



Adviesbureau Modelling en Optimalisatie

CORRELATIE-ONDERZOEK ZOUTE WATEREN

EVALUATIE CHEMISCHE MEETNET

Onderzoek verricht in opdracht van
Rijksinstituut voor Kust en Zee / RIKZ

Ir. E.C.J. van der Meulen

11 mei 1995

Samenvatting

In het kader van de evaluatie van het Chemische Meetnet van Rijksinstituut voor Kust en Zee (RIKZ) is een correlatie-onderzoek verricht. Het onderzoek betreft de gebieden de Westerschelde, de Noordzee bij Noordwijk, de Waddenzee West en de Eems-Dollard. Van een aantal waterkwaliteitsvariabelen is de correlatie onderzocht in de ruimte en in de tijd voor alle locaties in het meetnet. Door analyse van reeksen van een locatie met een andere dichtbij gelegen locatie van eenzelfde waterkwaliteitsvariabele is de afhankelijkheid in de ruimte bepaald. Door de reeks van een waterkwaliteitsvariabele op één bepaalde locatie te analyseren is de autocorrelatie, de afhankelijkheid in de tijd, bepaald.

Correlatie in de ruimte

Voor de verschillende gebieden is per stof en per meetlocatie bekeken in hoeverre andere locaties in het gebied de concentratie van een stof kunnen schatten. Een maat voor de correlatie in de ruimte is de verklaarde variantie. Een verklaarde variantie van 100% betekent een perfecte schatting van de concentratie. De verklaarde variantie per locatie is het gemiddelde van de verklaarde varianties over alle stoffen. Locaties met een hoge gemiddelde verklaarde variantie zijn locaties welke goed te schatten zijn uit andere locaties.

De locaties in het gebied de Westerschelde, de locaties dicht bij de kust bij Noordwijk en locaties in de Eems-Dollard hebben een gemiddelde verklaarde variantie groter dan 60%. Bij de Waddenzee West komen enkele locaties tot een verklaarde variantie van 43%.

Correlatie in de tijd

Voor de locatie in een gebied met de langste tijdreeksen voor de verschillende stoffen is bekeken hoe opeenvolgende metingen met elkaar zijn gecorreleerd. De correlatie in de tijd is bepaald door berekening van de autocorrelatiecoëfficiënten en ook door tussenliggende metingen te schatten met een autoregressief model. Tevens is bekeken in hoeverre de autocorrelatie afhankelijk is van seizoensinvloeden.

In de meeste gebieden is correlatie aangetoond, de autocorrelatiecoëfficiënt ligt tussen 0.1 en 0.7. Vooral de stof nitraat met nitriet is in alle gebieden goed te schatten met een eenvoudig autoregressief model. Door seizoensinvloeden te filteren worden de autocorrelaties in de regel kleiner, opeenvolgende metingen zijn nu meestal significant ongecorreleerd.

Inhoudsopgave

1	Inleiding	1
2	De locaties en de gebruikte meetgegevens	2
3	De methodiek	5
3.1	Data-voorbewerking	5
3.2	De verklaarde variantie per stof per locatie	6
3.3	De verklaarde variantie per locatie	7
3.4	Identificatie versus verificatie	7
3.5	De autocorrelatiecoëfficiënt	7
3.6	De relatieve schattingsfout per stof	8
3.7	De autocorrelatiecoëfficiënt na filtering van sezoensinvloeden	8
3.8	De autocorrelatiecoëfficiënt van de modelfouten van het locatiemodel	9
4	Resultaten	10
4.1	Resultaten voor het schatten van de correlatie in de ruimte ..	10
4.1.1	De Westerschelde	10
4.1.2	De Noordzee bij Noordwijk	14
4.1.3	De Waddenzee West	17
4.1.4	De Eems-Dollard	20
4.2	Resultaten voor het schatten van de correlatie in de tijd	23
4.2.1	De Westerschelde	23
4.2.2	De Noordzee bij Noordwijk	24
4.2.3	De Waddenzee West	25
4.2.4	De Eems-Dollard	26
5	Conclusies	28
6	Referenties/Literatuur	34
BIJLAGE :	Uitgebreide resultaten Westerschelde	35
BIJLAGE :	Uitgebreide resultaten Noordzee bij Noordwijk	45
BIJLAGE :	Uitgebreide resultaten Waddenzee West	55
BIJLAGE :	Uitgebreide resultaten Eems-Dollard	65
BIJLAGE :	Theoretische achtergronden	75

1 Inleiding

Dit statistisch onderzoek is uitgevoerd in het kader van een evaluatie van het chemische meetnet. Het onderzoek betreft de gebieden de Westerschelde, de Noordzee bij Noordwijk, de Waddenzee West en de Eems-Dollard. Van een aantal waterkwaliteitsvariabelen is het de vraag, wanneer deze niet meer gecorreleerd zijn in de ruimte en in de tijd. Door analyse van reeksen van een locatie met een andere dichtbij gelegen locatie van eenzelfde waterkwaliteitsvariabele is de afhankelijkheid in de ruimte te bepalen. Door de reeks van een waterkwaliteitsvariabele op één bepaalde locatie te analyseren is de autocorrelatie, de afhankelijkheid in de tijd, te bepalen.

Voor RIZA [3] en Rijkswaterstaat Directie Zuid-Holland [4] en [5] is een vergelijkbaar onderzoek uitgevoerd, waarbij vooral de correlatie in de ruimte van van locaties is onderzocht.

In hoofdstuk 2 wordt een overzicht gegeven van de gebruikte waterkwaliteitsgegevens en alle locaties van de 4 gebieden. In hoofdstuk 3 worden de methodes beschreven welke zijn gebruikt voor het bepalen van de afhankelijkheid in de ruimte en in de tijd. In hoofdstuk 4 worden de methodes toegepast op alle locaties van de 4 gebieden. In hoofdstuk 5 worden de resultaten samengevat, hierbij ligt het accent op het voorkomen van correlatie en niet zozeer op het waarom van de correlatie. In de bijlages 'Uitgebreide resultaten' zijn extra resultaten gegeven voor elk gebied. In de bijlage 'Theoretische achtergronden' worden achtergronden van de methodes toegelicht.

Dit onderzoek is uitgevoerd met begeleiding van O.C. Swertz en R.N.M. Duin van RIKZ, G.T.M. van Eck heeft commentaar geleverd op het concept-rapport.

2 De locaties en de gebruikte meetgegevens

Het onderzoek betreft waterkwaliteitsmetingen van 15 verschillende stoffen in de periode 1971 t/m 1994. Deze stoffen zijn bemeaten in de gebieden Westerschelde, de Noordzee bij Noordwijk, de Waddenzee West en de Eems-Dollard. In tabel 1 t/m 5 zijn de stoffen en de locaties per gebied gegeven. In de bijlage zijn per gebied extra tabellen gegeven van het aantal metingen en het aantal jaren dat er is gemeten per stof voor de verschillende locaties.

Tabel 1 : De onderzochte waterkwaliteitsvariabelen

nr	omschrijving grootheid	te hanteren code	eenheid
1	saliniteit	Salntt	PSU
2	zwevend stof	Zs	mg/l
3	chlorofyl- α	Chlfa	$\mu\text{g/l}$
4	nitraat plus nitriet	NO ₃ NO ₂	mg N/l
5	opgelost stikstof	N-nf	mg/l
6	ammonium	NH ₄	mg N/l
7	fosfaat	PO ₄	mg P/l
8	opgelost fosfor	P-nf	mg/l
9	fosfor	P	mg/l
10	opgelost cadmium	Cd-nf	$\mu\text{g/l}$
11	opgelost koper	Cu-nf	$\mu\text{g/l}$
12	opgelost arseen	As-nf	$\mu\text{g/l}$
13	cadmium	Cd	$\mu\text{g/l}$
14	koper	Cu	$\mu\text{g/l}$
15	arseen	As	$\mu\text{g/l}$

In de periode 1971 t/m 1994 zijn deze stoffen gemeten in de 4 gebieden de Westerschelde, de Noordzee bij Noordwijk, de Waddenzee West en de Eems-Dollard.

Tabel 2 : De onderzochte locaties in de Westerschelde

volg nr.	naam locatie	korte naam	verb. ¹ kappa getal	oude WORSRO-code
1	Bath	Bath	91	WS50
2	Lamswaarde	Lamswde	86	WS70
3	Zuidergat	Zuidgt	81	WS90
4	Hansweert	Hanswt	78	WS100
5	Hoedekenskerke	Hoedkkk	75	WS110
6	Terneuzen	Ternzn	68	WS130
7	Honte	Honte	63	WS140
8	Borsssele	Borssle	61	WS150
9	Vlissingen	Vlissgn	59	WS160
10	Wielingen	Wielgn	49	NZ400
11	Scheur west	Scheurwt	45	NZ390

Tabel 3 : De onderzochte locaties in de Noordzee bij Noordwijk

volg nr.	naam locatie	korte naam	verb. kappa getal	oude WORSRO-code
1	Noordwijk 1 km uit de kust	Nw01	110	NZRNW1
2	Noordwijk 2 km uit de kust	Nw02	110	NZRNW2
3	Noordwijk 4 km uit de kust	Nw04	110	NZRNW4
4	Noordwijk 10 km uit de kust	Nw10	110	NZRNW10
5	Noordwijk 20 km uit de kust	Nw20	110	NZRNW20
6	Noordwijk 30 km uit de kust	Nw30	110	NZRNW30
7	Noordwijk 50 km uit de kust	Nw50	110	NZRNW50
8	Noordwijk 70 km uit de kust	Nw70	110	NZRNW70

Tabel 4 : De onderzochte locaties in de Waddenzee West

volg nr.	naam locatie	korte naam	verb. kappa getal	oude WORSRO-code
1	Marsdiep noord	Marsdnd	197	WZ30
2	Doove Balg west	Doovbot	220	WZ110
3	Doove Balg oost	Doovbwt	259	WZ200
4	Vliestroom	Vliesm	238	WZ190
5	Blauwe Slenk oost	Blauwsot	250	WZ230
6	Harlingen	Harlgn	276	WZ310

Tabel 5 : De onderzochte locaties in de Eems-Dollard

volg nr.	naam locatie	korte naam	verb. kappa getal	oude WORSRO-code
1	Groote Gat noord	Grootgnd	341	ED30
2	Zeehavenkanaal	Zeehvknl	335	ED65
3	Oostfrieze Gaatje	Oostfsgje	327	ED90
4	Bocht van Watum	Bochtwtm	328	ED155
5	Bocht van Watum noord	Bochtwnd	314	ED170
6	Huibertgat oost	Huibgot	305	ED250

3 De methodiek

In dit hoofdstuk wordt de methodiek besproken voor het bepalen van de afhankelijkheid (correlatie) in de ruimte en in de tijd. Voordat deze afhankelijkheden bepaald kunnen worden, worden verschillende voorbereidingen op de data uitgevoerd.

Voor het bepalen van de **afhankelijkheid in de ruimte** wordt een locatiemodel toegepast. Bij een locatiemodel wordt de stof van de meetlocatie geschat met behulp van een andere meetlocatie. De afhankelijkheid wordt uitgedrukt in de **verklaarde variantie**.

Voor het bepalen van de **afhankelijkheid in de tijd** worden 4 verschillende methodes toegepast :

- de autocorrelatiecoëfficiënt wordt bepaald
- een autoregressief model wordt toegepast
- de autocorrelatiecoëfficiënt wordt bepaald na correctie van het seizoenseffect [6]
- de autocorrelatiecoëfficiënt wordt bepaald van de residuen van het locatiemodel

Bij het autoregressieve model worden metingen geschat door voorgaande metingen en de **relatieve schattingsfout** wordt berekend. Bij de correctie van het seizoenseffect worden de gegevens gecentreerd rond maandgemiddelde waarden en vervolgens wordt de **autocorrelatiecoëfficiënt** berekend.

Een uitgebreide uitleg van de methodiek is gegeven in de bijlage 'Theoretische achtergronden'.

3.1 Data-voorbewerking

Voor de analyse van de data worden de meetreeksen onderzocht op meetwaarden onder een detectiegrens en op uitbijters. Voor de identificatie van het locatiemodel dienen er voldoende gegevens te zijn, in dit onderzoek is aan deze conditie altijd voldaan.

Meetwaarden kleiner dan de detectiegrens krijgen de waarde $1/2$ maal de detectiegrens.

Uitbijters zijn uit de meetreeksen verwijderd en zijn niet gebruikt voor het identificeren van het locatiemodel. Het zijn die waarden welke groter of kleiner zijn dan 3 maal de standaard deviatie ten opzichte van het gemiddelde van de meetreeks. Uitbijters kunnen de modellering en daarmee de maat voor de verklaarde variantie vertekenen. In dit onderzoek is het van belang de gemiddelde niveaus zo goed mogelijk te schatten en toevallige extreem hoge concentraties verstoren dit. Voor een stof als bijvoorbeeld chlorofyl kan dit veroorzaken dat een voorjaarspiek niet gebruikt wordt.

De identificatie van het locatiemodel is aan de hand van gepaarde metingen van stoffen van locaties. Voor gepaarde metingen van de stoffen geldt dat er niet meer dan 5 dagen verschil is tussen de meetdag van de te schatten

locatie en de meetdag van de andere 'buur-locatie'.

Voor het bepalen van de correlatie in de tijd is het nodig dat het meetinterval in de reeksen steeds dezelfde is. Bij de meetreeksen in dit onderzoek is dit vaak niet het geval. Afhankelijk van het grootste meetinterval van een meetreeks van een stof is per maand of per kwartaal de mediaan gekozen als representatieve waarde voor een meetperiode. Het meetinterval voor de stoffen saliniteit t/m fosfor (zie tabel 1) wordt een maand en voor de stoffen opgelost cadmium t/m arseen een kwartaal.

3.2 De verklaarde variantie per stof per locatie

Als een locatie niet meer bemeten wordt, betekent dit informatieverlies. Maar mogelijk zijn de concentraties van de stoffen te schatten aan de hand van metingen van een andere dichtbij gelegen locatie. In een locatiemodel (zie 'Bijlage : Theoretische achtergronden') wordt nu de concentratie van een stof van een locatie geschat aan de hand van de concentratie van dezelfde stof van een in de buurt gelegen locatie.

De modelfout of het residu is het verschil tussen de werkelijk gemeten waarde en de modelschatting. Gemiddeld is de modelfout gelijk aan nul. De variantie van de modelfout is een maat voor de kwaliteit van de modelschattingen, hoe kleiner de variantie des te beter de modelschattingen.

Door het verschil van de variantie van de concentratie en de variantie van de modelfout te delen door de variantie in de concentratie, krijgen we een maat die per stof vergelijkbaar is, de verklaarde variantie is :

$$\text{verklaarde variantie} = (\text{var}(\text{concentratie}) - \text{var}(\text{modelfout})) / \text{var}(\text{concentratie}) * 100\%$$

Kan de concentratie goed geschat worden dan is de variantie in de modelfout relatief klein en daarmee is de verklaarde variantie groot.

Kan de concentratie niet geschat worden dan is de variantie van de modelfout gelijk aan de variantie van de te schatten concentratie en is de verklaarde variantie 0%. De verklaarde variantie is dus een waarde tussen 0% (geen relatie tussen beide locaties) en 100% (perfecte beschrijving).

In de rapporten [3], [4] en [5] is de grootheid informatieverlies gebruikt. Het informatieverlies is gelijk aan 100% minus de verklaarde variantie. Een ook veel gebruikte grootheid is de kruiscorrelatiecoëfficiënt (met lag 0). De kruiscorrelatiecoëfficiënt is gelijk aan de wortel van de verklaarde variantie gedeeld door 100%.

Er geldt :

$$\text{informatieverlies} = 100\% - \text{verklaarde variantie}$$

$$\text{kruiscorrelatiecoëfficiënt} = \sqrt{(\text{verklaarde variantie} / 100\%)}$$

3.3 De verklaarde variantie per locatie

De verklaarde variantie per locatie is een al of niet gewogen som van de verklaarde variantie per stof over alle 15 stoffen. Hierbij zijn verschillende keuzes mogelijk voor het wegen van de verklaarde varianties van de verschillende stoffen. Zijn alle 15 stoffen even belangrijk voor het bepalen van de verklaarde variantie dan voldoet het gemiddelde van de verklaarde varianties over alle 15 stoffen. Zijn sommige stoffen belangrijker of vaker bemeten dan dienen de informatieverliezen verschillend gewogen te worden. Minder belangrijke of minder vaak bemeten stoffen tellen dan minder zwaar mee in de verklaarde variantie van een locatie.

In dit rapport wordt voor alle locaties in de gebieden de gemiddelde verklaarde variantie over alle 15 stoffen bepaald en geanalyseerd. Als bepaalde stoffen niet bemeten zijn, is de verklaarde variantie op 0% gesteld. Op een aantal locaties, waar niet alle stoffen bemeten zijn, wordt hierdoor de gemiddelde verklaarde variantie onderschat.

3.4 Identificatie versus verificatie

Om de robuustheid te onderzoeken van een locatiemodel wordt een meetreeks verdeeld in een identificatie- en een verificatieperiode. Bij de identificatieperiode worden de modelparameters van het locatiemodel geschat. In de verificatieperiode wordt de modeluitkomsten van het locatiemodel vergeleken met de werkelijk gemeten concentraties. De schattingsfout, het verschil tussen de modeluitkomst en de werkelijk gemeten concentratie kan worden berekend. Indien de variantie van de schattingsfouten duidelijk verschilt met de variantie van de modelfouten in de identificatieperiode dan moet geconcludeerd worden, dat het geïdentificeerde locatiemodel niet universeel geldig is.

3.5 De autocorrelatiecoëfficiënt

Per gebied is de locatie gekozen met de langste tijdreeksen van de stoffen. Afhankelijk van de stof is per maand of per kwartaal de mediaan gekozen als representatieve waarde voor een meetperiode (zie 3.1). De autocorrelatiefunctie geeft per lag (in maanden) de autocorrelatiecoëfficiënt. De autocorrelatiecoëfficiënt geeft een maat voor de lineaire afhankelijk van opeenvolgende meetwaarden.

3.6 De relatieve schattingsfout per stof

Een vergelijkbare methode als het bepalen van de autocorrelatiecoëfficiënt, is het berekenen van de relatieve schattingsfout. Doordat opeenvolgende metingen met elkaar gecorreleerd zijn is het wellicht mogelijk de meetfrequentie te verlagen, zonder dat er veel informatie verloren gaat.

In elk gebied wordt de locatie gekozen met de meest frequent bemeten en langste tijdreeksen. Omdat de meetintervallen van de meetreeksen steeds verschillend zijn, wordt afhankelijk van de stof per maand of per kwartaal de mediaan gekozen als representatieve waarde voor een meetperiode (zie 3.1). Van deze tijdreeksen wordt voor elke stof de relatieve schattingsfout bepaald, waarbij tussenliggende metingen worden geschat. De schatting van tussenliggende metingen kan door lineaire interpolatie of met een autoregressief model.

Stel dat de meetfrequentie wordt verlaagd van eens per maand naar eens per 2 maanden. Bij lineaire interpolatie is de schatting van een tussenliggende meting het gemiddelde van een voorgaande en een volgende meting. Bij het meest eenvoudige autoregressieve model, het naïeve model, is de te schatten meting gelijk aan voorgaande meting. Door de schatting te vergelijken met de gemeten waarde is de relatieve schattingsfout te berekenen. De relatieve schattingsfout is als volgt gedefinieerd :

$$\text{relatieve schattingsfout} = (\Sigma \text{var(schattingsfout)} / \Sigma \text{var(concentratie)}) * 100\%$$

De relatieve schattingsfout is de som van de variantie van de schattingsfout over alle jaren gedeeld door de som van de variantie van alle gemeten concentraties over alle jaren. De schattingsfout is het verschil tussen de maandwaarde van metingen en de geschatte waarde door lineaire interpolatie. Door de jaarlijkse berekening van de varianties en sommatie over alle jaren, wordt de relatieve schattingsfout minder gevoelig voor trendbreuken.

Kunnen de tussenliggende concentraties goed geschat worden dan is de variantie in de schattingsfout relatief klein en daarmee is ook de relatieve schattingsfout klein. De relatieve schattingsfout is dan zeker kleiner dan 100%.

Kunnen de tussenliggende concentraties niet geschat worden dan is de variantie van de schattingsfout groter dan de variantie van de concentraties. De relatieve schattingsfout is dan zeker groter dan 100%.

3.7 De autocorrelatiecoëfficiënt na filtering van seizoensinvloeden

Bij de gewone autocorrelatiecoëfficiënt (3.4) en de relatieve schattingsfout (3.5) is er geen rekening gehouden met seizoensinvloeden. De concentraties van bepaalde stoffen kunnen daarentegen sterk bepaald zijn door het seizoen, waarin is gemeten. Door de maand- of kwartaalwaarden te centreren rond maand- of kwartaal gemiddelde waarden kunnen seizoensinvloeden genivelleerd worden.

