA. Deze methode leidt voor alledrie de modellen immers tot goede resultaten. Voor het OBL en MBLG model leidt deze methode tot betere resultaten en vergt ze bovendien minder rekentijd. Bij de kalibratie van het MBL model worden voor PSO en SIMPSA gelijkaardige resultaten bekomen. Toch gaat de voorkeur uit naar SIMPSA, omdat de kans op het vinden van een optimum groter is wanneer SIMPSA gebruikt wordt. Bovendien is SIMPSA sneller en een stuk gebruiksvriendelijker.

Met het oog op een verdere fijnstelling van de optimalisatieprocedure en ter vergelijking van de reeds gekalibreerde modellen onderling, worden de resultaten aan een statistische analyse onderworpen (Hoofdstuk 6). Hieruit kan besloten worden dat de modellen afwisselend beter of slechter presteren ten opzichte van elkaar, zij het dat het OBL model over het algemeen ondermaatse prestaties levert. Het wordt dus duidelijk dat desondanks het vinden van een betrouwbare optimalisatiemethode, de simulatieresultaten op bepaalde vlakken nog steeds aan verbetering toe zijn. Zo kan de reproductie van extreme neerslaggebeurtenissen nog verbeterd worden, samen met andere algemene statistische eigenschappen.

Om de reproductie van extreme neerslaggebeurtenissen te verbeteren wordt onderzocht wat de invloed van de introductie van het derde orde moment in de doelfunctie is op de extreme neerslag. Indien tijdens de kalibratie met deze eigenschap wordt rekening gehouden, wordt verondersteld dat de staart van de verdeling van de neerslag getrouwer wordt nagebootst door het model. De staart van de verdeling staat in voor het voorkomen van uitzonderlijk extreme neerslaggebeurtenissen. Deze hypothese wordt door de resultaten in Hoofdstuk 7 bevestigd. Het in gebruik nemen van het derde orde moment leidt vooral bij lagere aggregatieniveaus tot een aanzienlijk betere reproductie van de extreme neerslaggebeurtenissen. Een probleem dat echter optreedt is dat door de introductie van het derde orde moment bepaalde andere eigenschappen van de neerslag, zoals de gemiddelde neerslaghoogte minder goed worden benaderd door het model. In een poging om tegemoet te komen aan de opgetreden problemen worden in Hoofdstuk 8 twee alternatieve wegingsmethoden voor de verschillende momenten in de doelfunctie onderzocht. Een eerste methode weegt elk van de momenten op impliciete wijze, waardoor de bias die ontstaat bij de kalibratie wordt verkleind. Een tweede methode spitst zich specifiek toe op het teniet doen van de ontstane systematische afwijking van de gemiddelde neerslaghoogte tengevolge van de introductie van het derde orde moment in de doelfunctie. Hierbij worden de momenten gewogen in functie van hun relatieve bijdrage aan de fitness tijdens preliminaire kalibraties. De impliciete weging is niet in staat de systematische afwijking weg te werken. Wel kan besloten worden dat deze over het algemeen een positief effect heeft op de resultaten van het beschouwde model. De tweede methode is wel in staat om de systematische afwijking weg te werken. Het gevolg van deze methode is echter wel dat de voordelen van de introductie van het derde orde moment teniet gedaan worden.

Tot slot wordt in Hoofdstuk 9 een eerste poging gedaan om het TBL model te kalibreren. Controle van de beschikbare analytische momenten toont aan dat de uitdrukkingen voor de ZDP niet correct zijn. Verder onderzoek hieromtrent is dus aangewezen. Er kan uit de kalibratieresultaten voorlopig besloten worden dat het TBL model, op de ZDP bij hogere aggregatieniveaus na, de algemene statistische eigenschappen van de geobserveerde neerslag naar behoren benadert. Verder wordt bekomen dat de reproductie van extremen bij het gebruik van het TBL model nog verbeterd wordt.

Er kan dus gesteld worden dat het onderzoek grotendeels in zijn opzet geslaagd is. Er werd een zoekmethode geïdentificeerd die in staat is om de Bartlett-Lewis modellen op eenvoudige wijze naar behoren te kalibreren. Daarnaast werd de performantie, vooral de reproductie van extreme waarden, verbeterd door de ingebruikname van de impliciete weging van de doelfunctie en de incorporatie van het derde orde moment in de doelfunctie. Dat deze bevindingen extrapol eerbaar zijn naar andere modellen van het type Bartlett-Lewis werd bewezen aan de hand van een succesvolle kalibratie van het TBL model. Over de meerwaarde van het TBL model kan vooralsnog geen uitsluitsel geleverd worden, verder onderzoek is aan de orde.

## 10.2 Toekomstperspectieven

Tijdens het verloop van deze thesis is gebleken dat verschillende aspecten van de neerslagmodellering met de Bartlett-Lewis modellen bijzondere aandacht verdienen tijdens toekomstig onderzoek. Zo bleek de introductie van het derde orde moment in de doelfunctie een nefaste

invloed te hebben op de reproductie van enkele eerste en tweede orde momenten. Het oplossen van het probleem van deze systematische afwijking vormt dus een belangrijk aandachtspunt bij verder onderzoek. Hierbij dient bijzondere aandacht besteed te worden aan de oorzaak van de ontstane afwijking. Bij de kalibratie van het TBL model, waarbij het derde orde moment ook werd gebruikt in doelfunctie, bleek de systematische afwijking afwezig. Dit doet vermoeden dat de afwijking het gevolg is van het opnemen van te veel momenten in de doelfunctie. De vraag stelt zich vervolgens hoe dit kan worden opgelost. Verder kunnen de stopcriteria van de optimalisatiemethode verfijnd worden opdat betere oplossingen voor de doelfunctie worden bekomen, met langere rekentijd tot gevolg. Anderzijds zou nog meer aandacht besteed kunnen worden aan de weging van de verschillende momenten om alsnog op dergelijke wijze de bekomen afwijkingen te spreiden over de verschillende momenten.

In het huidige werk werd enkel gebruik gemaakt van de momentenmethode voor de kalibratie van de Bartlett-Lewis modellen. Deze methode wordt vooral gebruikt omdat andere methoden te rekenintensief zouden zijn. Met behulp van de steeds toenemende rekenkracht die voorhanden is zouden in het verleden afgedankte methoden kunnen herbekeken worden. Daarnaast zou in plaats van analytische momenten in de doelfunctie gebruik kunnen gemaakt worden van synthetische momenten. Het voordeel daarvan zou zijn dat met eigenschappen waarvoor geen analytische uitdrukking bestaat, kan worden rekening gehouden. Zo zou via *multi-objective* kalibratie kunnen rekening gehouden worden met de vooropgestelde voorwaardelijke probabiliteiten en extreme waarden, alsook de algemene statistische eigenschappen.

Wat betreft het TBL model is verder onderzoek nodig om te bepalen of de introductie van de truncatieparameter wel degelijk een oplossing biedt voor het probleem met de generatie van onrealistisch lange neerslagcellen dat wordt waargenomen bij de andere Bartlett-Lewis modellen. Er moet meer bepaald worden nagegaan of de met het TBL model gesimuleerde neerslag nog steeds leidt tot overdreven afvoer in rivieren. De gesimuleerde neerslag dient dus gebruikt te worden als input in een neerslag-afvoermodel om de invloed van de truncatieparameter concreet na te gaan. Verder dient de correcte uitdrukking voor de ZDP geïmplementeerd te worden in het kalibratieproces.



# Bibliografie

- Aarts, E. H. L. & Van Laarhoven, P. J. M. (1985). Statistical cooling - A general-approach to combinatorial optimization problems. *Philips Journal of Research*, 40(4):193–226.
- Aarts, E. H. L. & Van Laarhoven, P. J. M. (1987). *Simulated Annealing : Theory and Applications*. D. Reidel.
- Boughton, W. & Droop, O. (2003). Continuous simulation for design flood estimation - a review. *Environmental Modelling and Software*, 18(4):309–318.
- Box, M. J. (1965). A new method of constrained optimization and a comparison with other methods. *Computer Journal*, 8(1):42–52.
- Bras, R. L. & Rodriguez-Iturbe, I. (1976). Rainfall generation : A nonstationary time-varying multidimensional model. *Water Resources Research*, 12(3):450–456.
- Brutsaert, W. (2005). *Hydrology : an Introduction*. Cambridge University Press.
- Calenda, G. & Napolitano, F. (1999). Parameter estimation of Neyman-Scott processes for temporal point rainfall simulation. *Journal of Hydrology*, 225(1-2):45–66.
- Cameron, D., Beven, K. & Tawn, J. (2000). An evaluation of three stochastic rainfall models. *Journal of Hydrology*, 228(1-2):130–149.
- Cameron, D., Beven, K. & Tawn, J. (2001). Modelling extreme rainfall using a modified random pulse Bartlett-Lewis stochastic rainfall model (with uncertainty). *Advances in Water Resources*, 24(2):203–211.
- Cardoso, M. F., Salcedo, R. L. & De Azevedo, S. F. (1996). The Simplex-Simulated Annealing approach to continuous non-linear optimization. *Computers & Chemical Engineering*, 20(9):1065–1080.
- Chandler, R. E. & Wheeler, H. S. (1998a). Climate change detection using generalized linear models for rainfall - a case study from the West of Ireland. i, Preliminary analysis and modelling of rainfall occurrence. Technical Report 194, Department of Statistical Science, University College, London.
- Chandler, R. E. & Wheeler, H. S. (1998b). Climate change detection using generalized linear models for rainfall - a case study from the West of Ireland. ii, Modelling of rainfall events on wet days. Technical Report 195, Department of Statistical Science, University College, London.

- Cowpertwait, P. S. P. (1991). Further developments of the Neyman-Scott clustered point process for modeling rainfall. *Water Resources Research*, 27(7):1431–1438.
- Cowpertwait, P. S. P. (1994). A generalized point process model for rainfall. *Proceedings of the Royal Society of London Series A-Mathematical Physical and Engineering Sciences*, 447(1929):23–37.
- Cowpertwait, P. S. P. (1998). A Poisson-cluster model of rainfall : high-order moments and extreme values. *Proceedings of the Royal Society of London Series A-Mathematical Physical and Engineering Sciences*, 454(1971):885–898.
- Cowpertwait, P. S. P., Isham, V. & Onof, C. (2007). Point process models of rainfall : developments for fine-scale structure. *Proceedings of the Royal Society of London Series A-Mathematical Physical and Engineering Sciences*, 463(2086):2569–2587.
- Cox, D. R. & Isham, V. (1980). *Point Processes*. Chapman & Hall.
- De Jongh, I. L. M., Verhoest, N. E. C. & De Troch, F. P. (2006). Analysis of a 105-year time series of precipitation observed at Uccle, Belgium. *International Journal of Climatology*, 26:2023–2039.
- Demaree, G. (2003). The centennial recording raingauge of the Uccle Plateau: Its history, its data and its applications. *La Houille Blanche - Revue Internationale de l'Eau*, 4:95–102.
- Directive 2007/60/EC (2007). Directive 2007/60/EC of the European Parliament and of the Council of 23 October 2007 on the Assessment of flood risks. *Official Journal of the European Union*, L288:27–34.
- Donckels, B. (2009). *Optimal experimental design to discriminate among rival dynamic mathematical models*. Ph.D. thesis, Faculty of Bioscience Engineering, Ghent University.
- Engelbrecht, A. (2006). *Fundamentals of Computational Swarm Intelligence*. John Wiley.
- Entekhabi, D., Rodriguez-Iturbe, I. & Eagleson, P. S. (1989). Probabilistic representation of the temporal rainfall process by a Modified Neyman-Scott Rectangular Pulses Model : Parameter estimation and validation. *Water Resources Research*, 25(2):295–302.
- Favre, A. C., Musy, A. & Morgenthaler, S. (2004). Unbiased parameter estimation of the Neyman-Scott model for rainfall simulation with related confidence interval. *Journal of Hydrology*, 286(1-4):168–178.
- Feller, W. (1968). *An Introduction to Probability Theory and its Applications, vol. 1*. John Wiley.
- Foufoula-Georgiou, E. & Lettenmaier, D. (1987). A Markov renewal model for rainfall occurrences. *Water Resources Research*, 23(5):875–884.
- Gabriel, K. R. & Neumann, J. (1962). A Markov chain model for daily rainfall occurrence at Tel-Aviv. *Quarterly Journal of the Royal Meteorological Society*, 88(375):90–95.



- Gupta, V. & Waymire, E. (1994). A statistical analysis of mesoscale rainfall as a random cascade. *Journal of Applied Meteorology*, 32(2):251–267.
- Haan, C. T., Allen, D. M. & Street, J. O. (1976). A Markov chain model of daily rainfall. *Water Resources Research*, 12(3):443–449.
- Heneker, T. M., Lambert, M. F. & Kuczera, G. (2001). A point rainfall model for risk-based design. *Journal of Hydrology*, 247(1-2):54–71.
- Islam, S., Entekhabi, D., Bras, R. & Rodriguez-Iturbe, I. (1990). Parameter estimation and sensitivity analysis for the Modified Bartlett-Lewis rectangular pulses model of rainfall. *Journal of Geophysical Research*, 95(D3):2093–2100.
- Jiang, M., Luo, Y. P. & Yang, S. Y. (2007). Stochastic convergence analysis and parameter selection of the standard Particle Swarm Optimization algorithm. *Information Processing Letters*, 102(1):8–16.
- Kavvas, M. L. & Delleur, J. W. (1981). A stochastic cluster model of daily rainfall sequences. *Water Resources Research*, 17(4):1151–1160.
- Kennedy, J. & Eberhart, R. (1995). Particle swarm optimization. In *Proceedings of the IEEE International conference on neural networks, IV*, pp. 1942–1948. Piscataway, NJ: IEEE Service Center.
- Khaliq, M. N. & Cunnane, C. (1996). Modelling point rainfall occurrences with the Modified Bartlett-Lewis Rectangular Pulses Model. *Journal of Hydrology*, 180(1-4):109–138.
- Kirkpatrick, S. (1984). Optimization by simulated annealing - Quantitative studies. *Journal of Statistical Physics*, 34(5-6):975–986.
- Kirkpatrick, S., Gelatt, C. D. & Vecchi, M. P. (1983). Optimization by Simulated Annealing. *Science*, 220(4598):671–680.
- Kruskal, W. H. & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47(260):584–621.
- Lovejoy, S. & Schertzer, D. (1990). Multifractals, universality classes, and satellite and radar measurements of cloud and rain fields. *Journal of Geophysical Research*, 95(D3):2021–2031.
- Marien, J. L. & Vandewiele, G. L. (1986). A point rainfall generator with internal storm structure. *Water Resources Research*, 22(4):475–482.
- Mason, J. (1986). Numerical weather prediction. *Proceedings of the Royal Society of London Series A-Mathematical Physical and Engineering Sciences*, 407(1832):51–60.
- Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A. & Teller, E. (1953). Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21(6):1087–1092.
- Nelder, J. A. & Mead, R. (1965). A simplex-method for function minimization. *Computer Journal*, 7(4):308–313.

- Neyman, J. & Scott, E. L. (1958). A stochastic approach to problems of cosmology. *Journal of the Royal Statistical Society*, 20(1):1–43.
- Ntegeka, V. & Willems, P. (2008). Trends in multidecadal oscillations in rainfall extremes, based on a more than 100-year time series of 10 min rainfall intensities at Uccle, Belgium. *Water Resources Research*, 44(W07402):DOI: 10.1029/2007WR006471.
- Obeysekera, J. T. B., Tabios III, G. Q. & Salas, J. D. (1987). On parameter estimation of temporal rainfall models. *Water Resources Research*, 23(10):1837–1850.
- Onof, C. (1992). *Stochastic modelling of British rainfall using Poisson processes*. Ph.D. thesis, Imperial College, London.
- Onof, C. (2003). Mathematical expressions of generalised moments used in single-site rainfall models, internal report no. 8, defra project: Improved methods for national spatial-temporal rainfall and evaporation modelling for BSM.
- Onof, C. (2011). Modelling rainfall with a Bartlett-Lewis process. Third-order moments, proportion dry, and a truncated random parameter version. To be submitted for publication.
- Onof, C., Chandler, R. E., Kakou, A., Northrop, P., Wheeler, H. S. & Isham, V. (2000). Rainfall modelling using Poisson-cluster processes: a review of developments. *Stochastic Environmental Research and Risk Assessment*, 14(6):384–411.
- Onof, C. & Wheeler, H. S. (1993). Modelling of British rainfall using a random parameter Bartlett-Lewis Rectangular Pulse Model. *Journal of Hydrology*, 149(1-4):67–95.
- Onof, C. & Wheeler, H. S. (1994a). Improved fitting of the Bartlett-Lewis Rectangular Pulse Model for hourly rainfall. *Hydrological Sciences*, 39(6):663–680.
- Onof, C. & Wheeler, H. S. (1994b). Improvements to the modeling of British rainfall using a modified random parameter Bartlett-Lewis rectangular pulse model. *Journal of Hydrology*, 157(1-4):177–195.
- Onof, C., Wheeler, H. S. & Isham, V. (1994). Note on the analytical expression of the inter-event time characteristics for Bartlett-Lewis type rainfall models. *Journal of Hydrology*, 157(1-4):197–210.
- Press, W. H. & Teukolsky, S. A. (1991). Simulated Annealing over continuous spaces. *Computational Physics*, 5(4):426.
- Rodriguez-Iturbe, I., Cox, D. R. & Isham, V. (1987a). Some models for rainfall based on stochastic point processes. *Proceedings of the Royal Society of London Series A-Mathematical Physical and Engineering Sciences*, 410(1839):269–288.
- Rodriguez-Iturbe, I., Cox, D. R. & Isham, V. (1988). A point process model for rainfall : further developments. *Proceedings of the Royal Society of London Series A-Mathematical Physical and Engineering Sciences*, 417(1853):283–298.

- Rodriguez-Iturbe, I., Febres de Power, B. & Valdés, J. B. (1987b). Rectangular pulses point process models for rainfall: analysis of empirical data. *Journal of Geophysical Research*, 92(D8):9645–9656.
- Scheerlinck, K., Pauwels, V. R. N., Vernieuwe, H. & De Baets, B. (2009). Calibration of a water and energy balance model: Recursive parameter estimation versus particle swarm optimization. *Water Resources Research*, 45(W10422).
- Shi, Y. & Eberhart, R. (1998). A modified particle swarm optimizer. In *Proceedings of the IEEE International Conference on Evolutionary Computation*, pp. 69–73. Piscataway, NJ: IEEE Press.
- Smithers, J. C., Pegram, G. G. S. & Schulze, R. E. (2002). Design rainfall in South Africa using Bartlett-Lewis rectangular pulse rainfall models. *Journal of Hydrology*, 258(1-4):83–99.
- Spendley, W., Hext, G. R. & Himsforth, F. R. (1962). Sequential application of simplex designs in optimisation and evolutionary operation. *Technometrics*, 4(4):441–461.
- Stern, R. D. & Coe, R. (1984). A model-fitting analysis of daily rainfall data. *Journal of the Royal Statistical Society Series A-Statistics in Society*, 147(Part 1):1–34.
- Todorovich, P. & Woolhiser, D. A. (1975). Stochastic model of  $n$ -day precipitation. *Journal of Applied Meteorology*, 14(1):17–24.
- Trelea, I. (2003). The Particle Swarm Optimization algorithm: convergence analysis and parameter selection. *Information Processing Letters*, 85(6):317–325.
- Vandenbergh, S., Verhoest, N. E. C. & De Baets, B. (2010). Fitting bivariate copulas to the dependence structure between storm characteristics: A detailed analysis based on 105 year 10 min rainfall. *Water Resources Research*, 46(W01512).
- Velghe, P. A., Troch, P. A., De Troch, F. P. & Van de Velde, J. (1994). Evaluation of cluster-based rectangular pulses point process models for rainfall. *Water Resources Research*, 30(10):2847–2857.
- Vere-Jones, D. (1970). Stochastic models for earthquake occurrence. *Journal of the Royal Statistical Society*, 32:1–62.
- Verhoest, N. E. C., Troch, P. A. & De Troch, F. P. (1997). On the applicability of Bartlett-Lewis rectangular pulses models in the modeling of design storms at a point. *Journal of Hydrology*, 202(1-4):108–120.
- Verhoest, N. E. C., Vandenbergh, S., Cabus, P., Onof, C., Meca-Figueras, T. & Jameleddine, S. (2010). Are stochastic point rainfall models able to preserve extreme flood statistics? *Hydrological Processes*, 24(23):3439–3445.
- Verhoest, N. E. C., Velghe, T., Troch, P. A. & De Troch, F. P. (1995). Vergelijkende studie van cluster-gebaseerde puntneerslagmodellen met het oog op het bepalen van representatieve buien voor rioolontwerp berekeningen. Technical report, Laboratory of Hydrology and Water Management, University of Ghent, Ghent, Belgium.

- Waymire, E. & Gupta, V. (1981). The mathematical structure of rainfall representations; 1. A review of stochastic rainfall models. *Water Resources Research*, 17(5):1261–1272.
- Weisstein, E. W. (2011). Skewness. URL <http://mathworld.wolfram.com/Skewness.html>. Gelezen: 22/05/2011.
- Wheater, H., Isham, V. S., Chandler, R. E., Onof, C. J. & Stewart, E. J. (2006). Improved methods for national spatial-temporal rainfall and evaporation modelling for BSM. Technical Report F2105/TR, Department for Environment, Food and Rural Affairs, London.
- Wheater, H. S. (2002). Progress in and prospects for fluvial flood modelling. *Philosophical Transactions of the Royal Society A-Mathematical Physical and Engineering Sciences*, 360(1796):1409–1431.
- Wiser, E. H. (1964). *Persistence models in precipitation analysis*. Ph.D. thesis, North Carolina State University, Raleigh.

# Bijlage A

## Analytische modelbeschrijvingen

### A.1 Het originele Bartlett-Lewis model (OBL)

Bij het originele Bartlett-Lewis model manifesteren stormen zich volgens een Poisson proces met parameter  $\lambda$ . Elke stormoorsprong wordt gevolgd door een Poisson proces met parameter  $\beta$  voor de aanmaak van cel-oorsprongen. Na enige tijd, exponentieel verdeeld met parameter  $\gamma$ , worden geen oorsprongen meer aangemaakt. Bovendien wordt verondersteld dat met de stormoorsprong een cel verbonden is. Introduceren we de dimensieloze parameters  $\kappa = \beta/\eta$ ,  $\phi = \gamma/\eta$ , met  $1/\eta$  het gemiddelde van de exponentieel verdeelde celduren, dan volgt dat het aantal cellen per storm een geometrische verdeling met gemiddelde  $\mu_c = 1 + \kappa/\phi$  volgt (Rodriguez-Iturbe *et al.*, 1987a). De hoogte van elke cel, die de intensiteit van die cel voorstelt, volgt een exponentiële distributie met parameter  $1/\mu_x$ . De modelparameters voor het OBL model zijn bijgevolg:

$$\{\lambda, \beta, \gamma, \mu_x, \eta\}$$

Hieronder worden vervolgens de eerste en tweede orde eigenschappen, de ZDP en het derde orde moment van het geaggregeerde proces  $Y_i^{(h)}$ , met  $Y_i^{(h)}$  de cumulatieve neerslag over interval  $i$  met lengte  $h$ .

