

Seizoenverloop en trend in de chemische samenstelling van de neerslag te Lelystad

T.A. Buishand
J.H. Baard

Technische rapporten; TR-133

de bilt 1991

publicatienummer: Technische rapporten; TR-133 (FM)

postbus 201
3730 AE de bilt
wilhelminalaan 10
tel. (030) 206911
telex 47096

Afdeling Fysische Meteorologie

auteur: T.A Buishand, KNMI, De Bilt
J.H. Baard, ECN, Petten

U.D.C.: 551.577.13
(492.66)

ISSN: 0169-1708

© KNMI, De Bilt. Niets uit deze uitgave mag worden verveelvoudigd en / of openbaar gemaakt worden door middel van druk, fotocopie, microfilm, of op welke wijze dan ook zonder voorafgaande schriftelijk toestemming van het KNMI.

T.A. Buishand (KNMI) and J.H. Baard (ECN):

**SEASONAL VARIATION AND TREND IN THE CHEMICAL
COMPOSITION OF PRECIPITATION AT LELYSTAD (1983-
1985)**

Abstract

The chemical composition of precipitation at Lelystad (52°30' N, 5°25' E) has been sampled on a daily basis by means of wet-only collectors. A cluster analysis was applied to the concentrations of hydronium, ammonium, sodium, potassium, calcium, magnesium, sulfate, nitrate and chloride. This resulted in two main groups: one group with mainly precipitation from marine origin and another group containing nearly all samples from continental air masses. The temporal variation of the concentrations and the depositions of hydronium, ammonium, nitrate and sulfate was investigated in detail, both for the data in the marine group and the continental group, and for the case that there is no separation of the data into these groups.

To avoid problems with serial correlation, the statistical analysis was performed on the monthly concentrations and depositions instead of the original daily values. Temporal variation was described by fitting multiple regression models to the logarithms of concentrations and depositions. In these models the effects of a sinusoidal annual cycle, a linear trend and the precipitation quantity were considered simultaneously.

There was no indication of a linear trend during the measurement period. The effect of the precipitation quantity on the monthly concentrations and depositions was statistically significant in most cases. The most striking fact, however, was that the monthly values within the separate groups had a much weaker sinusoidal annual cycle than those of all precipitation samples. The stronger annual cycle in the latter case can largely be explained by the contribution of both groups to the average monthly concentrations. The peak concentrations in spring and summer are partly due to the relatively large contribution of polluted precipitation from continental origin in these seasons.

The systematic annual cycle of the wet-only samples at Lelystad differs markedly from that of the monthly bulk samples in the national network. This is not only due to the differences between bulk and wet-only samples, but also to the fact that the wet-only samples could not be analysed on days with little precipitation and to different sampling strategies during frost and snowfall.

Despite the fact that two distinct groups could be formed, the temporal variation of concentrations within these groups remains considerably large. As a result, the separation of the data into different groups can only improve the discrimination of long-term trends in situations that the systematic change in one group differs strongly from that in the other group.

SEIZOENVERLOOP EN TREND IN DE CHEMISCHE SAMENSTELLING VAN DE NEERSLAG TE LELYSTAD (1983-1985)

T.A. Buishand (KNMI) en J.H. Baard (ECN)

INHOUD

1	Inleiding	2
2	Onderverdeling van de dagmonsters met behulp van cluster-analyse	2
3	Bepaling seizoenverloop en trend met behulp van multi-pele regressie	5
4	Vergelijking resultaten ECN-gegevens met die uit het Landelijk Meetnet Regenwatersamenstelling	9
5	Mogelijkheden om de detecteerbaarheid van langjarige trends te verbeteren door splitsing in groepen	13
6	Conclusies	16
	Literatuur	17
	Appendices	
A	Resultaten van de regressie-analyse	19
B	Gestudentiseerde residuen	22
C	Een benadering voor het seizoenverloop van de depositie	25
D	Het effect van ontbrekende maanden op de detecteerbaarheid van langjarige trends	27

1 INLEIDING

Van februari 1983 tot en met december 1985 zijn door het ECN in Lelystad dagelijks metingen verricht van de chemische samenstelling van de neerslag met behulp van wet-only vangers. Om verdachte metingen te kunnen ontdekken werd gebruik gemaakt van acht neerslagvangers (Slanina e.a., 1990). Na eliminatie van uitschieters werd voor elke component per dag de gemiddelde concentratie bepaald. Zodoende werd een betrouwbaar bestand van dagwaarden van de concentraties van een groot aantal componenten verkregen. Het totale bestand bevat de gegevens van 268 regendagen (dagen met 0.5 mm of meer neerslag).