3.8 De autocorrelatiecoëfficiënt van de modelfouten van het locatiemodel

Indien in een locatiemodel de concentratie van een stof van een locatie goed geschat wordt aan de hand van de concentratie van dezelfde stof van een in de buurt gelegen locatie, dan zijn ook seizoeninvloeden goed beschreven. Zijn de modelfouten (residuen) sterk met elkaar gecorreleerd, dan betekent dit dat de meetfrequentie omlaag zou kunnen. Afhankelijk van de stof is per maand of per kwartaal de mediaan van de modelfouten gekozen als representatieve waarde voor een meetperiode (zie 3.1). De autocorrelatiefunctie geeft per lag (in maanden) de autocorrelatiecoëfficiënt van de modelfouten.

4 Resultaten

Voor het schatten van de correlatie in de ruimte is een eenvoudig lineair locatiemodel toegepast. De verklaarde variantie is een maat voor de correlatie in de ruimte. In 4.1 zijn tabellen met de resultaten voor de verschillende gebieden gegeven.

Voor het schatten van de (auto-) correlatie in de tijd zijn vier verschillende methoden toegepast :

- de berekening van de autocorrelatiecoëfficiënt
- door tussenliggende metingen te schatten met een autoregressief model
- berekening van de autocorrelatiecoëfficiënt na correctie van seizoensinvloeden
- door berekening van de autocorrelatiecoëfficiënt van de modelfouten van het locatiemodel.

De laatste twee methodes geven de correlatie in de tijd nadat seizoensinvloeden zijn genivelleerd. De methodes zijn toegepast op de locatie in een gebied met de langste tijdreeksen voor de verschillende stoffen. Van de meetgegevens zijn maand- en kwartaal-waarden gemaakt door de mediaan te nemen. In 4.2 zijn tabellen met de resultaten voor de verschillende gebieden gegeven van de gewone autocorrelatiecoëfficiënt, de overige methodes zijn gegeven in de bijlage 'Uitgebreide resultaten'.

4.1 Resultaten voor het schatten van de correlatie in de ruimte

Voor de gebieden de Westerschelde, de Noordzee bij Noordwijk, de Waddenzee West en de Eems-Dollard zijn in de volgende paragrafen de resultaten van de verklaarde varianties met een locatiemodel gegeven. In de grafieken worden de resultaten verduidelijkt. In de grafieken wordt ook de robuustheid van het locatiemodel aangetoond door een identificatieperiode te vergelijken met een verificatieperiode.

4.1.1 De Westerschelde

Voor elke locatie van de Westerschelde (tabel 2) is de verklaarde variantie bepaald. Met een locatiemodel zijn de concentraties van de 15 stoffen (tabel 1) van elke locatie geschat. In tabel 6 wordt de maximaal verklaarde variantie gegeven. Tevens is per locatie de locatie gegeven welke de maximaal verklaarde variantie geeft, het gemiddelde aantal gepaarde metingen van alle bemeten stoffen en het aantal bemeten stoffen.

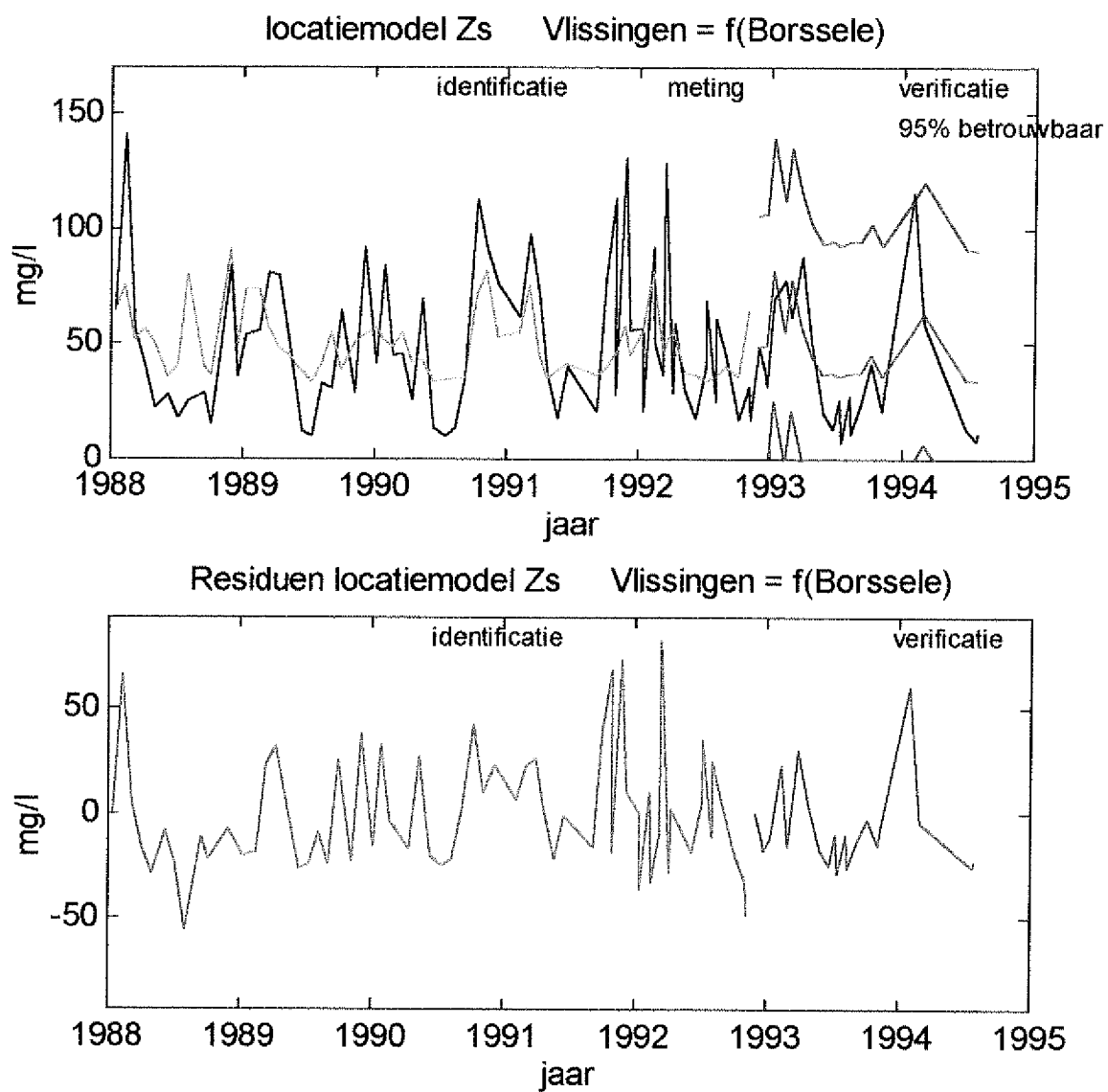
Tabel 6 : Max. verklaarde variantie (%) voor de locaties van de Westerschelde

Locatie	Verklaarde variantie	Beste buur-locatie	Gem. lengte	Aantal stoffen
Bath	71.7	Lamswaarde	114	15
Lamswaarde	75.4	Zuidergat	122	15
Zuidergat	77.0	Hansweert	121	15
Hansweert	76.9	Zuidergat	122	15
Hoedekenskerke	68.5	Hansweert	120	15
Terneuzen	72.5	Borssele	68	15
Honte	49.3	Borssele	96	9
Borssele	73.4	Terneuzen	68	15
Vlissingen	63.7	Borssele	74	15
Wielingen	61.3	Vlissingen	20	12
Scheur west	26.4	Borssele	16	8

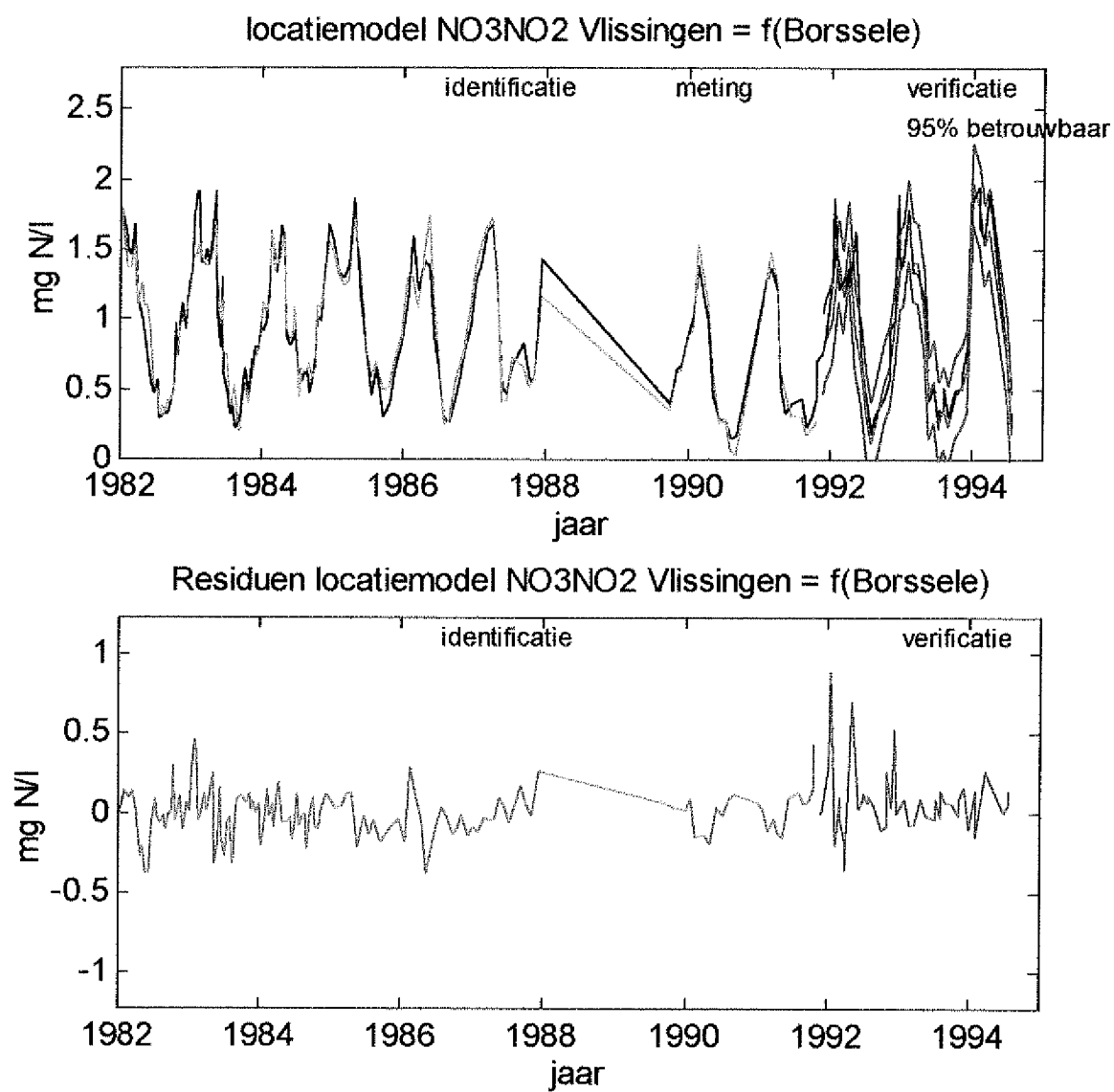
Een dichtbij gelegen locatie geeft steeds de maximale verklaarde variantie (tabel 3, bijlage). Dit komt overeen met wat fysisch verwacht kan worden. Bij de locaties Honte, Wielingen en Scheur west zijn niet alle stoffen bemeten. Bij deze locatie is de verklaarde variantie dan ook lager dan bij de overige locaties. De verklaarde variantie is immers op 0% gesteld als een stof niet bemeten is. Bij de locaties waar alle 15 stoffen zijn bemeten is de verklaarde variantie hoger dan 60%. Enkele stoffen met een hoge gemiddelde verklaarde variantie zijn Salntt, Chlfa, NO₃NO₂, N-nf, NH₄ (vooral landinwaarts), PO₄ en P-nf. De verklaarde variantie van de stof Cd-nf is zeer hoog (99%) bij de twee locaties Zuidergat en Hansweert.

In figuur 1 en 2 zijn de resultaten gegeven van het locatiemodel voor de stoffen Zs en NO₃NO₂. Per stof is weergegeven :

- Locatiemodel
 - identificatie (licht blauw). In de identificatie worden de modelparameters van het locatiemodel geschat a.h.v. de data (zwart).
 - verificatie (cyaan). In de verificatie worden, met het locatiemodel geschat in de identificatieperiode, de concentraties bij Vlissingen voorspeld.
 - 95% betrouwbaarheidsinterval (rood). Bij de voorspelde concentratie bij Vlissingen is ook een 95% betrouwbaarheidsgebied gegeven.
- Residuen
 - identificatie (groen). Het residue is het verschil tussen de gemeten concentratie en de schatting van het locatiemodel.
 - verificatie (magenta).



Figuur 1 : Residuen en gemeten versus locatiemodel zwevende stof van Vlissingen



Figuur 2 : Residuen en gemeten versus locatiemodel nitraat plus nitriet van Vlissingen

4.1.2 De Noordzee bij Noordwijk

Op verschillende afstanden voor de kust van Noordwijk zijn de verschillende stoffen bemeten. Voor elke locatie is de verklaarde variantie bepaald. Met een locatiemodel zijn de concentraties van de 15 stoffen (tabel 1) van elke locatie geschat. In tabel 7 wordt de maximaal verklaarde variantie gegeven. Tevens is per locatie de locatie gegeven welke de maximale verklaarde variantie geeft, het gemiddelde aantal gepaarde metingen van alle bemeten stoffen en het aantal bemeten stoffen.

Tabel 7 : Max. verklaarde variantie (%) voor de locaties bij Noordwijk

Locatie	Verklaarde variantie	Beste buur-locatie	Gem. lengte	Aantal stoffen
Noordwijk 1 km	77.2	Noordwijk 2 km	110	15
Noordwijk 2 km	76.9	Noordwijk 1 km	111	15
Noordwijk 4 km	76.3	Noordwijk 2 km	114	15
Noordwijk 10 km	59.5	Noordwijk 4 km	125	15
Noordwijk 20 km	54.8	Noordwijk 30 km	111	15
Noordwijk 30 km	56.4	Noordwijk 20 km	110	15
Noordwijk 50 km	47.3	Noordwijk 70 km	112	15
Noordwijk 70 km	46.7	Noordwijk 50 km	112	15

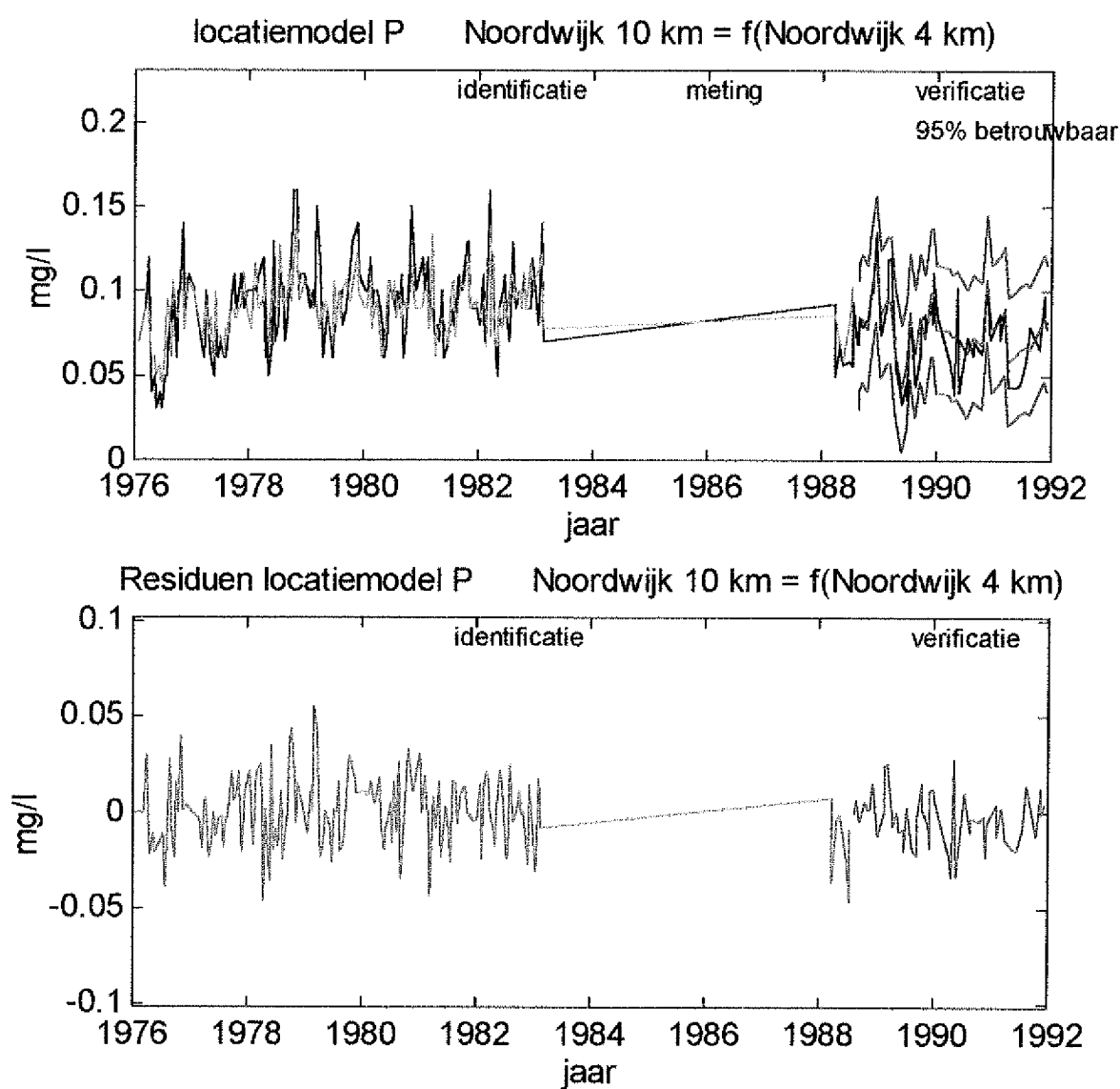
In tabel 7 geeft altijd een dichtbij gelegen locatie de maximale verklaarde variantie. Alle locaties zijn goed met elkaar vergelijkbaar, omdat op alle locaties alle stoffen goed bemeten zijn. Voor bijna alle stoffen is de verklaarde variantie dicht bij de kust het grootst (tabel 3, bijlage). In het bijzonder geldt dit voor de stof NH₄, bij de eerste twee locaties is de verklaarde variantie 90% en bij de laatste twee locaties, 50 en 70 km uit de kust, slechts 18%. Stoffen met de hoogste gemiddelde variantie zijn Salntt, NO₃NO₂, N-nf, PO₄, P-nf, Cd-nf en Cu-nf. Locaties dicht bij de kust zijn dus het beste te schatten. De verklaarde variantie varieert van 77% 1 km uit de kust tot 47% 70 km uit de kust.