#### A.1.1 OBL eerste en tweede orde eigenschappen

De eerste en tweede orde momenten worden gegeven door Rodriguez-Iturbe *et al.* (1987a).

$$E[Y_i^{(h)}] = \frac{\lambda\mu_x\mu_c h}{\eta} \quad (\text{A.1})$$

$$\begin{aligned} \text{var}[Y_i^{(h)}] &= 2\lambda\eta^{-2}\mu_c\mu_x^2 h \left(2 + \frac{\beta}{\gamma}\right) + 2\lambda\mu_c\mu_x^2 \frac{\beta}{\gamma^2(\gamma^2 - \eta^2)} (1 - e^{-\gamma h}) \\ &\quad - 2\lambda\mu_c\eta^{-3}\mu_x^2 \left(2 + \frac{\beta\gamma}{\gamma^2 - \eta^2}\right) (1 - e^{-\eta h}) \end{aligned} \quad (\text{A.2})$$

$$\begin{aligned} \text{cov}[Y_i^{(h)}, Y_{i+k}^{(h)}] &= \lambda\mu_c\eta^{-3}\mu_x^2 \left(2 + \frac{\beta\gamma}{\gamma^2 - \eta^2}\right) (1 - e^{-\eta h})^2 e^{-\eta(k-1)h} \\ &\quad - \lambda\mu_c \frac{\beta}{\gamma^2(\gamma^2 - \eta^2)} \mu_x^2 (1 - e^{-\gamma h})^2 e^{-\gamma(k-1)h} \end{aligned} \quad (\text{A.3})$$

### A.1.2 OBL Zero Depth Probability

De kans op een droog interval  $i$  met lengte  $h$  wordt gegeven door Rodriguez-Iturbe *et al.* (1987a):

$$P[Y_i^{(h)} = 0] = \exp \left( -\lambda(h + \mu_T) + \lambda G_p^*(0, 0) \frac{\gamma + \beta e^{-(\beta+\gamma)h}}{\beta + \gamma} \right) \quad (\text{A.4})$$

waarbij, ingeval  $\beta \ll \eta$  en  $\gamma \ll \eta$ ,  $\mu_T$  en  $G_p^*(0, 0)$  als volgt benaderd worden (Onof, 1992):

$$\begin{aligned} \mu_T \approx & \frac{1}{\gamma} \left( 1 + \frac{\gamma(\beta + \gamma/2)}{\eta^2} - \frac{\gamma(5\gamma\beta + \beta^2 + 2\gamma^2)}{4\eta^3} \right) \\ & + \frac{1}{\gamma} \left( \frac{\gamma(4\beta^3 + 31\beta^2\gamma + 99\beta\gamma^2 + 36\gamma^3)}{72\eta^4} \right) \end{aligned} \quad (\text{A.5})$$

$$G_p^*(0, 0) \approx \frac{1}{\gamma} \left( 1 - \frac{\beta + \gamma}{\eta} + \frac{3\beta\gamma + 2\gamma^2 + \beta^2}{2\eta^2} \right) \quad (\text{A.6})$$

Indien niet aan deze voorwaarden voor  $\beta$  en  $\gamma$  wordt voldaan kunnen deze termen als volgt benaderd worden (Onof, 2003):

$$\mu_T \approx \eta^{-1} \left( 1 + \phi \sum_{j=1}^M \frac{(-\kappa)^{j-1} (\kappa - j^2 - j)}{j(j+1)!} B(j+1, \phi) + \phi^{-1} \right) \quad (\text{A.7})$$

en

$$G_p^*(0, 0) \approx \eta^{-1} e^{-\kappa} \left( \sum_{j=0}^{M'} \frac{\kappa^j}{j!} B(j + \phi, 2) + \frac{\delta_{M'}(\kappa)}{(M' + \phi + 1)(M' + \phi + 2)} \right) \quad (\text{A.8})$$

met

$$\delta_{M'} = e^{\kappa} - \sum_{j=0}^{M'} \frac{\kappa^j}{j!} \quad (\text{A.9})$$

waarbij  $\kappa = \beta/\eta$  en  $\phi = \gamma/\eta$ . De waarden voor  $M$  en  $M'$  worden bij voorkeur voldoende groot gekozen om de benaderingsfout te beperken.

### A.1.3 OBL derde orde moment

Het derde centrale moment voor het OBL model wordt gegeven door Onof (2011):

$$E \left[ \left( Y_i^{(h)} - E[Y_i^{(h)}] \right)^3 \right] = \frac{\lambda \mu_c \sum_{k=1}^{k=8} P_k}{\phi^6 - 6\phi^4 + 9\phi^2 - 4} \quad (\text{A.10})$$

waarbij  $\mu_c = 1 + \kappa/\phi$  het gemiddeld aantal cellen per storm is, en de termen  $P_k$  worden gegeven door volgende uitdrukkingen.

$$\begin{aligned}
P_1 &= 6\eta^{-4}e^{-\eta h}\phi^2[\phi\kappa^2\mu_x^3(2\phi^4 - 7\phi^2 - 3\phi + 2) + 2\phi\mu_{x3}(\phi^6 - 6\phi^4 + 9\phi^2 - 4) \\
&\quad + \kappa\mu_x\mu_{x2}(4\phi^6 - 22\phi^4 - \phi^3 + 25\phi^2 + 4\phi - 4)] \\
P_2 &= 6\eta^{-3}e^{-\eta h}\phi^3h[\mu_x^3(\phi^6 - 6\phi^4 + 9\phi^2 - 4) + h\mu_x\mu_{x2}(\phi^2 - 1)(\phi^2 - 4)] \\
P_3 &= 6\eta^{-4}e^{-\eta\phi h}\kappa[\mu_x\mu_x^2(-\phi^5 + \phi^4 + 6\phi^3 - 4\phi^2 - 8\phi) \\
&\quad + 2\kappa\mu_x^3(\phi^5 - 3\phi^4 + 2\phi^3 - 14\phi^2 - 8)] \\
P_4 &= 6\eta^{-3}e^{-\eta\phi h}h\kappa^2\mu_x^3[\phi^3(\phi^2 - 5) - 4\phi] \\
P_5 &= \eta^{-4}[-12\phi^3\mu_{x3}(\phi^6 - 6\phi^4 + 9\phi^2 - 4) \\
&\quad + \kappa^2\mu_x^3(-9\phi^7 + 39\phi^5 + 18\phi^4 - 12\phi^3 - 84\phi^2 + 48) \\
&\quad - 3\phi\kappa\mu_x\mu_{x2}(7\phi^7 - 39\phi^5 - 2\phi^4 + 46\phi^3 + 12\phi^2 - 4\phi - 16)] \\
P_6 &= \eta^{-3}[(6h\phi^3\mu_{x3} + 12h\phi^2\kappa\mu_x\mu_{x2} + 6h\phi\kappa^2\mu_x^3)(\phi^6 - 6\phi^4 + 9\phi^2 - 4)] \\
P_7 &= 3\eta^{-4}e^{-2\eta h}\phi^4(1 - \phi^2)[\phi\kappa^2\mu_x^3 + \kappa\mu_x\mu_{x2}(\phi^2 - 4)] \\
P_8 &= 6\eta^{-4}e^{-(1+\phi)\eta h}\phi^2(2 - \phi)(\phi - 1)[\phi\kappa^2\mu_x^3\kappa\mu_x\mu_{x2}(\phi + 2)]
\end{aligned}$$

waarbij

$$\mu_{x2} = 2\mu_x^2$$

en

$$\mu_{x3} = 6\mu_x^3$$

#### A.1.4 Het gewijzigde Bartlett-Lewis model (MBL)

In de veronderstelling dat  $\eta$  variabel is van storm tot storm, worden de verhoudingen  $\kappa$  en  $\phi$  verondersteld constant te zijn.  $\eta$  is een toevalsveranderlijke met een gamma verdeling met parameters  $\alpha$  en  $\nu$ . Een fysische verantwoording voor deze aanname is dat stormen met een gemiddeld lagere intensiteit de neiging hebben om langer te duren (Rodriguez-Iturbe *et al.*, 1988). De modelparameters van het MBL model zijn de volgende:

$$\{\lambda, \kappa, \phi, \mu_x, \alpha, \nu\}$$

De eerste en tweede orde eigenschappen, de ZDP en het derde orde moment worden hieronder weergegeven.

#### A.1.5 MBL eerste en tweede orde eigenschappen

De eerste en tweede orde eigenschappen voor het MBL model van het geaggregeerde proces  $Y_i^{(h)}$ , met  $Y_i^{(h)}$  de cumulatieve neerslag over interval  $i$  met lengte  $h$  werden berekend door Islam *et al.*

(1990).

$$E[Y_i^{(h)}] = \lambda \mu_x \mu_c \frac{\nu}{\alpha - 1} h \quad (\text{A.11})$$

$$\begin{aligned} \text{var}[Y_i^{(h)}] &= \frac{2\nu^{2-\alpha} h}{\alpha - 2} \left( k_1 - \frac{k_2}{\phi} \right) - \frac{2\nu^{3-\alpha}}{(\alpha - 2)(\alpha - 3)} \left( k_1 - \frac{k_2}{\phi^2} \right) \\ &\quad + \frac{2}{(\alpha - 2)(\alpha - 3)} \left( k_1 (h + \nu)^{3-\alpha} - \frac{k_2}{\phi^2} (\phi h + \nu)^{3-\alpha} \right) \end{aligned} \quad (\text{A.12})$$

$$\begin{aligned} \text{cov}[Y_i^{(h)}, Y_{i+k}^{(h)}] &= \frac{k_1}{(\alpha - 2)(\alpha - 3)} [(h(k - 1) + \nu)^{3-\alpha} + (h(k + 1) + \nu)^{3-\alpha} - 2(h\kappa + \nu)^{3-\alpha}] \\ &\quad + \frac{k_2}{\phi^2(\alpha - 2)(\alpha - 3)} [2(\phi h k + \nu)^{3-\alpha} - (\phi h(k - 1) + \nu)^{3-\alpha} - (\phi h(k + 1) + \nu)^{3-\alpha}] \end{aligned} \quad (\text{A.13})$$

met

$$k_1 = \lambda \mu_c \mu_x^2 \left( 2 + \frac{\kappa \phi}{\phi^2 - 1} \right) \frac{\nu^\alpha}{\alpha - 1} \quad (\text{A.14})$$

$$k_2 = \frac{\lambda \mu_c \mu_x^2 \kappa}{\phi^2 - 1} \left( \frac{\nu^\alpha}{\alpha - 1} \right) \quad (\text{A.15})$$

### A.1.6 MBL Zero Depth Probability

Indien  $\kappa \ll 1$  en  $\phi \ll 1$  kan de ZDP van het MBL model benaderd worden door onderstaande vergelijking (Onof, 1992):

$$\begin{aligned} P[Y_i^{(h)} = 0] &\approx \exp \left\{ -\lambda h - \frac{\lambda \nu}{\phi(\alpha - 1)} \left[ 1 + \phi \left( \kappa + \frac{\phi}{2} \right) - \frac{1}{4} \phi (5\phi \kappa + \kappa^2 + 2\phi^2) + \frac{1}{72} \phi (4\kappa^3 + 31\kappa^2 \phi + 99\kappa \phi^2 + 36\phi^3) \right] \right. \\ &\quad \left. + \frac{\lambda \nu}{\phi(\alpha - 1)} \left( 1 - \kappa - \phi + \frac{3}{2} \kappa \phi + \phi^2 + \frac{1}{2} \kappa^2 \right) \left[ \frac{\phi}{\phi + \kappa} + \frac{\kappa}{\phi + \kappa} \left( \frac{\nu}{\nu + (\kappa + \phi)h} \right)^{\alpha - 1} \right] \right\} \end{aligned} \quad (\text{A.16})$$

Indien echter niet voldaan wordt aan de voorwaarden omtrent kleine waarden voor  $\kappa$  en  $\phi$  wordt de proportie droge intervallen als volgt benaderd (Onof, 2003):

$$P[Y_i^{(h)} = 0] \approx \exp \left[ -\lambda(h + \hat{\mu}_{T_M}) + \frac{\lambda \nu e^{-\kappa}}{\alpha - 1} \times \frac{\phi + \kappa \left( \frac{\nu}{\nu + (\kappa + \phi)h} \right)^{\alpha - 1}}{\phi + \kappa} \hat{I}_{M'} \right] \quad (\text{A.17})$$

$$(\text{A.18})$$

met

$$\hat{\mu}_{T_M} = \frac{\nu}{\alpha - 1} \left( 1 + \phi \sum_{j=1}^M \frac{(-\kappa)^{(j-1)} (\kappa - j^2 - j)}{j(j+1)!} B(j+1, \phi) + \phi^{-1} \right) \quad (\text{A.19})$$

en

$$\hat{I}_{M'} = \sum_{j=0}^{M'} \frac{\kappa^j}{j!} B(j + \phi, 2) + \frac{\delta_{M'}(\kappa)}{(M' + \phi + 1)(M' + \phi + 2)} \quad (\text{A.20})$$

met

$$\delta_{M'}(\kappa) = e^\kappa - \sum_{j=0}^{M'} \frac{\kappa^j}{j!} \quad (\text{A.21})$$

waarbij  $M$  en  $M'$  groot genoeg moeten worden genomen voor een goede benadering.



### A.1.7 MBL derde orde moment

Het derde orde moment voor het MBL model werd afgeleid door Onof (2011) en wordt hieronder weergegeven:

$$\mathbb{E} \left[ \left( Y_i^{(h)} - \mathbb{E}[Y_i^{(h)}] \right)^3 \right] = \frac{\lambda \mu_c \sum_{k=1}^{k=8} Q_k}{\phi^6 - 6\phi^4 + 9\phi^2 - 4} \quad (\text{A.22})$$

waarbij de termen  $Q_k$  gegeven worden door onderstaande vergelijkingen:

$$\begin{aligned} Q_1 &= 6T(4, h)\phi^2[\phi\kappa^2\mu_x^3(2\phi^4 - 7\phi^2 - 3\phi + 2) + 2\phi\mu_{x3}(\phi^6 - 6\phi^4 + 9\phi^2 - 4) \\ &\quad + \kappa\mu_x\mu_{x2}(4\phi^6 - 22\phi^4 - \phi^3 + 25\phi^2 + 4\phi - 4)] \\ Q_2 &= 6T(3, h)\phi^3h[\mu_x^3(\phi^6 - 6\phi^4 + 9\phi^2 - 4) + h\mu_x\mu_{x2}(\phi^2 - 1)(\phi^2 - 4)] \\ Q_3 &= 6T(4, \phi h)\kappa[\mu_x\mu_x^2(-\phi^5 + \phi^4 + 6\phi^3 - 4\phi^2 - 8\phi) \\ &\quad + 2\kappa\mu_x^3(\phi^5 - 3\phi^4 + 2\phi^3 - 14\phi^2 - 8)] \\ Q_4 &= 6T(3, \phi h)h\kappa^2\mu_x^3[\phi^3(\phi^2 - 5) - 4\phi] \\ Q_5 &= T(4, 0)[-12\phi^3\mu_{x3}(\phi^6 - 6\phi^4 + 9\phi^2 - 4) \\ &\quad + \kappa^2\mu_x^3(-9\phi^7 + 39\phi^5 + 18\phi^4 - 12\phi^3 - 84\phi^2 + 48) \\ &\quad - 3\phi\kappa\mu_x\mu_{x2}(7\phi^7 - 39\phi^5 - 2\phi^4 + 46\phi^3 + 12\phi^2 - 4\phi - 16)] \\ Q_6 &= T(3, 0)[(6\eta h\phi^3\mu_{x3} + 12h\phi^2\kappa\mu_x\mu_{x2} + 6h\phi\kappa^2\mu_x^3)(\phi^6 - 6\phi^4 + 9\phi^2 - 4)] \\ Q_7 &= 3T(4, 2h)\phi^4(1 - \phi^2)[\phi\kappa^2\mu_x^3 + \kappa\mu_x\mu_{x2}(\phi^2 - 4)] \\ Q_8 &= 6T(4, (1 + \phi)h)\phi^2(2 - \phi)(\phi - 1)[\phi\kappa^2\mu_x^3\kappa\mu_x\mu_{x2}(\phi + 2)] \end{aligned}$$

met

$$T(k, \alpha) = \frac{\nu^\alpha}{(\nu + a)^{\alpha-k}} \frac{\Gamma(\alpha - k)}{\Gamma(\alpha)}$$

## A.2 Het gewijzigde Bartlett-Lewis Gamma model (MBLG)

Het gebruik van de Gamma distributie voor het randomiseren van de gemiddelde celduur bij verschillende stormen levert een betere reproductie van de ZDP en tweede orde karakteristieken op (Rodriguez-Iturbe *et al.*, 1988; Entekhabi *et al.*, 1989; Onof & Wheeler, 1993; Velghe *et al.*, 1994). Niettegenstaande is er voor de reproductie van extreme waarden nog steeds ruimte voor verbetering (Onof & Wheeler, 1994b).

Onof & Wheeler (1994b) ontwikkelen met het oog op een betere reproductie van extreme waarden het MBLG model. Het gebruik van een twee parameter Gamma distributie met vormparameter  $p$  en schaalparameter  $\delta$  in plaats van een exponentiële distributie van de celintensiteit zou leiden tot een verbeterde representatie van extreme waarden. De zwaardere staart van de gamma verdeling zou hiervoor kunnen instaan, zoals reeds werd geopperd door Rodriguez-Iturbe *et al.* (1987a).

De zeven modelparameters voor het MBLG model zijn de volgende:

$$\{\lambda, \kappa, \phi, \alpha, \nu, p, \delta\}$$

In onderstaande secties worden de eerste en tweede orde eigenschappen en de ZDP weergegeven.

### A.2.1 MBLG eerste en tweede orde eigenschappen

De eerste en tweede orde eigenschappen voor het MBLG model worden gegeven door Onof & Wheeler (1994b):

$$E[Y_i^{(h)}] = \lambda \mu_c \mu_x \frac{\nu}{\alpha - 1} h \quad (\text{A.23})$$

$$\begin{aligned} \text{var}[Y_i^{(h)}] &= 2A_1 [(\alpha - 3)h\nu^{2-\alpha} - \nu^{3-\alpha} + (\nu + h)^{3-\alpha}] \\ &\quad - 2A_2 [\phi(\alpha - 3)h\nu^{2-\alpha} - \nu^{3-\alpha} + (\nu + \phi h)^{3-\alpha}] \end{aligned} \quad (\text{A.24})$$

$$\begin{aligned} \text{cov}[Y_i^{(h)}, Y_{i+k}^{(h)}] &= A_1 \{[\nu + (k+1)h]^{3-\alpha} - 2(\nu + \kappa h)^{3-\alpha} + [\nu + (k-1)h]^{3-\alpha}\} \\ &\quad - A_2 \{[\nu + (k+1)\phi h]^{3-\alpha} - 2(\nu + \kappa \phi h)^{3-\alpha} + [\nu + (k-1)\phi h]^{3-\alpha}\} \end{aligned} \quad (\text{A.25})$$

met

$$A_1 = \frac{\lambda \mu_c \nu^\alpha}{\delta^2 (\alpha - 1)(\alpha - 2)(\alpha - 3)} \left[ p(p+1) + \frac{\kappa \phi p^2}{\phi^2 - 1} \right] \quad (\text{A.26})$$

$$A_2 = \frac{\lambda \mu_c \mu_x^2 \kappa \nu^\alpha}{\phi^2 (\phi^2 - 1)(\alpha - 1)(\alpha - 2)(\alpha - 3)} \quad (\text{A.27})$$

### A.2.2 MBLG Zero Depth Probability

De kans op droge intervallen is identiek aan deze bij het MBL model, wordt voor de volledigheid hieronder herhaald. Indien  $\kappa \ll 1$  en  $\phi \ll 1$  kan de ZDP van het MBLG model benaderd worden door onderstaande vergelijking (Onof, 1992):

$$\begin{aligned} P[Y_i^{(h)} = 0] &\approx \exp \left\{ -\lambda h - \frac{\lambda \nu}{\phi(\alpha-1)} \left[ 1 + \phi \left( \kappa + \frac{\phi}{2} \right) - \frac{1}{4} \phi (5\phi \kappa + \kappa^2 + 2\phi^2) + \frac{1}{72} \phi (4\kappa^3 + 31\kappa^2 \phi + 99\kappa \phi^2 + 36\phi^3) \right] \right. \\ &\quad \left. + \frac{\lambda \nu}{\phi(\alpha-1)} \left( 1 - \kappa - \phi + \frac{3}{2} \kappa \phi + \phi^2 + \frac{1}{2} \kappa^2 \right) \left[ \frac{\phi}{\phi + \kappa} + \frac{\kappa}{\phi + \kappa} \left( \frac{\nu}{\nu + (\kappa + \phi)h} \right)^{\alpha-1} \right] \right\} \end{aligned} \quad (\text{A.28})$$

Indien echter niet voldaan wordt aan de voorwaarden omtrent kleine waarden voor  $\kappa$  en  $\phi$  wordt de proportie droge intervallen als volgt benaderd (Onof, 2003):

$$P[Y_i^{(h)} = 0] \approx \exp \left[ -\lambda(h + \hat{\mu}_{T_M}) + \frac{\lambda \nu e^{-\kappa}}{\alpha - 1} \times \frac{\phi + \kappa \left( \frac{\nu}{\nu + (\kappa + \phi)h} \right)^{\alpha-1}}{\phi + \kappa} \hat{I}_{M'} \right] \quad (\text{A.29})$$

$$(\text{A.30})$$

met

$$\hat{\mu}_{T_M} = \frac{\nu}{\alpha - 1} \left( 1 + \phi \sum_{j=1}^M \frac{(-\kappa)^{(j-1)} (\kappa - j^2 - j)}{j(j+1)!} B(j+1, \phi) + \phi^{-1} \right) \quad (\text{A.31})$$

en

$$\hat{I}_{M'} = \sum_{j=0}^{M'} \frac{\kappa^j}{j!} B(j + \phi, 2) + \frac{\delta_{M'}(\kappa)}{(M' + \phi + 1)(M' + \phi + 2)} \quad (\text{A.32})$$

met

$$\delta_{M'}(\kappa) = e^\kappa - \sum_{j=0}^{M'} \frac{\kappa^j}{j!} \quad (\text{A.33})$$

waarbij  $M$  en  $M'$  groot genoeg moeten worden genomen voor een goede benadering.

### A.3 Het Truncated Bartlett-Lewis model (TBL)

Het TBL model werd ontwikkeld door Onof (2011) om het voorkomen van onrealistisch lange cellen tegen te gaan (Verhoest *et al.*, 2010). De parameters van het TBL model zijn dezelfde als deze van het MBL model, met dat verschil dat er een zevende parameter bijkomt die een restrictie van  $\eta$  oplegt. Op die manier wordt verhinderd dat extreem lage waarden voor  $\eta$  getrokken worden uit de Gamma distributie, wat ervoor zorgt dat de gemiddelde celduur tijdens een storm ( $1/\eta$ ) en de gemiddelde duur van de storm  $(\eta\phi)^{-1}$  niet te hoog is en het gemiddeld aantal celorsprongen per storm ( $\eta\kappa$ ) niet onrealistisch klein wordt. Er wordt dus vermeden dat er zich onrealistisch lange stormen met slechts enkele cellen van gemiddelde intensiteit en lange duur voordoen die door de statistische en extreme waarden analyse niet opgemerkt worden. Het aantal parameters wordt verhoogd tot zeven (ten opzichte van het MBL model) door de intrede van de truncatie parameter  $\epsilon$ .

$$\{\lambda, \kappa, \phi, \mu_x, \alpha, \nu, \epsilon\}$$

Onderstaande secties geven de eerste en tweede orde eigenschappen van het TBL model, alsook de ZDP en de derde orde eigenschap, zoals beschreven door Onof (2011).