Op het dagwaarden bestand werd een cluster-analyse uitgevoerd. Dit leidde tot een onderverdeling van de dagmonsters in twee groepen: één groep van mariene oorsprong en een andere groep met hoofdzakelijk continentale luchtaanvoer tijdens de monsternamen. Deze splitsing van het databestand in twee min of meer homogene groepen kan tot een verbetering van de detecteerbaarheid van langjarige trends leiden. Dit is onderzocht voor de componenten hydro-nium, ammonium, nitraat en sulfaat*. Hiervan wordt in dit rapport verslag gedaan.

In par. 2 wordt de cluster-analyse van de meetgegevens te Lelystad besproken. Daarna wordt in par. 3 ingegaan op de bepaling van het seizoenverloop en de trend met behulp van regressie-analyse. De resultaten voor de ECN-gegevens worden in par. 4 vergeleken met die voor de meetreeksen uit het Landelijk Meetnet Regenwatersamenstelling (LMR). Tenslotte wordt in par. 5 nagegaan in hoeverre langjarige trends beter onderkend kunnen worden bij een onderverdeling in groepen.

2 ONDERVERDELING VAN DE DAGMONSTERS MET BEHULP VAN CLUSTER-ANALYSE

Cluster-analyse is de benaming van een aantal statistische technieken, die tot doel hebben een verzameling gegevens te splitsen in een aantal min of meer homogene groepen (de clusters). De bruikbaarheid van deze technieken om regenwatermonsters te classificeren is door Slanina e.a. (1983) op basis van 26 wet-only regenwatermonsters te Den Helder onderzocht met behulp van het programmapakket CLUSTAN-1C. Dit pakket is ook gebruikt bij de meetgegevens van Lelystad. Eerst zal worden ingegaan op een aantal bijzonderheden van de gevolgde procedure.

* Bij sulfaat is gecorrigeerd voor de bijdrage van zeezout-aerosolen.

Daarna zullen enkele eigenschappen van de gevonden clusters worden besproken.

De concentraties van de verschillende componenten voor regendag k kunnen worden weergegeven door de vector $\mathbf{x}_k = (x_{k1}, \dots, x_{kn})$. Bij het regenwateronderzoek in Lelystad is uitgegaan van $n = 9$ componenten: hydronium, ammonium, natrium, kalium, calcium, magnesium, sulfaat, nitraat en chloride. Elke vector \mathbf{x}_k vormt een punt in de n -dimensionale ruimte. Punten die dicht bij elkaar gelegen zijn, geven dagmonsters weer die een sterke overeenkomst vertonen. De Euclidische afstand tussen twee punten wordt gegeven door:

$$d(\mathbf{x}_j, \mathbf{x}_k) = \sqrt{(x_{j1} - x_{k1})^2 + \dots + (x_{jn} - x_{kn})^2} \quad (1)$$

Deze afstandsmaat is alleen zinvol indien de verschillende componenten goed vergelijkbaar zijn. Bij concentraties van chemische componenten in regenwatermonsters is dit echter bepaald niet het geval. Sommige componenten vertonen een veel grotere spreiding dan andere, en deze componenten bepalen dan in sterke mate de Euclidische afstand. In plaats van de oorspronkelijke concentratie x_{ki} is daarom bij dit onderzoek uitgegaan van de gestandaardiseerde concentratie:

$$\tilde{x}_{ki} = (x_{ki} - \bar{x}_{.i}) / s_i \quad (2)$$

met $\bar{x}_{.i}$ het gemiddelde en s_i de steekproefstandaardafwijking van de i -de component. Behalve een afstandsmaat tussen punten, heeft men bij cluster-analyse ook een criterium voor de afstand tussen groepen van punten nodig. Bij de gestandaardiseerde concentraties van de dagmonsters in Lelystad is uitgegaan van de Euclidische afstand tussen de zwaartepunten van de groepen (zwaartepuntsverbindinglijn of "centroid"), vanwege het feit dat dit criterium bij het eerdere onderzoek van Slanina e.a. (1983) de beste resultaten gaf.