In figuur 3 en 4 zijn de resultaten gegeven van het locatiemodel van de stoffen P en N-nf. Per stof is weergegeven :

- Locatiemodel

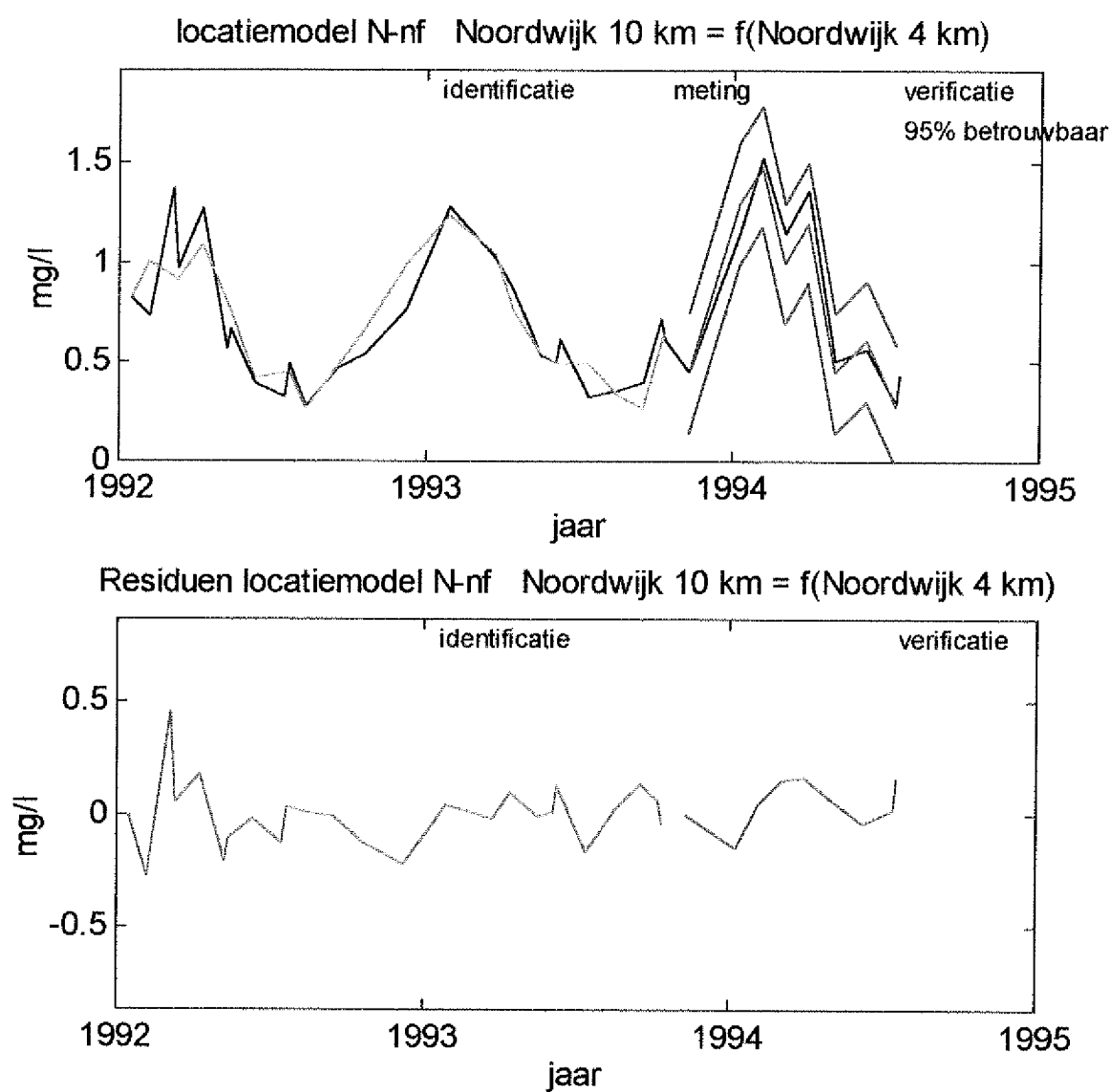
- identificatie (licht blauw). In de identificatie worden de modelparameters van het locatiemodel geschat a.h.v. de data (zwart).
- verificatie (cyaan). In de verificatie worden, met het locatiemodel geschat in de identificatieperiode, de concentraties bij Noordwijk (10 km) voorspeld.
- 95% betrouwbaarheidsinterval (rood). Bij de voorspelde concentratie bij Noordwijk (10 km) is ook een 95% betrouwbaarheidsgebied gegeven.

- 95% betrouwbaarheidsinterval (rood). Bij de voorspelde concentratie bij Noordwijk (10 km) is ook een 95% betrouwbaarheidsgebied gegeven.
- Residuen
 - identificatie (groen). Het residue is het verschil tussen de gemeten concentratie en de schatting van het locatiemodel.
 - verificatie (magenta).



Figuur 3 : Residuen en gemeten versus locatiemodel fosfor van Noordwijk 10 km

Uit de grafiek blijkt dat in de periode 1983 t/m 1987 P niet is bemeaten.



Figuur 4 : Residuen en gemeten versus locatiemodel opgelost stikstof bij Noordwijk 10 km

4.1.3 De Waddenzee West

Voor elke locatie is de verklaarde variantie bepaald. Met een locatiemodel zijn de concentraties van de 15 stoffen (tabel 1) van elke locatie geschat. In tabel 8 wordt de maximale verklaarde variantie gegeven. Tevens is per locatie de locatie gegeven welke de maximale verklaarde variantie geeft, het gemiddelde aantal gepaarde metingen van alle bemeten stoffen en het aantal bemeten stoffen.

Tabel 8 : Max. verklaarde variantie (%) voor de locaties op de Waddenzee West

Locatie	Verklaarde variantie	Beste buur-locatie	Gem. lengte	Aantal stoffen
Marsdiep noord	42.8	Vliestroom	120	15
Doove Balg west	35.1	Doove Balg oost	150	9
Doove Balg oost	38.1	Vliestroom	104	15
Vliestroom	43.4	Marsdiep noord	119	15
Blauwe Slenk oost	31.8	Harlingen	158	9
Harlingen	30.8	Blauwe Slenk oost	186	9

De verklaarde varianties zijn duidelijk lager dan bij de vorige twee gebieden. Stoffen met de hoogste gemiddelde variantie zijn NO₃NO₂ en N-nf (tabel 3, bijlage). De locatie Marsdiep noord is vergelijkbaar met locatie Vliestroom vanwege de positie t.o.v. de Noordzee. De overige locaties liggen bij elkaar in de buurt. Bij 3 locaties zijn 9 stoffen bemeten en bij de andere 3 locaties alle 15 stoffen. De verklaarde varianties van de 3 locaties waar minder stoffen bemeten zijn zijn kleiner (tot 35%) dan bij de locaties waar alle stoffen bemeten zijn (tot 43%).

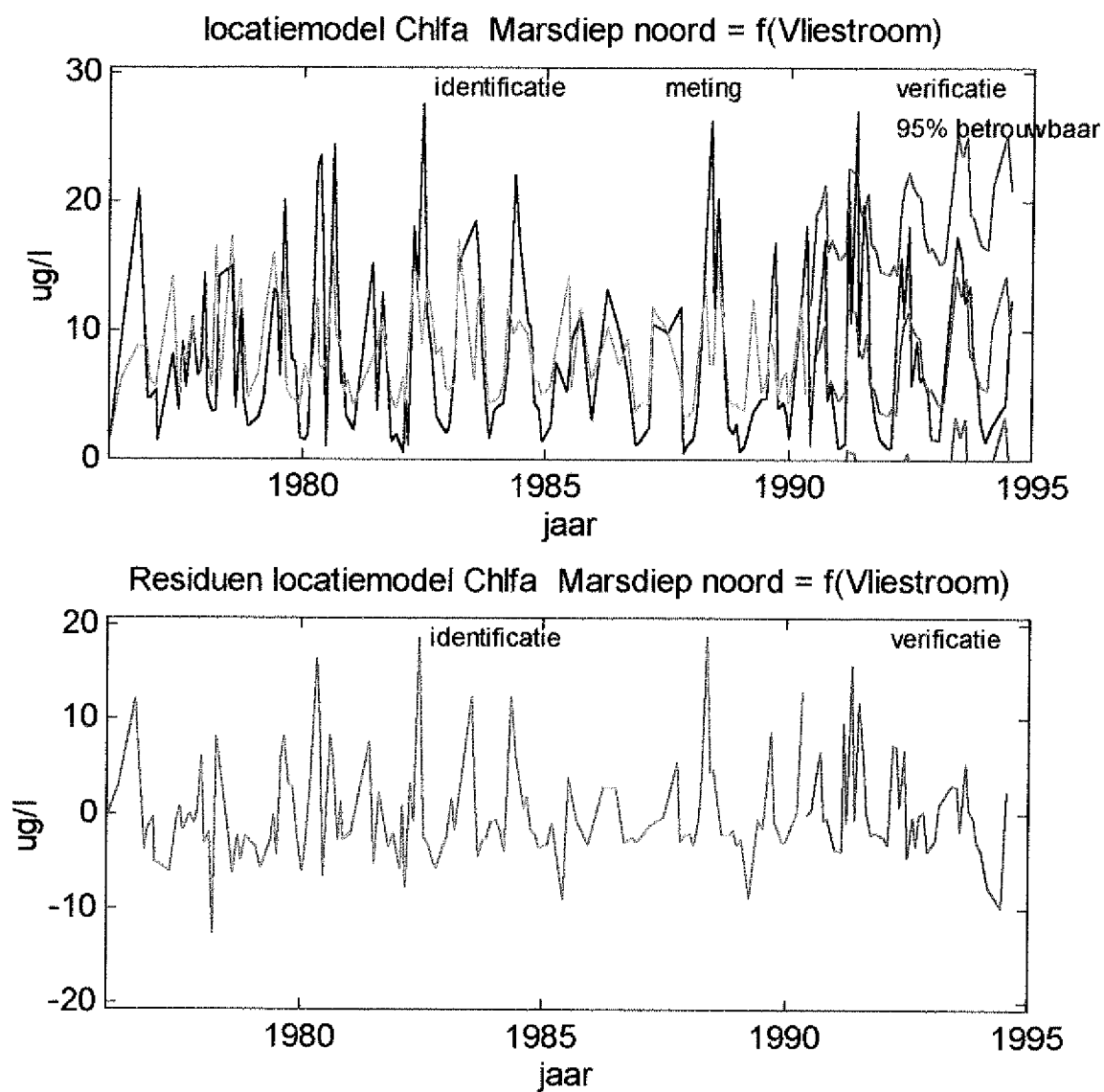
In figuur 5 en 6 zijn de resultaten gegeven van het locatiemodel voor de stoffen Chlfa en As. Per stof is weergegeven :

- Locatiemodel

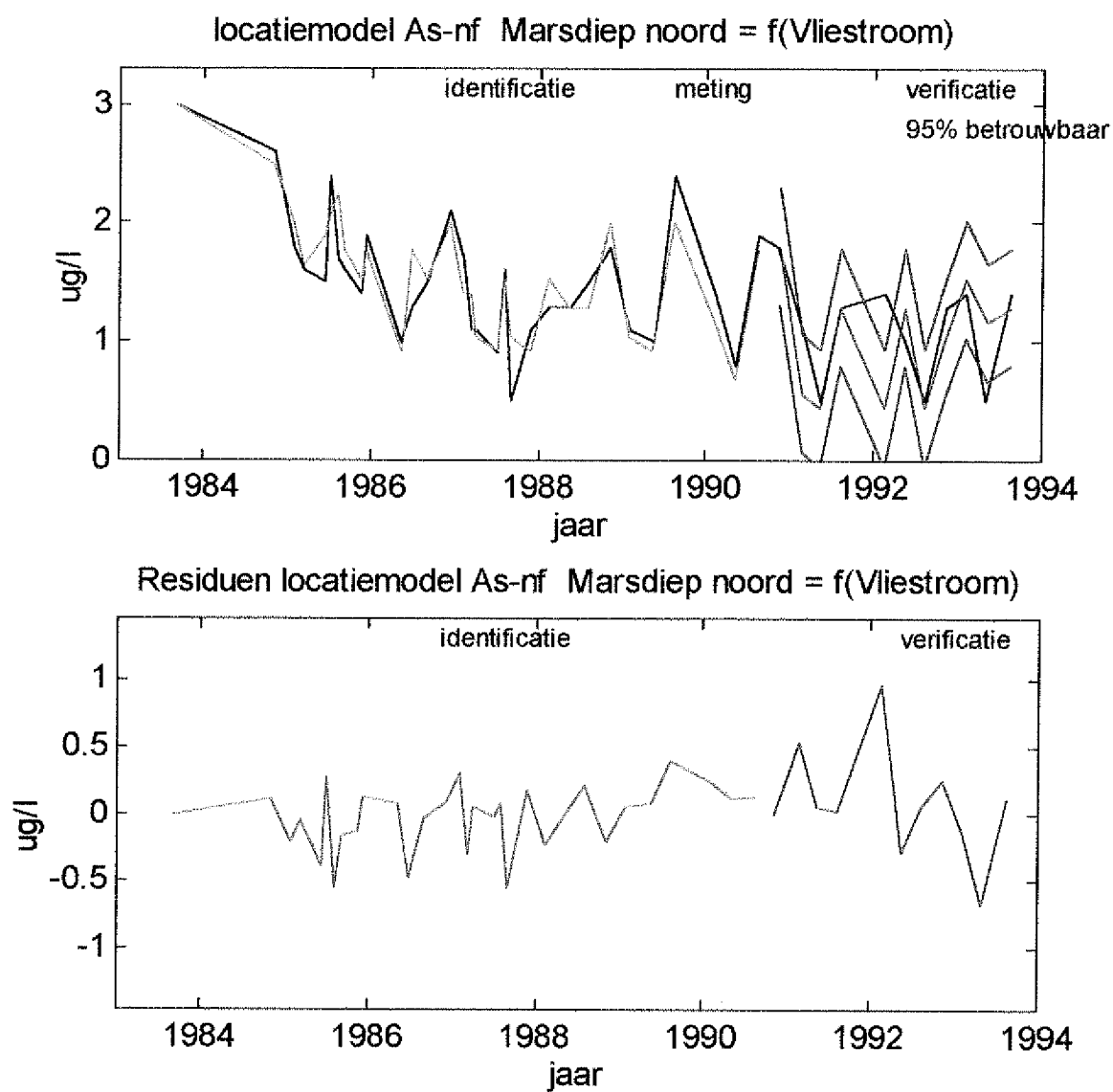
- identificatie (licht blauw). In de identificatie worden de modelparameters van het locatiemodel geschat a.h.v. de data (zwart).
- verificatie (cyaan). In de verificatie worden, met het locatiemodel geschat in de identificatieperiode, de concentraties bij Marsdiep noord voorspeld.
- 95% betrouwbaarheidsinterval (rood). Bij de voorspelde concentratie bij Marsdiep noord is ook een 95% betrouwbaarheidsgebied gegeven.

- Residuen

- identificatie (groen). Het residue is het verschil tussen de gemeten concentratie en de schatting van het locatiemodel.
- verificatie (magenta).



Figuur 5 : Residuen en gemeten versus locatiemodel Chlorofyl- α bij Marsdiep noord



Figuur 6 : Residuen en gemeten versus locatiemodel opgelost arseen bij Marsdiep noord

4.1.4 De Eems-Dollard

Voor elke locatie in de Eems-Dollard is de verklaarde variantie bepaald. Met een locatiemodel zijn de concentraties van de 15 stoffen (tabel 1) van elke locatie geschat. In tabel 9 wordt de maximale verklaarde variantie gegeven. Tevens is per locatie de locatie gegeven welke de maximale verklaarde variantie geeft, het gemiddelde aantal gepaarde metingen van alle bemeaten stoffen en het aantal bemeaten stoffen.

Tabel 9 : Max. verklaarde variantie (%) voor de locaties op de Eems-Dollard

Locatie	Verklaarde variantie	Beste buur-locatie	Gem. lengte	Aantal stoffen
Groote Gat noord	61.4	Zeehavenkanaal	42	15
Zeehavenkanaal	64.6	Bocht van Watum	41	15
Oostfrieze Gaatje	74.6	Bocht van Watum	42	15
Bocht van Watum	73.9	Oostfrieze Gaatje	42	15
Bocht van Watum n.	61.1	Bocht van Watum	43	15
Huibertsgat oost	38.6	Bocht van Watum n.	214	15

Bij alle locaties zijn alle stoffen bemeaten. Stoffen met de hoogste gemiddelde verklaarde variantie zijn : Salntt, NO₃NO₂, N-nf, PO₄, P-nf en Cu-nf. De locatie Huibertsgat oost is het meest bemeaten en is het slechts uit andere locaties te schatten, maar ligt ook ver verwijderd van de overige locaties. De centraal in de Eems-Dollard gelegen locatie Bocht van Watum is vaak de beste buur-locatie. In de Eems-Dollard geldt niet dat een locatie het beste door een dichtbij gelegen buur-locatie geschat wordt. De verklaarde variantie voor de locaties in de Eems-Dollard is groter dan 60%.

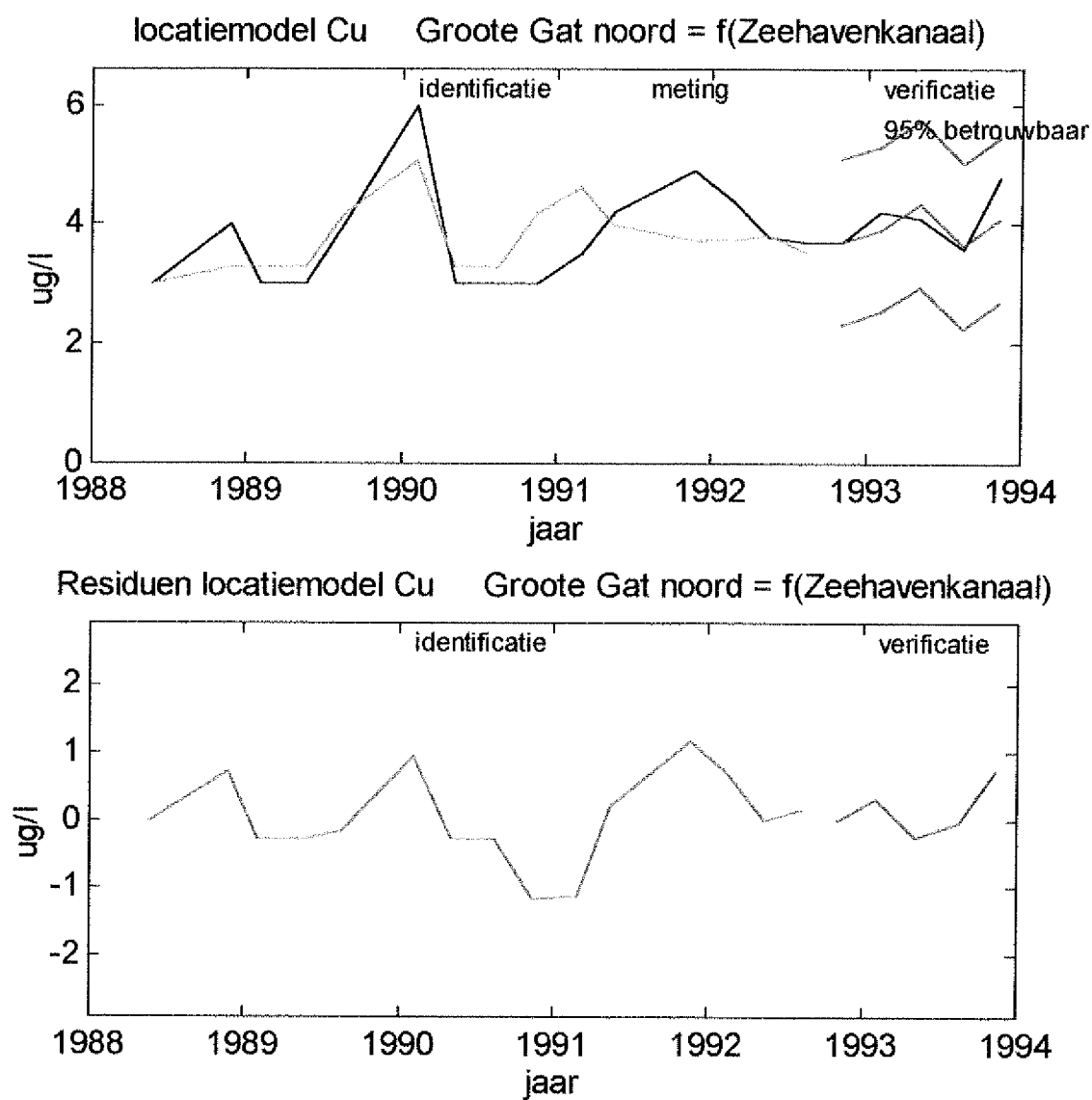
In figuur 7 en 8 zijn de resultaten gegeven van het locatiemodel voor de stoffen Cu en P-nf. Per stof is weergegeven :

- Locatiemodel

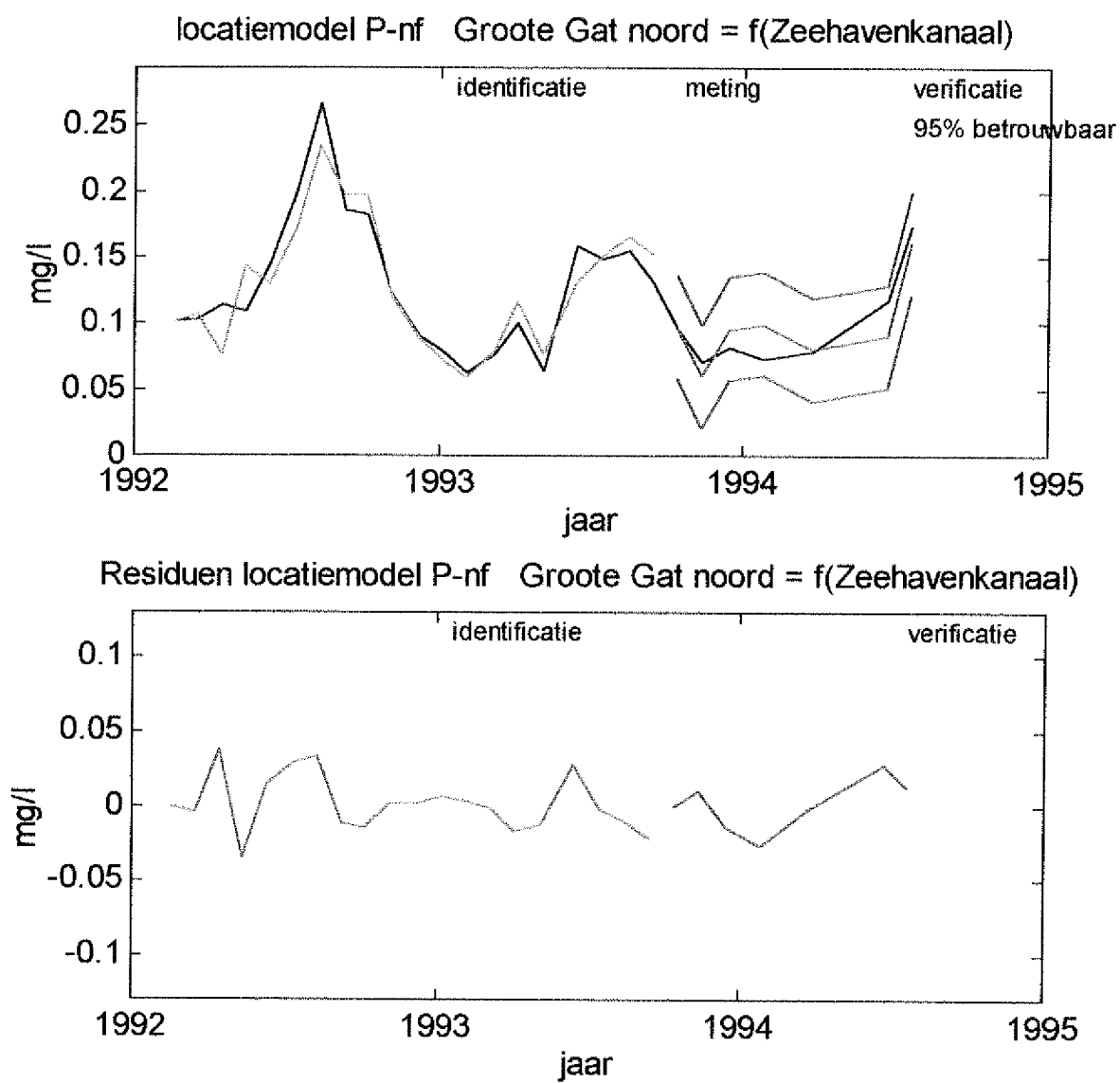
- identificatie (licht blauw). In de identificatie worden de modelparameters van het locatiemodel geschat a.h.v. de data (zwart).
- verificatie (cyaan). In de verificatie worden, met het locatiemodel geschat in de identificatieperiode, de concentraties bij Groote Gat noord voorspeld.
- 95% betrouwbaarheidsinterval (rood). Bij de voorspelde concentratie bij Groote Gat noord is ook een 95% betrouwbaarheidsgebied gegeven.