#### A.3.1 TBL eerste en tweede orde eigenschappen

De eerste en tweede orde eigenschappen van het TBL model worden als volgt gedefinieerd (Onof, 2011):

$$E[Y_i^{(h)}] = T(\epsilon, 1, 0)\lambda h\mu_x\mu_c \quad (\text{A.34})$$

$$\begin{aligned} \text{var}[Y_i^{(h)}] &= 2\lambda\mu_c \left[ T(\epsilon, 2, 0)h \left( \mu_{x2} + \mu_x^2 \frac{\kappa}{\phi} \right) + T(\epsilon, 3, 0) \left( \frac{\kappa\mu_x^2}{\phi^2 - 1} \left( \frac{1}{\phi^2} - \phi \right) - \mu_{x2} \right) \right. \\ &\quad \left. - T(\epsilon, 3, \phi h) \left( \frac{\kappa\mu_x^2}{\phi^2(\phi^2 - 1)} \right) - T(\epsilon, 3, h) \left( \frac{\kappa\phi\mu_x^2}{\phi^2 - 1} \right) \right] \end{aligned} \quad (\text{A.35})$$

$$\begin{aligned} \text{cov}[Y_i^{(h)}, Y_{i+k}^{(h)}] &= \lambda\mu_c \left[ \mu_{x2} + \frac{\kappa\phi\mu_x^2}{\phi^2 - 1} \right] [T(\epsilon, 3, (k-1)h) + T(\epsilon, 3, (k+1)h) - 2T(\epsilon, 3, kh)] \\ &\quad - \lambda\mu_c \frac{\kappa\mu_x^2}{\phi^2(\phi^2 - 1)} [T(\epsilon, 3, \phi(k-1)h) + T(\epsilon, 3, \phi(k+1)h) - 2T(\epsilon, 3, \phi kh)] \end{aligned} \quad (\text{A.36})$$

met

$$T(\epsilon, k, a) \approx \frac{\frac{\Gamma(\alpha-k)}{(a+\nu)^{\alpha-k}} - \epsilon^{\alpha-k} \sum_{n=0}^{N-1} (-1)^n \frac{(\alpha+\nu)^n \epsilon^n}{(\alpha-k+n)n!}}{\frac{\Gamma(\alpha)}{(\nu)^\alpha} - \epsilon \sum_{n=0}^{N-1} (-1)^n \frac{\nu^n \epsilon^n}{(\alpha+n)n!}} \quad (\text{A.37})$$

en

$$\mu_{x2} = 2\mu_x^2 \quad (\text{A.38})$$

### A.3.2 TBL Zero Depth Probability

De kans dat in cel  $i$  van lengte  $h$  geen neerslag valt wordt gegeven door de ZDP (Onof, 2011):

$$\begin{aligned} P[Y_i^{(h)} = 0] &= \exp \left\{ -\lambda \left[ h + \left( \frac{1}{\phi} + \phi f(\kappa, \phi) \right) T(\epsilon, -1, 0) \right] \right. \\ &\quad \left. + \frac{\lambda e^{-\kappa} g(\kappa, \phi)}{\phi + \kappa} [\phi T(\epsilon, -1, 0) + \kappa T(\epsilon, -1, (\kappa + \phi)h)] \right\} \end{aligned} \quad (\text{A.39})$$

met

$$f(\kappa, \phi) \approx \frac{1}{\phi} + \sum_{j=1}^M \frac{(-\kappa)^{j-1} (\kappa - j^2 - j)}{j(j+1)!} B(j+1, \phi) \quad (\text{A.40})$$

en

$$g(\kappa, \phi) \approx \sum_{j=0}^M \frac{\kappa^j}{j!} B(j + \phi, 2) + \frac{e^{\kappa} - \sum_{j=0}^M \frac{\kappa^j}{j!}}{(M + \phi + 1)(M + \phi + 2)} \quad (\text{A.41})$$

waarbij  $M$  groot genoeg dient gekozen te worden voor een goede benadering. Voor richtlijnen met betrekking tot de keuze van  $M$  wordt verwezen naar Onof (2011).

### A.3.3 TBL derde orde moment

Het centrale derde orde moment voor het TBL wordt gegeven door Onof (2011):

$$\mathbb{E} \left[ \left( Y_i^{(h)} - \mathbb{E}[Y_i^{(h)}] \right)^3 \right] = \frac{\lambda \mu_c \sum_{k=1}^{k=8} R_k}{\phi^6 - 6\phi^4 + 9\phi^2 - 4} \quad (\text{A.42})$$

waarbij de termen  $R_k$  gegeven worden door onderstaande vergelijkingen.

$$\begin{aligned} R_1 &= 6T(\epsilon, 4, h)\phi^2[\phi\kappa^2\mu_x^3(2\phi^4 - 7\phi^2 - 3\phi + 2) + 2\phi\mu_{x3}(\phi^6 - 6\phi^4 + 9\phi^2 - 4) \\ &\quad + \kappa\mu_x\mu_{x2}(4\phi^6 - 22\phi^4 - \phi^3 + 25\phi^2 + 4\phi - 4)] \\ R_2 &= 6T(\epsilon, 3, h)\phi^3h[\mu_x^3(\phi^6 - 6\phi^4 + 9\phi^2 - 4) + h\mu_x\mu_{x2}(\phi^2 - 1)(\phi^2 - 4)] \\ R_3 &= 6T(\epsilon, 4, \phi h)\kappa[\mu_x\mu_x^2(-\phi^5 + \phi^4 + 6\phi^3 - 4\phi^2 - 8\phi) \\ &\quad + 2\kappa\mu_x^3(\phi^5 - 3\phi^4 + 2\phi^3 - 14\phi^2 - 8)] \\ R_4 &= 6T(\epsilon, 3, \phi h)h\kappa^2\mu_x^3[\phi^3(\phi^2 - 5) - 4\phi] \\ R_5 &= T(\epsilon, 4, 0)[-12\phi^3\mu_{x3}(\phi^6 - 6\phi^4 + 9\phi^2 - 4) \\ &\quad + \kappa^2\mu_x^3(-9\phi^7 + 39\phi^5 + 18\phi^4 - 12\phi^3 - 84\phi^2 + 48) \\ &\quad - 3\phi\kappa\mu_x\mu_{x2}(7\phi^7 - 39\phi^5 - 2\phi^4 + 46\phi^3 + 12\phi^2 - 4\phi - 16)] \\ R_6 &= T(\epsilon, 3, 0)[(6\eta h\phi^3\mu_{x3} + 12h\phi^2\kappa\mu_x\mu_{x2} + 6h\phi\kappa^2\mu_x^3)(\phi^6 - 6\phi^4 + 9\phi^2 - 4)] \\ R_7 &= 3T(\epsilon, 4, 2h)\phi^4(1 - \phi^2)[\phi\kappa^2\mu_x^3 + \kappa\mu_x\mu_{x2}(\phi^2 - 4)] \\ R_8 &= 6T(\epsilon, 4, (1 + \phi)h)\phi^2(2 - \phi)(\phi - 1)[\phi\kappa^2\mu_x^3\kappa\mu_x\mu_{x2}(\phi + 2)] \end{aligned}$$

Waarbij  $T(\epsilon, k, a)$  gegeven wordt door Vergelijking A.37.

**Bijlage B**

**Tabellen**

**Tabel B.1:** Historische momenten voor Ukkel (1898-2002).

Maand	AN	Avg (mm)	Var (mm <sup>2</sup> )	Cov (mm <sup>2</sup> )	Cor(1) (-)	Cor(2) (-)	Cor(3) (-)	ZDP (-)	Skew (-)
Jan									
	10 min	0.01532	0.00573	0.00376	0.65698	0.51637	0.44005	0.92383	12.72454
	0.5 u	0.04596	0.03843	0.02365	0.61550	0.43165	0.34336	0.88015	9.11847
	1 u	0.09191	0.12483	0.06973	0.55858	0.35701	0.25999	0.84423	7.33889
	6 u	0.55147	2.14465	0.68751	0.32057	0.12678	0.07549	0.66651	4.72563
	12 u	1.10295	5.81539	1.37390	0.23625	0.09787	0.10308	0.53948	4.09157
	24 u	2.20590	14.67946	3.15068	0.21463	0.10359	0.03190	0.38402	3.16911
Feb									
	10 min	0.01368	0.00498	0.00325	0.65285	0.50709	0.42708	0.93233	11.64976
	0.5 u	0.04104	0.03311	0.02012	0.60766	0.42749	0.34084	0.89361	8.56160
	1 u	0.08207	0.10627	0.05991	0.56378	0.35194	0.25760	0.86140	7.01956
	6 u	0.49244	1.81923	0.57463	0.31586	0.12581	0.10272	0.70051	4.60354
	12 u	0.98487	4.89085	1.16938	0.23909	0.13086	0.12447	0.57858	3.66148
	24 u	1.96975	12.32244	2.95642	0.23992	0.15098	0.10692	0.41990	2.89834
Mar									
	10 min	0.01348	0.00575	0.00352	0.61244	0.47018	0.39855	0.93661	14.95142
	0.5 u	0.04045	0.03689	0.02155	0.58407	0.41419	0.32138	0.90058	9.47763
	1 u	0.08091	0.11723	0.06373	0.54366	0.33660	0.24507	0.87094	7.62738
	6 u	0.48543	1.93105	0.62729	0.32485	0.12408	0.09495	0.71444	4.87967
	12 u	0.97086	5.22495	1.27722	0.24445	0.11098	0.07858	0.60338	4.04449
	24 u	1.94172	13.02736	2.73853	0.21021	0.11894	0.11546	0.45253	2.94370
Apr									
	10 min	0.01331	0.00667	0.00380	0.56997	0.39121	0.32024	0.94148	16.40199
	0.5 u	0.03993	0.04071	0.02069	0.50832	0.33502	0.26613	0.90601	10.62923
	1 u	0.07986	0.12311	0.05835	0.47394	0.29901	0.21410	0.87439	8.25055
	6 u	0.47918	1.88099	0.55843	0.29688	0.12436	0.08275	0.71143	4.69328
	12 u	0.95837	4.93116	1.16226	0.23570	0.11970	0.08054	0.59651	3.76689
	24 u	1.91673	12.33683	2.69366	0.21834	0.10289	0.06247	0.45746	3.05255
Mei									
	10 min	0.01388	0.01423	0.00781	0.54850	0.34013	0.25283	0.95145	28.52766
	0.5 u	0.04164	0.08317	0.03689	0.44358	0.23849	0.17035	0.92232	19.40366
	1 u	0.08328	0.23526	0.09432	0.40092	0.18305	0.12089	0.89549	14.27870
	6 u	0.49969	3.02546	0.68129	0.22519	0.08205	0.06316	0.75284	6.94533
	12 u	0.99937	7.54009	1.33098	0.17652	0.08968	0.02928	0.64255	5.17709
	24 u	1.99874	17.28832	3.11053	0.17992	0.03113	0.03054	0.49002	3.79164
Jun									
	10 min	0.01524	0.01801	0.00942	0.52306	0.30529	0.22468	0.95211	28.96479
	0.5 u	0.04573	0.10284	0.04177	0.40619	0.21173	0.16632	0.92201	20.46491
	1 u	0.09145	0.29180	0.10172	0.34860	0.19396	0.12589	0.89472	15.28024
	6 u	0.54871	3.68783	0.86965	0.23582	0.06940	0.04330	0.74849	7.41782
	12 u	1.09741	9.30122	1.44568	0.15543	0.05874	0.03718	0.64238	6.09695
	24 u	2.19483	21.10057	2.96013	0.14029	0.05594	0.03537	0.49841	4.59554

**Tabel B.1:** Historische momenten voor Ukkel (1898-2002) (vervolg).

Maand	AN	Avg (mm)	Var (mm <sup>2</sup> )	Cov (mm <sup>2</sup> )	Cor(1) (-)	Cor(2) (-)	Cor(3) (-)	ZDP (-)	Skew (-)
Jul									
	10 min	0.01776	0.03258	0.01682	0.51613	0.26406	0.17743	0.95163	33.81639
	0.5 u	0.05328	0.18794	0.06217	0.33079	0.16891	0.11044	0.92137	26.08646
	1 u	0.10657	0.51242	0.14264	0.27836	0.11609	0.08540	0.89341	19.56643
	6 u	0.63940	5.31901	1.07559	0.20222	0.07407	0.04789	0.74409	7.61418
	12 u	1.27880	12.96461	2.03046	0.15662	0.06579	0.01821	0.63886	5.61871
	24 u	2.55760	29.97458	3.96276	0.13220	0.02148	0.04888	0.49309	3.84328
Aug									
	10 min	0.01601	0.02405	0.01171	0.48717	0.27619	0.20215	0.95676	28.31798
	0.5 u	0.04802	0.13075	0.05159	0.39460	0.20953	0.14862	0.92966	17.48358
	1 u	0.09605	0.36368	0.12694	0.34903	0.17482	0.11581	0.90341	12.84827
	6 u	0.57630	4.41489	1.00845	0.22842	0.07348	0.05551	0.75799	6.84232
	12 u	1.15260	11.02947	1.80300	0.16347	0.06192	0.03001	0.65069	5.65852
	24 u	2.30519	25.36604	3.61744	0.14261	0.04562	0.06167	0.50077	4.31955
Sep									
	10 min	0.01470	0.01567	0.00830	0.52955	0.31634	0.23581	0.95080	32.66932
	0.5 u	0.04410	0.09009	0.03803	0.42212	0.23907	0.18399	0.92040	21.13152
	1 u	0.08820	0.25403	0.09838	0.38730	0.21096	0.13954	0.89263	13.61904
	6 u	0.52917	3.26487	0.83222	0.25490	0.10172	0.05654	0.74484	6.24810
	12 u	1.05835	8.07200	1.72055	0.21315	0.11462	0.08644	0.63317	4.62278
	24 u	2.11670	19.46112	4.09514	0.21043	0.09873	0.04921	0.48413	3.45279
Okt									
	10 min	0.01586	0.01019	0.00615	0.60345	0.43608	0.36223	0.93873	18.69057
	0.5 u	0.04759	0.06361	0.03539	0.55641	0.36983	0.29150	0.90375	11.61336
	1 u	0.09518	0.20083	0.10012	0.49855	0.30198	0.22014	0.87385	10.03645
	6 u	0.57108	3.08663	1.04943	0.33999	0.14607	0.10386	0.71935	5.15415
	12 u	1.14215	8.39627	2.22488	0.26498	0.14340	0.08617	0.60968	4.41269
	24 u	2.28430	21.47647	5.25431	0.24465	0.12662	0.04329	0.45745	3.46940
Nov									
	10 min	0.01733	0.00769	0.00504	0.65465	0.50020	0.42676	0.92311	12.48816
	0.5 u	0.05198	0.05115	0.03103	0.60667	0.44325	0.35625	0.88008	8.72036
	1 u	0.10396	0.16424	0.09485	0.57748	0.37005	0.26571	0.84429	6.96541
	6 u	0.62376	2.93013	0.94437	0.32230	0.13029	0.11479	0.66794	4.48033
	12 u	1.24752	7.64682	2.07266	0.27105	0.14655	0.09757	0.55000	3.52833
	24 u	2.49505	19.63558	4.93389	0.25127	0.12409	0.09108	0.38603	2.90976
Dec									
	10 min	0.01710	0.00698	0.00476	0.68149	0.54032	0.47056	0.92102	11.05278
	0.5 u	0.05129	0.04780	0.03041	0.63621	0.45862	0.35962	0.87769	8.27027
	1 u	0.10258	0.15714	0.09127	0.58085	0.35943	0.24967	0.84304	7.20774
	6 u	0.61545	2.76081	0.84518	0.30614	0.12032	0.09226	0.66521	5.46360
	12 u	1.23091	7.31097	1.72238	0.23559	0.14719	0.11492	0.53810	4.31139
	24 u	2.46181	18.02008	4.61034	0.25584	0.13979	0.05743	0.37112	3.09097

**Tabel B.2:** Voorwaardelijke distributiekarakteristieken voor Ukkel (1898-2002).

Maand	AN	vwd 1 (-)	vwd 2 (mm)	vwd 3 (mm <sup>2</sup> )	vwd 4 (mm)	vwd 5 (mm <sup>2</sup> )	vwd 6 (mm)	vwd 7 (mm <sup>2</sup> )	vwd 8 (mm)	vwd 9 (mm <sup>2</sup> )
Jan										
	10 min	0.8976	0.2011	0.0378	0.1384	0.0182	0.1281	0.0094	0.2341	0.0450
	0.5 u	0.8442	0.3834	0.1913	0.2483	0.1058	0.1992	0.0533	0.4413	0.2167
	1 u	0.7961	0.5900	0.5075	0.3908	0.2428	0.2980	0.1520	0.6792	0.6002
	6 u	0.5382	1.6537	4.6089	1.2796	2.5071	1.1161	2.1974	1.8875	5.7818
	12 u	0.3809	2.3950	9.5346	1.9456	5.9411	1.6665	5.4268	2.6309	11.2630
	24 u	0.2283	3.5811	18.9080	2.6207	10.1170	2.4879	9.9755	3.9062	21.4740
Feb										
	10 min	0.9089	0.2022	0.0355	0.1417	0.0171	0.1303	0.0093	0.2343	0.0424
	0.5 u	0.8614	0.3857	0.1782	0.2657	0.1142	0.2109	0.0659	0.4379	0.1971
	1 u	0.8171	0.5921	0.4647	0.4084	0.2507	0.3216	0.1810	0.6785	0.5420
	6 u	0.5776	1.6442	4.1812	1.3099	2.5675	1.1255	2.4947	1.8777	5.1760
	12 u	0.4198	2.3371	8.4470	2.0499	7.1137	1.8130	6.3257	2.5105	9.1845
	24 u	0.2632	3.3955	16.4030	2.6508	11.4960	2.4533	10.6200	3.6721	17.9640
Mar										
	10 min	0.9152	0.2127	0.0483	0.1487	0.0261	0.1331	0.0152	0.2454	0.0565
	0.5 u	0.8701	0.4069	0.2220	0.2845	0.1472	0.2176	0.0805	0.4610	0.2455
	1 u	0.8290	0.6269	0.5661	0.4604	0.3768	0.3332	0.2037	0.7070	0.6376
	6 u	0.5988	1.6999	4.6985	1.2977	2.7985	1.1282	2.6793	1.9741	5.8112
	12 u	0.4521	2.4478	9.5600	1.9293	6.6086	1.7527	7.0174	2.7685	11.1230
	24 u	0.2978	3.5467	18.1060	3.0371	15.8440	2.3619	12.1310	3.7431	18.8560
Apr										
	10 min	0.9207	0.2274	0.0653	0.1659	0.0411	0.1420	0.0182	0.2614	0.0754
	0.5 u	0.8738	0.4248	0.2696	0.3233	0.2110	0.2384	0.1177	0.4779	0.2921
	1 u	0.8297	0.6358	0.6267	0.4985	0.4834	0.3481	0.2490	0.7118	0.6899
	6 u	0.5946	1.6605	4.5573	1.4176	3.4766	1.0687	1.9970	1.8260	5.2277
	12 u	0.4486	2.3752	8.8576	2.1508	7.5296	1.6627	4.9880	2.5051	9.5854
	24 u	0.3014	3.5329	17.0320	2.7951	11.4640	2.4065	11.0740	3.8311	18.9850
Mei										
	10 min	0.9342	0.2859	0.2155	0.1987	0.1388	0.1520	0.0372	0.3341	0.2513
	0.5 u	0.8952	0.5361	0.8057	0.4256	0.7408	0.2704	0.2139	0.5954	0.8306
	1 u	0.8567	0.7969	1.6827	0.6784	1.6134	0.4275	0.6479	0.8670	1.7107
	6 u	0.6392	2.0217	9.1658	1.8664	9.7455	1.4851	6.0136	2.1538	8.6401
	12 u	0.4896	2.7958	16.0750	2.7164	16.0040	2.3263	12.8020	2.8566	16.1420
	24 u	0.3282	3.9192	26.3780	3.6901	26.3000	3.0205	15.7500	4.0252	26.4250
Jun										
	10 min	0.9347	0.3182	0.2796	0.2266	0.2012	0.1594	0.0455	0.3705	0.3169
	0.5 u	0.8942	0.5863	1.0018	0.4972	0.9942	0.2883	0.2565	0.6356	0.9993
	1 u	0.8550	0.8687	2.0968	0.7825	2.3306	0.4684	0.7196	0.9209	1.9483
	6 u	0.6393	2.1817	11.1030	2.0493	10.0380	1.5439	6.9212	2.2831	11.9010
	12 u	0.4921	3.0687	19.9640	2.8366	15.0270	2.4049	12.6850	3.2383	23.5000
	24 u	0.3350	4.3758	32.5320	4.1350	25.4110	3.1476	22.2420	4.4950	35.9730



**Tabel B.2:** Voorwaardelijke distributiekarakteristieken voor Ukkel (1898-2002) (vervolg).

Maand	AN	vwd 1 (-)	vwd 2 (mm)	vwd 3 (mm <sup>2</sup> )	vwd 4 (mm)	vwd 5 (mm <sup>2</sup> )	vwd 6 (mm)	vwd 7 (mm <sup>2</sup> )	vwd 8 (mm)	vwd 9 (mm <sup>2</sup> )
Jul										
	10 min	0.9342	0.3672	0.5453	0.2519	0.3140	0.1689	0.0749	0.4322	0.6641
	0.5 u	0.8932	0.6777	1.9672	0.5609	1.7318	0.3355	0.5876	0.7430	2.0873
	1 u	0.8528	0.9998	3.9147	0.8910	4.0657	0.5310	1.4292	1.0668	3.8107
	6 u	0.6352	2.4985	16.1430	2.3511	17.9650	1.7666	10.0700	2.6077	14.7730
	12 u	0.4895	3.5410	27.8950	3.3769	30.0150	2.5004	16.3710	3.6570	26.3860
	24 u	0.3233	5.0455	46.5890	4.9691	48.1340	3.4989	33.4470	5.0885	45.8690
Aug										
	10 min	0.9412	0.3702	0.4250	0.2523	0.3208	0.1683	0.0675	0.4365	0.4714
	0.5 u	0.9029	0.6828	1.4255	0.5494	1.4254	0.3182	0.4011	0.7646	1.4081
	1 u	0.8644	0.9944	2.8721	0.8602	2.7005	0.5359	1.1887	1.0852	2.9685
	6 u	0.6478	2.3813	13.9470	2.2128	12.2940	1.6940	7.5510	2.5220	15.2930
	12 u	0.4993	3.2996	24.4970	3.2415	23.1640	2.6277	16.0960	3.3441	25.5290
	24 u	0.3408	4.6175	40.1420	4.4583	31.8950	3.6923	28.9700	4.6926	44.0490
Sep										
	10 min	0.9331	0.2987	0.2337	0.1972	0.1110	0.1508	0.0430	0.3555	0.2933
	0.5 u	0.8927	0.5540	0.8494	0.4028	0.5852	0.2704	0.4069	0.6348	0.9719
	1 u	0.8530	0.8214	1.7638	0.6413	1.5073	0.4086	0.9150	0.9267	1.8839
	6 u	0.6335	2.0739	9.5938	1.7111	7.4688	1.4305	6.2874	2.3545	11.0620
	12 u	0.4825	2.8852	16.7380	2.5751	14.7240	2.0668	11.2080	3.1012	18.0390
	24 u	0.3172	4.1031	29.5800	3.4459	24.0900	2.6248	15.5150	4.4168	31.9200
Okt										
	10 min	0.9181	0.2589	0.1035	0.1691	0.0570	0.1408	0.0197	0.3046	0.1209
	0.5 u	0.8735	0.4944	0.4400	0.3171	0.2415	0.2253	0.1226	0.5759	0.5101
	1 u	0.8324	0.7545	1.0946	0.5206	0.7203	0.3269	0.2410	0.8691	1.2382
	6 u	0.6036	2.0349	8.0211	1.6267	5.7916	1.2211	3.8177	2.3214	9.3902
	12 u	0.4560	2.9262	16.2940	2.3695	11.3860	1.8550	8.1841	3.2874	19.1560
	24 u	0.2916	4.2103	31.4810	2.9430	15.1430	2.6059	14.8650	4.7685	37.6760
Nov										
	10 min	0.8976	0.2253	0.0532	0.1465	0.0193	0.1327	0.0100	0.2644	0.0654
	0.5 u	0.8441	0.4335	0.2612	0.2720	0.1248	0.2071	0.0680	0.5028	0.3038
	1 u	0.7962	0.6676	0.6785	0.4339	0.3346	0.3051	0.2009	0.7722	0.7969
	6 u	0.5419	1.8784	6.4680	1.4836	4.7797	1.1137	2.5453	2.1200	7.3493
	12 u	0.3829	2.7723	12.7680	2.1987	8.7052	1.9136	7.5405	3.1107	14.8620
	24 u	0.2169	4.0638	25.6090	3.3090	21.2610	2.9331	17.1560	4.3502	26.9770
Dec										
	10 min	0.8949	0.2165	0.0453	0.1423	0.0169	0.1299	0.0097	0.2532	0.0553
	0.5 u	0.8422	0.4193	0.2365	0.2610	0.1076	0.2027	0.0513	0.4840	0.2747
	1 u	0.7953	0.6535	0.6411	0.4314	0.3121	0.3151	0.1861	0.7506	0.7540
	6 u	0.5341	1.8383	5.9991	1.4677	3.7592	1.1837	3.0430	2.0775	7.2991
	12 u	0.3701	2.6648	12.0080	2.1525	9.6675	1.8875	6.3855	2.9586	13.1190
	24 u	0.2127	3.9146	22.9700	2.8116	14.1370	2.7705	13.8050	4.2887	25.4170