De onderverdeling in groepen is in dit onderzoek tot stand gebracht door middel van hiërarchische fusie. Bij deze methode vormt in de beginsituatie elk regenwatermonster een éénpunts-cluster. Vervolgens worden stap voor stap steeds twee clusters (twee punten, een punt en een reeds eerder gevormde groep van punten, of twee groepen van punten) samengevoegd en wel die clusters die het dichtste bij elkaar liggen. Bij de gegevens van Lelystad worden in de laatste fase twee groepen verkregen, die ongeveer evenveel regenwatermonsters bevatten. De ene

groep (104 monsters) bestaat vrijwel uitsluitend uit neerslagmonsters van mariene oorsprong, terwijl de andere groep (164 monsters) vrijwel alle neerslagmonsters bij continentale luchtaanvoer omvat. Een trajectoriën-analyse voor de jaren 1983, 1984 (Janssen, 1988) laat zien dat in de tweede groep de oorsprong van de neerslag sterker varieert dan in de eerste groep.

In tabel 1 zijn de gemiddelde (volume gewogen) concentraties van hydronium, ammonium, nitraat en sulfaat gegeven. Voor alle vier componenten is de gemiddelde concentratie voor de continentale groep een factor 2 à 3 maal zo hoog als die voor de mariene groep.

Tabel 1. Gemiddelde concentratie ($\mu\text{mol/l}$)

Component	Marien	Continentaal	Alle monsters
Hydronium	16.2	43.9	30.8
Ammonium	43.1	108.1	77.2
Nitraat	19.9	57.4	39.6
Sulfaat	19.7	50.1	35.6

Een nuttige indicator voor de oorsprong van de neerslag is de verhouding tussen de ammonium- en natriumconcentraties (Asman e.a., 1981). Ammonium is een component van continentale emissies, terwijl natrium vrijwel uitsluitend van zeezout-aerosol afkomstig is. Figuur 1 laat de frequentieverdeling van de NH_4/Na -verhouding zien voor zowel de mariene als de continentale groep. Zoals verwacht mocht worden, verschillen deze frequentieverdelingen sterk; in de mariene groep is de NH_4/Na -verhouding laag, terwijl voor de continentale groep het tegenovergestelde geldt. In een gering aantal gevallen wordt in de mariene groep echter een zeer hoge NH_4/Na -verhouding aangetroffen. Evenals bij de andere regenwatermonsters in de mariene groep, zijn bij deze regenwatermonsters de concentraties van hydronium, nitraat en sulfaat in het algemeen laag. De sterk afwijkende NH_4/Na -verhouding is mogelijk het gevolg van het laatste traject boven land. Monsters in de continentale groep met een lage NH_4/Na -verhouding zijn in veel gevallen afkomstig uit sterk verontreinigde mariene luchtmassa's (bijvoorbeeld als gevolg van emissies in Engeland).