- Residuen

- identificatie (groen). Het residue is het verschil tussen de gemeten concentratie en de schatting van het locatiemodel.
- verificatie (magenta).



Figuur 7 : Residuen en gemeten versus locatiemodel koper bij Groote Gat noord



Figuur 8 : Residuen en gemeten versus locatiemodel opgelost fosfor bij Groote Gat noord

4.2 Resultaten voor het schatten van de correlatie in de tijd

Voor elk gebied wordt per stof de langste tijdreeks genomen. Deze tijdreeksen worden geconverteerd naar reeksen met eenzelfde meetinterval. Afhankelijk van de stof is per maand of per kwartaal de mediaan gekozen als representatieve waarde voor een meetperiode. Voor de stoffen salntt t/m P (zie tabel 1) hebben de autocorrelatiecoëfficiënten betrekking op maandwaarden en bij de stoffen Cd-nf t/m As op kwartaalwaarden. Op deze nieuwe reeksen worden de methodieken toegepast om de afhankelijkheid in de tijd te kunnen bepalen. In de tabellen 10 t/m 13 is voor de verschillende gebieden de autocorrelatiecoëfficiënt bepaald. In de bijlage zijn de resultaten gegeven van de bepaling van de relatieve schattingsfout bij het gebruik van het naïeve model, de autocorrelatiecoëfficiënt na correctie op seizoensinvloeden en de autocorrelatiecoëfficiënt van de modelfouten van het locatiemodel.

4.2.1 De Westerschelde

Vlissingen is in de Westerschelde de locatie met de langste tijdreeksen. In tabel 10 zijn autocorrelatiecoëfficiënten gegeven voor alle stoffen als functie van het tijdsverschil tussen twee metingen. Heeft toets de waarde 0 dan wordt de nulhypothese dat opeenvolgende waarnemingen ongecorrleerd zijn met een betrouwbaarheid van 99%, niet verworpen. Heeft toets de waarde 1 dan wordt de alternatieve hypothese, opeenvolgende waarnemingen zijn gecorrleerd, geaccepteerd. Hoe hoger de autocorrelatiecoëfficiënt en hoe kleiner het aantal waarnemingen des te groter de kans dat de nulhypothese wordt verworpen.

Tabel 10 : Autocorrelatiecoëfficiënten voor alle stoffen bij Vlissingen

Stof	1 mnd	toets	2 mnd	toets	3 mnd	toets	4 mnd	toets
Salntt	0.64	1	0.50	1	0.27	1	0.06	0
Zs	0.46	1	0.20	0	-0.06	0	-0.15	0
Chlfa	0.38	1	0.14	0	0.05	0	-0.12	0
NO3NO2	0.70	1	0.40	1	0.06	0	-0.22	1
N-nf	0.58	1	0.29	0	-0.11	0	-0.37	0
NH4	0.48	1	0.35	1	0.26	1	0.27	1
PO4	0.51	1	0.35	1	0.21	1	0.01	0
P-nf	0.41	1	0.13	0	0.13	0	-0.05	0
P	0.59	1	0.29	1	0.02	0	-0.16	0

Stof	3 mnd	toets	6 mnd	toets	9 mnd	toets	jaar	toets
Cd-nf	0.10	0	0.12	0	0.24	1	0.13	0
Cu-nf	0.30	1	0.19	0	0.33	1	0.20	0
As-nf	0.10	0	-0.33	0	0.18	0	0.57	1
Cd	0.58	1	0.43	1	0.42	1	0.41	1
Cu	0.31	1	0.16	0	0.20	0	0.30	0
As	0.11	0	-0.34	0	0.03	0	0.40	1

De nulhypothese dat opeenvolgende metingen ongecorrleerd zijn bij een lag van 1 maand, wordt verworpen bij de stoffen salntt t/m P. Bij een lag van 1 maand is de autocorrelatiecoëfficiënt groter dan 0.6 bij de stoffen salntt en NO₃NO₂. Bij de stoffen salntt t/m P neemt de autocorrelatiecoëfficiënt af als de lag groter wordt. Van de onderste 6 stoffen wordt de nulhypothese dat opeenvolgende metingen ongecorrleerd zijn bij een lag van 3 maand niet verworpen voor de stoffen Cd-nf, As-nf en As. Bij de stoffen As-nf en As is de autocorrelatiecoëfficiënt negatief bij een lag van 6 maanden en hoog bij een lag van 1 jaar, dit duidt op een periodiciteit van een jaar.

4.2.2 De Noordzee bij Noordwijk

De langste tijdreeksen heeft de meetlocatie 10 km uit de kust bij Noordwijk. In tabel 11 zijn autocorrelatiecoëfficiënten gegeven voor alle stoffen als functie van het tijdsverschil tussen twee metingen. Heeft toets de waarde 0 dan wordt de nulhypothese dat de waarnemingen ongecorrleerd zijn met een betrouwbaarheid van 99%, niet verworpen. Heeft toets de waarde 1 dan wordt de alternatieve hypothese, opeenvolgende waarnemingen zijn gecorrleerd, geaccepteerd. Hoe hoger de autocorrelatiecoëfficiënt en hoe kleiner het aantal waarnemingen des te groter de kans dat de nulhypothese wordt verworpen.

Tabel 11 : Autocorrelatiecoëfficiënten voor alle stoffen bij de meetlocatie 10 km uit de kust bij Noordwijk

Stof	1 mnd	toets	2 mnd	toets	3 mnd	toets	4 mnd	toets
Salntt	0.26	1	0.17	0	0.12	0	0.11	0
Zs	0.19	0	0.14	0	0.05	0	-0.14	0
Chlfa	0.30	1	0.09	0	-0.04	0	-0.19	0
NO ₃ NO ₂	0.59	1	0.24	1	-0.03	0	-0.29	1
N-nf	0.58	1	0.24	0	-0.07	0	-0.36	0
NH ₄	0.28	1	0.10	0	-0.03	0	-0.07	0
PO ₄	0.58	1	0.35	1	0.15	0	0.03	0

P-nf	0.51	1	0.25	0	0.05	0	-0.12	0
P	0.29	1	0.20	0	0.07	0	0.11	0
Stof	3 mnd	toets	6 mnd	toets	9 mnd	toets	jaar	toets
Cd-nf	0.36	1	0.14	0	0.08	0	-0.25	0
Cu-nf	-0.07	0	-0.02	0	-0.33	0	0.20	0
As-nf	0.07	0	-0.29	0	-0.22	0	0.03	0
Cd	0.03	0	-0.74	1	-0.20	0	0.64	1
Cu	0.52	1	0.55	1	0.45	0	0.34	0
As	0.24	0	0.00	0	-0.13	0	0.00	0

Bij de stoffen salntt t/m P, uitgezonderd Zs, wordt bij een lag van 1 maand de nulhypothese verworpen dat opeenvolgende metingen ongecorrleerd zijn. Bij een lag van 1 maand is de autocorrelatiecoëfficiënt het hoogst bij de stoffen NO₃NO₂, N-nf, PO₄ en P-nf. Bij de stoffen Salntt t/m P neemt de autocorrelatiecoëfficiënt af als de lag groter wordt. Bij een lag van 4 maanden is de autocorrelatiecoëfficiënt vaak negatief. Bij de onderste 6 stoffen zijn de autocorrelatiecoëfficiënten van Cd-nf en Cu het grootst en wordt de nulhypothese bij een lag van 3 maanden verworpen. Bij de stof Cd is de autocorrelatiecoëfficiënt bij een lag van 6 maanden erg klein en bij een lag van 1 jaar erg groot, wat duidt op een periodiciteit van een jaar.

4.2.3 De Waddenzee West

De locatie Marsdiep noord heeft de langste tijdreeksen. In tabel 12 zijn autocorrelatiecoëfficiënten gegeven voor alle stoffen als functie van het tijdsverschil tussen twee metingen. Heeft toets de waarde 0 dan wordt de nulhypothese dat de waarnemingen ongecorrleerd zijn met een betrouwbaarheid van 99%, niet verworpen. Heeft toets de waarde 1 dan wordt de alternatieve hypothese, opeenvolgende waarnemingen zijn gecorreleerd, geaccepteerd. Hoe hoger de autocorrelatiecoëfficiënt en hoe kleiner het aantal waarnemingen des te groter de kans dat de nulhypothese wordt verworpen.

Tabel 12 : Autocorrelatiecoëfficiënten voor alle stoffen bij Marsdiep noord

Stof	1 mnd	toets	2 mnd	toets	3 mnd	toets	4 mnd	toets
Salntt	0.15	0	0.11	0	0.04	0	0.17	0
Zs	0.37	1	0.26	1	0.30	1	0.21	1
Chlfa	0.35	1	0.22	1	0.00	0	-0.20	1
NO3NO2	0.43	1	0.17	0	0.04	0	-0.17	0
N-nf	0.33	0	0.13	0	0.07	0	-0.23	0
NH4	0.43	1	0.25	1	0.30	1	0.32	1
PO4	0.56	1	0.30	1	0.19	1	0.13	0
P-nf	0.41	0	0.31	0	0.14	0	-0.04	0
P	0.14	0	0.10	0	0.17	0	0.08	0
Stof	3 mnd	toets	6 mnd	toets	9 mnd	toets	jaar	toets
Cd-nf	0.40	1	0.54	1	0.51	1	0.37	1
Cu-nf	0.10	0	0.13	0	0.23	0	0.25	0
As-nf	0.22	0	-0.06	0	0.14	0	0.21	0
Cd	0.71	1	0.58	1	0.58	1	0.61	1
Cu	0.36	1	0.14	0	0.23	0	0.20	0
As	0.21	0	0.08	0	0.17	0	0.32	0

Van de stoffen salntt t/m P wordt de nulhypothese dat opeenvolgende metingen ongecorrleerd zijn alleen bij de stoffen Salntt, N-nf, P-nf en P, niet verworpen. Bij een lag van 1 maand is de autocorrelatiecoëfficiënt het hoogst bij de stoffen NO3NO2, NH4 en PO4. Bij de onderste 6 stoffen zijn de autocorrelatiecoëfficiënten van Cd-nf, Cd en Cu het hoogst en wordt de nulhypothese bij een lag van 3 maanden verworpen. Alleen bij de stof As-nf is de autocorrelatiecoëfficiënt bij een lag van 6 maanden klein en bij een lag van 1 jaar weer groot, wat duidt op een periodiciteit van een jaar.

4.2.4 De Eems-Dollard

De locatie Groote Gat noord heeft de langste tijdreeksen. In tabel 12 zijn autocorrelatiecoëfficiënten gegeven voor alle stoffen als functie van het tijdsverschil tussen twee metingen. Heeft toets de waarde 0 dan wordt de nulhypothese dat de waarnemingen ongecorrleerd zijn met een betrouwbaarheid van 99%, niet verworpen. Heeft toets de waarde 1 dan wordt de alternatieve hypothese, opeenvolgende waarnemingen zijn gecorrleerd, geaccepteerd. Hoe hoger de autocorrelatiecoëfficiënt en hoe kleiner het aantal waarnemingen des te groter de kans dat de nulhypothese wordt verworpen.

Tabel 13 : Autocorrelatiecoëfficiënten voor alle stoffen bij Groote Gat noord

Stof	1 mnd	toets	2 mnd	toets	3 mnd	toets	4 mnd	toets
Salntt	0.66	1	0.4	1	0.16	0	-0.09	0
Zs	0.02	0	0.05	0	0.08	0	0.00	0
Chlfa	0.35	1	0.24	1	0.00	0	-0.06	0
NO3NO2	0.44	1	0.28	1	0.05	0	-0.19	1
N-nf	0.68	1	0.37	0	-0.04	0	-0.38	0
NH4	0.57	1	0.31	1	0.13	0	0.09	0
PO4	0.61	1	0.35	1	0.07	0	-0.21	1
P-nf	0.67	1	0.38	0	0.03	0	-0.23	0
P	0.12	0	0.11	0	0.05	0	0.01	0
Stof	3 mnd	toets	6 mnd	toets	9 mnd	toets	jaar	toets
Cd-nf	0.44	1	0.43	1	0.38	1	0.30	1
Cu-nf	0.06	0	-0.10	0	0.13	0	0.52	1
As-nf	0.17	0	-0.34	0	-0.05	0	0.21	0
Cd	0.12	0	0.24	0	-0.04	0	0.12	0
Cu	0.06	0	-0.01	0	0.31	1	0.22	0
As	0.01	0	0.09	0	0.01	0	0.05	0

Van de stoffen salntt t/m P worden alleen bij de stoffen Zs en P de nulhypothese dat opeenvolgende metingen ongecorrleerd zijn, niet verworpen. Bij een lag van 1 maand is de autocorrelatiecoëfficiënt bij de stoffen salntt, N-nf, PO4 en P-nf groter dan 0.6. Bij de stoffen salntt t/m P neemt de autocorrelatiecoëfficiënt af als de lag groter wordt. Bij een lag van 4 maanden is de autocorrelatiecoëfficiënt vaak negatief. Van de onderste 6 stoffen wordt alleen voor Cd-nf de nulhypothese dat opeenvolgende metingen ongecorrleerd zijn, niet verworpen. Bij de stof As-nf is de autocorrelatiecoëfficiënt negatief bij een lag van 6 maanden en hoog bij een lag van 1 jaar, dit duidt op een periodiciteit van een jaar.

5 Conclusies

Voor de gebieden de Westerschelde, de Noordzee bij Noordwijk, de Waddenzee West en de Eems-Dollard is de afhankelijkheid (correlatie) in de ruimte en in de tijd berekend.

Per locatie in een gebied is de gemiddelde verklaarde variantie berekend met een locatiemodel. Bij een locatiemodel wordt de stof van een locatie geschat met behulp van een andere locatie. De verklaarde variantie is een maat voor de afhankelijkheid in de ruimte.

Voor het schatten van de afhankelijkheid in de tijd wordt de relatieve schattingsfout berekend door het schatten van tussenliggende metingen in de meetreeks van een stof met lineaire interpolatie. Tevens is de gewone autocorrelatiecoëfficiënt bepaald en de autocorrelatiecoëfficiënt na het toepassen van de methode Duin voor het filteren van seizoensinvloeden.

Afhankelijkheid in de ruimte

De verklaarde variantie is berekend met een eenvoudig locatiemodel. De verklaarde variantie is het gemiddelde van de verklaarde variantie van alle 15 stoffen. Is een stof niet bemeten dan is de verklaarde variantie op 0% gesteld.

De Westerschelde

Er zijn 11 meetlocaties op de Westerschelde. Bij alle locaties, uitgezonderd de locaties Honte (9), Wielingen (12) en Scheur west (8), zijn alle 15 stoffen bemeten. Enkele stoffen met een hoge gemiddelde verklaarde variantie zijn Salntt, Chlfa, NO₃NO₂, N-nf, NH₄ (vooral landinwaarts), PO₄ en P-nf. Een dichtbij gelegen buur-locatie geeft steeds de maximaal verklaarde variantie. En hoe verder de buur-locatie is verwijderd des te kleiner de verklaarde variantie. De maximaal verklaarde variantie is hoog en daarmee ook de afhankelijkheid in de ruimte. Bij de locaties waar alle 15 stoffen zijn bemeten is de verklaarde variantie hoger dan 60%.

De Noordzee bij Noordwijk

Er zijn 8 meetlocaties op verschillende afstanden van de kust bij Noordwijk. Op alle locaties zijn alle 15 stoffen bemeten. Stoffen met hoge gemiddelde variantie zijn Salntt, NO₃NO₂, N-nf, PO₄, P-nf, Cd-nf en Cu-nf. Voor bijna alle stoffen is de verklaarde variantie dicht bij de kust het grootst en daarmee zijn dus de locaties dicht bij de kust het beste te schatten. De meest dichtbij gelegen buur-locatie geeft steeds de maximaal verklaarde variantie. De verklaarde variantie varieert van 77% 1 km uit de kust tot 47% 70 km uit de kust.

De Waddenzee West

Er zijn 6 meetlocaties bemeten op de Waddenzee West. Op 3 locaties zijn alle 15 stoffen bemeten en op 3 locaties slechts 9 stoffen. Stoffen met de hoogste gemiddelde variantie zijn NO₃NO₂ en N-nf. De verklaarde varianties zijn duidelijk lager dan bij de vorige twee gebieden. De locatie Marsdiep noord is vergelijkbaar met locatie Vliestroom vanwege de positie t.o.v. de Noordzee. De overige locaties liggen bij elkaar in de buurt. De verklaarde variantie van de 3 locaties waar minder stoffen bemeten zijn, is kleiner (tot 35%) dan bij de locaties waar alle stoffen bemeten zijn (tot 43%).

De Eems-Dollard

Op de Eems-Dollard zijn 6 locaties bemeten. Alle locaties zijn alle 15 stoffen bemeten. Stoffen met de hoogste gemiddelde verklaarde variantie zijn Salntt, NO₃NO₂, N-nf, PO₄, P-nf en Cu-nf. De locatie Huibertsgat oost is het slechts uit andere locaties te schatten, maar ligt ook ver verwijderd van de overige locaties. In de Eems-Dollard geldt niet dat een locatie het best door een dichtbij gelegen buur-locatie geschat wordt. De verklaarde variantie voor de locaties in de Eems-Dollard is groter dan 60%.

De locaties in het gebied Westerschelde, de locaties dicht bij de kust bij Noordwijk en locaties in de Eems-Dollard hebben een gemiddelde verklaarde variantie groter dan 60%. Bij de Waddenzee komen enkele locaties tot 43%.

Afhankelijkheid in de tijd

Voor het schatten van de afhankelijkheid in de tijd zijn de gewone autocorrelatiecoëfficiënt, de relatieve schattingsfout, de autocorrelatiecoëfficiënt na correctie van seizoensinvloeden en de autocorrelatiecoëfficiënt van de modelfouten van het locatiemodel berekend. Voor elk gebied is de locatie gekozen met de langste tijdreeksen. De gegevens van de langste meetreeksen zijn afhankelijk van de meetfrequentie van een stof geconverteerd naar maand- en kwartaalwaarden.

De Westerschelde

De locatie met de langste tijdreeksen is Vlissingen. Bij de stoffen salntt t/m P, met maandwaarden, wordt bij een lag van 1 maand de nulhypothese verworpen dat opeenvolgende metingen ongecorrleerd zijn. Bij de stoffen Cd-nf, As-nf en As, met kwartaalwaarden, wordt de nulhypothese niet verworpen.

Door filtering van seizoensinvloeden worden autocorrelatiecoëfficiënten kleiner. Bij de stoffen Chlfa, N-nf en P-nf bij een lag van 1 maand en voor Cd-nf, As-nf en As bij een lag van 3 maanden wordt de nulhypothese niet verworpen.

Uit de berekening van de relatieve schattingsfout blijkt, dat de nulhypothese dat het model zinvol is voor salntt, NO₃NO₂, N-nf, P-nf en P niet wordt verworpen.