**Tabel B.3:** Voorwaardelijke karakteristieken inzake neerslagintensiteit voor Ukkel (1898-2002).

Maand	AN	DP <sub>0.5</sub>	DP <sub>1</sub>	DP <sub>2</sub>	DP <sub>3</sub>	DP <sub>4</sub>	DP <sub>5</sub>
Jan							
	10 min	0.9946	0.9993	0.9999	1.0000	1.0000	1.0000
	0.5 u	0.9677	0.9902	0.9987	0.9997	0.9999	0.9999
	1 u	0.9367	0.9714	0.9921	0.9978	0.9992	0.9996
	6 u	0.7849	0.8486	0.9079	0.9405	0.9605	0.9765
	12 u	0.6688	0.7436	0.8238	0.8734	0.9080	0.9376
	24 u	0.5063	0.5942	0.6863	0.7496	0.8009	0.8495
Feb							
	10 min	0.9950	0.9993	0.9999	1.0000	1.0000	1.0000
	0.5 u	0.9707	0.9913	0.9989	0.9997	0.9999	1.0000
	1 u	0.9429	0.9738	0.9934	0.9980	0.9995	0.9998
	6 u	0.8021	0.8600	0.9147	0.9467	0.9665	0.9790
	12 u	0.6956	0.7649	0.8418	0.8877	0.9175	0.9408
	24 u	0.5413	0.6219	0.7123	0.7828	0.8246	0.8651
Mar							
	10 min	0.9947	0.9992	0.9999	1.0000	1.0000	1.0000
	0.5 u	0.9713	0.9906	0.9985	0.9996	0.9998	1.0000
	1 u	0.9452	0.9739	0.9925	0.9976	0.9991	0.9997
	6 u	0.8153	0.8673	0.9159	0.9462	0.9660	0.9775
	12 u	0.7149	0.7782	0.8442	0.8857	0.9180	0.9439
	24 u	0.5748	0.6464	0.7260	0.7825	0.8227	0.8633
Apr							
	10 min	0.9942	0.9989	0.9998	0.9999	1.0000	1.0000
	0.5 u	0.9719	0.9903	0.9982	0.9994	0.9998	0.9999
	1 u	0.9474	0.9741	0.9922	0.9973	0.9990	0.9995
	6 u	0.8150	0.8683	0.9206	0.9498	0.9668	0.9775
	12 u	0.7071	0.7743	0.8478	0.8940	0.9235	0.9421
	24 u	0.5622	0.6387	0.7241	0.7841	0.8318	0.8651
Mei							
	10 min	0.9931	0.9979	0.9994	0.9997	0.9998	0.9999
	0.5 u	0.9749	0.9887	0.9965	0.9985	0.9992	0.9995
	1 u	0.9551	0.9750	0.9896	0.9951	0.9974	0.9985
	6 u	0.8427	0.8820	0.9213	0.9462	0.9621	0.9739
	12 u	0.7492	0.7982	0.8532	0.8932	0.9184	0.9386
	24 u	0.6065	0.6645	0.7389	0.7911	0.8298	0.8661
Jun							
	10 min	0.9918	0.9972	0.9992	0.9997	0.9998	0.9999
	0.5 u	0.9736	0.9872	0.9954	0.9979	0.9988	0.9993
	1 u	0.9535	0.9734	0.9877	0.9936	0.9963	0.9977
	6 u	0.8327	0.8767	0.9179	0.9414	0.9594	0.9698
	12 u	0.7362	0.7875	0.8465	0.8829	0.9102	0.9306
	24 u	0.5984	0.6508	0.7219	0.7721	0.8143	0.8483

**Tabel B.3:** Voorwaardelijke karakteristieken inzake neerslagintensiteit voor Ukkel (1898-2002) (vervolg).

Maand	AN	DP <sub>0.5</sub>	DP <sub>1</sub>	DP <sub>2</sub>	DP <sub>3</sub>	DP <sub>4</sub>	DP <sub>5</sub>
Jul							
	10 min	0.9908	0.9966	0.9988	0.9994	0.9996	0.9998
	0.5 u	0.9721	0.9860	0.9944	0.9970	0.9983	0.9989
	1 u	0.9515	0.9709	0.9861	0.9922	0.9951	0.9968
	6 u	0.8323	0.8704	0.9085	0.9345	0.9522	0.9631
	12 u	0.7372	0.7851	0.8369	0.8711	0.8982	0.9174
	24 u	0.5972	0.6559	0.7204	0.7668	0.8018	0.8304
Aug							
	10 min	0.9911	0.9968	0.9990	0.9995	0.9997	0.9998
	0.5 u	0.9742	0.9863	0.9946	0.9972	0.9984	0.9990
	1 u	0.9564	0.9730	0.9863	0.9924	0.9954	0.9969
	6 u	0.8405	0.8815	0.9187	0.9415	0.9569	0.9664
	12 u	0.7452	0.7986	0.8505	0.8854	0.9103	0.9273
	24 u	0.6061	0.6654	0.7330	0.7807	0.8187	0.8504
Sep							
	10 min	0.9921	0.9977	0.9994	0.9997	0.9999	0.9999
	0.5 u	0.9734	0.9877	0.9961	0.9982	0.9990	0.9994
	1 u	0.9535	0.9738	0.9888	0.9942	0.9967	0.9980
	6 u	0.8411	0.8803	0.9191	0.9430	0.9573	0.9698
	12 u	0.7473	0.7995	0.8497	0.8881	0.9100	0.9324
	24 u	0.6108	0.6714	0.7365	0.7867	0.8235	0.8562
Okt							
	10 min	0.9918	0.9981	0.9997	0.9999	0.9999	1.0000
	0.5 u	0.9681	0.9865	0.9969	0.9988	0.9995	0.9997
	1 u	0.9441	0.9695	0.9876	0.9949	0.9977	0.9988
	6 u	0.8192	0.8654	0.9124	0.9365	0.9545	0.9664
	12 u	0.7246	0.7811	0.8399	0.8763	0.9037	0.9260
	24 u	0.5859	0.6522	0.7250	0.7770	0.8120	0.8412
Nov							
	10 min	0.9924	0.9988	0.9998	1.0000	1.0000	1.0000
	0.5 u	0.9628	0.9864	0.9977	0.9995	0.9998	0.9999
	1 u	0.9328	0.9655	0.9888	0.9963	0.9985	0.9994
	6 u	0.7882	0.8437	0.8994	0.9314	0.9516	0.9660
	12 u	0.6770	0.7419	0.8113	0.8595	0.8932	0.9179
	24 u	0.5191	0.5946	0.6765	0.7394	0.7854	0.8257
Dec							
	10 min	0.9929	0.9990	0.9999	1.0000	1.0000	1.0000
	0.5 u	0.9636	0.9876	0.9979	0.9994	0.9998	0.9999
	1 u	0.9320	0.9671	0.9897	0.9966	0.9987	0.9994
	6 u	0.7805	0.8402	0.8988	0.9310	0.9531	0.9694
	12 u	0.6630	0.7353	0.8085	0.8610	0.8956	0.9223
	24 u	0.4949	0.5807	0.6676	0.7355	0.7834	0.8267

**Tabel B.4:** Parameters, bekomen met SIMPSA, met bijbehorende fitness voor de verschillende maanden voor het OBL model.

		<i>fitness</i>	$\lambda$	$\beta$	$\gamma$	$\mu_x$	$\eta$
Jan	SIMPSA	8.39E-09	0.01977	2.42439	0.11333	1.16896	5.63091
	Verhoest <i>et al.</i> (1995)	9.32E-12	0.02256	1.37291	0.13026	0.78086	2.21912
Feb	SIMPSA	1.12E-07	0.01798	2.05842	0.10032	1.16616	5.49648
	Verhoest <i>et al.</i> (1995)	2.98E-10	0.01605	1.19224	0.10898	0.78845	1.95310
Mrt	SIMPSA	1.87E-08	0.01744	2.28602	0.11466	1.43162	6.46060
	Verhoest <i>et al.</i> (1995)	2.07E-12	0.01781	1.41145	0.10577	0.90306	2.35172
Apr	SIMPSA	1.07E-07	0.01782	1.72289	0.10676	1.82392	6.97736
	Verhoest <i>et al.</i> (1995)	4.26E-16	0.01598	1.51887	0.10944	0.99482	3.15626
Mei	SIMPSA	1.08E-08	0.01540	0.93394	0.11315	3.95685	6.77255
	Verhoest <i>et al.</i> (1995)	5.56E-10	0.01182	0.12040	0.02547	2.53722	1.84259
Jun	SIMPSA	1.18E-08	0.01639	1.16780	0.14486	4.66411	7.57192
	Verhoest <i>et al.</i> (1995)	1.39E-11	0.01823	0.50564	0.11685	2.54774	2.32332
Jul	SIMPSA	9.41E-09	0.01653	0.73205	0.13658	7.41204	7.30627
	Verhoest <i>et al.</i> (1995)	2.54E-12	0.01504	0.14857	0.04650	3.98859	2.59499
Aug	SIMPSA	9.74E-09	0.01505	1.06247	0.14025	6.23702	8.38003
	Verhoest <i>et al.</i> (1995)	2.85E-10	0.01831	0.25022	0.11583	4.12242	2.84264
Sep	SIMPSA	2.70E-08	0.01435	0.94319	0.09888	4.20894	7.21620
	Verhoest <i>et al.</i> (1995)	1.41E-12	0.01028	0.31689	0.04458	1.98336	1.98253
Okt	SIMPSA	2.89E-08	0.01380	1.78615	0.09655	2.22968	6.30267
	Verhoest <i>et al.</i> (1995)	1.35E-10	0.01191	0.46545	0.05795	1.10153	1.32052
Nov	SIMPSA	2.80E-08	0.01770	2.09044	0.09614	1.41542	5.48124
	Verhoest <i>et al.</i> (1995)	3.53E-10	0.01793	0.36206	0.05313	0.90730	1.14009
Dec	SIMPSA	1.63E-08	0.01867	1.85993	0.09308	1.26938	4.84693
	Verhoest <i>et al.</i> (1995)	4.87E-12	0.01930	0.73522	0.09090	0.84365	1.50765

**Tabel B.5:** Parameters, bekomen met SIMPSA, met bijbehorende fitness voor de verschillende maanden voor het MBL model.

		<i>fitness</i>	$\lambda$	$\kappa$	$\phi$	$\mu_x$	$\alpha$	$\nu$
Jan	SIMPSA	3.80E-07	0.03665	0.96826	0.01942	1.21210	2.66110	0.06755
	Verhoest <i>et al.</i> (1995)	6.24E-08	0.03669	0.97145	0.01888	1.22169	2.65683	0.06476
Feb	SIMPSA	5.18E-06	0.03324	0.93401	0.01675	1.24262	2.58450	0.05544
	Verhoest <i>et al.</i> (1995)	4.78E-04	0.03337	0.92673	0.01450	1.30815	2.57115	0.04552
Mrt	SIMPSA	3.38E-06	0.03023	0.73961	0.01790	1.58698	2.77684	0.07080
	Verhoest <i>et al.</i> (1995)	7.32E-04	0.03035	0.74867	0.01635	1.63984	2.75567	0.06101
Apr	SIMPSA	2.63E-08	0.03003	0.57360	0.01622	2.18762	2.74000	0.05816
	Verhoest <i>et al.</i> (1995)	6.92E-05	0.03004	0.57430	0.01608	2.19564	2.73791	0.05733
Mei	SIMPSA	6.76E-05	0.02671	0.21910	0.02813	4.25879	2.88322	0.15699
	Verhoest <i>et al.</i> (1995)	3.44E-04	0.02525	0.14117	0.02125	4.25149	3.22418	0.22589
Jun	SIMPSA	8.26E-05	0.02586	0.19145	0.02669	5.01071	3.47365	0.21372
	Verhoest <i>et al.</i> (1995)	2.88E-04	0.02511	0.15161	0.02325	4.98078	3.87868	0.27995
Jul	SIMPSA	3.43E-04	0.02676	0.14587	0.03028	8.00817	3.20524	0.18871
	Verhoest <i>et al.</i> (1995)	2.30E-04	0.02617	0.11663	0.02608	7.97147	3.43343	0.22738
Aug	SIMPSA	8.84E-05	0.02627	0.17624	0.02528	7.02055	3.03551	0.13301
	Verhoest <i>et al.</i> (1995)	1.42E-04	0.02534	0.12585	0.02079	6.88912	3.35176	0.18361
Sep	SIMPSA	2.96E-05	0.02542	0.13090	0.01695	4.65823	2.85401	0.15835
	Verhoest <i>et al.</i> (1995)	2.45E-05	0.02543	0.13113	0.01697	4.65825	2.85323	0.15817
Okt	SIMPSA	2.84E-08	0.02953	0.50285	0.02330	2.42497	2.42762	0.08401
	Verhoest <i>et al.</i> (1995)	6.70E-04	0.02951	0.49923	0.02333	2.42421	2.42916	0.08489
Nov	SIMPSA	1.66E-07	0.03632	0.87424	0.02056	1.45862	2.45134	0.06543
	Verhoest <i>et al.</i> (1995)	9.16E-04	0.03635	0.87820	0.02026	1.46366	2.44968	0.06388
Dec	SIMPSA	2.93E-07	0.03792	1.03706	0.02395	1.20522	2.44538	0.07324
	Verhoest <i>et al.</i> (1995)	7.02E-04	0.03795	1.04513	0.02350	1.20833	2.44324	0.07098

**Tabel B.6:** Parameters, bekomen met SIMPSA, met bijbehorende fitness voor de verschillende maanden voor het MBLG model.

<i>fitness</i>		$\lambda$	$\kappa$	$\phi$	$\alpha$	$\nu$	$p$	$\delta$	
Jan	SIMPSA	8.39E-09	0.03519	0.36746	0.02884	2.81116	0.17500	4.73050	2.40662
	Verhoest <i>et al.</i> (1995)	9.73E-05	0.03510	0.37699	0.02885	2.82746	0.17883	4.01645	2.11184
Feb	SIMPSA	1.12E-07	0.03092	0.27349	0.02572	2.81075	0.20235	5.37340	2.63491
	Verhoest <i>et al.</i> (1995)	3.29E-04	0.03086	0.26963	0.02574	2.81952	0.20485	5.80974	2.82263
Mrt	SIMPSA	1.87E-08	0.02872	0.28282	0.02368	3.03635	0.18323	4.12857	1.70696
	Verhoest <i>et al.</i> (1995)	6.59E-05	0.02867	0.29760	0.02361	3.05414	0.18596	3.21234	1.40223
Apr	SIMPSA	1.07E-07	0.02774	0.17573	0.01945	3.14369	0.19316	4.21374	1.32687
	Verhoest <i>et al.</i> (1995)	4.93E-04	0.02743	0.19801	0.01917	3.26184	0.21002	2.44561	0.88354
Mei	SIMPSA	1.08E-08	0.02563	0.11475	0.02145	3.06054	0.20224	1.51791	0.29212
	Verhoest <i>et al.</i> (1995)	2.09E-03	0.02494	0.17239	0.02074	3.37806	0.24843	0.68241	0.19875
Jun	SIMPSA	1.18E-08	0.02594	0.13744	0.02663	3.36400	0.20360	1.96799	0.29541
	Verhoest <i>et al.</i> (1995)	4.15E-04	0.02539	0.23657	0.02511	3.77577	0.25615	0.60385	0.16122
Jul	SIMPSA	9.41E-09	0.02657	0.05365	0.02493	3.03546	0.18288	7.20407	0.50654
	Verhoest <i>et al.</i> (1995)	5.87E-04	0.02494	0.18476	0.02237	4.39360	0.36077	0.38456	0.08858
Aug	SIMPSA	9.74E-09	0.02634	0.08489	0.02430	2.89169	0.12592	6.45218	0.52750
	Verhoest <i>et al.</i> (1995)	4.84E-04	0.02546	0.19795	0.02211	3.36958	0.17956	0.57104	0.11418
Sep	SIMPSA	2.70E-08	0.02568	0.07150	0.01624	2.68476	0.13976	4.41579	0.57486
	Verhoest <i>et al.</i> (1995)	9.74E-05	0.02494	0.14277	0.01602	2.96134	0.17482	0.75994	0.18985
Okt	SIMPSA	2.89E-08	0.02818	0.26649	0.02188	2.50948	0.12917	1.61667	0.53859
	Verhoest <i>et al.</i> (1995)	8.86E-07	0.02826	0.28046	0.02195	2.50223	0.12763	1.44293	0.50136
Nov	SIMPSA	2.80E-08	0.03422	0.33924	0.02586	2.56969	0.16192	2.37726	1.14082
	Verhoest <i>et al.</i> (1995)	7.13E-05	0.03426	0.31898	0.02594	2.56372	0.16108	2.99219	1.35034
Dec	SIMPSA	1.63E-08	0.03506	0.30843	0.02901	2.60285	0.20980	2.75685	1.43418
	Verhoest <i>et al.</i> (1995)	8.01E-05	0.03514	0.29604	0.02917	2.58952	0.20709	3.29130	1.63793

**Tabel B.7:** Parameters, bekomen met SIMPSA, met bijbehorende fitness voor de verschillende maanden voor het MBL model, gekalibreerd met inbegrip van het derde centrale moment.

	<i>fitness</i>	$\lambda$	$\kappa$	$\phi$	$\mu_x$	$\alpha$	$\nu$
Jan	2.63E-02	0.03409	0.28851	0.01737	1.83830	2.67989	0.12911
Feb	1.29E-02	0.02888	0.24213	0.01568	1.61174	2.86244	0.18912
Mrt	2.73E-02	0.02738	0.20526	0.01378	2.25394	2.94575	0.14792
Apr	2.48E-02	0.02634	0.14611	0.01142	2.78826	3.08331	0.15209
Mei	5.55E-02	0.02103	0.03719	0.00604	6.80813	3.16073	0.15679
Jun	5.07E-02	0.02530	0.07528	0.01340	8.08038	3.14649	0.13016
Jul	5.60E-02	0.02288	0.02820	0.00795	12.37300	3.71675	0.20059
Aug	2.94E-02	0.02420	0.05432	0.01073	9.76384	3.25694	0.13994
Sep	8.04E-02	0.01148	0.01762	0.00138	8.02897	3.00349	0.12262
Okt	3.55E-02	0.01775	0.05566	0.00341	3.81720	2.49185	0.11146
Nov	1.98E-02	0.02946	0.16003	0.01071	2.15483	2.56299	0.15013
Dec	1.34E-02	0.03171	0.20547	0.01466	1.78565	2.59165	0.18172

**Tabel B.8:** Parameters, bekomen met SIMPSA, met bijbehorende fitness voor de verschillende maanden voor het MBL model, met impliciete weging.

	<i>fitness</i>	$\lambda$	$\kappa$	$\phi$	$\mu_x$	$\alpha$	$\nu$
Jan	1.57E-04	0.03616	0.88435	0.02522	1.14347	2.72359	0.10629
Feb	2.66E-05	0.03316	0.92806	0.01801	1.21609	2.59349	0.06173
Mrt	2.35E-05	0.03012	0.72359	0.01907	1.55632	2.79729	0.07963
Apr	7.96E-05	0.02947	0.50760	0.01998	2.00514	2.84171	0.09422
Mei	2.58E-04	0.02779	0.36406	0.03352	4.29642	2.67116	0.09831
Jun	1.97E-04	0.02234	0.08601	0.01546	4.86214	8.65265	0.98246
Jul	7.94E-04	0.02213	0.05014	0.01338	7.66332	9.83826	1.17098
Aug	2.16E-04	0.02184	0.06188	0.01227	6.55801	6.70205	0.63328
Sep	1.46E-04	0.01771	0.05206	0.00715	4.39724	8.67316	1.05080
Okt	2.06E-06	0.02960	0.51622	0.02324	2.42231	2.42352	0.08145
Nov	2.46E-06	0.03629	0.86720	0.02103	1.45078	2.45425	0.06801
Dec	1.61E-06	0.03790	1.03262	0.02424	1.20278	2.44674	0.07466

**Tabel B.9:** Parameters, bekomen met SIMPSA, met bijbehorende fitness voor de verschillende maanden voor het MBL model, gekalibreerd met inbegrip van het derde centrale moment, met impliciete weging.

	<i>fitness</i>	$\lambda$	$\kappa$	$\phi$	$\mu_x$	$\alpha$	$\nu$
Jan	4.99E-02	0.03424	0.30162	0.01722	1.84577	2.72002	0.12639
Feb	2.49E-02	0.02899	0.24878	0.01576	1.61188	2.89015	0.18865
Mrt	5.19E-02	0.02751	0.21310	0.01382	2.26091	3.00199	0.14819
Apr	4.72E-02	0.02649	0.15161	0.01156	2.79624	3.13925	0.15324
Mei	1.04E-01	0.02218	0.04334	0.00707	6.85876	3.22137	0.15598
Jun	9.50E-02	0.02556	0.08251	0.01399	8.15266	3.20459	0.12882
Jul	1.05E-01	0.02354	0.03203	0.00884	12.47442	3.80043	0.20089
Aug	5.55E-02	0.02441	0.05773	0.01115	9.81123	3.29292	0.13954
Sep	1.49E-01	0.01466	0.02050	0.00201	8.16190	3.04337	0.12138
Okt	6.68E-02	0.02002	0.06547	0.00439	3.83265	2.51882	0.11039
Nov	3.78E-02	0.03005	0.17242	0.01130	2.15430	2.59072	0.14839
Dec	2.59E-02	0.03208	0.21658	0.01508	1.78416	2.61070	0.17907

**Tabel B.10:** Parameters, bekomen met SIMPSA, met bijbehorende fitness voor de verschillende maanden voor het MBL model, gewogen volgens de relatieve bijdrage aan de fitness.