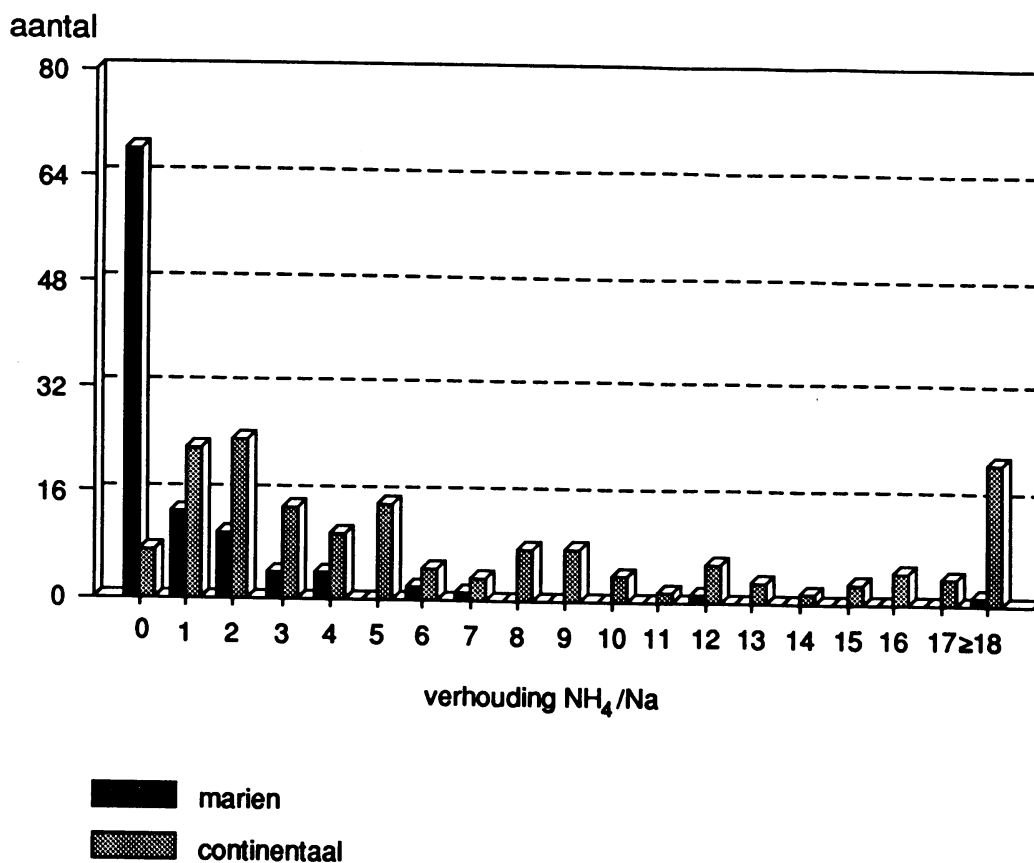


Fig. 1 Frequentieverdeling van de verhouding van de ammonium- en natriumconcentraties van dagmonsters voor de mariene en de continentale groep

Cluster-analyse leidt dus tot een groep regenwatermonsters van mariene oorsprong met relatief lage concentraties van hydronium, ammonium, nitraat en sulfaat. Daarnaast is er een tweede groep regenwatermonsters die van meer verontreinigde lucht afkomstig is. De vervuiling is voor een belangrijk deel afkomstig van diverse landen op het Europese continent (Janssen, 1988). Er is echter ook een bijdrage van sterk verontreinigde lucht van mariene oorsprong.

3 BEPALING SEIZOENVERLOOP EN TREND MET BEHULP VAN MULTIPLE REGRESSIE

De statistische analyse werd uitgevoerd op basis van maandwaarden. Hierbij is per maand gemiddeld over de mariene regendagen enerzijds en de continentale regendagen anderzijds. De overgang van dagwaarden op maandwaarden leidt voor een onderzoek naar langjarige trends

nauwelijks tot verlies aan informatie. Het voordeel van het werken met maandwaarden is dat men geen rekening hoeft te houden met autocorrelatie. Verder zal bij maandwaarden in het algemeen beter aan de normaliteitsveronderstelling voor de toets op trend zijn voldaan.

Evenals in Buishand e.a. (1988) en Buishand en Van Montfort (1989) is het tijdsverloop van concentraties en deposities beschreven door een multipel regressiemodel. Het model werd toegepast op de natuurlijke logaritmen van de maandwaarden. Voor de concentraties is het model van de vorm:

$$\ln C_i = a_0 + a_1 \cos(2\pi i/N - \phi_1) + b i/N + c/\sqrt{h_i} + e_i, \quad (3)$$

en voor de deposities:

$$\ln D_i = a_0 + a_1 \cos(2\pi i/N - \phi_1) + b i/N + c \sqrt{h_i} + e_i, \quad (4)$$

met $N = 12$. Hierin zijn C_i , D_i en h_i de concentratie (volume-gewogen), de depositie en de neerslaghoeveelheid van de i -de maand. De maandindex wordt gegeven door $i = 12(j - 1983) + m$, met j het jaartal en m het volgnummer van de kalendermaand, waarop de concentratie, de depositie en de neerslaghoeveelheid betrekking hebben. De cosinusterm in de rechter leden stelt een systematische jaarlijkse gang voor met amplitude a_1 . De fasehoek ϕ_1 bepaalt het tijdstip waarop de bijdrage van deze term maximaal is. Als gerekend wordt in graden in plaats van radialen, dan geeft ϕ_1 ongeveer het dagnummer aan waarop de cosinusterm zijn maximum bereikt. De term $b i/N$ stelt een langjarige trend voor. De regressiecoëfficiënt b geeft de relatieve verandering van de concentratie of depositie weer in een tijdvak van een jaar. Door de termen $c/\sqrt{h_i}$ en $c \sqrt{h_i}$ wordt de invloed van de neerslaghoeveelheid op de concentraties en deposities beschreven. De term e_i stelt een fout voor (residu). Verondersteld wordt dat de e_i 's onderling onafhankelijke, normaal verdeelde kansvariabelen zijn met verwachting nul en alle met dezelfde standaardafwijking σ_e .

In tegenstelling tot bovengenoemde publikaties zijn er maanden waarvoor geen metingen beschikbaar zijn. Voor de bepaling van de regressiecoëfficiënten is dit geen bezwaar.

Per component werd een regressie-analyse uitgevoerd voor:

- (M) de concentratie en de depositie in de mariene groep,
- (C) de concentratie en de depositie in de continentale groep, en
- (M&C) de concentratie en de depositie van alle regendagen.

Bij de regressie-analyse werd gebruik gemaakt van het pakket SPSS/PC⁺. Een overzicht van de resultaten wordt gegeven in Appendix A.