De Noordzee bij Noordwijk

De locatie met de langste tijdreeksen is 10 km uit de kust bij Noordwijk. Bij de stoffen salntt t/m P, uitgezonderd Zs, wordt bij een lag van 1 maand de nulhypothese verworpen dat opeenvolgende metingen ongecorrleerd zijn. Bij de stoffen met kwartaalwaarden, Cd-nf t/m As zijn de autocorrelatie-coëfficiënten van Cd-nf en Cu het grootst en wordt de nulhypothese bij een lag van 3 maanden verworpen.

Doordat seizoensinvloeden zijn weggefilterd worden de autocorrelatie-coëfficiënten meestal kleiner. Dit geldt vooral voor de stoffen met maandwaarden, salntt t/m P, en in het bijzonder voor de stoffen NO₃NO₂ en N-nf. Bij de stoffen salntt bij een lag van 1 maand, PO₄ bij alle gegeven lags en Cu bij een lag van 3 maanden en 6 maanden wordt de nulhypothese verworpen.

Uit de berekening van de relatieve schattingsfout blijkt, dat de nulhypothese dat het model zinvol is alleen voor de stoffen salntt, NO₃NO₂, N-nf, PO₄ en P niet wordt verworpen.

De Waddenzee West

De locatie met de langste tijdreeksen is Marsdiep noord. Van de stoffen salntt t/m P wordt de nulhypothese dat opeenvolgende metingen ongecorrleerd zijn alleen bij de stoffen Salntt, N-nf, P-nf en P, niet verworpen. Bij de 6 stoffen met kwartaalwaarden zijn de autocorrelatiecoëfficiënten van Cd-nf, Cd en Cu het hoogst en wordt de nulhypothese bij een lag van 3 maanden verworpen.

Doordat seizoensinvloeden zijn weggefilterd worden de autocorrelatie-coëfficiënten meestal kleiner. Dit geldt vooral voor de eerste negen stoffen Salntt t/m P, i.h.b. voor Chlfa, NO₃NO₂ en N-nf. Bij de stoffen Zs, NH₄ en PO₄ voor een lag van 1 maand en voor Cd-nf, Cd en Cu voor een lag van 3 maanden wordt de nulhypothese niet verworpen.

Uit de berekening van de relatieve schattingsfout blijkt, dat de nulhypothese dat het model zinvol is, alleen voor NO₃NO₂ bij een lag van 1 maand en Cd-nf bij een lag van 9 maanden niet wordt verworpen.

De Eems-Dollard

De locatie met de langste tijdreeksen is Groote Gat noord. Van de stoffen salntt t/m P worden alleen bij de stoffen Zs en P de nulhypothese dat opeenvolgende metingen ongecorrleerd zijn, niet verworpen. Van de stoffen Cd-nf t/m As wordt alleen voor Cd-nf de nulhypothese verworpen.

Doordat seizoensinvloeden zijn weggefilterd worden de autocorrelatie-coëfficiënten kleiner. Dit geldt vooral voor de eerste negen stoffen Salntt t/m P en in het bijzonder voor de stoffen NO₃NO₂, N-nf en PO₄. Bij de stoffen salntt, NO₃NO₂, NH₄, PO₄ en P-nf voor een lag van 1 maand en voor Cd-nf voor een lag van 3 maanden wordt de nulhypothese verworpen.

Uit de berekening van de relatieve schattingsfout blijkt, dat de nulhypothese dat het model zinvol is, alleen bij salntt, NO₃NO₂, N-nf, NH₄, PO₄ en P-nf niet wordt verworpen.

In de verschillende gebieden wordt de nulhypothese dat de stoffen met maandwaarden, salntt t/m P, ongecorrleerd zijn met een betrouwbaarheid van

99%, meestal verworpen. Uitzonderingen zijn Zs bij Noordwijk 10 km, salntt, N-nf, P-nf en P bij Marsdiep noord en Zs en P in de Eems-Dollard. Bij de stoffen met kwartaalwaarden, Cd-nf t/m As komen soms erg hoge autocorrelaties voor, maar soms zijn de autocorrelaties ook dusdanig klein dat de nulhypothese niet verworpen wordt.

Kleine relatieve schattingsfouten komen overeen met hoge autocorrelatiecoëfficiënten. Bij de stof NO₃NO₂ wordt de nulhypothese dat het naïeve model zinvol is, nooit verworpen.

Door seizoensinvloeden te filteren worden de autocorrelaties in de regel kleiner en i.h.b. bij de stoffen met maandwaarden. Bij de Westerschelde wordt de nulhypothese niet meer verworpen voor Chlfa, N-nf en P-nf. Bij de overige gebieden geldt dat de nulhypothese nu in de regel niet meer wordt verworpen. Uitzonderingen zijn bij Noordwijk 10 km, salntt bij een lag van 1 maand, PO₄ bij alle gegeven lags en Cu bij een lag van 3 maanden en 6 maanden. Uitzonderingen bij Marsdiep noord zijn Zs, NH₄ en PO₄ voor een lag van 1 maand en Cd-nf, Cd en Cu voor een lag van 3 maanden. En uitzonderingen bij Groote Gat noord in de Eems-Dollard zijn salntt, NO₃NO₂, NH₄, PO₄ en P-nf voor een lag van 1 maand en voor Cd-nf voor een lag van 3 maanden.

De autocorrelatiecoëfficiënten van de modelfouten zijn in de regel dusdanig klein dat de nulhypothese dat de modelfouten ongecorrleerd zijn nooit verworpen wordt.

Discussie

Voor het schatten van de afhankelijkheid in de ruimte is een eenvoudig lineair model gebruikt. Het is duidelijk dat als er een 'beter' model gebruikt zou worden de verklaarde variantie en daarmee ook de afhankelijkheid in de ruimte alleen maar groter kan worden. In dit rapport is voor het eenvoudigste model gekozen, lineair en in wezen slechts met één te schatten modelparameter. Het model is robuust en gemakkelijk interpreteerbaar. Beter zou kunnen zijn een wiskundig model met meerdere relevante co-variabelen, waardoor de meetgegevens beter beschreven worden. Beter zou ook kunnen zijn een model waarbij een grotere fysische werkelijkheid ingebed is. Daarbij kan gedacht worden aan het meemodelleren van de saliniteit en seizoensinvloeden.

Voor elk gebied kan een gebiedsmodel gemaakt worden, waarin de concentraties van alle locaties geschat kan worden a.h.v. enkele locaties. M.a.w. wat zou de verklaarde variantie zijn, indien in elk gebied 2 of 3 locaties gekozen zouden worden, waarmee de concentraties van de stoffen in de overige locaties geschat zou worden.

Vergelijking met eerder onderzoek

In de rapporten [3], [4] en [5] is eerder correlatie-onderzoek uitgevoerd op locaties in een meetnet. De onderzoeken in de rijksbinnenwateren, het benedenrivierengebied in Zuid-Holland en de zoute wateren zijn echter niet direct vergelijkbaar, aangezien de onderzochte waterkwaliteitsvariabelen verschillend zijn geweest. De variabelen welke de onderzoeken gemeenschappelijk hebben zijn cadmium en chlorofyl- α . In tabel 14 en tabel 15 wordt de verklaarde variantie van deze stoffen gegeven bij enkele locaties.

Tabel 14 : Verklaarde variantie (%) van Cadmium voor enkele locaties

Locatie	Verklaarde variantie	Buur-locatie	Aantal metingen
Kampen	80.2	Lobith	70
Lith	32.6	Eysden	67
Y1 (IJsselmeer)	55.9	Y23	41
Kinderdijk	59.1	Eilbrien	35
Vuren	73.0	Lobith	141
Vlissingen	72	Borssele	25
Noordwijk 10 km	68	Noordwijk 4 km	22
Marsdiep noord	48	Vliestroom	89
Groote Gat noord	53	Zeehavenkanaal	21

Tabel 15 : Verklaarde variantie (%) van Chlorofyl- α voor enkele locaties

Locatie	Verklaarde variantie	Buur-locatie	Aantal metingen
Kampen	81.5	Lobith	130
Lith	55.9	Eysden	148
Y1 (IJsselmeer)	30.2	Y23	83
Kinderdijk	94.3	Eilbrien	118
Vuren	90.0	Lobith	269
Vlissingen	81	Borssele	73
Noordwijk 10 km	63	Noordwijk 4 km	228
Marsdiep noord	32	Vliestroom	165
Groote Gat noord	54	Zeehavenkanaal	67

De tabellen geven een indruk hoe de verklaarde variantie zich verhouden tussen de verschillende locaties. Uit onderzoek [3] bleek dat op de Rijn- en Rijntakken de locaties goed vanuit Lobith te schatten waren. In het benedenrivierengebied in Zuid-Holland waren dichtbij gelegen locaties goed met elkaar gecorreleerd. Zo niet op de Maas en op het IJsselmeer, waarbij de relatie tussen locaties ook snel afnam, naarmate de afstand toenam. Voor de zoute wateren geldt dat dichtbij elkaar gelegen locaties in het gebied Westerschelde, dicht bij de kust bij Noordwijk en in de Eems-Dollard goed met elkaar gecorreleerd zijn. De correlaties tussen locaties op de Waddenzee west zijn een stuk slechter.

6 Referenties/Literatuur

[1]

G.E.P.Box & G.M. Jenkins (1976)
Time Series Analysis - forecasting and control
Holden-Day, California

[2]

L. Ljung (1987)
System Identification- Theory for the User
Prentice Hall, Englewood Cliffs, N.J.

[3]

H.C. Klavers et al. (1992)
Optimalisatie routinematig onderzoek waterkwaliteit binnenwateren
deel 3 Statistisch Onderzoek p 38-40., RIZA

[4]

E.C.J. van der Meulen & W.A.J.M. Swart
Onderzoek van de meetlocatie Papendrecht
Rijkswaterstaat Directie Zuid-Holland, AMO-AWS

[5]

E.C.J. van der Meulen & W.A.J.M. Swart
Locatie-onderzoek van het beneden rivierengebied
Rijkswaterstaat Directie Zuid-Holland, AMO-AWS

[6]

R.N.M. Duin
Trendanalyse jaargegevens DIN, DIP en DINDIP t.b.v. vervolgnota NW3
Rijkswaterstaat Kust en Zee / RIKZ / OS 94.151x

[7]

P.K. Baggelaar
A protocol for trend analysis of the input of pollutants to the North Sea
Rijkswaterstaat Kust en Zee / RIKZ, KIWA

BIJLAGE : Uitgebreide resultaten Westerschelde

Inhoudsopgave

1	Gebruikte meetgegevens van de Westerschelde	36
	Tabel 1 : Aantal metingen per stof voor de locaties	36
	Tabel 2 : Aantal jaren dat er is gemeten per stof voor alle locaties	37
2	Resultaten met het locatiemodel voor de Westerschelde	38
	Tabel 3 : Gemiddelde verklaarde variantie (%) voor de locaties . .	38
	Tabel 4 : Verklaarde variantie (%) per stof voor de locaties	39
	Tabel 5 : F-toets voor de locaties van de Westerschelde	40
	Tabel 6 : Aantal gepaarde metingen per stof voor de locaties	41
3	Resultaten van berekening van de autocorrelatiecoëfficiënten	42
	Tabel 7 : De relatieve schattingsfout met de naïeve methode	42
	Tabel 8 : De autocorrelatiecoëfficiënt na correctie op seizoensinvloeden	43
	Tabel 9 : Autocorrelatiecoëfficiënt van de modelfouten bij het locatiemodel	44

1 Gebruikte meetgegevens van de Westerschelde

In tabel 1 is gegeven het totaal aantal metingen wat er in de periode 1971 t/m 1994 is verricht op een stof bij een locatie in de Westerschelde. In tabel 2 is gegeven het aantal jaren waarin metingen zijn verricht.

Tabel 1 : Aantal metingen per stof voor de locaties

	B a t h	L a m s w d e	Z u i d g t	H a n s w t	H o e d k k	T e r n z n	H o n t e	B o r s s l e	V l i s s g n	W i e l g n	S c h e u r w t
Salntt	76	114	78	176	80	160	82	81	201	27	26
Zs	82	120	80	218	80	198	82	81	246	29	29
Chlfa	81	170	79	261	79	242	79	76	291	28	29
NO3NO2	403	471	462	490	465	468	467	174	512	29	29
N-nf	30	30	30	63	30	31	31	32	90	28	28
NH4	427	495	487	511	489	490	490	198	535	29	29
PO4	427	456	448	496	450	452	452	199	499	29	29
P-nf	32	32	32	64	29	31	32	32	90	28	28
P	52	218	50	378	50	298	51	50	299	1	2
Cd-nf	25	84	32	173	25	135	0	25	190	9	0
Cu-nf	25	52	32	136	25	100	0	25	167	9	0
As-nf	25	43	22	41	24	42	0	25	85	9	0
Cd	25	83	23	165	25	126	0	25	185	0	0
Cu	25	51	23	143	25	92	0	25	149	9	0
As	25	43	23	69	25	70	0	25	89	0	0

Tabel 2 : Aantal jaren dat er is gemeten per stof voor alle locaties

	B a t h	L a m s w d e	Z u i d g t	H a n s w t	H o e d k k k	T e r n z n	H o n t e	B o r s s l e	V l i s s g n	W i e l g n	S c h e u r w t
Salntt	7	10	7	13	7	13	7	7	13	4	4
Zs	7	10	7	13	7	13	7	7	13	4	4
Chlfa	7	13	7	16	7	16	7	7	16	4	4
NO3NO2	21	23	23	23	23	23	23	12	23	4	4
N-nf	3	3	3	4	3	3	3	3	4	4	4
NH4	22	24	24	24	24	24	24	13	24	4	4
PO4	22	23	23	24	23	23	23	13	23	4	4
P-nf	3	3	3	4	3	3	3	3	4	4	4
P	5	12	5	19	4	15	4	4	15	1	2
Cd-nf	7	15	7	24	7	19	0	7	23	3	0
Cu-nf	7	11	7	19	7	15	0	7	20	3	0
As-nf	7	10	6	10	7	10	0	7	12	3	0
Cd	7	15	6	23	7	18	0	7	23	0	0
Cu	7	11	6	20	7	14	0	7	19	3	0
As	7	10	6	12	7	12	0	7	12	0	0

2 Resultaten met het locatiemodel voor de Westerschelde

In tabel 6 van de hoofdtekst is de maximale gemiddelde verklaarde variantie gegeven als concentraties van stoffen worden geschat door een andere locatie. Bij de Westerschelde is dat steeds een in de buurt gelegen locatie. In tabel 3 van deze bijlage wordt de gemiddelde verklaarde variantie gegeven van de locatie, indien de stoffen geschat worden uit de andere buur-locaties.

Tabel 3 : Gemiddelde verklaarde variantie (%) voor de locaties

	Buur-locatie										
	B a t h	L a m s w d e	Z u i d g t	H a n s w t	H o e d k k k	T e r n z n	H o n t e	B o r s s l e	V l i s s g n	W i e l g n	S c h e u r w t
Bath		72	65	59	46	39	26	32	34	32	19
Lamswde	73		75	72	57	54	29	47	38	32	21
Zuidgt	65	75		77	65	63	33	48	44	42	19
Hanswt	59	72	77		69	62	36	56	40	34	20
Hoedkkk	47	56	64	69		68	36	55	51	38	20
Ternzn	38	54	62	61	67		45	73	53	44	19
Honte	27	29	34	36	37	46		49	42	27	21
Borssle	33	45	47	55	55	73	49		65	43	25
Vlissgn	36	36	43	40	50	50	39	64		57	23
Wielgn	30	34	40	32	37	40	26	41	61		24
Scheurwt	20	23	20	20	21	22	22	26	25	23	

De locaties in bovenstaande tabel zijn gegeven in de volgorde van Antwerpen tot aan de Noordzee.

In tabel 4 is per stof de verklaarde variantie gegeven bij elke locatie, waarbij de locatie van tabel 3, welke de maximale verklaarde variantie geeft, de stoffen schat. Is de verklaarde variantie 0%, dan zijn er niet genoeg gepaarde metingen geweest voor het identificeren van het locatiemodel.

Tabel 4 : Verklaarde variantie (%) per stof voor de locaties

	B a t h	L a m s w d e	Z u i d g t	H a n s w t	H o e d k k k	T e r n z n	H o n t e	B o r s s l e	V l i s s g n	W i e l g n	S c h e u r w t
Salntt	96	99	99	98	96	84	92	83	73	83	45
Zs	55	59	32	32	43	63	51	64	30	43	17
Chlfa	74	70	84	84	67	61	89	61	81	82	64
NO3NO2	93	91	90	89	82	91	93	91	89	84	52
N-nf	98	97	97	97	74	95	94	95	84	86	68
NH4	97	93	97	97	89	36	75	36	61	79	42
PO4	92	94	92	92	80	86	88	86	78	79	51
P-nf	90	90	87	88	87	76	82	76	53	80	58
P	64	65	58	59	66	73	75	74	51	0	0
Cd-nf	28	40	99	99	63	80	0	81	76	94	0
Cu-nf	32	42	51	52	80	66	0	66	49	70	0
As-nf	57	74	56	55	48	73	0	76	67	82	0
Cd	68	74	68	68	64	79	0	80	72	0	0
Cu	77	81	76	75	59	61	0	63	48	58	0
As	54	62	71	67	32	64	0	67	42	0	0

Stoffen met een hoge gemiddelde verklaarde varianties zijn Salntt, Chlfa, NO3NO2, N-nf, NH4 (vooral landinwaarts), PO4 en P-nf. De verklaarde variantie van de stof Cd-nf is zeer hoog (99%) bij de locaties Zuidergat en Hansweert!

De statistische F-toets (zie bijlage : 'Theoretische achtergronden') geeft de kans dat het informatieverlies kleiner is dan 100% op toeval berust. Is de kans kleiner dan 0.01 dan wordt de nulhypothese dat het gebruik van het locatiemodel significant zinvol is met een betrouwbaarheid van 99%, niet verworpen. In de lege cellen van de tabel zijn er niet voldoende metingen voor het identificeren van het locatiemodel.

In tabel 5 is per stof de F-toets toegepast bij elke locatie, waarbij de locatie van tabel 3, welke de maximale verklaarde variantie geeft, de stoffen schat.

Tabel 5 : F-toets voor de locaties van de Westerschelde

	B a t h	L a m s w d e	Z u i d g t	H a n s w t	H o e d k k k	T e r n z n	H o n t e	B o r s s l e	V l i s s g n	W i e l g n	S c h e u r w t
Salntt	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.01
Zs	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.20
Chlfa	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
NO3NO2	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
N-nf	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
NH4	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.01
PO4	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
P-nf	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
P	.00	.00	.00	.00	.00	.00	.00	.00	.00		
Cd-nf	.01	.00	.00	.00	.00	.00		.00	.00	.00	
Cu-nf	.01	.00	.00	.00	.00	.00		.00	.00	.03	
As-nf	.00	.00	.00	.00	.00	.00		.00	.00	.01	
Cd	.00	.00	.00	.00	.00	.00		.00	.00		
Cu	.00	.00	.00	.00	.00	.00		.00	.00	.07	
As	.00	.00	.00	.00	.01	.00		.00	.00		

Alleen bij de locatie Hoedekenskerke bij de stof As, bij de locatie Wielingen bij de stoffen Cu-nf, As-nf en Cu en bij de locatie Scheur west bij de stoffen salntt, Zs en NH4 wordt de nulhypothese dat het locatiemodel zinvol is, verworpen.

In tabel 6 zijn de aantallen gepaarde metingen gegeven op basis waarvan de verklaarde variantie van tabel 4 is bepaald.

Tabel 6 : Aantal gepaarde metingen per stof voor de locaties

	B a t h	L a m s w d e	Z u i d g t	H a n s w t	H o e d k k	T e r n z n	H o n t e	B o r s s l e	V l i s s g n	W i e l g n	S c h e u r w t
Salntt	76	75	78	79	75	82	78	81	94	26	15
Zs	79	78	78	78	76	79	73	79	90	25	16
Chlfa	74	66	68	68	60	69	69	69	73	25	15
NO3NO2	400	461	460	462	461	174	173	174	187	29	17
N-nf	28	28	28	28	29	31	31	31	45	26	16
NH4	415	471	450	452	438	158	167	158	172	28	17
PO4	421	443	438	440	437	195	195	195	209	29	16
P-nf	32	32	31	31	29	31	32	31	43	25	16
P	49	49	50	50	49	49	48	49	49	0	0
Cd-nf	25	23	26	26	25	25	0	25	25	7	0
Cu-nf	24	22	23	23	25	25	0	25	24	7	0
As-nf	25	22	21	21	23	24	0	24	24	7	0
Cd	21	21	22	22	25	25	0	25	25	0	0
Cu	23	21	23	23	25	25	0	25	24	7	0
As	24	22	22	22	25	25	0	25	24	0	0

Bij de locaties Wielingen en vooral Scheur west zijn de aantallen gepaarde metingen het kleinst.