	<i>fitness</i>	$\lambda$	$\kappa$	$\phi$	$\mu_x$	$\alpha$	$\nu$
Jan	3.15E-05	0.03668	1.01132	0.02070	1.17304	2.73688	0.07440
Feb	2.86E-03	0.03335	1.27358	0.03303	0.98156	11.41594	0.66017
Mar	1.48E-05	0.03035	0.77151	0.01735	1.58963	2.79030	0.06602
Apr	2.92E-05	0.03003	0.59178	0.01737	2.12743	2.84633	0.06584
Mei	5.41E-03	0.01987	0.06904	0.01021	3.97289	6.46439	0.75687
Jun	6.66E-03	0.02179	0.08729	0.01395	4.61078	7.70285	0.85810
Jul	1.88E-02	0.01800	0.03106	0.00307	9.39110	2.20955	0.06959
Aug	9.07E-03	0.02077	0.06006	0.01112	6.05292	10.96253	1.21542
Sep	5.14E-03	0.01857	0.06055	0.00824	4.14388	10.10715	1.27572
Okt	1.74E-05	0.02947	0.50450	0.02467	2.37988	2.49351	0.09446
Nov	2.95E-03	0.03638	1.12701	0.04269	1.18137	12.08700	0.97878
Dec	3.52E-05	0.03795	1.08554	0.02565	1.16685	2.52961	0.08179

**Tabel B.11:** Parameters, bekomen met SIMPSA, met bijbehorende fitness voor de verschillende maanden voor het MBL model, gekalibreerd met inbegrip van het derde centrale moment, gewogen volgens de relatieve bijdrage aan de fitness.

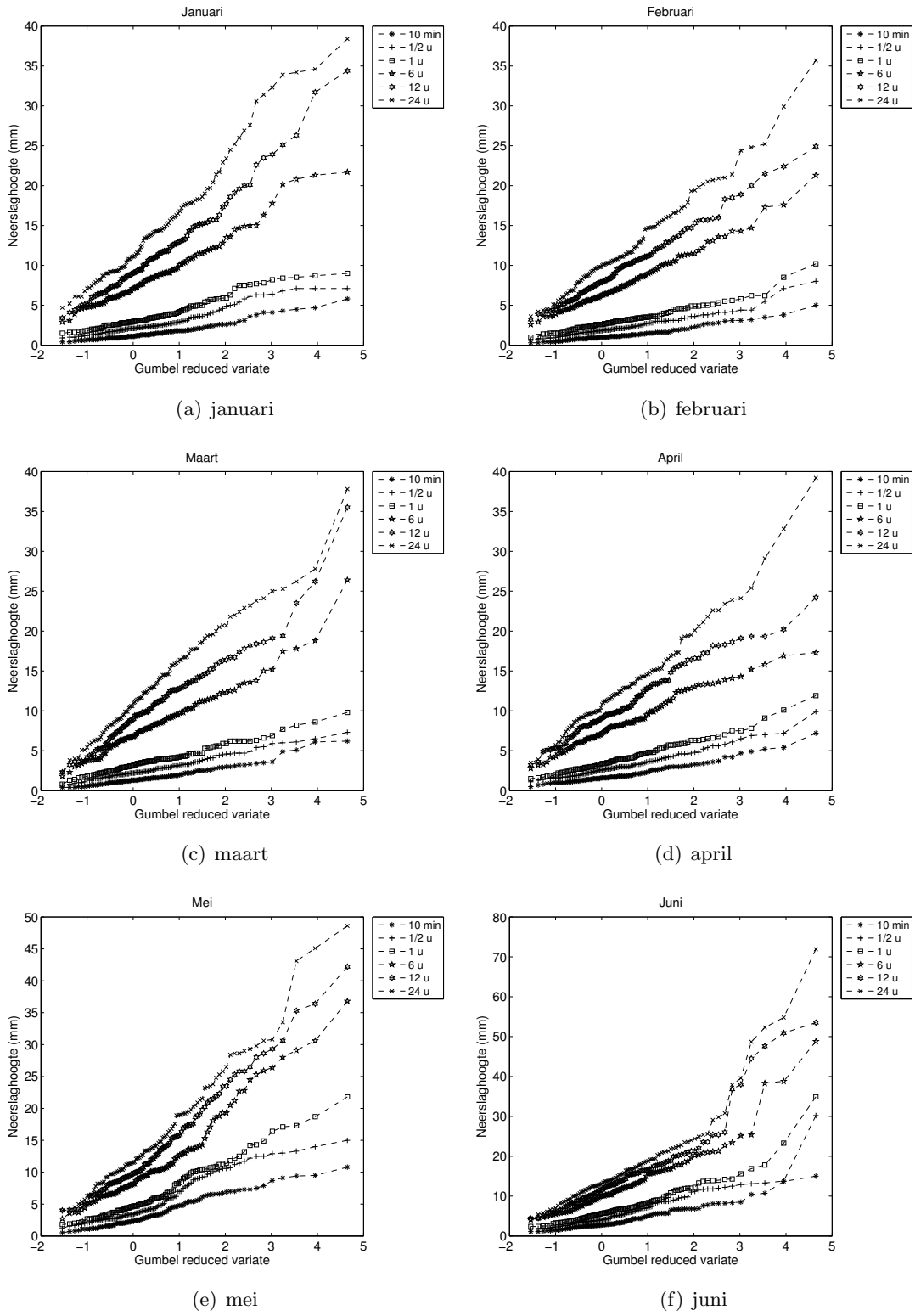
	<i>fitness</i>	$\lambda$	$\kappa$	$\phi$	$\mu_x$	$\alpha$	$\nu$
Jan	8.67E-02	0.03252	0.35847	0.02178	1.30513	3.00350	0.24630
Feb	5.20E-02	0.02496	0.20874	0.01526	1.29049	4.91980	0.67601
Mrt	9.57E-02	0.02385	0.20349	0.01511	1.58150	5.78861	0.70465
Apr	9.26E-02	0.02270	0.14702	0.01230	1.96895	7.88053	0.94226
Mei	1.47E-01	0.02419	0.10674	0.01611	4.44697	3.32902	0.23420
Jun	1.59E-01	0.02174	0.07620	0.01393	4.97136	12.12078	1.44286
Jul	1.56E-01	0.02201	0.04765	0.01273	7.87484	9.76754	1.12744
Aug	1.14E-01	0.02063	0.05132	0.01050	6.62050	12.96446	1.41769
Sep	2.08E-01	0.01855	0.05340	0.00755	4.53487	7.22399	0.80208
Okt	1.26E-01	0.02075	0.11356	0.00973	2.51881	4.04473	0.43429
Nov	6.86E-02	0.03146	0.29204	0.01802	1.61173	2.64396	0.19424
Dec	4.79E-02	0.03145	0.26860	0.01902	1.42328	2.72577	0.25952



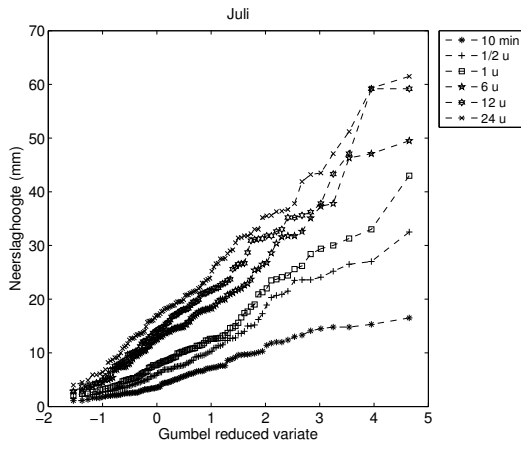


**Bijlage C**

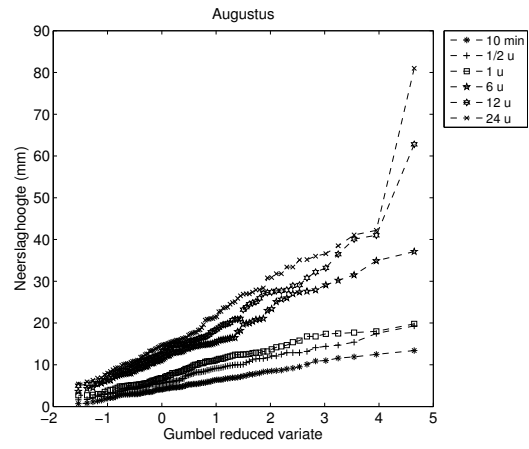
**Figuren**



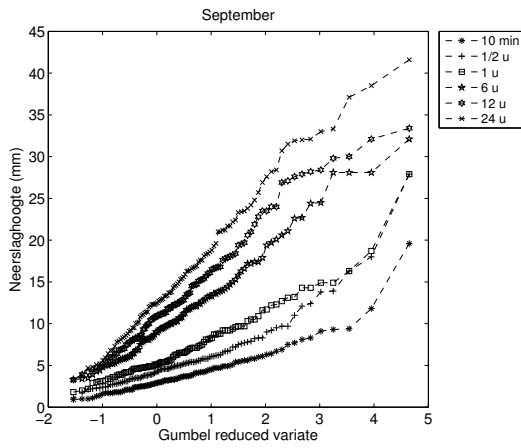
Figuur C.1: Extreme waarden diagramma voor maandelijkse maxima.



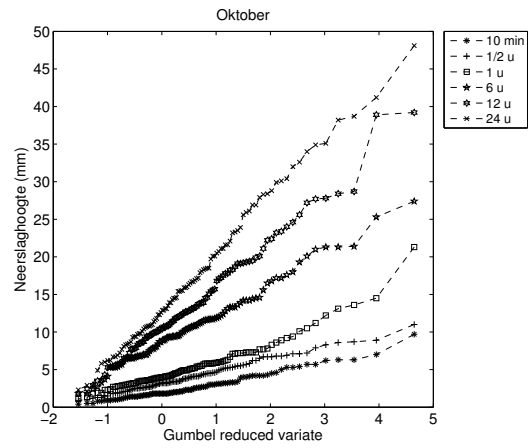
(g) juli



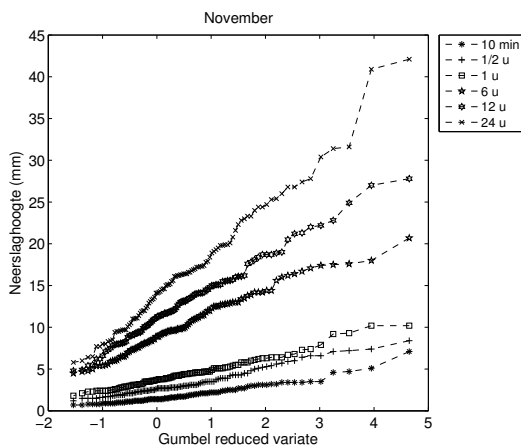
(h) augustus



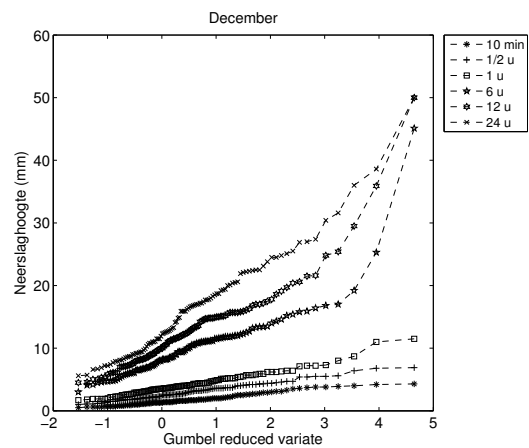
(i) september



(j) oktober

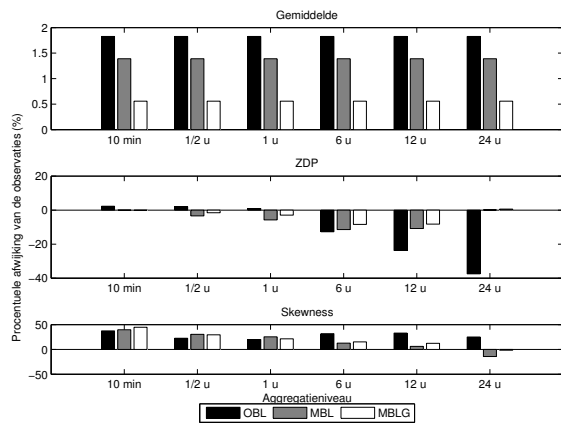


(k) november

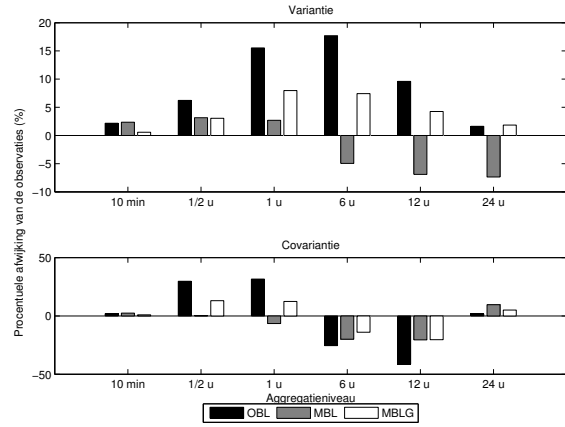


(l) december

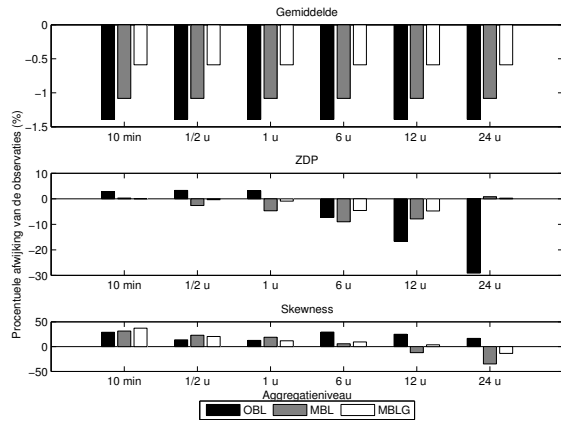
**Figuur C.1:** Extreme waarden diagramma voor maandelijkse maxima (vervolg).



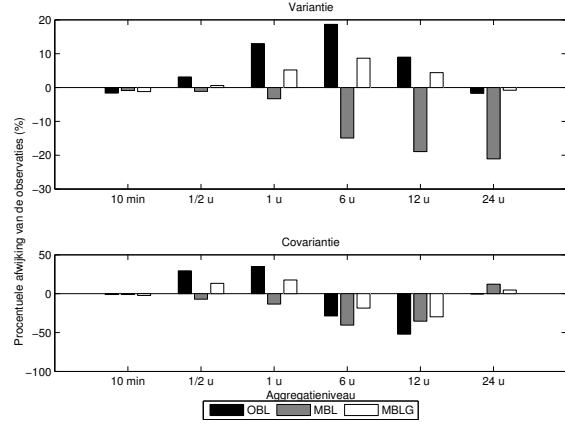
(a) januari



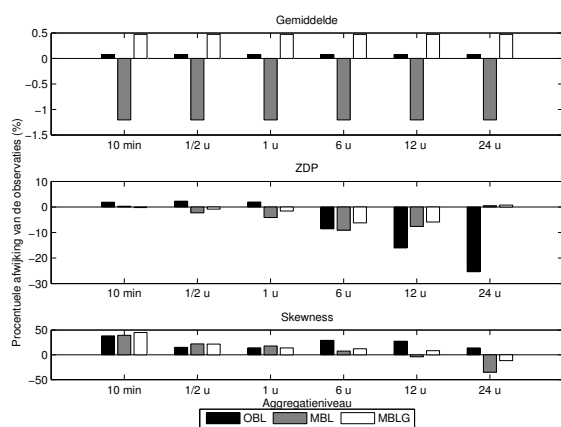
(b) januari



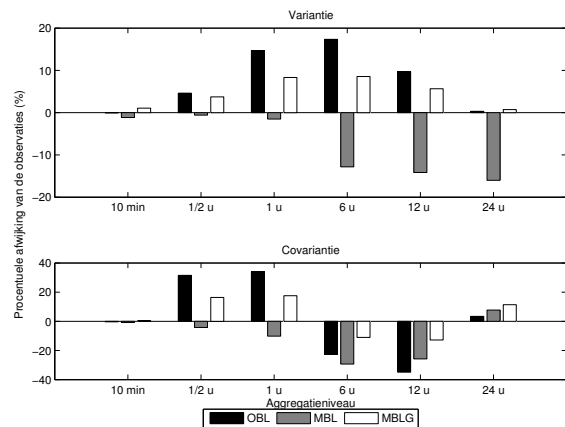
(c) februari



(d) februari

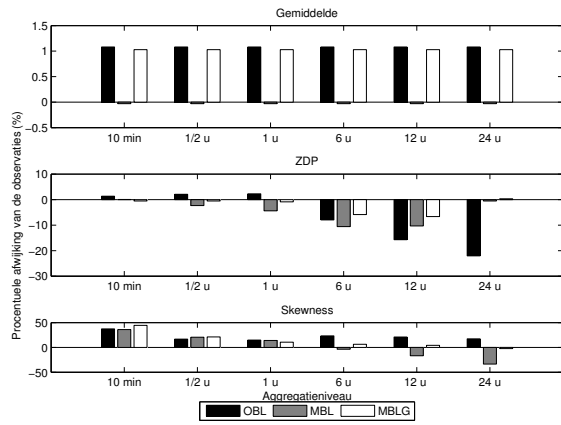


(e) maart

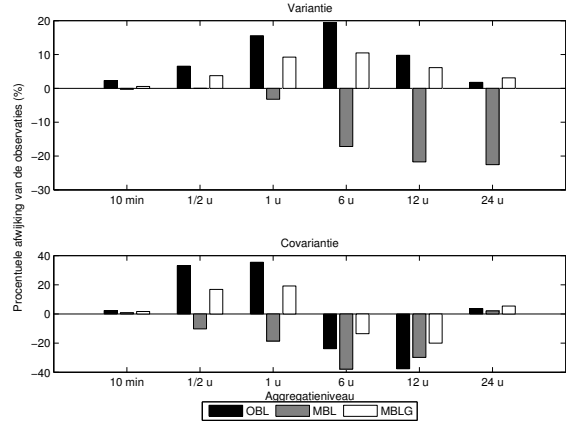


(f) maart

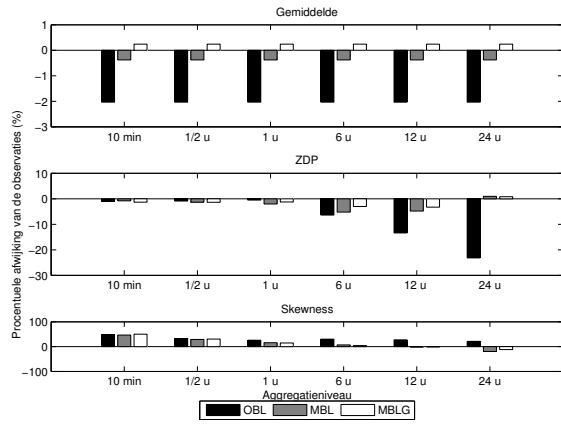
**Figuur C.2:** Procentuele afwijking modelvoorspellingen van de historische waarden (januari-maart).



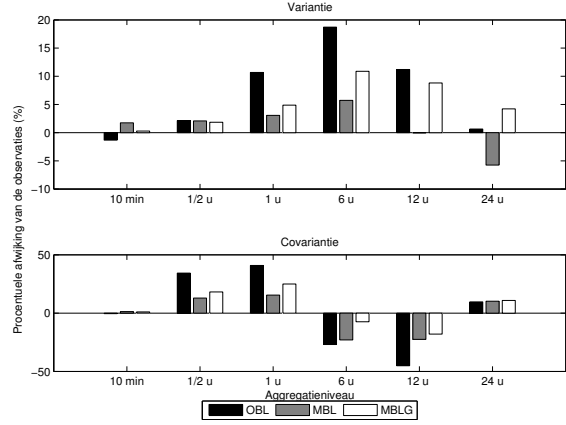
(g) april



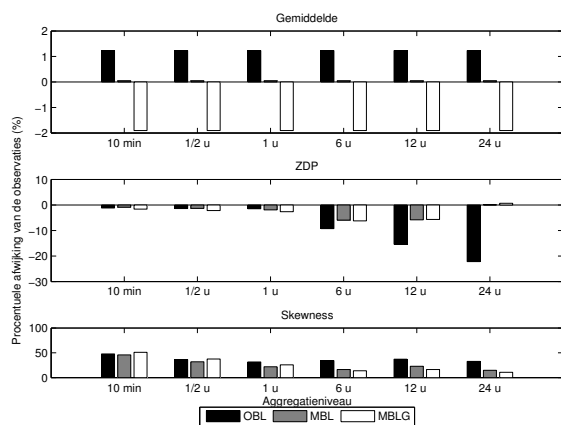
(h) april



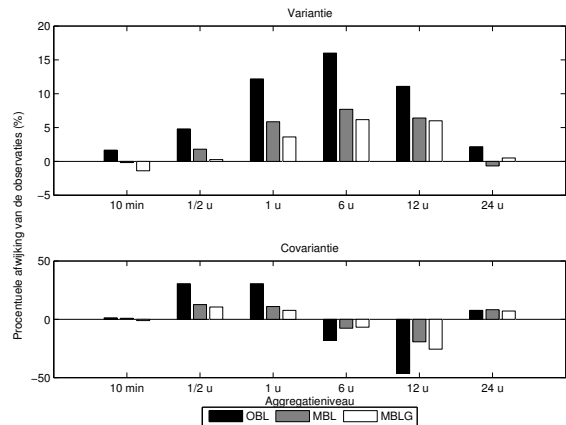
(i) mei



(j) mei

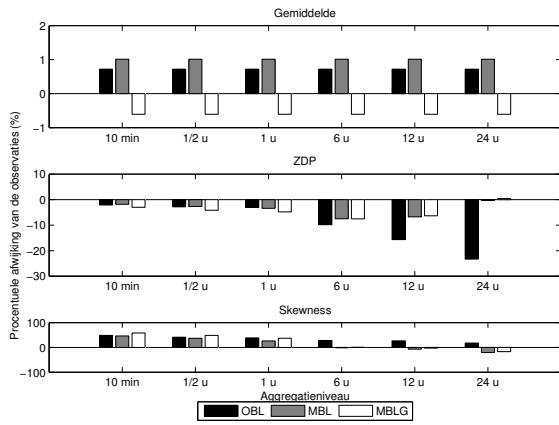


(k) juni

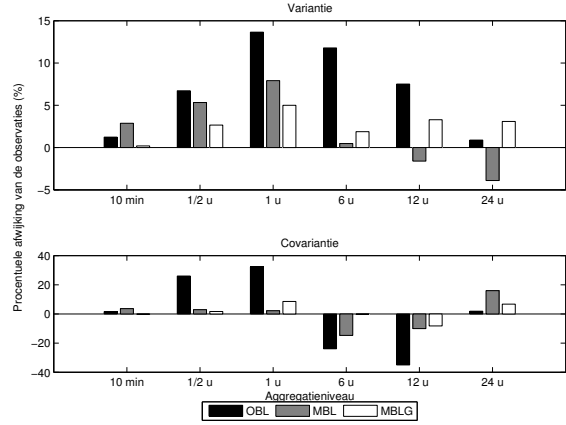


(l) juni

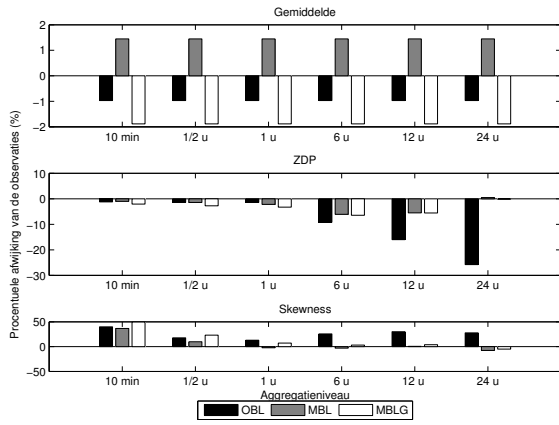
Figuur C.2: Procentuele afwijking modelvoorspellingen van de historische waarden (april-juni).