Voor februari 1983 is de waarde van hydronium een uitschieter in de gevallen (C) en (M&C), zowel bij de concentraties als bij de deposities. Het gestudentiseerde residu* ligt voor deze maand tussen -3.6 en -3.3. Voor het geval (M&C) is het tijdsverloop van de concentratie van de hydroniumionen weergegeven in figuur 2. Er zijn geen M-waarnemingen in februari 1983, zodat de neerslaghoeveelheid in de gevallen (C) en (M&C) hetzelfde is (9.85 mm). De vrij geringe neerslaghoeveelheid in deze maand gaat samen met een zeer lage hydroniumconcentratie (2.54 $\mu\text{mol/l}$). De resultaten in Appendix A laten zien dat weglating van deze maand in de regressie-analyse tot een vrij aanzienlijke reductie van de residuele standaardafwijking $\hat{\sigma}_e$ leidt. Verder blijkt dat een significant stijgende trend bij hydronium in het geval (M&C) voor een belangrijk deel moet worden toegeschreven aan de negatieve uitschieter in februari 1983. In alle andere situaties wordt geen significante trend gevonden.

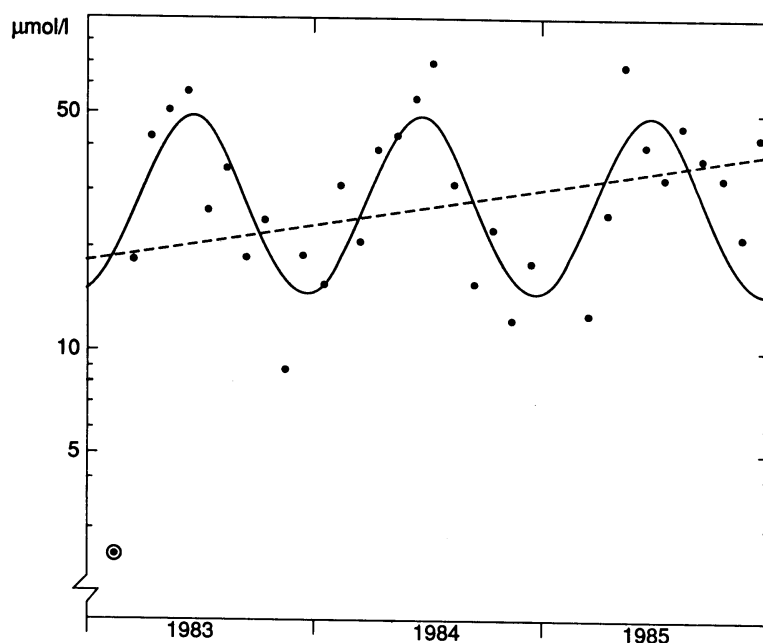


Fig. 2 Tijdsverloop van de concentratie van de hydroniumionen. De punten hebben betrekking op de gemiddelde waarde over alle regendagen (M&C) binnen de desbetreffende maand. De waarde voor februari 1983 is omcirkeld. De volle lijn stelt het geschatte sinusoïdale seizoenverloop voor; de streeplijn geeft de geschatte lineaire trend weer

* Op het begrip gestudentiseerd residu en de statistische significantie van de afwijkende waarde voor hydronium in februari 1983 wordt nader ingegaan in Appendix B.

Bij de toepassing van het regressiemodel op alle gegevens, geval (M&C), is voor alle vier componenten de amplitude a_1 van de cosinusterm in de modellen (3) en (4) significant bij een betrouwbaarheidsdrempel van 1%. De significantie van de amplitude daalt sterk bij splitsing van het materiaal in twee groepen. In de mariene groep is zelfs voor geen enkele component de amplitude statistisch significant. Bij de concentraties leidt splitsing in groepen er tevens toe dat de significantie van de coëfficiënt c voor de invloed van de neerslag sterk afneemt.

Voor de verklaring van de afzwakking van de cosinusterm bij splitsing in groepen is het nodig om het seizoenverloop van de neerslaghoeveelheid te bekijken. Figuur 3 laat zien dat de neerslaghoeveelheid in beide groepen een duidelijk seizoenverloop heeft. In de mariene groep bereikt de neerslagterm $c\sqrt{h_i}$ in model (4) zijn maximum in november, terwijl dat voor de continentale groep in juni het geval is. Blijkbaar komen in de eerste groep relatief veel najaarsregens voor, terwijl de tweede groep meer gedomineerd wordt door zomerbuien.