3 Resultaten van berekening van de autocorrelatiecoëfficiënten

Gegeven zijn de relatieve schattingsfout met een autoregressief model, de autocorrelatiecoëfficiënt na correctie voor seizoensinvloeden en de autocorrelatiecoëfficiënt van de modelfouten van het locatiemodel. Voor de stoffen salntt t/m P (zie tabel 1 van de hoofdtekst) betreffen de resultaten maandwaarden en bij de stoffen Cd-nf t/m As kwartaalwaarden. Gegeven zijn de resultaten bij een lag van 1, 2, 3 en 4 maanden bij de stoffen salntt t/m P en bij een lag van 3, 6, 9 en 12 maanden bij de stoffen Cd-nf t/m As. De meest bemeten locatie is Vlissingen.

In tabel 7 is de relatieve schattingsfout bepaald bij toepassing van de naïeve schattingsmethode. Bij de naïeve is de schatting van de concentratie gelijk aan de voorgaande meetwaarde. Getoetst is met een betrouwbaarheid van 99% of het naïeve model zinvol is, of de relatieve schattingsfout significant kleiner is dan 100%. Heeft toets de waarde 1, dan wordt de nulhypothese niet verworpen.

Tabel 7 : De relatieve schattingsfout met de naïeve methode

Stof	1 mnd	toets	2 mnd	toets	3 mnd	toets	4 mnd	toets
Salntt	77	0	122	1	182	1	236	1
Zs	108	1	161	1	202	1	241	1
Chlfa	118	1	180	1	198	1	239	1
NO3NO2	61	0	122	1	193	1	255	1
N-nf	82	0	124	1	204	1	251	1
NH4	134	1	173	1	214	1	228	1
PO4	117	1	151	1	186	1	230	1
P-nf	93	0	151	1	165	1	203	1
P	86	0	142	1	195	1	232	1
Stof	3 mnd	toets	6 mnd	toets	9 mnd	toets	jaar	toets
Cd-nf	229	1	164	1	209	1	190	1
Cu-nf	188	1	217	1	177	1	198	1
As-nf	172	1	289	1	176	1	71	0
Cd	156	1	209	1	202	1	194	1
Cu	182	1	254	1	229	1	176	1
As	224	1	278	1	242	1	100	1

Alleen bij de stoffen salntt, NO3NO2, N-nf, P-nf en P wordt de nulhypothese dat het naïeve model zinvol is (bij een lag van 1 maand) met een 99%

betrouwbaarheid, niet verworpen. Door de periodiciteit van een jaar is de relatieve schattingsfout van As-nf bij een lag van 12 maanden ook kleiner dan 100%; de nulhypothese wordt ook hier niet verworpen. Kleine relatieve schattingsfouten komen overeen met hoge autocorrelatiecoëfficiënten van tabel 10 in de hoofdtekst.

In tabel 8 zijn de autocorrelatiecoëfficiënten gegeven na correctie op seizoensinvloeden. Door de stoffen te centreren rond maand- en kwartaal gemiddelde waarden worden seizoensinvloeden genivelleerd. Met een betrouwbaarheid van 99% wordt getoetst of opeenvolgende concentraties in de tijd niet gecorreleerd zijn. Heeft toets de waarde 0, dan wordt de nulhypothese niet verworpen. Heeft toets de waarde 1 dan wordt de alternatieve hypothese, opeenvolgende concentraties zijn gecorreleerd, geaccepteerd. Hoe hoger de autocorrelatiecoëfficiënt en hoe kleiner het aantal waarnemingen des te groter de kans dat de nulhypothese wordt verworpen.

Tabel 8 : De autocorrelatiecoëfficiënt na correctie op seizoensinvloeden

Stof	1 mnd	toets	2 mnd	toets	3 mnd	toets	4 mnd	toets
Salntt	0.52	1	0.49	1	0.37	1	0.27	1
Zs	0.28	1	0.09	0	-0.06	0	0.02	0
Chlfa	0.16	0	0.07	0	0.13	0	0.03	0
NO3NO2	0.54	1	0.32	1	0.18	1	0.14	0
N-nf	0.12	0	0.04	0	-0.19	0	-0.01	0
NH4	0.49	1	0.42	1	0.34	1	0.35	1
PO4	0.39	1	0.32	1	0.34	1	0.26	1
P-nf	0.15	0	-0.09	0	0.26	0	0.23	0
P	0.31	1	0.14	0	0.12	0	0.22	1
Stof	3 mnd	toets	6 mnd	toets	9 mnd	toets	jaar	toets
Cd-nf	0.09	0	0.10	0	0.22	1	0.13	0
Cu-nf	0.31	1	0.22	0	0.35	1	0.18	0
As-nf	0.19	0	-0.18	0	0.20	0	0.47	1
Cd	0.59	1	0.44	1	0.43	1	0.41	1
Cu	0.35	1	0.26	0	0.23	0	0.26	0
As	0.24	0	-0.05	0	0.08	0	0.09	0

Doordat seizoensinvloeden zijn weggefilterd zijn de autocorrelatiecoëfficiënten in tabel 8 meestal lager dan in tabel 10 van de hoofdtekst. Dit geldt vooral voor de eerste negen stoffen Salntt t/m P en in het bijzonder voor de stoffen N-nf en P-nf. De autocorrelatiecoëfficiënten van de stoffen Salntt, NO3NO2, PO4 en P

zijn het grootst voor lag 1 in beide tabellen. Bij de stoffen Chlfa, N-nf en P-nf voor een lag van 1 maand en voor Cd-nf, As-nf en As voor een lag van 3 maanden wordt de nulhypothese niet verworpen.

In tabel 9 zijn de autocorrelatiecoëfficiënten gegeven van de modelfouten van het locatiemodel van Vlissingen met Borssele. Met een betrouwbaarheid van 99% wordt getoetst of de modelfouten gecorreleerd zijn. Heeft toets de waarde 0, dan wordt de nulhypothese niet verworpen.

Tabel 9 : Autocorrelatiecoëfficiënt van de modelfouten bij het locatiemodel

Stof	1 mnd	toets	2 mnd	toets	3 mnd	toets	4 mnd	toets
Salntt	0.26	0	0.19	0	0.11	0	-0.07	0
Zs	0.21	0	0.09	0	-0.08	0	-0.02	0
Chlfa	-0.07	0	0.08	0	0.02	0	0.07	0
NO3NO2	0.04	0	0.06	0	-0.04	0	0.14	0
N-nf	-0.23	0	0.03	0	-0.20	0	-0.01	0
NH4	-0.04	0	0.05	0	-0.05	0	-0.04	0
PO4	0.01	0	0.05	0	0.09	0	0.06	0
P-nf	-0.13	0	0.25	0	-0.11	0	-0.04	0
P	0.26	0	0.27	0	0.01	0	0.01	0
Stof	3 mnd	toets	6 mnd	toets	9 mnd	toets	jaar	toets
Cd-nf	0.00	0	0.00	0	-0.13	0	0.00	0
Cu-nf	-0.16	0	-0.31	0	0.34	0	0.19	0
As-nf	-0.08	0	-0.23	0	-0.10	0	0.25	0
Cd	-0.07	0	-0.08	0	0.08	0	-0.01	0
Cu	-0.19	0	-0.02	0	0.07	0	-0.11	0
As	-0.11	0	-0.08	0	0.00	0	0.06	0

Bij alle stoffen en voor alle gegeven lags is de autocorrelatiecoëfficiënt van de modelfouten dusdanig klein dat de nulhypothese nooit verworpen wordt.

BIJLAGE : Uitgebreide resultaten Noordzee bij Noordwijk

Inhoudsopgave

1	Gebruikte meetgegevens bij de Noordzee bij Noordwijk	46
	Tabel 1 : Aantal metingen per stof voor de locaties	46
	Tabel 2 : Aantal jaren dat er is gemeten per stof voor de locaties	47
2	Analyse van de resultaten met een locatiemodel voor Noordwijk	48
	Tabel 3 : Gemiddelde verklaarde variantie (%) voor de locaties . .	48
	Tabel 4 : Verklaarde variantie (%) per stof voor de locaties	49
	Tabel 5 : F-toets voor de locaties bij Noordwijk	50
	Tabel 6 : Aantal gepaarde metingen per stof voor de locaties	51
3	Resultaten van berekening autocorrelatie	52
	Tabel 7 : De relatieve schattingsfout met de naïeve methode	52
	Tabel 8 : De autocorrelatiecoëfficiënt na correctie op seizoensinvloeden	53
	Tabel 9 : Autocorrelatiecoëfficiënt van de modelfouten bij het locatiemodel	54

1 Gebruikte meetgegevens bij de Noordzee bij Noordwijk

In tabel 1 is gegeven het totaal aantal metingen wat er in de periode 1971 t/m 1994 is verricht op een stof bij de meetlocaties van Noordwijk. In tabel 2 is gegeven het aantal jaren waarin metingen zijn verricht.

Tabel 1 : Aantal metingen per stof voor de locaties

	Nw01	Nw02	Nw04	Nw10	Nw20	Nw30	Nw50	Nw70
Salntt	229	348	241	429	338	241	240	255
Zs	233	362	243	444	344	239	238	250
Chlfa	226	266	238	393	251	238	237	251
NO3NO2	212	337	222	340	323	220	218	233
N-nf	30	71	30	131	56	30	30	55
NH4	234	362	245	448	347	244	244	260
PO4	232	360	242	447	345	240	241	257
P-nf	30	71	30	130	56	30	30	53
P	196	197	206	311	206	198	206	194
Cd-nf	22	128	22	60	89	22	22	88
Cu-nf	22	125	22	61	86	22	22	89
As-nf	21	96	22	59	50	22	21	60
Cd	22	103	22	61	62	22	22	88
Cu	22	102	22	61	60	22	22	87
As	22	98	22	61	56	22	22	60

Tabel 2 : Aantal jaren dat er is gemeten per stof voor de locaties

	Nw01	Nw02	Nw04	Nw10	Nw20	Nw30	Nw50	Nw70
Salntt	16	20	16	16	20	16	16	16
Zs	16	20	16	16	20	16	16	16
Chlfa	16	16	16	16	16	16	16	16
NO3NO2	15	19	15	15	19	15	15	15
N-nf	3	4	3	4	4	3	3	4
NH4	16	20	16	16	20	16	16	16
PO4	16	20	16	16	20	16	16	16
P-nf	3	4	3	4	4	3	3	4
P	13	13	13	13	13	13	13	13
Cd-nf	6	15	6	7	14	6	6	10
Cu-nf	6	15	6	7	14	6	6	10
As-nf	6	12	6	7	11	6	6	7
Cd	6	12	6	7	12	6	6	10
Cu	6	12	6	7	12	6	6	10
As	6	12	6	7	11	6	6	7

2 Analyse van de resultaten met een locatiemodel voor Noordwijk

In tabel 7 van de hoofdtekst is de maximale gemiddelde verklaarde variantie gegeven als concentraties van stoffen worden geschat door een andere locatie. Bij Noordwijk is dat steeds een in de buurt gelegen locatie. In tabel 3 van deze bijlage wordt de gemiddelde verklaarde variantie gegeven van de locatie, indien de stoffen geschat worden uit de andere buur-locaties.

Tabel 3 : Gemiddelde verklaarde variantie (%) voor de locaties

Buur-locatie								
	Nw01	Nw02	Nw04	Nw10	Nw20	Nw30	Nw50	Nw70
Nw01		77	66	53	39	26	20	22
Nw02	77		76	57	37	30	23	20
Nw04	67	76		61	44	32	23	28
Nw10	52	55	60		53	33	23	24
Nw20	39	40	43	55		55	30	30
Nw30	26	30	32	35	56		43	36
Nw50	20	23	22	24	31	43		47
Nw70	22	21	26	26	32	35	47	

De locatie Nw01 staat voor de meetlocatie bij Noordwijk 1 km van de kustlijn.
De locatie Nw70 staat voor de meetlocatie bij Noordwijk 70 km van de kustlijn.

In tabel 4 is per stof de verklaarde variantie gegeven bij elke locatie, waarbij de locatie van tabel 3 welke de maximale verklaarde variantie, geeft de stoffen schat.

Tabel 4 : Verklaarde variantie (%) per stof voor de locaties

	Nw01	Nw02	Nw04	Nw10	Nw20	Nw30	Nw50	Nw70
Salntt	95	95	93	50	44	45	63	63
Zs	30	30	38	24	21	22	25	25
Chlfa	71	71	84	63	40	40	30	31
NO3NO2	98	98	94	78	72	72	80	80
N-nf	98	98	96	85	77	77	75	75
NH4	90	90	88	55	27	27	18	18
PO4	94	94	87	71	61	61	74	74
P-nf	81	76	78	80	72	72	52	53
P	49	49	60	54	40	38	47	47
Cd-nf	91	91	82	71	61	60	71	69
Cu-nf	85	85	77	54	46	56	56	56
As-nf	82	82	59	52	50	52	11	10
Cd	78	79	87	68	74	73	39	38
Cu	61	61	65	38	84	99	22	17
As	55	56	58	48	53	52	45	45

Voor bijna alle stoffen is de verklaarde variantie dicht bij de kust het grootst. In het bijzonder geldt dit voor de stof NH₄, bij de eerste twee locaties is de verklaarde variantie 90% en bij de laatste twee locaties, 50 en 70 km uit de kust, slechts 18%. Stoffen met hoge gemiddelde variantie zijn Salntt, NO₃NO₂, N-nf, PO₄, P-nf, Cd-nf en Cu-nf.

De statistische F-toets (zie bijlage : 'Theoretische achtergronden') geeft de kans dat het informatieverlies kleiner is dan 100% op toeval berust. Is de kans kleiner dan 0.01 dan wordt de nulhypothese dat het gebruik van het locatiemodel significant zinvol is met een betrouwbaarheid van 99%, niet verworpen.

Tabel 5 : F-toets voor de locaties bij Noordwijk

F-test	Nw01	Nw02	Nw04	Nw10	Nw20	Nw30	Nw50	Nw70
Salntt	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zs	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Chlfa	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
NO3NO2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
N-nf	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
NH4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
PO4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
P-nf	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
P	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cd-nf	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cu-nf	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
As-nf	0.00	0.00	0.00	0.00	0.00	0.00	0.29	0.34
Cd	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cu	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.11
As	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Alleen bij de locaties 50 km en 70 km uit de kust van Noordwijk en bij de stoffen As-nf en Cu wordt de nulhypothese dat het locatiemodel zinvol is, verworpen.

In tabel 6 zijn de aantallen gepaarde metingen gegeven op basis waarvan de verklaarde variantie van tabel 4 is bepaald.

Tabel 6 : Aantal gepaarde metingen per stof voor de locaties

	Nw01	Nw02	Nw04	Nw10	Nw20	Nw30	Nw50	Nw70
Salntt	221	221	230	253	227	227	228	228
Zs	208	209	212	229	209	209	208	208
Chlfa	191	192	204	228	189	189	203	203
NO3NO2	207	208	214	228	206	206	213	213
N-nf	29	30	29	34	29	29	30	30
NH4	225	226	237	253	219	219	229	229
PO4	228	229	237	263	233	233	234	234
P-nf	29	30	29	35	30	30	29	29
P	184	184	191	214	191	191	186	186
Cd-nf	22	22	22	22	22	21	21	21
Cu-nf	22	22	22	22	22	21	19	19
As-nf	21	21	22	22	22	21	20	20
Cd	22	22	22	22	22	21	21	21
Cu	22	22	22	22	22	21	21	21
As	22	22	22	21	20	19	21	21

3 Resultaten van berekening autocorrelatie

Gegeven zijn de relatieve schattingsfout met een autoregressief model, de autocorrelatiecoëfficiënt na correctie voor seizoensinvloeden en de autocorrelatiecoëfficiënt van de modelfouten van het locatiemodel. Voor de stoffen salntt t/m P (zie tabel 1 van de hoofdtekst) betreffen de resultaten maandwaarden en bij de stoffen Cd-nf t/m As kwartaalwaarden. Gegeven zijn de resultaten bij een lag van 1, 2, 3 en 4 maanden bij de stoffen salntt t/m P en bij een lag van 3, 6, 9 en 12 maanden bij de stoffen Cd-nf t/m As. De meest bemeten locatie is 10 km uit de kust bij Noordwijk.

In tabel 7 is de relatieve schattingsfout bepaald bij toepassing van de naïeve schattingsmethode. Bij de naïeve is de schatting van de concentratie gelijk aan de voorgaande meetwaarde. Getoetst is met een betrouwbaarheid van 99% of het naïeve model zinvol is, of de relatieve schattingsfout significant kleiner is dan 100%. Heeft toets de waarde 1, dan wordt de nulhypothese niet verworpen.

Tabel 7 : De relatieve schattingsfout met de naïeve methode

Stof	1 mnd	toets	2 mnd	toets	3 mnd	toets	4 mnd	toets
Salntt	154	1	158	1	166	1	164	1
Zs	184	1	155	1	175	1	215	1
Chlfa	116	1	153	1	182	1	200	1
NO3NO2	74	0	138	1	188	1	234	1
N-nf	67	0	134	1	195	1	250	1
NH4	133	1	170	1	205	1	221	1
PO4	96	0	150	1	192	1	218	1
P-nf	95	0	151	1	184	1	221	1
P	154	1	145	1	189	1	171	1
Stof	3 mnd	toets	6 mnd	toets	9 mnd	toets	jaar	toets
Cd-nf	101	1	151	1	192	1	293	1
Cu-nf	221	1	204	1	262	1	136	1
As-nf	232	1	245	1	208	1	156	1
Cd	169	1	267	1	161	1	31	0
Cu	662	1	612	1	662	1	712	1
As	195	1	240	1	243	1	214	1

Bij de stoffen salntt, NO3NO2, N-nf, PO4 en P wordt de nulhypothese dat het naïeve model zinvol is (bij een lag van 1 maand) met een 99% betrouwbaarheid, niet verworpen. Door de periodiciteit van een jaar is de relatieve schattingsfout

van Cd ruim kleiner dan 100%; de nulhypothese wordt dus ook hier niet verworpen.

In tabel 8 zijn de autocorrelatiecoëfficiënten gegeven na correctie op seizoensinvloeden. Door de stoffen te centreren rond maand- en kwartaal gemiddelde waarden worden seizoensinvloeden genivelleerd. Met een betrouwbaarheid van 99% wordt getoetst of opeenvolgende concentraties in de tijd ongecorreleerd zijn. Heeft toets de waarde 0, dan wordt de nulhypothese niet verworpen. Heeft toets de waarde 1 dan wordt de alternatieve hypothese, opeenvolgende concentraties zijn gecorreleerd, geaccepteerd. Hoe hoger de autocorrelatiecoëfficiënt en hoe kleiner het aantal waarnemingen des te groter de kans dat de nulhypothese wordt verworpen.

Tabel 8 : De autocorrelatiecoëfficiënt na correctie op seizoensinvloeden

Stof	1 mnd	toets	2 mnd	toets	3 mnd	toets	4 mnd	toets
Salntt	0.24	1	0.17	0	0.16	0	0.18	0
Zs	0.10	0	0.05	0	0.07	0	-0.09	0
Chlfa	-0.01	0	-0.03	0	-0.04	0	-0.12	0
NO3NO2	0.10	0	-0.14	0	0.00	0	0.11	0
N-nf	0.07	0	-0.28	0	-0.19	0	-0.09	0
NH4	0.17	0	0.11	0	0.00	0	0.01	0
PO4	0.44	1	0.36	1	0.34	1	0.40	1
P-nf	0.39	0	0.29	0	0.28	0	0.25	0
P	0.17	0	0.22	0	0.14	0	0.23	1
Stof	3 mnd	toets	6 mnd	toets	9 mnd	toets	jaar	toets
Cd-nf	0.02	0	0.03	0	0.16	0	-0.07	0
Cu-nf	0.13	0	-0.16	0	-0.20	0	-0.02	0
As-nf	0.11	0	-0.05	0	-0.25	0	-0.34	0
Cd	0.01	0	-0.39	0	-0.22	0	0.23	0
Cu	0.62	1	0.60	1	0.51	0	0.31	0
As	0.32	0	0.21	0	-0.11	0	-0.22	0

Doordat seizoensinvloeden zijn weggefilterd zijn de autocorrelatiecoëfficiënten in tabel 8 meestal lager dan in tabel 11 van de hoofdtekst. Dit geldt vooral voor de eerste negen stoffen Salntt t/m P en in het bijzonder voor de stoffen NO3NO2 en N-nf. De autocorrelatiecoëfficiënten van de stof Cd is in beide tabellen voor alle lags hoog. Bij de stoffen salntt bij een lag van 1 maand, PO4 bij alle gegeven lags en Cu bij een lag van 3 en 6 maanden wordt de nulhypothese verworpen.

In tabel 9 zijn de autocorrelatiecoëfficiënten gegeven van de modelfouten van het locatiemodel van Noordwijk 10 km met Noordwijk 4 km. Met een betrouwbaarheid van 99% wordt getoetst of de modelfouten gecorreleerd zijn. Heeft toets de waarde 0, dan wordt de nulhypothese niet verworpen.