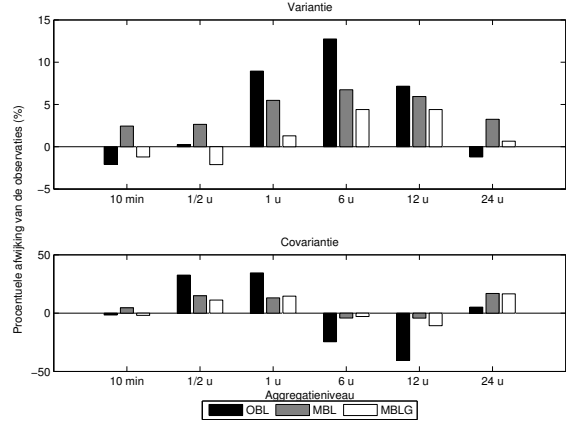
(m) juli



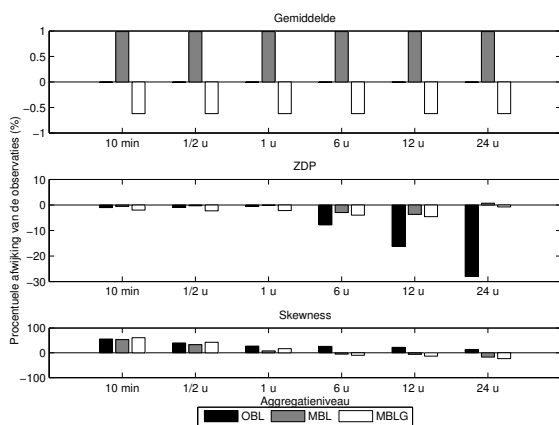
(n) juli



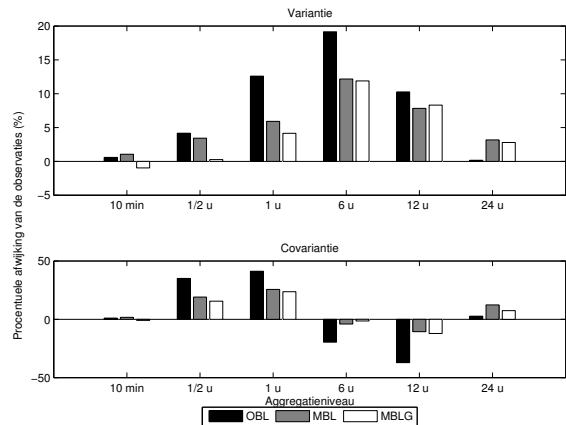
(o) augustus



(p) augustus

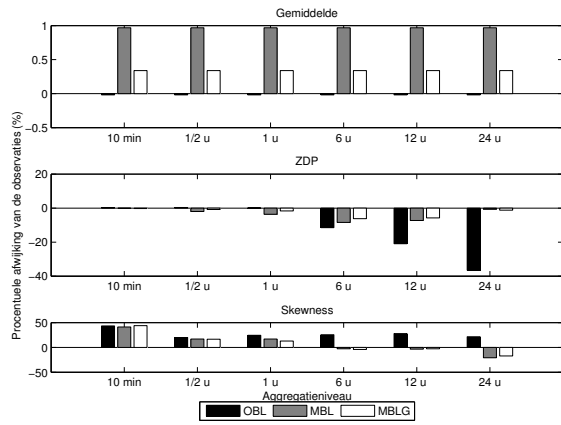


(q) september

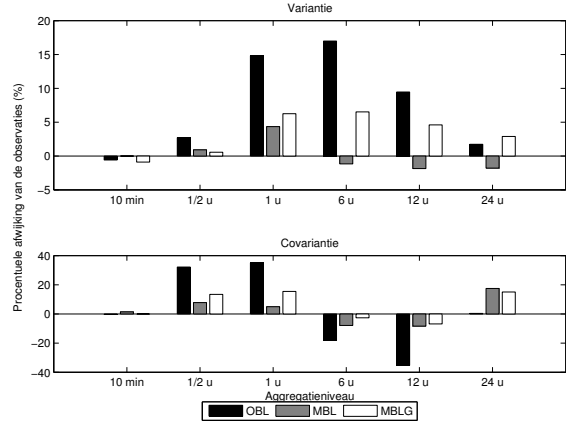


(r) september

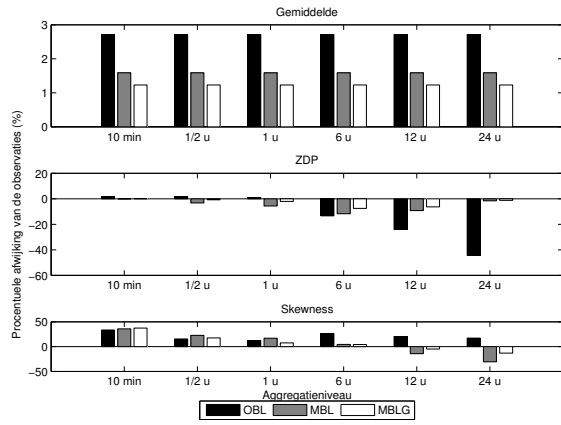
Figuur C.2: Procentuele afwijking modelvoorspellingen van de historische waarden (juli-september).



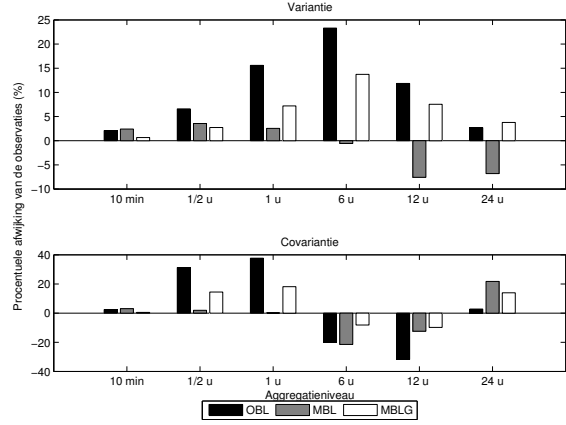
(s) oktober



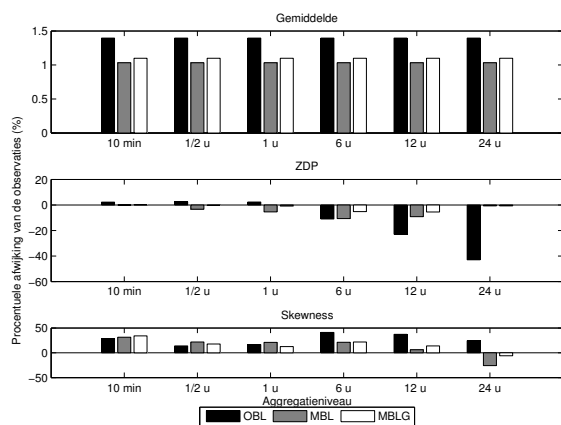
(t) oktober



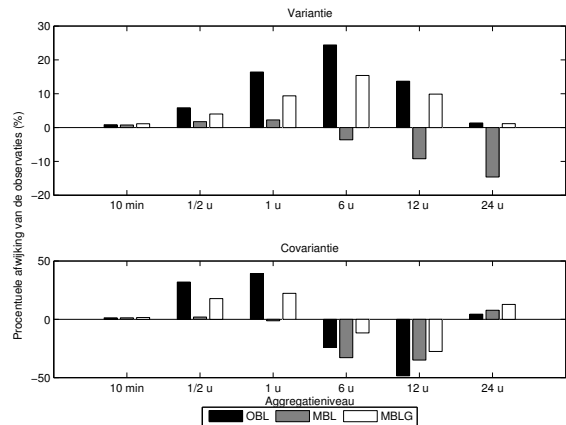
(u) november



(v) november

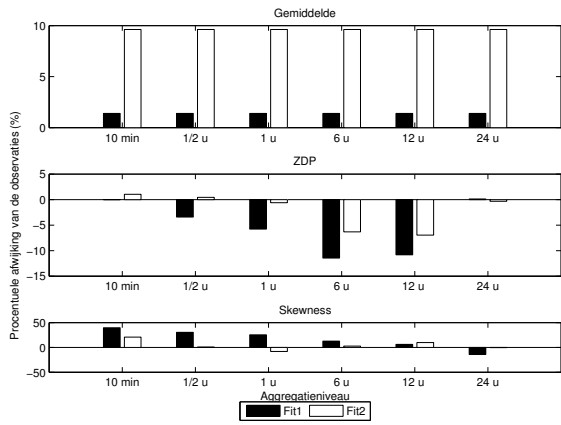


(w) december

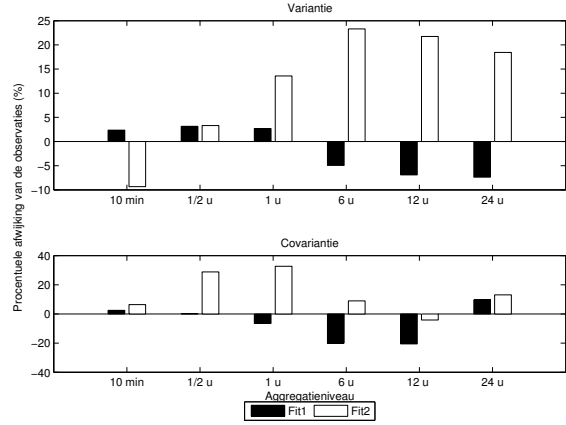


(x) december

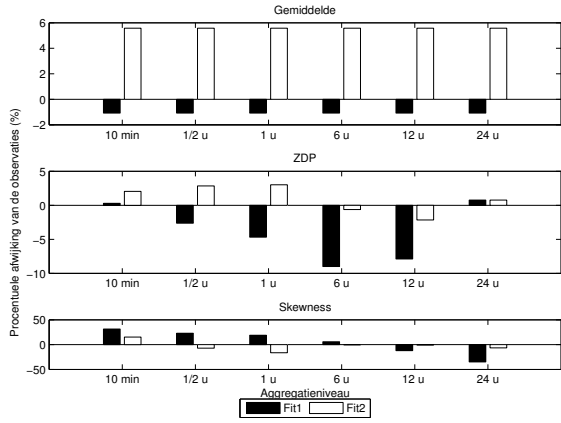
**Figuur C.2:** Procentuele afwijking modelvoorspellingen van de historische waarden (oktober-december).



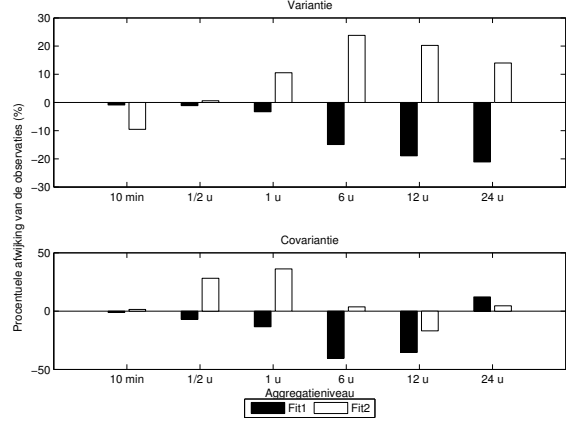
(a) januari



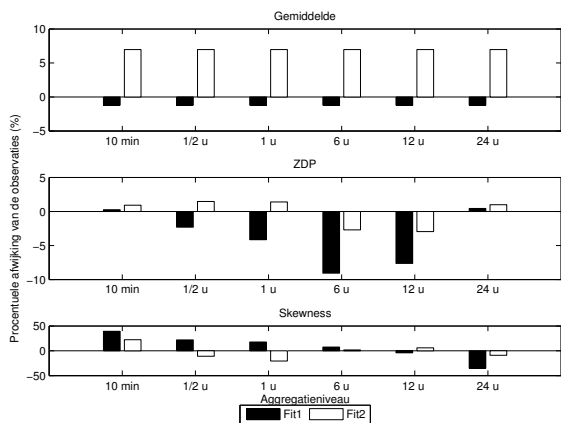
(b) januari



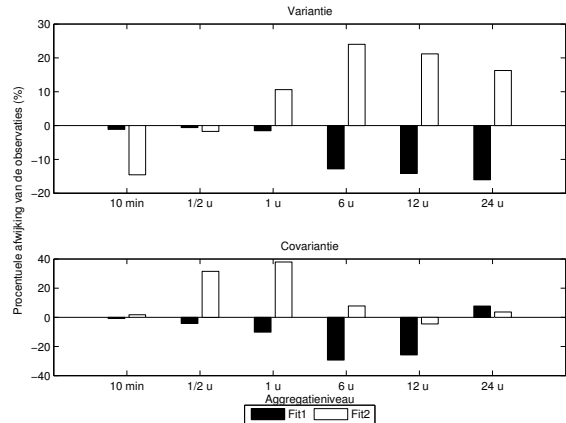
(c) februari



(d) februari



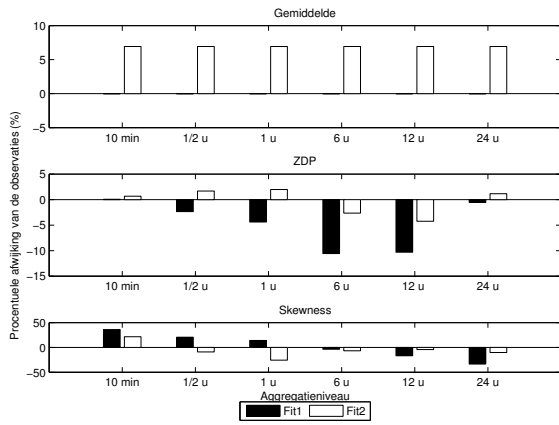
(e) maart



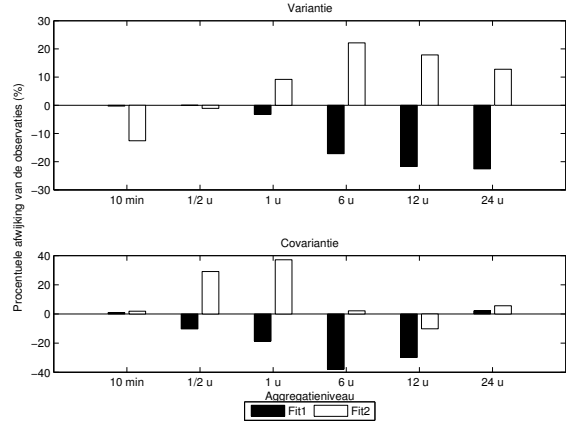
(f) maart

Figuur C.3: Procentuele afwijking modelvoorspellingen van de historische waarden (januari-maart).

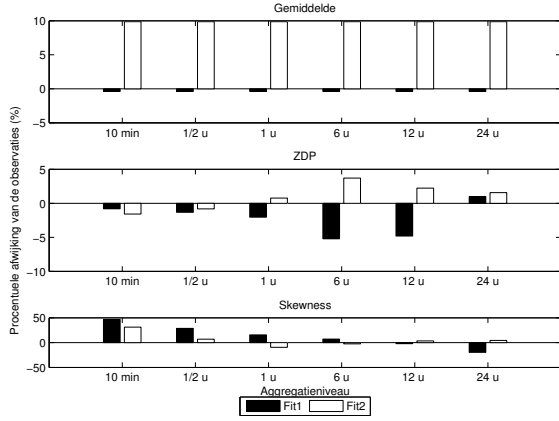




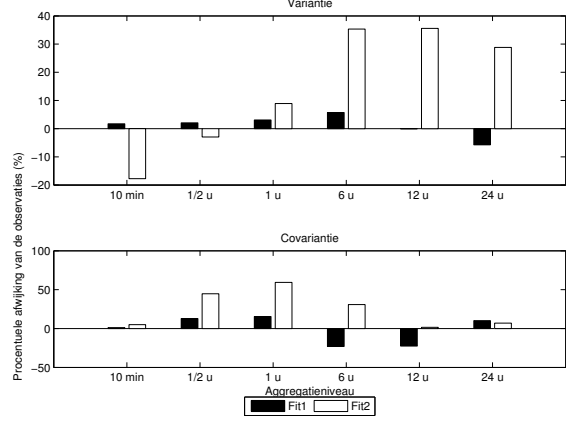
(g) april



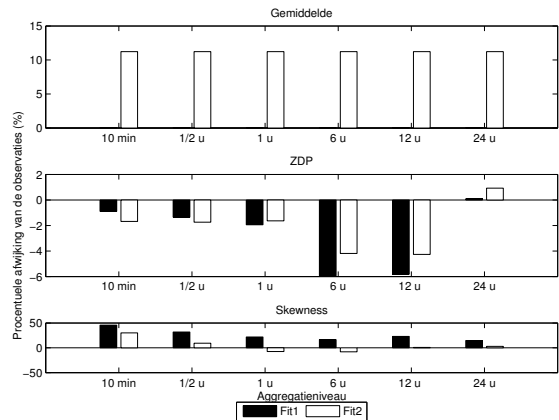
(h) april



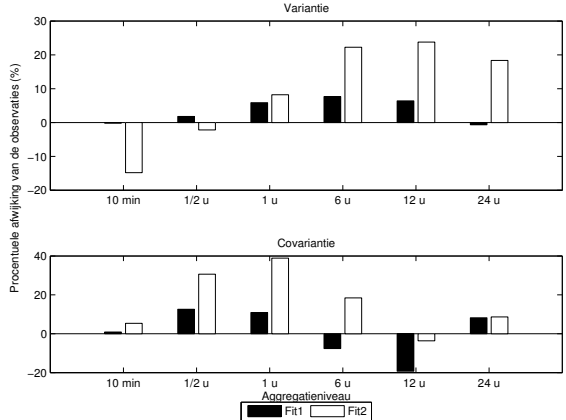
(i) mei



(j) mei

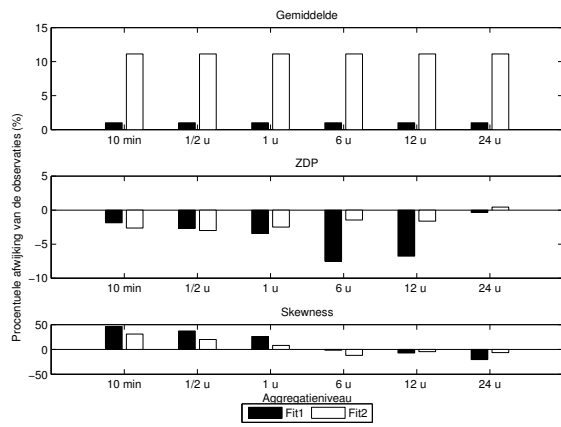


(k) juni

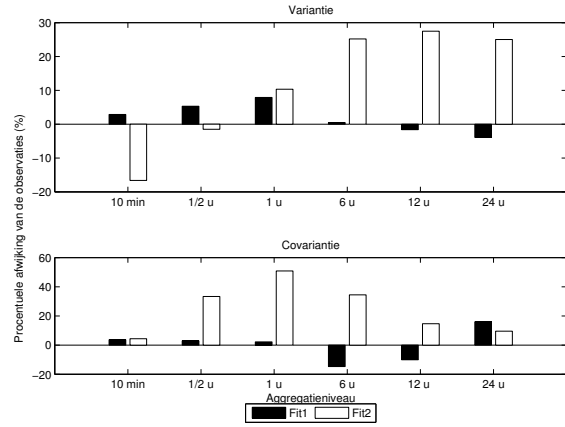


(l) juni

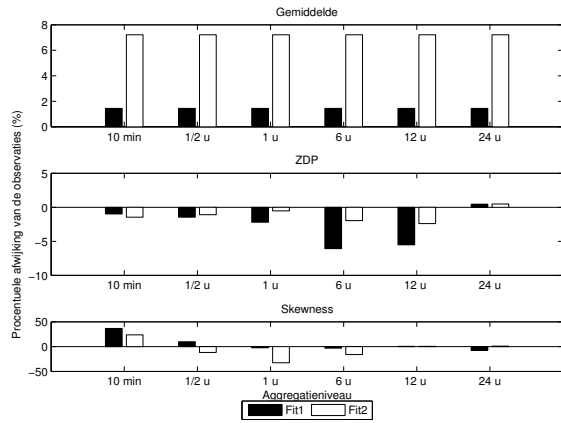
Figuur C.3: Procentuele afwijking modelvoorspellingen van de historische waarden (april-juni).



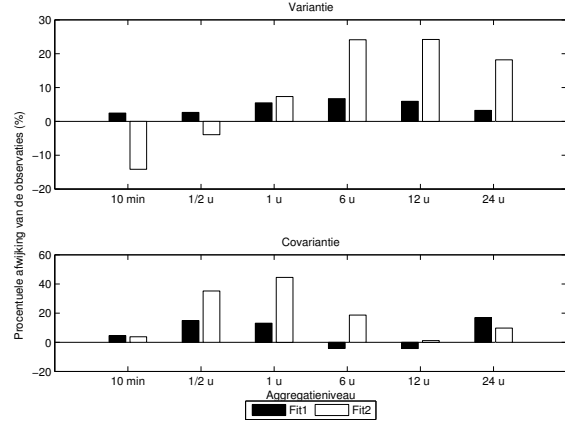
(m) juli



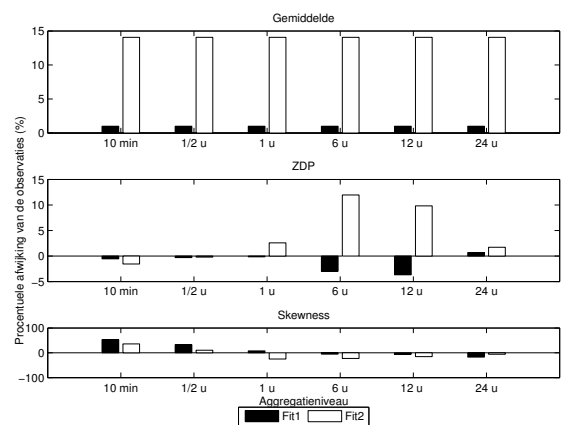
(n) juli



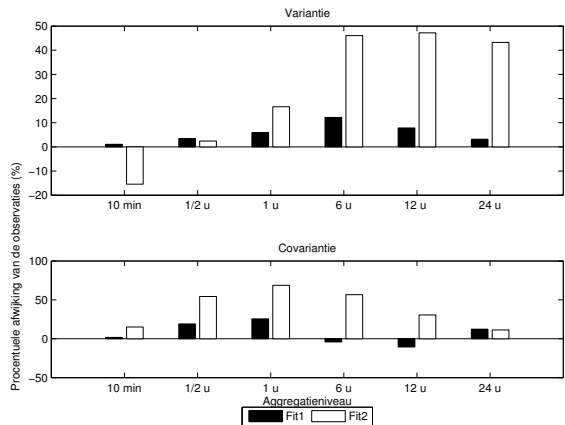
(o) augustus



(p) augustus

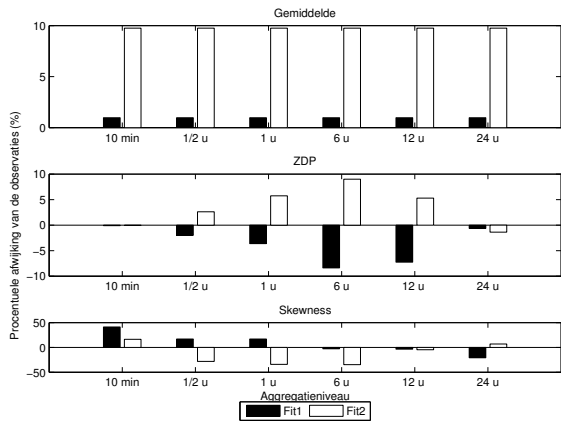


(q) september

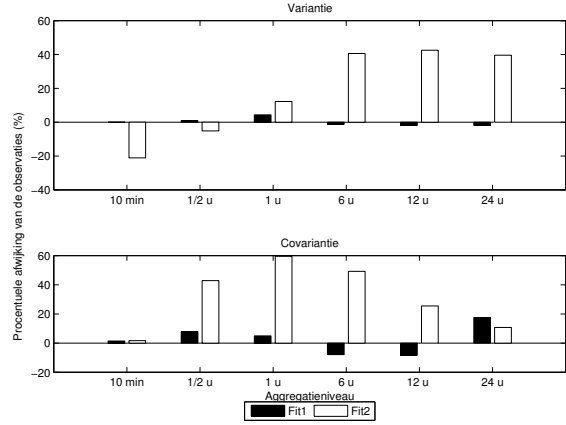


(r) september

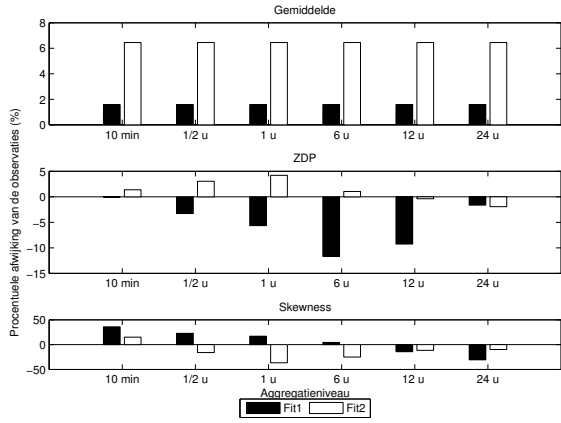
Figuur C.3: Procentuele afwijking modelvoorspellingen van de historische waarden (juli-september).