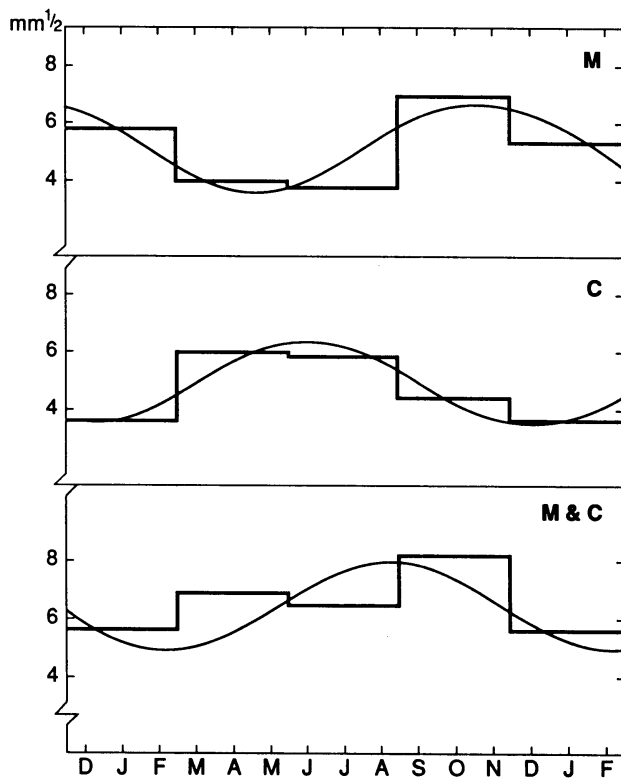


Fig. 3 Maandgemiddelden van de wortel van de neerslaghoeveelheid. De rechte lijnen hebben betrekking op het werkelijke gemiddelde van $\sqrt{h_i}$ in een bepaald kwartaal. De vloeiende lijn geeft de sinusoidale benadering van het seizoenverloop volgens vergelijking (C1) in Appendix C weer

De mariene groep levert een relatief kleine bijdrage tot de totale depositie (ongeveer 25%, zie par. 5) en ook zijn de verschillen tussen de seizoengemiddelden in deze groep klein ten opzichte van de gemiddelde depositie van alle regenwatermonsters. De mariene groep is daarom betrekkelijk onbelangrijk voor het seizoenverloop van de totale depositie. Bij de maandwaarden van de regenwatermonsters in de continentale groep bereikt de neerslaghoeveelheid zijn maximum op ongeveer dezelfde tijd als de consinusterm bij de maandwaarden van alle regenwatermonsters. Een belangrijk deel van het seizoenverloop van de deposities kan daarom verklaard worden door de neerslagterm in deze groep. Appendix C gaat verder in op het seizoenverloop van de depositie.

Voor sulfaat en ammonium is er nauwelijks een seizoenverloop van de concentratie in de mariene en de continentale groep. Zo is de gemiddelde sulfaatconcentratie het gehele jaar door ongeveer 20 $\mu\text{mol/l}$ voor de mariene regendagen, terwijl deze voor de continentale regendagen ongeveer 50 $\mu\text{mol/l}$ bedraagt. De relatief hoge sulfaatconcentraties in het voor- en najaar in de situatie (M&C) kunnen nu voor een belangrijk deel worden toegeschreven aan het feit dat in deze periode veel neerslag van continentale regendagen afkomstig is. Voor het najaar en de winter zijn de gemiddelde sulfaatconcentraties relatief laag vanwege de grotere bijdrage van de mariene regendagen. Bij hydronium en ammonium wordt dit effect versterkt door het seizoenverloop van de concentratie in de continentale groep.

Een belangrijk deel van de jaarlijkse gang in de concentraties en deposities kan dus verklaard worden door het feit dat de bijdrage van de verschillende brongebieden van seizoen tot seizoen verandert. Dit is in overeenstemming met de bevindingen van Davies e.a. (1990) omtrent het seizoenverloop van de sulfaatconcentraties in Eskdalemuir (Schotland). De meteorologische invloed op de concentraties en deposities gaat dus aanzienlijk verder dan de neerslaghoeveelheid waarmee de modellen (3) en (4) rekenen.