Tabel 9 : Autocorrelatiecoëfficiënt van de modelfouten bij het locatiemodel

Stof	1 mnd	toets	2 mnd	toets	3 mnd	toets	4 mnd	toets
Salntt	0.10	0	0.07	0	0.06	0	-0.01	0
Zs	0.01	0	0.00	0	-0.04	0	0.05	0
Chlfa	0.05	0	0.00	0	0.08	0	0.03	0
NO3NO2	-0.09	0	0.00	0	-0.01	0	-0.02	0
N-nf	-0.16	0	-0.21	0	0.00	0	-0.17	0
NH4	0.23	1	0.10	0	0.04	0	-0.15	0
PO4	0.27	1	0.15	0	0.07	0	-0.08	0
P-nf	0.01	0	0.01	0	-0.24	0	-0.13	0
P	0.06	0	-0.04	0	-0.05	0	0.01	0
Stof	3 mnd	toets	6 mnd	toets	9 mnd	toets	jaar	toets
Cu-nf	-0.01	0	0.03	0	-0.22	0	-0.38	0
As-nf	0.03	0	0.02	0	0.32	0	0.19	0
Cd	-0.13	0	0.00	0	-0.28	0	-0.40	0
Cu	-0.15	0	0.12	0	0.16	0	-0.08	0
As	0.09	0	-0.14	0	0.15	0	0.10	0
As	-0.11	0	-0.08	0	0.00	0	0.06	0

Bij de stoffen NH4 en PO4 bij een lag van 1 maand wordt de nulhypothese verworpen.

BIJLAGE : Uitgebreide resultaten Waddenzee west

Inhoudsopgave

1	Gebruikte meetgegevens voor Waddenzee west	56
	Tabel 1 : Aantal metingen per stof voor de locaties	56
	Tabel 2 : Aantal jaren dat er is gemeten per stof voor de locaties	57
2	Resultaten met een locatiemodel voor de Waddenzee west	58
	Tabel 3 : Gemiddelde verklaarde variantie (%) voor de locaties . .	58
	Tabel 4 : Verklaarde variantie (%) per stof voor de locaties	59
	Tabel 5 : F-toets voor de locaties op de Waddenzee West	60
	Tabel 6 : Aantal gepaarde metingen per stof voor de locaties	61
3	Resultaten van berekening autocorrelatie	62
	Tabel 7 : De relatieve schattingsfout met de naïeve methode	62
	Tabel 8 : De autocorrelatiecoëfficiënt na correctie op seizoensinvloeden	63
	Tabel 9 : Autocorrelatiecoëfficiënt van de modelfouten bij het locatiemodel	64

1 Gebruikte meetgegevens voor Waddenzee west

In tabel 1 is gegeven het totaal aantal metingen wat er in de periode 1971 t/m 1994 is verricht op een stof bij de meetlocaties in de Waddenzee west. In tabel 2 is gegeven het aantal jaren waarin metingen zijn verricht.

Tabel 1 : Aantal metingen per stof voor de locaties

	Marsdnd	Doovbot	Doovbwt	Vliesm	Blauwsot	Harlgn
Salntt	214	196	197	195	197	226
Zs	268	242	232	239	239	310
Chlfa	218	166	169	195	130	231
NO3NO2	265	226	212	216	216	293
N-nf	53	28	27	28	28	26
NH4	290	250	234	242	244	319
PO4	244	227	221	218	219	274
P-nf	55	27	28	27	28	27
P	242	215	208	215	214	289
Cd-nf	146	0	52	132	0	0
Cu-nf	146	0	52	132	0	0
As-nf	56	0	35	52	0	0
Cd	135	0	50	121	0	0
Cu	136	0	51	120	0	0
As	58	0	40	56	0	0

Tabel 2 : Aantal jaren dat er is gemeten per stof voor de locaties

	Marsdnd	Doovbot	Doovbwt	Vliesm	Blauwsot	Harlgn
Salntt	18	18	18	18	18	18
Zs	22	22	22	22	22	22
Chlfa	19	16	16	19	13	19
NO3NO2	23	21	21	21	21	21
N-nf	4	3	3	3	3	3
NH4	24	22	22	22	22	22
PO4	20	20	20	20	20	20
P-nf	4	3	3	3	3	3
P	21	19	19	19	19	19
Cd-nf	23	0	12	23	0	0
Cu-nf	23	0	12	23	0	0
As-nf	11	0	10	11	0	0
Cd	22	0	12	23	0	0
Cu	22	0	12	22	0	0
As	11	0	10	11	0	0

2 Resultaten met een locatiemodel voor de Waddenzee west

In tabel 8 van de hoofdtekst is de maximale gemiddelde verklaarde variantie gegeven als concentraties van stoffen worden geschat door een andere locatie. In tabel 3 van deze bijlage wordt de gemiddelde verklaarde variantie gegeven van de locatie, indien de stoffen geschat worden uit de andere buurlocaties.

Tabel 3 : Gemiddelde verklaarde variantie (%) voor de locaties

Buur-locatie						
	Marsdnd	Doovbot	Doovbwt	Vliesm	Blauwsot	Harlgn
Marsdnd		32	37	43	20	16
Doovbot	31		35	28	29	26
Doovbwt	37	35		38	31	29
Vliesm	43	29	39		30	22
Blauwsot	20	29	30	29		32
Harlgn	15	26	28	21	31	

In tabel 4 is per stof de verklaarde variantie gegeven bij elke locatie, waarbij de locatie van tabel 3, welke de maximale verklaarde variantie geeft, de stoffen schat. Is de verklaarde variantie 0%, dan zijn er niet genoeg gepaarde metingen geweest voor het identificeren van het locatiemodel.

Tabel 4 : Verklaarde variantie (%) per stof voor de locaties

	Marsdnd	Doovbot	Doovbwt	Vliesm	Blauwsot	Harlgn
Salntt	12	54	30	12	68	69
Zs	20	43	30	20	14	16
Chlfa	32	16	9	32	43	43
NO3NO2	58	80	51	58	78	71
N-nf	72	76	61	70	89	89
NH4	39	69	37	36	56	52
PO4	64	75	47	64	54	48
P-nf	39	73	20	51	65	65
P	11	39	15	11	9	10
Cd-nf	72	0	33	72	0	0
Cu-nf	34	0	54	36	0	0
As-nf	75	0	28	70	0	0
Cd	48	0	63	51	0	0
Cu	21	0	68	22	0	0
As	44	0	26	45	0	0

Stoffen met een hoge gemiddelde verklaarde variantie zijn NO3NO2 en N-nf.

De statistische F-toets (zie bijlage : "Theoretische achtergronden") geeft de kans dat het informatieverlies kleiner is dan 100% op toeval berust. Is de kans kleiner dan 0.01 dan wordt de nulhypothese dat het gebruik van het locatiemodel significant zinvol is met een betrouwbaarheid van 99%, niet verworpen. In de lege cellen van de tabel zijn er niet voldoende metingen voor het identificeren van het locatiemodel.

In tabel 5 is per stof de F-toets toegepast bij elke locatie, waarbij de locatie van tabel 3, welke de maximale verklaarde variantie geeft, de stoffen schat.

Tabel 5 : F-toets voor de locaties op de Waddenzee West

F-test	Marsdnd	Doovbot	Doovbwt	Vliesm	Blauwsot	Harlgn
Salntt	0.00	0.00	0.00	0.00	0.00	0.00
Zs	0.00	0.00	0.00	0.00	0.00	0.00
Chlfa	0.00	0.00	0.00	0.00	0.00	0.00
NO3NO2	0.00	0.00	0.00	0.00	0.00	0.00
N-nf	0.00	0.00	0.00	0.00	0.00	0.00
NH4	0.00	0.00	0.00	0.00	0.00	0.00
PO4	0.00	0.00	0.00	0.00	0.00	0.00
P-nf	0.00	0.00	0.04	0.00	0.00	0.00
P	0.00	0.00	0.00	0.00	0.00	0.00
Cd-nf	0.00		0.00	0.00		
Cu-nf	0.00		0.00	0.00		
As-nf	0.00		0.01	0.00		
Cd	0.00		0.00	0.00		
Cu	0.00		0.00	0.00		
As	0.00		0.01	0.00		

Alleen bij de locatie Doove Balg west bij de stoffen P-nf, As-nf en As wordt de nulhypothese dat het locatiemodel zinvol is, verworpen.

In tabel 6 zijn de aantallen gepaarde metingen gegeven op basis waarvan de verklaarde variantie van tabel 4 is bepaald.

Tabel 6 : Aantal gepaarde metingen per stof voor de locaties

	Marsdnd	Doovbot	Doovbwt	Vliesm	Blauwsot	Harlgn
Salntt	172	193	186	172	197	226
Zs	189	195	194	187	221	261
Chlfa	165	139	138	165	116	116
NO3NO2	173	186	176	172	208	254
N-nf	24	25	26	24	26	26
NH4	195	203	199	194	232	274
PO4	197	211	201	196	204	244
P-nf	26	27	26	26	24	24
P	168	168	174	166	195	246
Cd-nf	119	0	43	117	0	0
Cu-nf	106	0	48	105	0	0
As-nf	42	0	30	42	0	0
Cd	89	0	44	86	0	0
Cu	94	0	48	93	0	0
As	38	0	32	38	0	0

3 Resultaten van berekening autocorrelatie

Gegeven zijn de relatieve schattingsfout met een autoregressief model, de autocorrelatiecoëfficiënt na correctie voor seizoensinvloeden en de autocorrelatiecoëfficiënt van de modelfouten van het locatiemodel. Voor de stoffen salntt t/m P (zie tabel 1 van de hoofdtekst) betreffen de resultaten maandwaarden en bij de stoffen Cd-nf t/m As kwartaalwaarden. Gegeven zijn de resultaten bij een lag van 1, 2, 3 en 4 maanden bij de stoffen salntt t/m P en bij een lag van 3, 6, 9 en 12 maanden bij de stoffen Cd-nf t/m As. De meest bemeten locatie is Marsdiep noord.

In tabel 7 is de relatieve schattingsfout bepaald bij toepassing van de naïeve schattingsmethode. Bij de naïeve is de schatting van de concentratie gelijk aan de voorgaande meetwaarde. Getoetst is met een betrouwbaarheid van 99% of het naïeve model zinvol is, of de relatieve schattingsfout significant kleiner is dan 100%. Heeft toets de waarde 1, dan wordt de nulhypothese niet verworpen.

Tabel 7 : De relatieve schattingsfout met de naïeve methode

Stof	1 mnd	toets	2 mnd	toets	3 mnd	toets	4 mnd	toets
Salntt	181	1	199	1	214	1	177	1
Zs	174	1	222	1	185	1	213	1
Chlfa	122	1	135	1	169	1	243	1
NO3NO2	89	0	158	1	211	1	276	1
N-nf	92	1	147	1	238	1	387	1
NH4	164	1	192	1	211	1	200	1
PO4	106	1	169	1	196	1	217	1
P-nf	138	1	172	1	178	1	278	1
P	175	1	201	1	176	1	199	1
Stof	3 mnd	toets	6 mnd	toets	9 mnd	toets	jaar	toets
Cd-nf	123	1	138	1	56	0	283	1
Cu-nf	241	1	193	1	179	1	164	1
As-nf	166	1	196	1	141	1	130	1
Cd	150	1	209	1	195	1	125	1
Cu	177	1	195	1	190	1	189	1
As	185	1	252	1	210	1	147	1

Alleen bij de stof NO3NO2 bij een lag van 1 maand en bij de stof Cd-nf bij een lag van 9 maanden wordt de nulhypothese dat het naïeve model zinvol is met een 99% betrouwbaarheid, niet verworpen.

In tabel 8 zijn de autocorrelatiecoëfficiënten gegeven na correctie op seizoensinvloeden. Door de stoffen te centreren rond maand- en kwartaal gemiddelde waarden worden seizoensinvloeden genivelleerd. Met een betrouwbaarheid van 99% wordt getoetst of opeenvolgende concentraties in de tijd niet gecorreleerd zijn. Heeft toets de waarde 0, dan wordt de nulhypothese niet verworpen. Heeft toets de waarde 1 dan wordt de alternatieve hypothese, opeenvolgende concentraties zijn gecorreleerd, geaccepteerd. Hoe hoger de autocorrelatiecoëfficiënt en hoe kleiner het aantal waarnemingen des te groter de kans dat de nulhypothese wordt verworpen.

Tabel 8 : De autocorrelatiecoëfficiënt na correctie op seizoensinvloeden

Stof	1 mnd	toets	2 mnd	toets	3 mnd	toets	4 mnd	toets
Salntt	0.12	0	0.06	0	0.05	0	0.21	1
Zs	0.35	1	0.27	1	0.31	1	0.26	1
Chlfa	0.04	0	0.18	0	0.02	0	-0.04	0
NO3NO2	0.13	0	-0.09	0	0.07	0	0.08	0
N-nf	-0.05	0	-0.25	0	0.18	0	0.04	0
NH4	0.44	1	0.27	1	0.33	1	0.35	1
PO4	0.50	1	0.37	1	0.38	1	0.41	1
P-nf	0.32	0	0.32	0	0.24	0	0.20	0
P	0.13	0	0.08	0	0.16	0	0.10	0
Stof	3 mnd	toets	6 mnd	toets	9 mnd	toets	jaar	toets
Cd-nf	0.41	1	0.54	1	0.5	1	0.38	1
Cu-nf	0.11	0	0.13	0	0.23	0	0.24	0
As-nf	0.33	0	0.07	0	0.17	0	0.09	0
Cd	0.72	1	0.59	1	0.60	1	0.60	1
Cu	0.35	1	0.17	0	0.24	0	0.18	0
As	0.27	0	0.18	0	0.22	0	0.25	0

Doordat seizoensinvloeden zijn weggefilterd zijn de autocorrelatiecoëfficiënten in tabel 8 meestal lager dan in tabel 10 van de hoofdtekst. Dit geldt vooral voor de eerste negen stoffen Salntt t/m P, i.h.b. voor Chlfa, NO3NO2 en N-nf. Bij de stoffen Zs, NH4 en PO4 voor een lag van 1 maand en voor Cd-nf, Cd en Cu voor een lag van 3 maanden wordt de nulhypothese niet verworpen.

In tabel 9 zijn de autocorrelatiecoëfficiënten gegeven van de modelfouten van het locatiemodel van Marsdiep noord met Vliestroom. Met een betrouwbaarheid van 99% wordt getoetst of de modelfouten gecorreleerd zijn. Heeft toets de waarde 1, dan wordt de nulhypothese niet verworpen.

Tabel 9 : Autocorrelatiecoëfficiënt van de modelfouten bij het locatiemodel

Stof	1 mnd	toets	2 mnd	toets	3 mnd	toets	4 mnd	toets
Salntt	0.04	0	0.08	0	-0.03	0	0.17	0
Zs	0.15	0	0.13	0	0.16	0	0.09	0
Chlfa	0.14	0	0.13	0	-0.02	0	-0.07	0
NO3NO2	0.07	0	-0.03	0	0.01	0	0.02	0
N-nf	0.4	0	0.21	0	-0.08	0	-0.15	0
NH4	0.07	0	0.06	0	-0.02	0	0.01	0
PO4	0.06	0	0.14	0	0.01	0	0.03	0
P-nf	0.05	0	-0.06	0	-0.09	0	0.03	0
P	0.04	0	-0.01	0	0.18	1	0.15	0
Stof	3 mnd	toets	6 mnd	toets	9 mnd	toets	jaar	toets
Cd-nf	-0.07	0	0.06	0	0.14	0	-0.02	0
Cu-nf	0.12	0	-0.15	0	0.05	0	0.09	0
As-nf	-0.22	0	-0.15	0	0.12	0	0.18	0
Cd	-0.02	0	0.07	0	0.16	0	-0.03	0
Cu	0.12	0	-0.08	0	0.04	0	0.05	0
As	-0.03	0	-0.15	0	0.33	0	0.31	0

Alleen bij de stof P voor lag 3 wordt de nulhypothese verworpen.

BIJLAGE : Uitgebreide resultaten Eems-Dollard

Inhoudsopgave

1	Gebruikte meetgegevens van de Eems-Dollard	66
	Tabel 1 : Aantal metingen per stof voor de locaties	66
	Tabel 2 : Aantal jaren dat er is gemeten per stof voor de locaties	67
2	Resultaten met een locatiemodel voor de Eems-Dollard	68
	Tabel 3 : Gemiddelde verklaarde variantie (%) voor de locaties . .	68
	Tabel 4 : Verklaarde variantie (%) per stof voor de locaties	69
	Tabel 5 : F-toets voor de locaties	70
	Tabel 6 : Aantal gepaarde metingen per stof voor de locaties	71
3	Resultaten van berekening autocorrelatie	72
	Tabel 7 : De relatieve schattingsfout met de naïeve methode	72
	Tabel 8 : De autocorrelatiecoëfficiënt na correctie op seizoensinvloeden	73
	Tabel 9 : Autocorrelatiecoëfficiënt van de modelfouten bij het locatiemodel	74

1 Gebruikte meetgegevens van de Eems-Dollard

In tabel 1 is gegeven het totaal aantal metingen wat er in de periode 1971 t/m 1994 is verricht op de verschillende stoffen in de Eems-Dollard. In tabel 2 is gegeven het aantal jaren waarin metingen zijn verricht.

Tabel 1 : Aantal metingen per stof voor de locaties

	Grootgnd	Zeehvknl	Oostfsgje	Bochtwtm	Bochtwnd	Huibgot
Salntt	409	76	384	78	389	400
Zs	533	76	416	78	505	461
Chlfa	448	75	335	75	437	434
NO3NO2	540	40	517	41	531	453
N-nf	56	28	28	30	30	56
NH4	571	75	560	76	573	486
PO4	566	75	548	76	563	481
P-nf	56	28	28	30	30	55
P	512	47	440	47	531	430
Cd-nf	151	23	23	22	86	67
Cu-nf	152	23	23	22	88	67
As-nf	65	23	23	22	63	47
Cd	150	23	23	22	87	67
Cu	0	23	23	22	88	67
As	68	23	23	22	69	51

Tabel 2 : Aantal jaren dat er is gemeten per stof voor de locaties

	Grootgnd	Zeehvknl	Oostfsgje	Bochtwtm	Bochtwnd	Huibgot
Salntt	18	7	18	7	18	18
Zs	22	7	18	7	22	22
Chlfa	19	7	14	7	19	19
NO3NO2	23	4	23	4	23	22
N-nf	4	3	3	3	3	4
NH4	24	7	24	7	24	24
PO4	24	7	24	7	24	23
P-nf	4	3	3	3	3	4
P	21	4	17	4	21	20
Cd-nf	18	6	6	6	12	12
Cu-nf	18	6	6	6	12	12
As-nf	11	6	6	6	11	10
Cd	18	6	6	6	12	12
Cu	0	6	6	6	12	12
As	11	6	6	6	11	10

2 Resultaten met een locatiemodel voor de Eems-Dollard

In tabel 9 van de hoofdtekst is de maximale gemiddelde verklaarde variantie gegeven als concentraties van stoffen worden geschat door een andere locatie. In tabel 3 wordt de gemiddelde verklaarde variantie gegeven van een locatie, indien de stoffen geschat worden uit de andere buur-locaties.

Tabel 3 : Gemiddelde verklaarde variantie (%) voor de locaties

Buur-locatie

	Grootgnd	Zeehvknl	Oostfsgje	Bochtwtm	Bochtwnd	Huibgot
Grootgnd		58	49	52	40	23
Zeehvknl	60		64	65	57	33
Oostfsgje	51	65		75	55	34
Bochtwtm	54	64	74		61	36
Bochtwnd	40	57	54	61		38
Huibgot	23	32	33	34	39	

In tabel 4 is per stof de verklaarde variantie gegeven bij elke locatie, waarbij de locatie van tabel 3, welke de maximale verklaarde variantie geeft, de stoffen schat.

Tabel 4 : Verklaarde variantie (%) per stof voor de locaties

	Grootgnd	Zeehvknl	Oostfsgje	Bochtwtm	Bochtwnd	Huibgot
Salntt	91	76	75	75	78	60
Zs	28	22	40	40	17	4
Chlfa	54	55	74	74	76	36
NO3NO2	93	91	95	95	97	59
N-nf	94	94	95	95	96	78
NH4	43	51	79	78	76	48
PO4	80	77	84	84	92	42
P-nf	85	91	94	94	75	18
P	29	19	47	48	12	15
Cd-nf	58	69	66	65	31	60
Cu-nf	82	82	88	87	63	38
As-nf	57	83	76	76	84	48
Cd	53	57	74	75	34	31
Cu	0	33	60	63	46	22
As	28	71	72	58	40	20

Stoffen met de hoogste gemiddelde verklaarde variantie zijn Salntt, NO3NO2, N-nf, PO4, P-nf en Cu-nf. De verklaarde varianties bij de locatie Huibertsgat oost zijn kleiner dan elders.

De statistische F-toets (zie bijlage : 'Theoretische achtergronden') geeft de kans dat het informatieverlies kleiner is dan 100% op toeval berust. Is de kans kleiner dan 0.01 dan wordt de nulhypothese dat het gebruik van het locatiemodel significant zinvol is met een betrouwbaarheid van 99%, niet verworpen. In de lege cellen van de tabel zijn er niet voldoende metingen voor het identificeren van het locatiemodel.