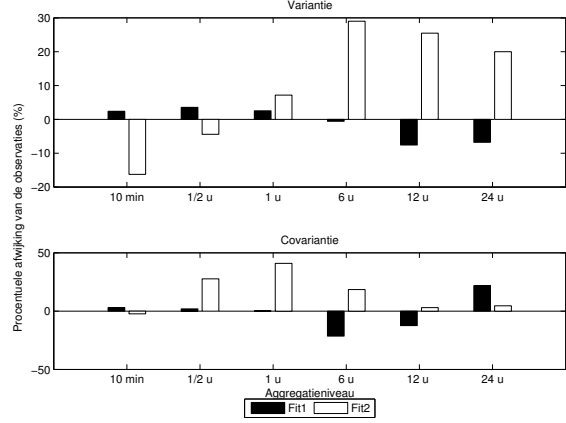
(s) oktober



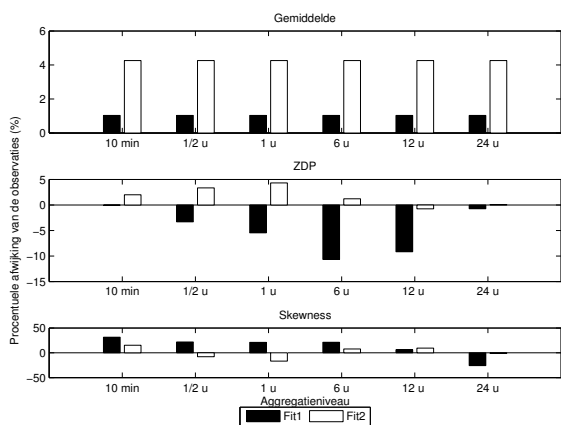
(t) oktober



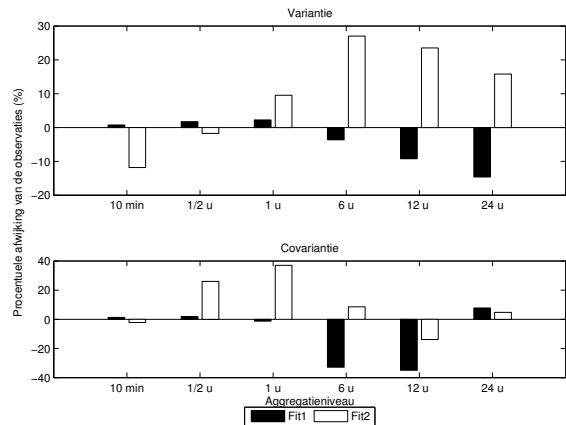
(u) november



(v) november

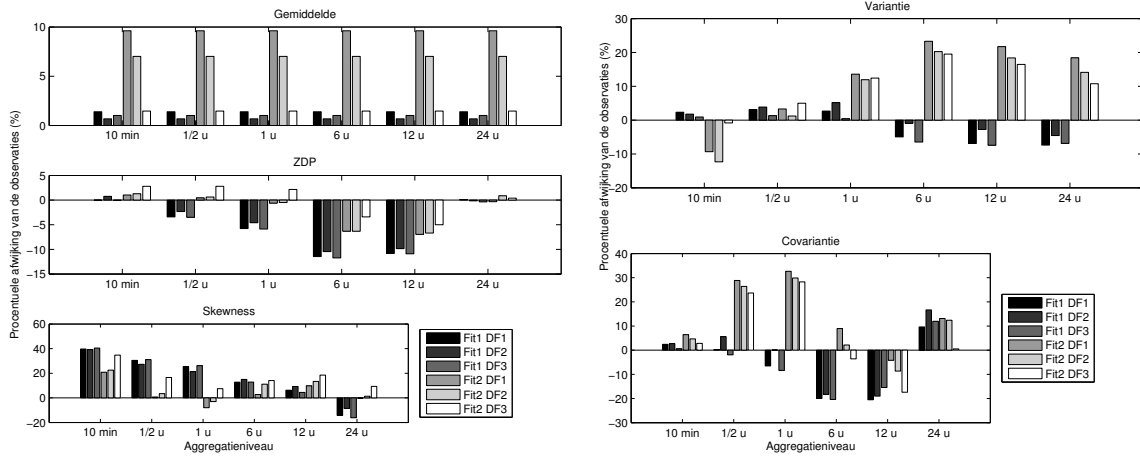


(w) december



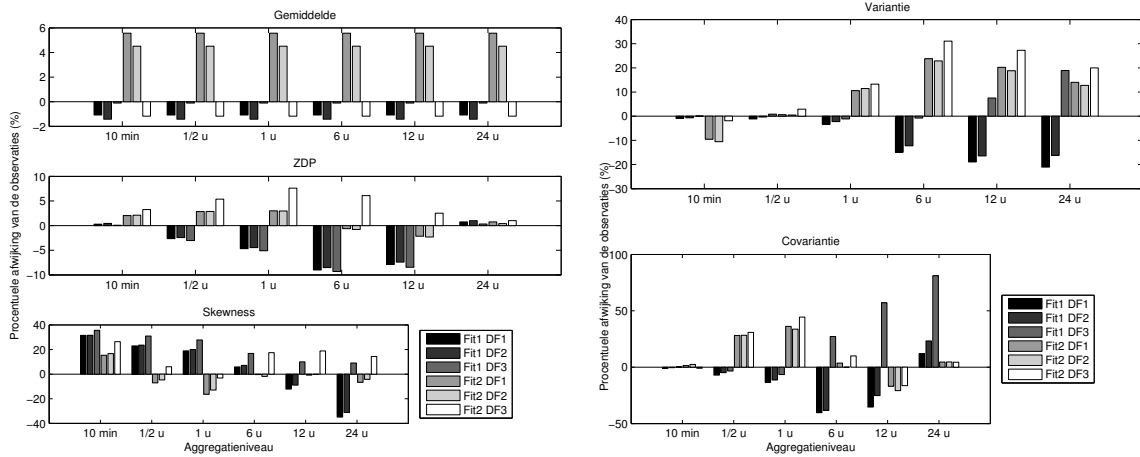
(x) december

**Figuur C.3:** Procentuele afwijking modelvoorspellingen van de historische waarden (oktober-december).



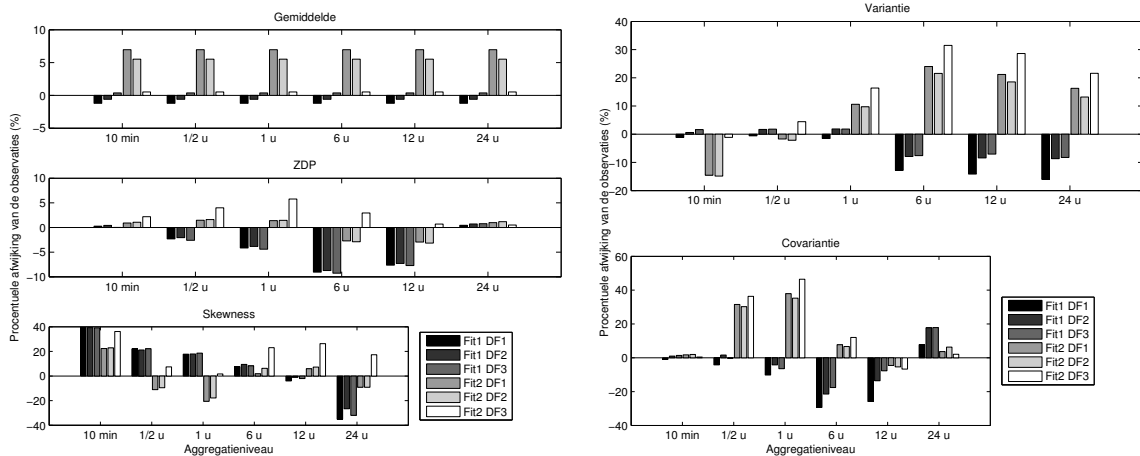
(a) januari

(b) januari



(c) februari

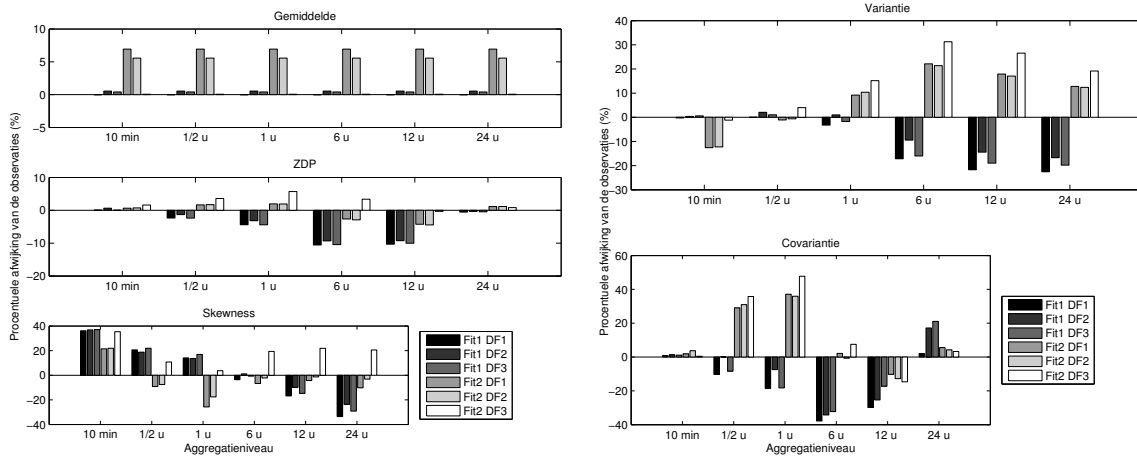
(d) februari



(e) maart

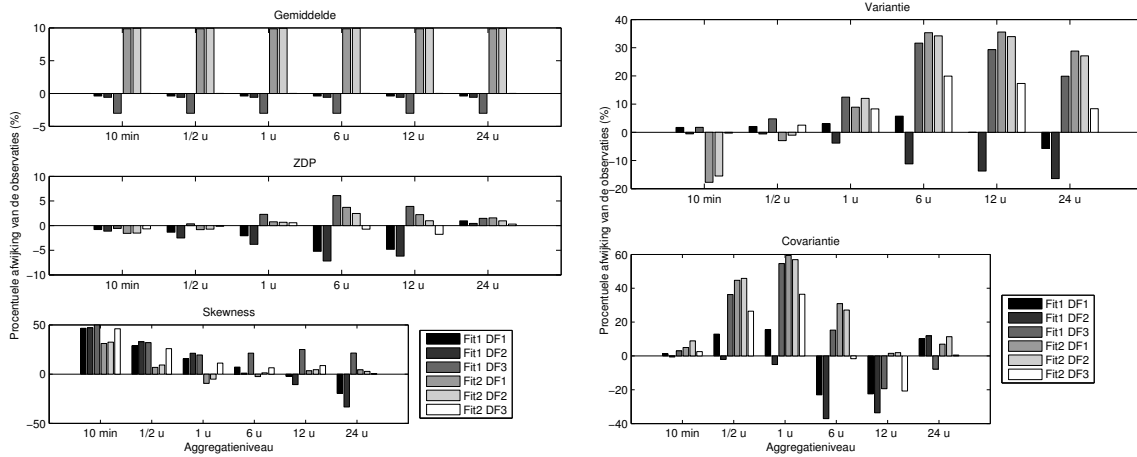
(f) maart

Figuur C.4: Procentuele afwijking modelvoorspellingen van de historische waarden (januari-maart).



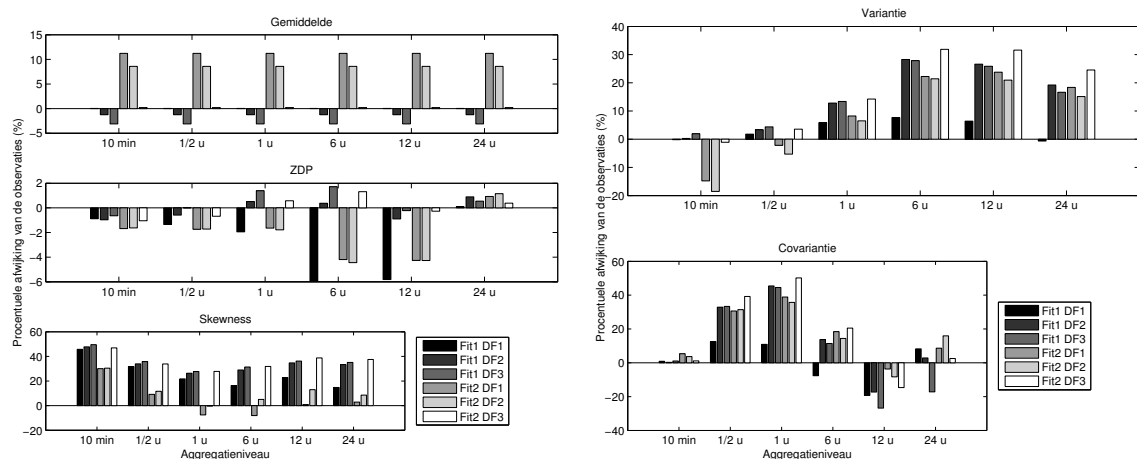
(g) april

(h) april



(i) mei

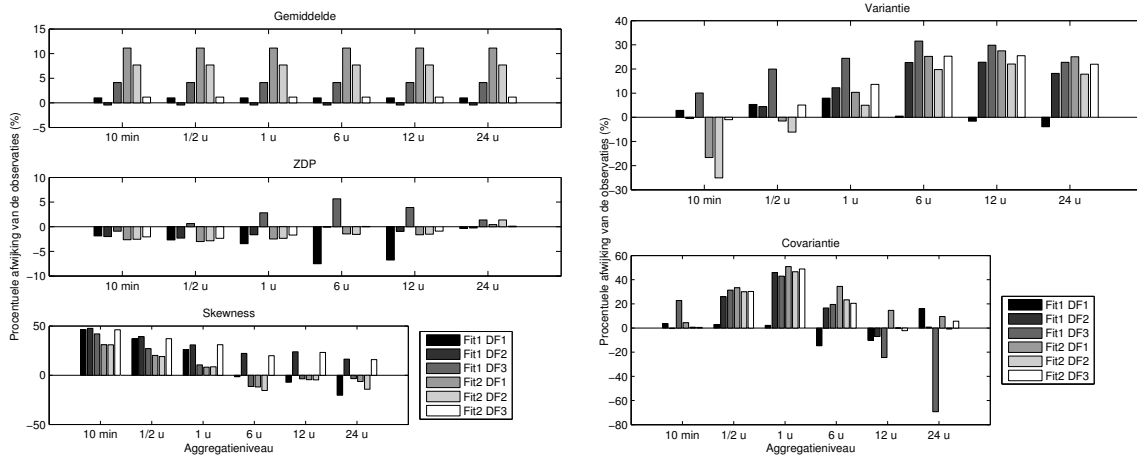
(j) mei



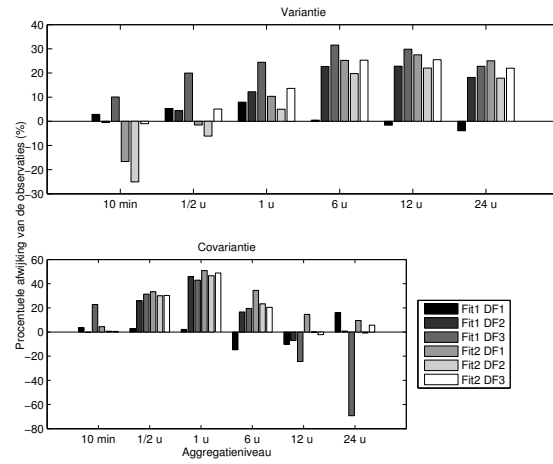
(k) juni

(l) juni

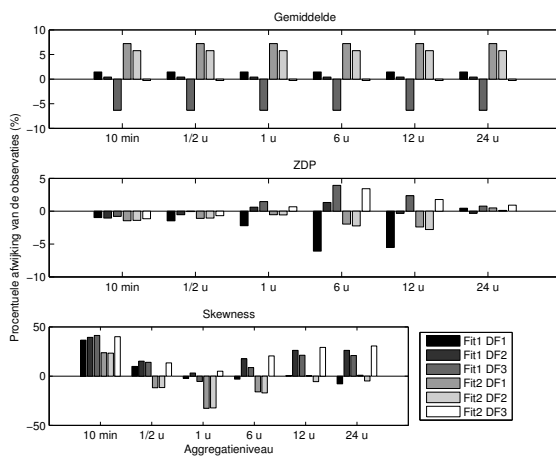
Figuur C.4: Procentuele afwijking modelvoorspellingen van de historische waarden (april-juni).



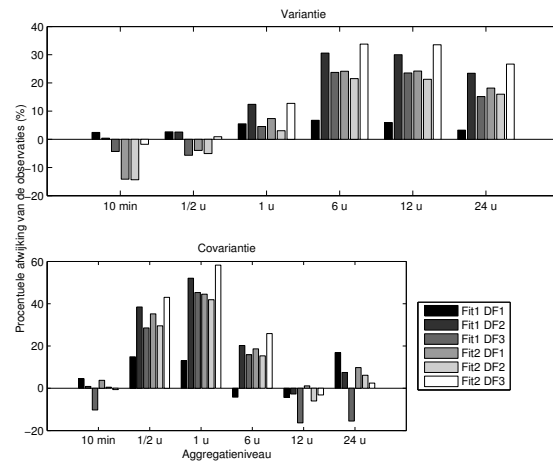
(m) juli



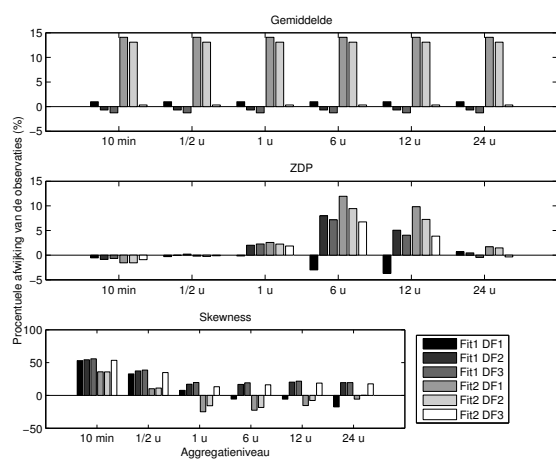
(n) juli



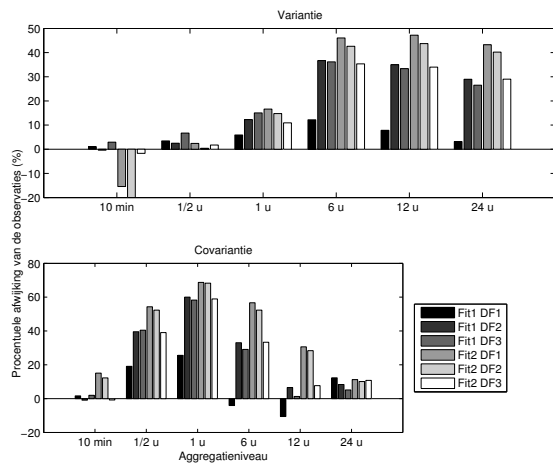
(o) augustus



(p) augustus

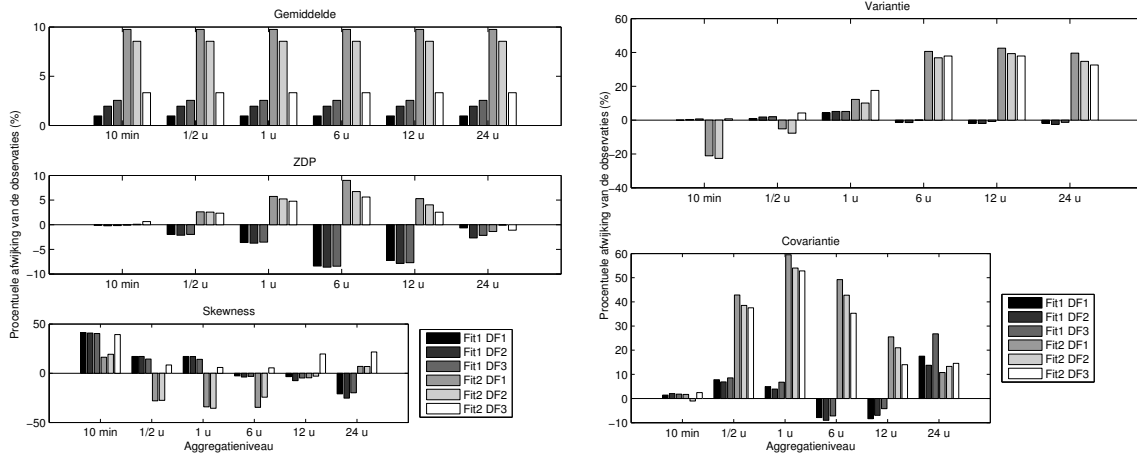


(q) september

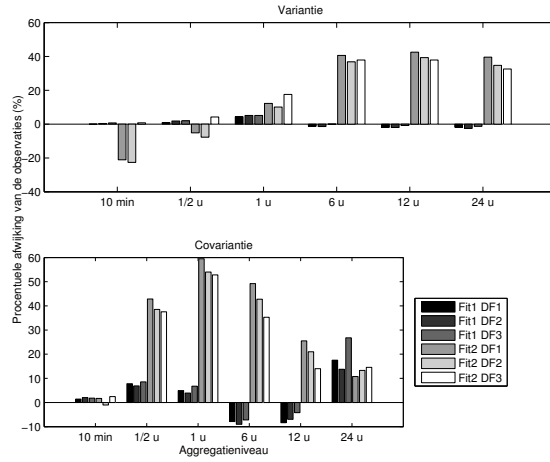


(r) september

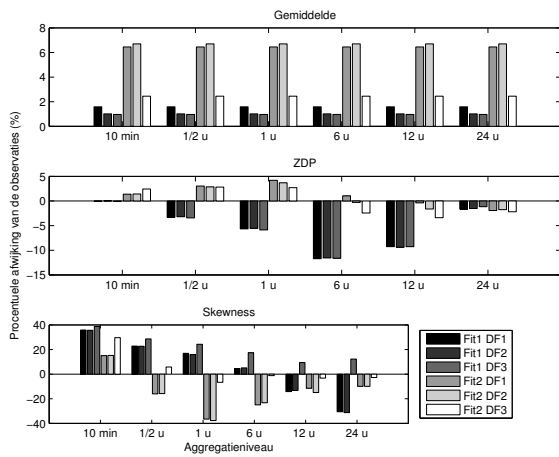
Figuur C.4: Procentuele afwijking modelvoorspellingen van de historische waarden (juli-september).