4 VERGELIJKING RESULTATEN ECN-GEGEVENS MET DIE UIT HET LANDELIJK MEETNET REGENWATERSAMENSTELLING

In het Landelijk Meetnet Regenwatersamenstelling (LMR) is tot april 1988 niet met wet-only, maar met open vangers gemeten. Een ander verschil met het ECN is dat de regenwatermonsters op maandbasis zijn geanalyseerd in plaats van op dagbasis. Verder is bij onderzoek naar trend

in de meetgegevens van het LMR over een aantal stations gemiddeld. In Buishand e.a. (1988) werd het regressiemodel (3) toegepast op de maandgemiddelde concentraties van 12 stations, terwijl in Buishand en Van Montfort (1989) over 8 stations werd gemiddeld. Voor de hier beschouwde componenten lopen de resultaten van beide publikaties weinig uiteen. In deze paragraaf is zoveel als mogelijk gebruik gemaakt van de resultaten in Buishand e.a. (1988) voor het 7-jarige tijdvak 1978-1984. Om een beter inzicht te krijgen in de oorzaak van de verschillen tussen de getallen in Buishand e.a. (1988) en de resultaten van de ECN-gegevens, is het regressiemodel (3) ook toegepast op de concentraties van het LMR-meetpunt Lelystad voor dezelfde periode als die waarin door het ECN gemeten werd.

In tabel 2 zijn de parameters van het seizoenverloop van de ECN-gegevens vergeleken met die van het LMR. Voor alle vier componenten is bij de ECN-gegevens de amplitude groter dan bij het landgemiddelde van de LMR-stations, terwijl de cosinusterm in model (3) later zijn maximum bereikt. Voor het LMR-meetpunt Lelystad komt bij hydronium de amplitude overeen met die van het landgemiddelde en is aanzienlijk lager dan die van de ECN-gegevens; bij de andere componenten wijkt de amplitude minder sterk af van die van de ECN-gegevens. Het gevonden seizoenverloop bij het LMR-meetpunt Lelystad is duidelijk verschoven ten opzichte van dat bij de ECN-gegevens.

Hoewel in tabel 2 bij hydronium de verschillen in fasehoek groot zijn, moet rekening worden gehouden met het feit dat de standaardafwijking van $\hat{\phi}_1$ voor de meetgegevens van het LMR groot is ($> 20^\circ$). Het maximum voor sulfaat valt bij de ECN-gegevens in mei. Een dergelijke vrij late piek voor sulfaat is ook gevonden door Berge (1988) voor drie stations in Noorwegen en door Skeffington (1984) en Martin en Barber (1984) voor twee meetpunten in Engeland. Daarentegen komt de positie van de piek voor de sulfaatconcentraties bij de LMR-gegevens overeen met die van de gegevens van het European Air Chemistry Network in Rohde en Granat (1984)*. Bij nitraat wijkt de waarde van $\hat{\phi}_1$ voor het LMR-meetpunt vrij sterk af van de waarden voor het landgemiddelde en de ECN-gegevens.

* Vanwege het effect van de neerslaghoeveelheid kan er een verschil zijn tussen het tijdstip waarvoor de concentratie maximaal is en de datum waarop de cosinusterm in vergelijking (3) zijn maximum bereikt. Het tijdstip van het maximum in Berge (1988) heeft betrekking op de geschatte fasehoek in een soortgelijk regressiemodel als vergelijking (3); bij Skeffington (1984), Martin en Barber (1984) en Rohde en Granat (1984) is gekeken naar de maand waarvoor de gemiddelde concentratie maximaal is. Voor de ECN-gegevens is de gemiddelde (volume-gewogen) sulfaatconcentratie $43.0 \mu\text{mol/l}$ voor de lentemaanden maart, april, mei en $46.3 \mu\text{mol/l}$ voor de zomermaanden juni, juli, augustus.

Tabel 2. Geschatte waarden van de paramters a_1 en ϕ_1 in het regressiemodel (3) bij een land-gemiddelde op basis van 12 stations in het Landelijk Meetnet Regenwatersamenstelling (LMR) voor het tijdvak 1978-1984 (Buishand e.a., 1988) en bij de gegevens van Lelystad voor het tijdvak 1983-1985. De waarden van \hat{a}_1 en $\hat{\phi}_1$ hangen niet af van de eenheden waarin C_i en h_i zijn uitgedrukt

Component	\hat{a}_1			$\hat{\phi}_1$ (°)		
	Landgemid- delde LMR	Lelystad LMR	Lelystad ECN	Landgemid- delde LMR	Lelystad LMR	Lelystad ECN
Hydronium	0.20	0.17*	0.50*	133	168*	186*
Ammonium	0.18	0.29	0.28	103	99	161
Nitraat	0.31	0.38	0.50	162	137	178
Sulfaat	0.10	0.20	0.31	115	133	169

* Met weglating van februari 1983 (uitschieter bij ECN-gegevens) en juli 1983, november 1983 (regenwatermonster van de LMR regenvanger is basisch in deze maanden).