In tabel 5 is per stof de F-toets toegepast bij elke locatie, waarbij de locatie van tabel 3, welke de maximale verklaarde variantie geeft, de stoffen schat.

Tabel 5 : F-toets voor de locaties

	Grootgnd	Zeehvknl	Oostfsgje	Bochtwtm	Bochtwnd	Huibgot
Salntt	0.00	0.00	0.00	0.00	0.00	0.00
Zs	0.00	0.00	0.00	0.00	0.00	0.00
Chlfa	0.00	0.00	0.00	0.00	0.00	0.00
NO3NO2	0.00	0.00	0.00	0.00	0.00	0.00
N-nf	0.00	0.00	0.00	0.00	0.00	0.00
NH4	0.00	0.00	0.00	0.00	0.00	0.00
PO4	0.00	0.00	0.00	0.00	0.00	0.00
P-nf	0.00	0.00	0.00	0.00	0.00	0.05
P	0.00	0.01	0.00	0.00	0.03	0.00
Cd-nf	0.00	0.00	0.00	0.00	0.01	0.00
Cu-nf	0.00	0.00	0.00	0.00	0.00	0.00
As-nf	0.00	0.00	0.00	0.00	0.00	0.00
Cd	0.00	0.00	0.00	0.00	0.01	0.00
Cu	0.00	0.01	0.00	0.00	0.00	0.00
As	0.02	0.00	0.00	0.00	0.00	0.00

Alleen bij de stoffen P en Cu bij Zeehavenkanaal, de stoffen P, Cd-nf en Cd bij Bocht van Watum noord en P-nf bij Huibertsgat oost wordt de nulhypothese dat het locatiemodel zinvol is, verworpen.

In tabel 6 zijn de aantallen gepaarde metingen gegeven op basis waarvan de verklaarde variantie van tabel 4 is bepaald.

Tabel 6 : Aantal gepaarde metingen per stof voor de locaties

	Grootgnd	Zeehvknl	Oostfsgje	Bochtwtm	Bochtwnd	Huibgot
Salntt	76	75	76	76	78	371
Zs	70	71	71	71	77	399
Chlfa	67	66	68	68	66	351
NO3NO2	40	39	39	39	41	416
N-nf	27	28	28	28	30	28
NH4	68	69	71	71	74	438
PO4	75	74	74	74	75	444
P-nf	27	28	28	28	29	27
P	45	44	46	46	46	402
Cd-nf	23	22	22	22	22	58
Cu-nf	22	21	21	21	21	61
As-nf	23	22	22	22	20	42
Cd	21	19	21	21	21	63
Cu	0	22	21	21	22	64
As	22	22	22	22	21	48

3 Resultaten van berekening autocorrelatie

Gegeven zijn de relatieve schattingsfout met een autoregressief model, de autocorrelatiecoëfficiënt na correctie voor seizoensinvloeden en de autocorrelatiecoëfficiënt van de modelfouten van het locatiemodel. Voor de stoffen salntt t/m P (zie tabel 1 van de hoofdtekst) betreffen de resultaten maandwaarden en bij de stoffen Cd-nf t/m As kwartaalwaarden. Gegeven zijn de resultaten bij een lag van 1, 2, 3 en 4 maanden bij de stoffen salntt t/m P en bij een lag van 3, 6, 9 en 12 maanden bij de stoffen Cd-nf t/m As. De meest bemeten locatie is Groote Gat noord.

In tabel 7 is de relatieve schattingsfout bepaald bij toepassing van de naïeve schattingsmethode. Bij de naïeve is de schatting van de concentratie gelijk aan de voorgaande meetwaarde. Getoetst is met een betrouwbaarheid van 99% of het naïeve model zinvol is, of de relatieve schattingsfout significant kleiner is dan 100%. Heeft toets de waarde 1, dan wordt de nulhypothese niet verworpen.

Tabel 7 : De relatieve schattingsfout met de naïeve methode

Stof	1 mnd	toets	2 mnd	toets	3 mnd	toets	4 mnd	toets
Salntt	72	0	133	1	186	1	239	1
Zs	185	1	186	1	171	1	202	1
Chlfa	141	1	155	1	204	1	208	1
NO3NO2	100	1	156	1	198	1	235	1
N-nf	50	0	105	1	183	1	258	1
NH4	83	0	146	1	180	1	233	1
PO4	85	0	117	1	176	1	223	1
P-nf	58	0	126	1	165	1	219	1
P	178	1	184	1	218	1	206	1
Stof	3 mnd	toets	6 mnd	toets	9 mnd	toets	jaar	toets
Cd-nf	185	1	146	1	196	1	236	1
Cu-nf	204	1	230	1	167	1	98	1
As-nf	173	1	275	1	209	1	126	1
Cd	245	1	209	1	284	1	228	1
Cu	202	1	221	1	137	1	152	1
As	223	1	172	1	204	1	191	1

Alleen bij de stoffen salntt, NO3NO2, N-nf, NH4, PO4 en P-nf wordt de nulhypothese dat het naïeve model zinvol is (bij een lag van 1 maand) met een

99% betrouwbaarheid, niet verworpen. De relatieve schattingsfout van Cu-nf is kleiner dan 100% voor een lag van een jaar, maar de nulhypothese wordt hier verworpen.

In tabel 8 zijn de autocorrelatiecoëfficiënten gegeven na correctie op seizoensinvloeden. Door de stoffen te centreren rond maand- en kwartaal gemiddelde waarden worden seizoensinvloeden genivelleerd. Met een betrouwbaarheid van 99% wordt getoetst of opeenvolgende concentraties in de tijd niet gecorreleerd zijn. Heeft toets de waarde 0, dan wordt de nulhypothese niet verworpen. Heeft toets de waarde 1 dan wordt de alternatieve hypothese, opeenvolgende concentraties zijn gecorreleerd, geaccepteerd. Hoe hoger de autocorrelatiecoëfficiënt en hoe kleiner het aantal waarnemingen des te groter de kans dat de nulhypothese wordt verworpen.

Tabel 8 : De autocorrelatiecoëfficiënt na correctie op seizoensinvloeden

Stof	1 mnd	toets	2 mnd	toets	3 mnd	toets	4 mnd	toets
Salntt	0.58	1	0.39	1	0.30	1	0.20	1
Zs	-0.01	0	0.05	0	0.10	0	0.02	0
Chlfa	0.14	0	0.15	0	0.02	0	0.14	0
NO3NO2	0.21	1	0.23	1	0.11	0	0.04	0
N-nf	0.39	0	0.31	0	-0.19	0	-0.11	0
NH4	0.43	1	0.31	1	0.24	1	0.15	0
PO4	0.35	1	0.26	1	0.22	1	0.15	0
P-nf	0.74	1	0.62	1	0.56	1	0.39	0
P	0.06	0	0.06	0	0.06	0	0.09	0
Stof	3 mnd	toets	6 mnd	toets	9 mnd	toets	jaar	toets
Cd-nf	0.41	1	0.44	1	0.43	1	0.37	1
Cu-nf	0.12	0	0.16	0	0.24	0	0.24	0
As-nf	0.35	0	0.05	0	-0.09	0	-0.22	0
Cd	0.14	0	0.22	0	-0.03	0	0.11	0
Cu	0.08	0	0.01	0	0.34	1	0.18	0
As	0.19	0	0.07	0	0.13	0	-0.14	0

Doordat seizoensinvloeden zijn weggefilterd zijn de autocorrelatiecoëfficiënten in tabel 8 meestal lager dan in tabel 10 van de hoofdtekst. Dit geldt vooral voor de eerste negen stoffen Salntt t/m P en in het bijzonder voor de stoffen NO3NO2, N-nf en PO4. De autocorrelatiecoëfficiënten van de stoffen salntt, NO3NO2 en P-nf zijn het grootst voor lag 1 in beide tabellen. Bij de stoffen salntt, NO3NO2, NH4, PO4, P-nf voor een lag van 1 maand en voor Cd-nf voor

een lag van 3 maanden wordt de nulhypothese verworpen.

In tabel 9 zijn de autocorrelatiecoëfficiënten gegeven van de modelfouten van het locatiemodel van Vlissingen met Borssele. Met een betrouwbaarheid van 99% wordt getoetst of de modelfouten gecorreleerd zijn. Heeft toets de waarde 0, dan wordt de nulhypothese niet verworpen.

Tabel 9 : Autocorrelatiecoëfficiënt van de modelfouten bij het locatiemodel

Stof	1 mnd	toets	2 mnd	toets	3 mnd	toets	4 mnd	toets
Salntt	-0.04	0	0.05	0	-0.08	0	0.21	0
Zs	0.10	0	0.03	0	0.13	0	-0.10	0
Chlfa	0.21	0	0.07	0	-0.03	0	-0.08	0
NO3NO2	-0.10	0	0.03	0	0.00	0	0.01	0
N-nf	0.11	0	0.10	0	-0.32	0	0.03	0
NH4	0.25	0	0.17	0	0.05	0	-0.03	0
PO4	0.24	0	-0.04	0	0.00	0	-0.01	0
P-nf	0.02	0	-0.12	0	-0.05	0	0.17	0
P	0.12	0	0.07	0	-0.03	0	0.05	0
Stof	3 mnd	toets	6 mnd	toets	9 mnd	toets	jaar	toets
Cd-nf	0.00	0	0.00	0	0.11	0	0.00	0
Cu-nf	0.38	0	-0.04	0	0.09	0	0.38	0
As-nf	0.21	0	-0.09	0	0.18	0	-0.07	0
Cd	0.04	0	-0.47	0	-0.09	0	0.29	0
Cu	0.30	0	-0.03	0	-0.28	0	-0.36	0
As	-0.19	0	-0.24	0	0.09	0	0.26	0

Bij alle stoffen en voor alle gegeven lags is de autocorrelatiecoëfficiënt van de modelfouten dusdanig klein dat de nulhypothese nooit verworpen wordt.

BIJLAGE : Theoretische achtergronden

Inhoudsopgave

1	Beschrijving ARX-model	76
2	Kruiscorrelatie en tijdvertraging	77
3	Het locatiemodel	78
4	Toetsing van de kwaliteit van een locatiemodel	79
5	Ratio verklaarde variantie	79
6	De F-toets	80
7	Betrouwbaarheidsinterval voor een voorspelling	81
8	Autocorrelatie	81
9	Literatuur	83

1 Beschrijving ARX-model

De concentratiemetingen van de verschillende meetlocaties zijn tijdreeksen. Om twee tijdreeksen aan elkaar te relateren zijn zgn. ARX-modellen geschikt. Een ARX-model behoort tot de klasse van de lineaire tijdreeksmodellen en worden beschreven in [1] en [2]. ARX staat voor Auto-Regressive-eXogeneous en is een combinatie van auto-regressieve modellen waarin voorgaande waarnemingen worden meegemodelleerd en exogene modellen waarin een tijdreeks als functie van een andere tijdreeks wordt beschreven.

De formulering voor een ARX-model is:

$$Y(k) + a_1 Y(k-1) + \dots + a_{na} Y(k-na) = b_1 U(k-nk) + b_2 U(k-nk-1) + \dots + b_{nb} U(k-nk-nb+1) + W(k) \quad (1)$$

waarin :

$Y(k)$	:	het uitgangssignaal, de concentratie van een stof in de te beschrijven locatie
$U(k)$	:	het ingangssignaal, de concentratie van een stof in een andere bemeten locatie
$W(k)$	:	modelverstoring met verwachtingswaarde nul
a_i, b_i	:	modelparameters
na	:	orde van autoregressie : aantal uitgangssignalen op vroegere tijdstippen die in het model worden meegenomen
nb	:	orde van de modelingang : aantal ingangssignalen die in het model worden meegenomen
nk	:	tijdvertraging tussen de modelingang en -uitgang
k	:	(discrete) tijdparameter

De orde van het ARX-model wordt bepaald door de parameters na , nb en nk .

2 Kruiscorrelatie en tijdvertraging

Met correlatie-technieken is het mogelijk de waarden van na , nb en nk voor een ARX-model te bepalen.

De kruiscorrelatie tussen twee (scalaire) grootheden is gedefinieerd als :

$$\rho(y,u) = \frac{COV(y,u)}{\sqrt{VAR(y)VAR(u)}} \quad (2)$$

Zijn zowel Y als U zwak-stationair zijn, d.w.z. beide processen hebben een constant gemiddelde en de autocorrelatie tussen twee tijdstippen is slechts afhankelijk van het tijdsverschil. De correlatiefunctie van Y en U is dan een functie van het tijdsverschil tussen beide processen :

$$\rho_{YU}(s) = \frac{COV(Y(k),U(k-s))}{\sqrt{VAR(Y(k))VAR(U(k-s))}} \quad (3)$$

met:

$$\begin{aligned} COV(Y(k),U(k-s)) &= E(Y(k)-EY(k))(U(k-s)-EU(k-s)) \\ VAR(Y(k)) &= E(Y(k)-EY(k))^2 \\ VAR(U(k-s)) &= E(U(k-s)-EU(k-s))^2 \end{aligned} \quad (4)$$

Een schatting voor de tijdvertraging tussen de uitgang Y en de ingang U van een systeem is dan :

$$nk = \max_s(\rho_{YU}(s)) \quad (5)$$

Voor de keuze van de orde van autoregressie na en de modelingang nb geeft de correlatiefunctie geen eenduidig antwoord. De keuze van na en nb wordt voornamelijk bepaald door de kwaliteit van de modelbeschrijving of door fysische overwegingen.

3 Het locatiemodel

Het locatiemodel beschrijft de relatie van de concentratie van een stof tussen twee locaties. Zo kan bijvoorbeeld de concentratie van Fosfor (P) bij Vlissingen goed voorspeld worden uit de concentratie van P bij Borssele. Het locatiemodel behoort tot de klasse van ARX-modellen, met $na=0$, $nk=0$ en $nb=1$ [3].

Beschouw het locatiemodel :

$$Y(k) = cU(k) + d + W(k) \quad (6)$$

met :

$Y(k)$	:	de concentratie van een stof op de afhankelijke locatie
$U(k)$	:	de concentratie van een stof op de onafhankelijke locatie
c, d	:	modelparameters
$W(k)$	:	modelruis $N(0, \sigma_w^2)$

Als we de schattingen voor het gemiddelde en de varianties en covarianties schrijven als :

$$\begin{aligned} m_Y &= \frac{1}{N} \sum_{k=1}^N Y(k) \\ m_U &= \frac{1}{N} \sum_{k=1}^N U(k) \\ s_{UY} &= \frac{1}{N-1} \sum_{k=1}^N (U(k) - m_U)(Y(k) - m_Y) \\ s_U^2 &= \frac{1}{N-1} \sum_{k=1}^N (U(k) - m_U)^2 \end{aligned} \quad (7)$$

dan geldt voor de modelparameters c en d :

$$\begin{aligned} c &= \frac{s_{UY}}{s_U^2} \\ d &= m_Y - cm_U \end{aligned} \quad (8)$$

4 Toetsing van de kwaliteit van een locatiemodel

Om de kwaliteit van een locatiemodel te valideren worden er in dit onderzoek twee methodieken toegepast.

In de identificatieperiode worden de onbekende modelparameters c van het locatiemodel geschat. In de verificatieperiode wordt de modeluitkomsten van het locatiemodel vergeleken met de werkelijk gemeten concentraties. De schattingsfout, het verschil tussen de modeluitkomst en de werkelijk gemeten concentratie kan worden berekend.

Door verschillende verificatieperiodes te kiezen wordt het geïdentificeerde locatiemodel statistisch gevalideerd. Indien de schattingsfouten in de verschillende verificatieperiodes onderling duidelijk verschillen in dynamiek en absolute grootte dan moet geconcludeerd worden, dat het geïdentificeerde locatiemodel niet universeel geldig is.

5 Ratio verklaarde variantie

Om te onderzoeken hoe goed een locatiemodel de gemeten concentratie beschrijft is de ratio VV , de verklaarde variantie, geïntroduceerd. De verklaarde variantie geeft aan in welke mate de betrouwbaarheid van een meting afneemt als deze wordt vervangen door zijn schatting op basis van een andere locatie :

$$VV = (1 - \frac{s_w^2}{s_y^2}).100\% \quad (9)$$

De verklaarde variantie is 100% minus het informatieverlies (R).

$$R = \frac{s_w^2}{s_y^2}.100\% \quad (10)$$

Het informatieverlies geeft de verhouding aan van de variantie van de modelruis en de variantie van de te beschrijven concentratie. Des te kleiner het informatieverlies of des te groter de verklaarde variantie des te beter is de beschrijving door het locatiemodel.

De boven gedefinieerde ratio voor de verklaarde variantie heeft alles te maken met de bij regressie-modellen gebruikte correlatie-coëfficiënt r^2 , de schatting van de correlatie ρ zie vgl. (3) tussen de waarnemingen van de afhankelijke locatie Y en de waarnemingen van de onafhankelijke locatie U :

$$r = \frac{s_{UY}}{s_U s_Y} \quad (11)$$

Het verband wordt gegeven door :

$$VV = (1 - \frac{s_w^2}{s_Y^2}).100\% = r^2.100\% \quad (12)$$

6 De F-toets

Voor het toetsen van de hypothese of het zinvol is om een model met meer parameters toe te passen, wordt de volgende grootte geïntroduceerd :

$$X = \frac{[chisq(q) - chisq(p)] / (p-q)}{chisq(p) / (N-p)} \quad (13)$$

waarin :

$$chisq(p) = (N - 1) \cdot \frac{s_{wp}^2}{s_Y^2} \quad (14)$$

met :

X	:	F-verdeeld met p-q, N-p vrijheidsgraden
N	:	aantal gebruikte metingen voor het schatten van de modelparameters van de twee modellen
s_{wp}^2	:	variantie van de modelruis met p parameters
p, q	:	aantal geschatte parameters bij de modellen ($p > q$)

De grootte X is F-verdeeld en de waarde $P(X|p-q, N-p)$ is de kans dat de reductie in de $chisq(p)$ van het model met meer parameters, toeval is. Om te

toetsen of het eenvoudige locatiemodel (6) met een betrouwbaarheid van 99% significant is (of $VV > 0\%$, of $R < 100\%$), moet gelden :

$$P(X|1, N-1) < 0.01 \quad (15)$$

In (15) wordt het model van (6) met de twee parameters, c en d, tegen het model met alleen de parameter c getoetst.

7 Betrouwbaarheidsinterval voor een voorspelling

De standaarddeviatie voor een nieuwe voorspelling horend bij een zekere u_0 ziet er als volgt uit :

$$s_{Y(u_0)} = s \sqrt{1 + \frac{1}{N} + \frac{(u_0 - \bar{u})^2}{(N-1) \cdot s_u^2}} \quad (16)$$

waarbij s de standaarddeviatie van de residuen is, N het aantal waarnemingen en s_u^2 de standaarddeviatie van de bij het model gebruikte u-waarden. Een 95 % betrouwbaarheidsinterval volgt uit bovenstaande standaarddeviatie :

$$Y - 1.96 \cdot s_{Y(u_0)} < Y < Y + 1.96 \cdot s_{Y(u_0)} \quad (17)$$

8 Autocorrelatie

In het onderzoek voor RIKZ is naast de correlatie in de ruimte ook geanalyseerd de correlatie in de tijd, de autocorrelatie. De autocorrelatiefunctie als functie van het tijdverschil s is als volgt gedefinieerd :

$$\rho_Y(s) = \frac{COV(Y(k), Y(k-s))}{VAR(Y(k))} \quad (18)$$

met:

Y is een zwak-stationair proces

Er geldt [2] :

$$\begin{aligned} \sqrt{(N)} \cdot \rho_Y(s) &\in AsN(0,P) \\ \sqrt{(N)} \cdot P &= \sum_{k=-\infty}^{k=\infty} \rho_Y(k)^2 \end{aligned} \quad (19)$$

Stel $Y(k)$ en $Y(k-s)$ zijn niet gecorreleerd dan geldt :

$$\begin{aligned} \rho_Y(s) &= 0, \text{ als } s > 0 \\ \rho_Y(s) &= 1, \text{ als } s = 0 \\ \sqrt{(N)} \cdot \rho_Y(s) &\in AsN(0,1) \end{aligned} \quad (20)$$

De hypothese (met een 95% betrouwbaarheid) dat $Y(k)$ en $Y(k+s)$ niet gecorreleerd zijn wordt verworpen als :

$$|\rho_Y(s)| > \frac{1.96}{\sqrt{(N)}} \quad (21)$$

9 Literatuur

[1]

G.E.P.Box & G.M. Jenkins (1976)
Time Series Analysis - forecasting and control
Holden-Day, California

[2]

L. Ljung (1987)
System Identification- Theory for the User
Prentice Hall, Englewood Cliffs, N.J.

[3]

H.C. Klavers et al. (1992)
Optimalisatie routinematig onderzoek waterkwaliteit binnenwateren
deel 3 Statistisch Onderzoek p 38-40.