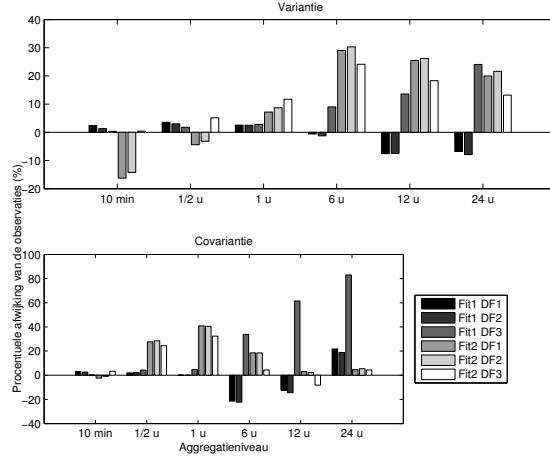
(s) oktober



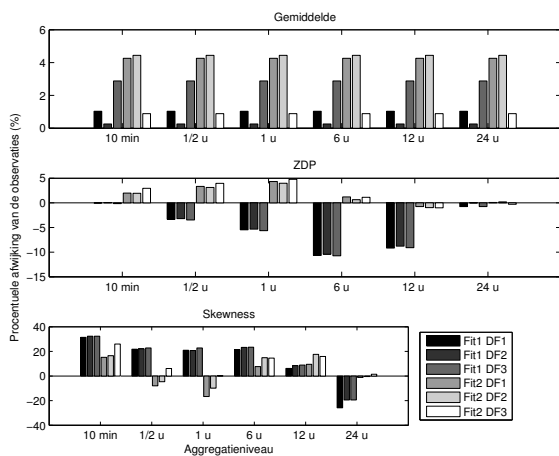
(t) oktober



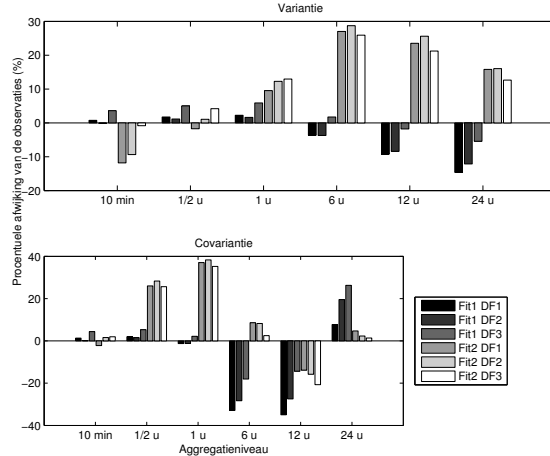
(u) november



(v) november

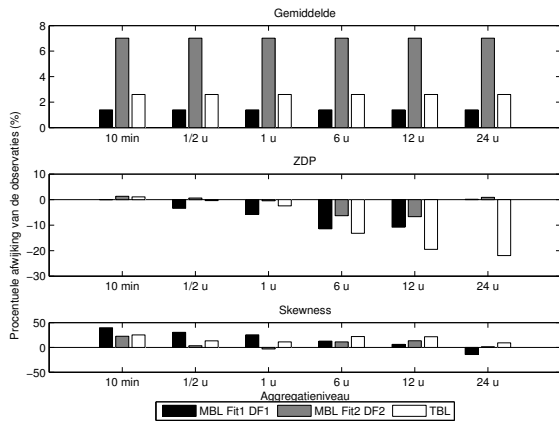


(w) december

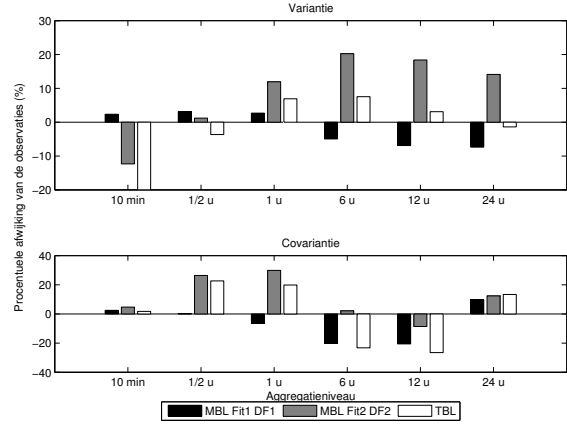


(x) december

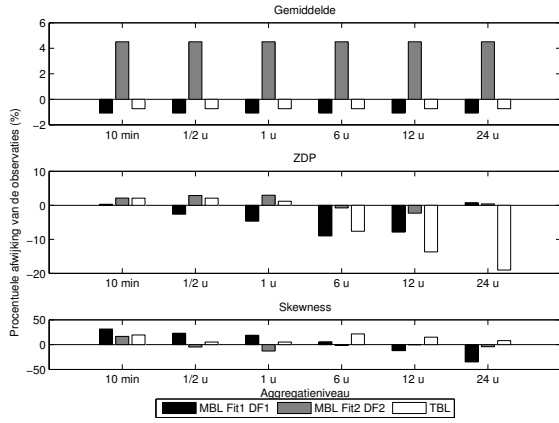
Figuur C.4: Procentuele afwijking modelvoorspellingen van de historische waarden (oktober-december).



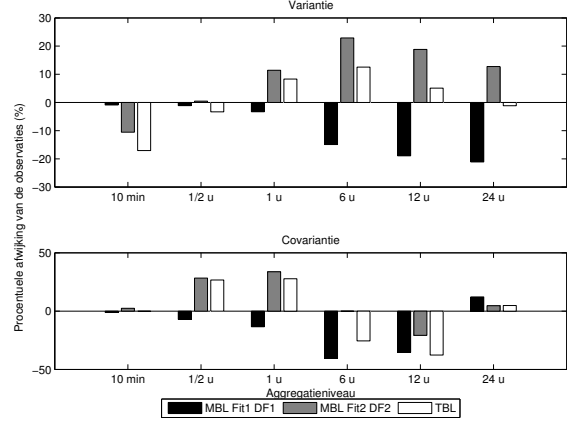
(a) januari



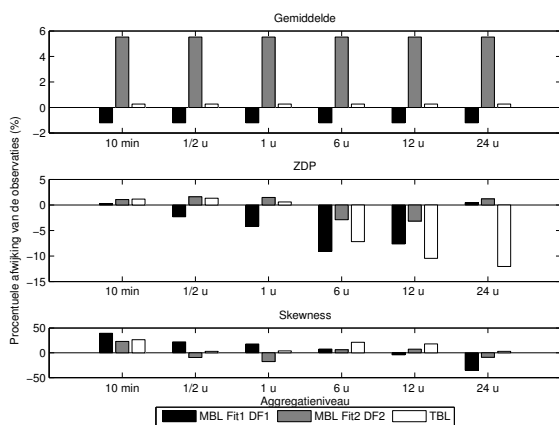
(b) januari



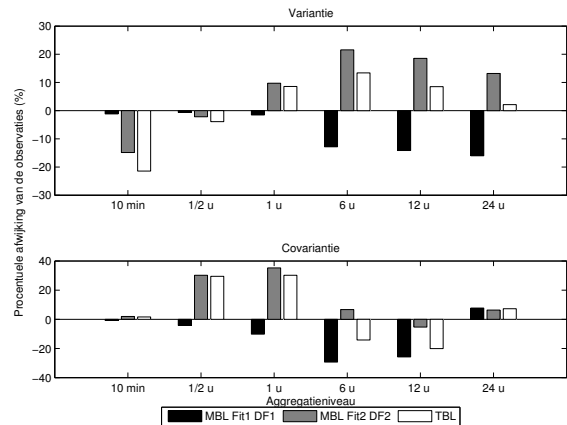
(c) februari



(d) februari



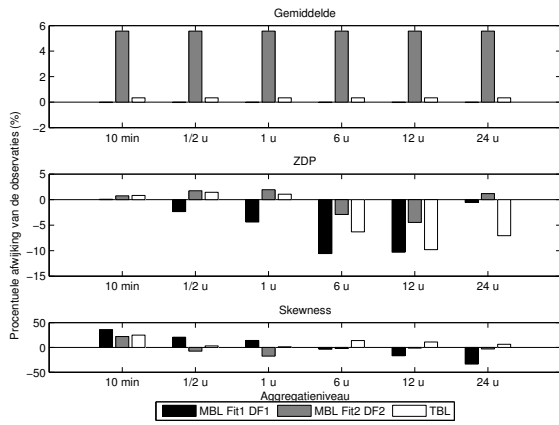
(e) maart



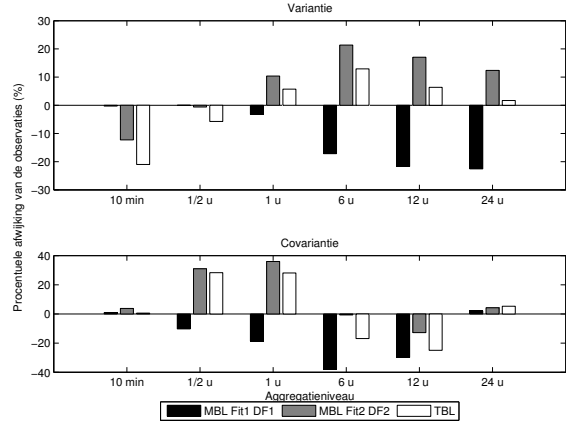
(f) maart

Figuur C.5: Procentuele afwijking modelvoorspellingen van de historische waarden (januari-maart).

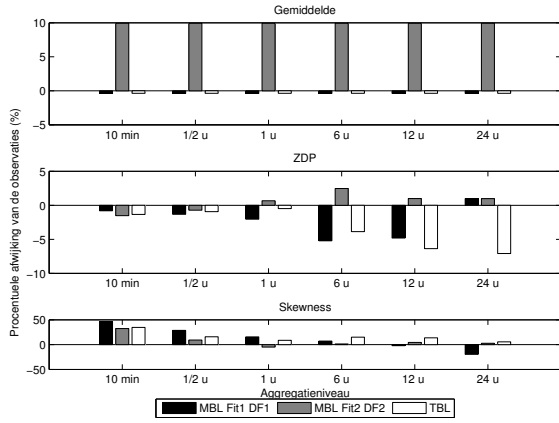




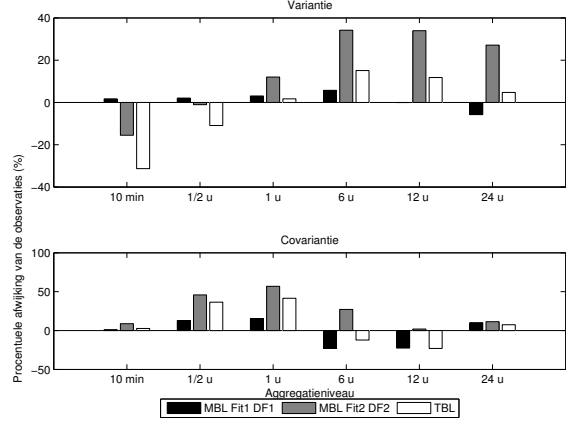
(g) april



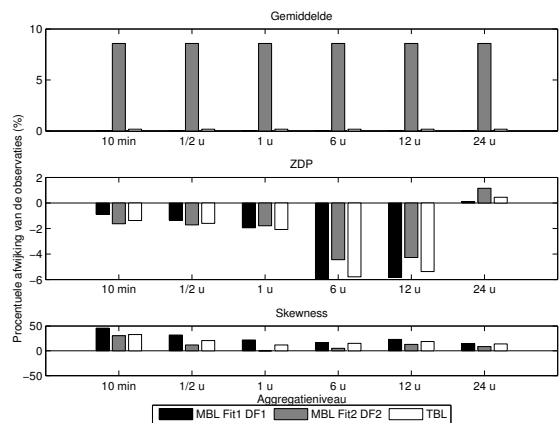
(h) april



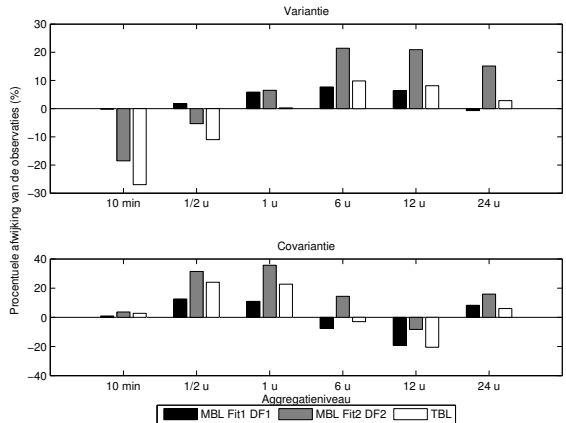
(i) mei



(j) mei

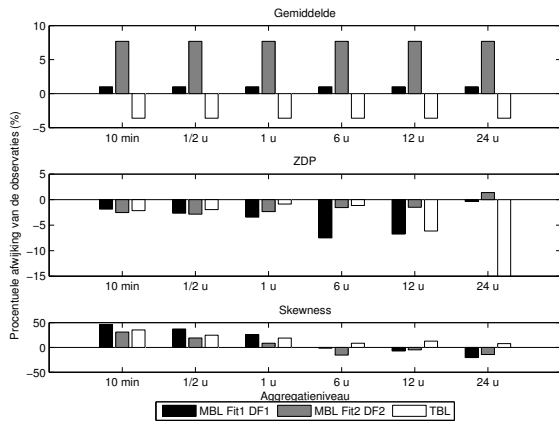


(k) juni

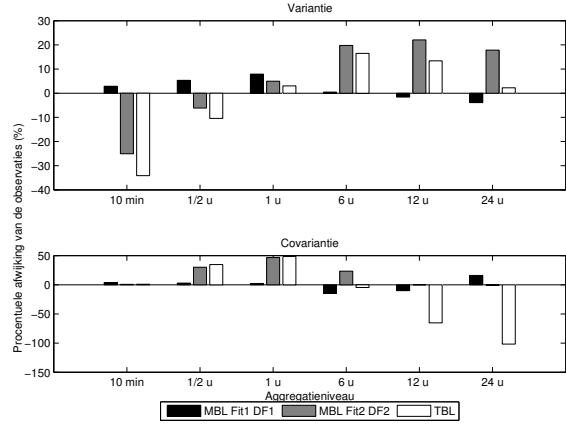


(l) juni

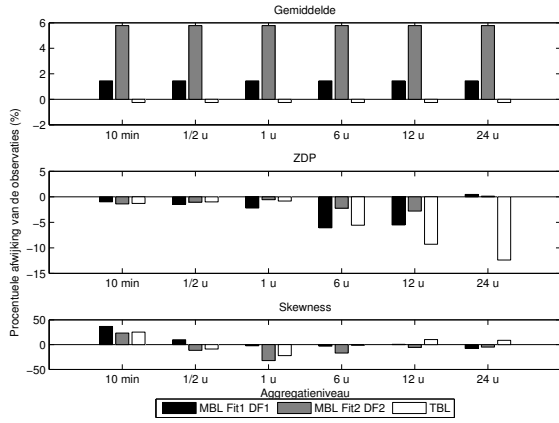
Figuur C.5: Procentuele afwijking modelvoorspellingen van de historische waarden (april-juni).



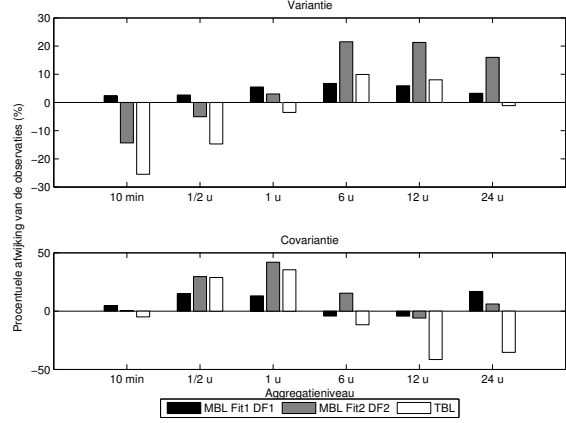
(m) juli



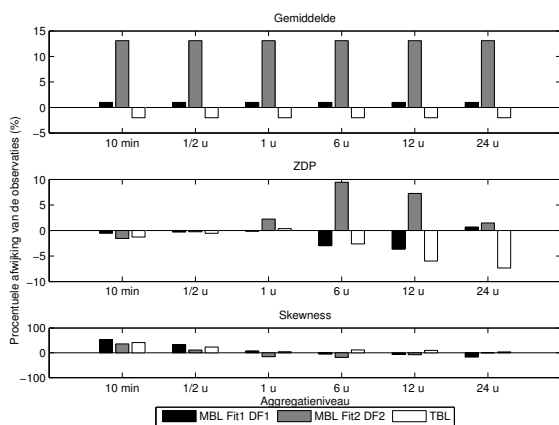
(n) juli



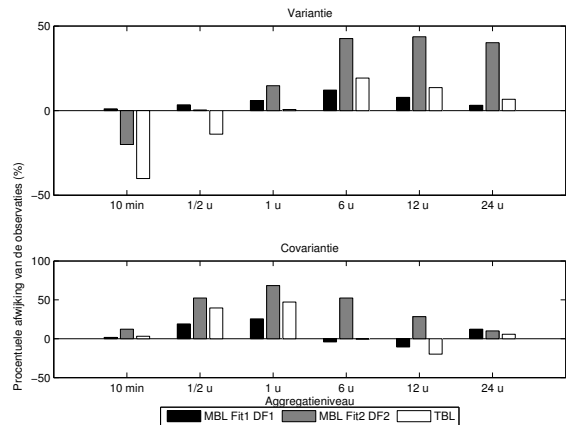
(o) augustus



(p) augustus

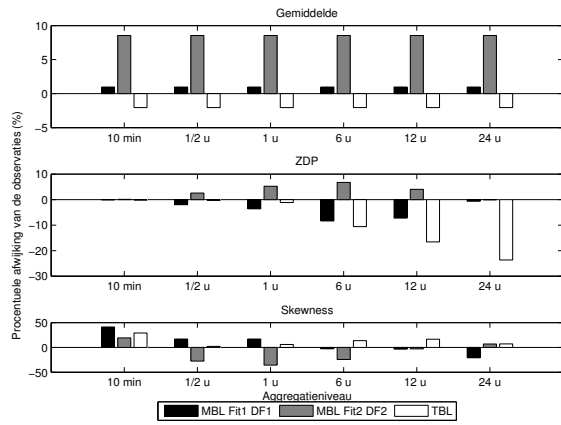


(q) september

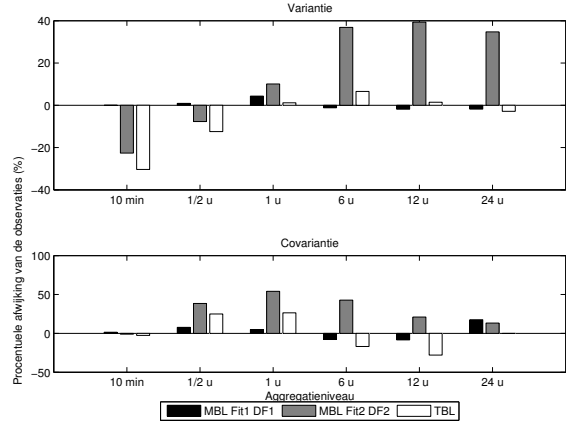


(r) september

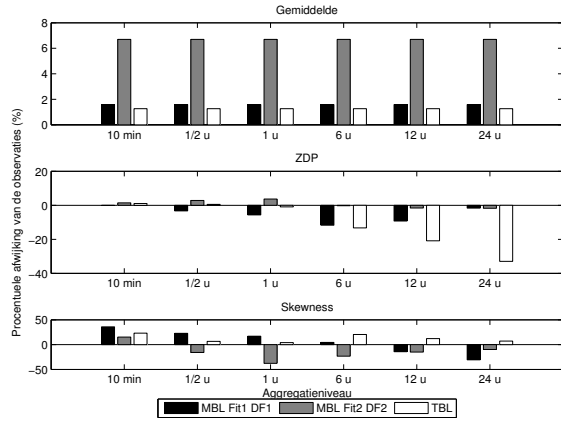
**Figuur C.5:** Procentuele afwijking modelvoorspellingen van de historische waarden (juli-september).



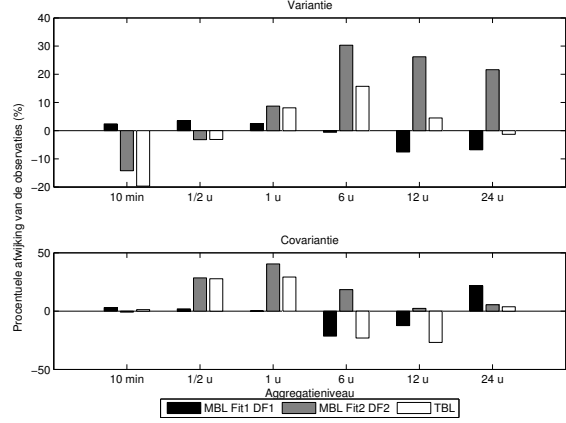
(s) oktober



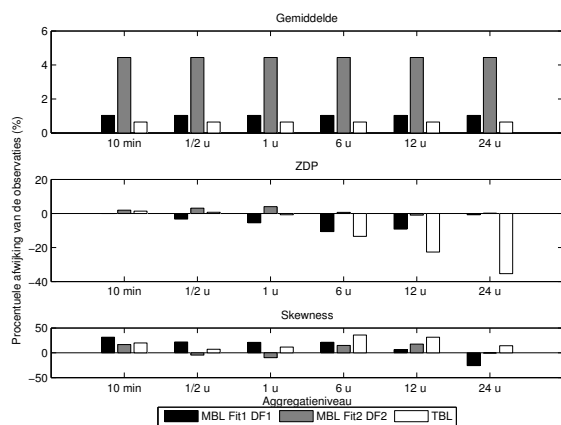
(t) oktober



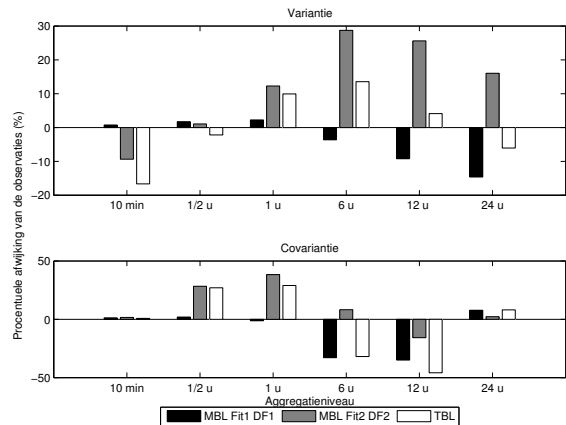
(u) november



(v) november



(w) december



(x) december

**Figuur C.5:** Procentuele afwijking modelvoorspellingen van de historische waarden (oktober-december).