Het verschil in amplitude bij hydronium en de verschuiving van het tijdstip van het maximum van de cosinusterm in vergelijking (3) bij de andere componenten wijst op een effect van meetmethoden. Behalve het verschil tussen wet-only en open vangers, zijn er verschillen doordat bij de ECN-vangers de neerslagmonsters voor regendagen met een neerslaghoeveelheid kleiner dan 0.5 mm niet geanalyseerd werden en de bemonstering gestaakt werd op dagen met vorst en sneeuwval.

In tabel 3 zijn de geschatte waarden van de residuele standaardafwijking gegeven. De waarde van $\hat{\sigma}_e$ is voor hydronium bij de ECN-gegevens kleiner dan bij de LMR-gegevens. Een verlaging van σ_e leidt tot een betere detecteerbaarheid van langjarige trends. De lagere residuele standaardafwijking bij de ECN-gegevens is waarschijnlijk het gevolg van het gebruik van wet-only vangers in plaats van open vangers, en van een vrij stringente controle op uitschieters. Bij de andere drie componenten komt de waarde van $\hat{\sigma}_e$ voor de ECN-gegevens redelijk overeen

met die voor het LMR-meetpunt Lelystad. Bij het landgemiddelde van de LMR-stations is de residuele standaardafwijking altijd lager dan bij het individuele LMR-meetpunt Lelystad. Naast een variantie verlagend effect door het middelen over stations, wordt de reductie van $\hat{\sigma}_e$ voor het landgemiddelde mede veroorzaakt door het opnemen van de neerslaghoeveelheid als een verklarende variabele in het regressiemodel (Buishand, 1984). Zonder deze verklarende variabele zou het middelen over stations bij ammonium, nitraat en sulfaat slechts een beperkt effect hebben vanwege de vrij sterke ruimtelijke correlatie (Buishand, 1986).

Tabel 3. Geschatte waarden van de residuele standaardafwijking σ_e in het regressiemodel (3) bij een landgemiddelde op basis van 12 LMR stations voor het tijdvak 1978-1984 en bij de gegevens van Lelystad voor het tijdvak 1983-1985

Component	Landgemiddelde LMR	Lelystad LMR	Lelystad ECN
Hydronium	0.51	0.80*	0.36*
Ammonium	0.19	0.26	0.28
Nitraat	0.18	0.26	0.31
Sulfaat	0.15	0.22	0.26

* Als in tabel 2 met weglating van februari, juli en november 1983.

Droge depositie kan bij open vangers tot extreem hoge concentraties leiden in regenwatermonsters van relatief droge maanden. De concentraties bij de open vangers in het LMR vertonen dan ook een aanzienlijk grotere variatie dan de concentraties bij de wet-only vangers van het ECN. Een gevolg hiervan is dat in regressiemodel (3) de parameter c en de multiële correlatiecoëfficiënt R lager zijn dan bij de LMR-gegevens, zoals tabel 4 laat zien. Hydronium is in deze tabel buiten beschouwing gelaten omdat zowel voor de ECN- als de LMR-gegevens de parameter c bij deze component niet significant van nul verschilt. Vergelijking van de resultaten in Appendix A met die in Buishand en Van Montfort (1989) leert dat voor de deposities het tegenovergestelde geldt.

Tabel 4. Geschatte waarden van de parameter c in het regressiemodel (3) en het kwadraat van de multiële correlatiecoëfficiënt R^2 bij een landgemiddelde op basis van 12 LMR stations voor het tijdvak 1978-1984 en bij de ECN-gegevens van Lelystad voor het tijdvak 1983-1985

Component	$\hat{c}(\text{mm}^{1/2})$		R^2	
	Landgemid- delde LMR	Lelystad ECN	Landgemid- delde LMR	Lelystad ECN
Ammonium	7.59	2.21	0.75	0.53
Nitraat	7.30	1.38	0.84	0.62
Sulfaat	6.89	2.14	0.81	0.59

Vergelijking van de getallen in de laatste kolommen van de tabellen A1 en A2 in Appendix A leert dat bij de logaritmen van de concentraties de residuele standaardafwijking in het algemeen wat lager is dan bij de logaritmen van de deposities. Een trend is daardoor beter te detecteren bij de logaritmen van de concentraties. Voor de gegevens van het LMR waren de regressiemodellen (3) en (4) daarentegen ongeveer gelijkwaardig na weglating van één zeer uitzonderlijke maand (Buishand en Van Montfort, 1989).

5 MOGELIJKHEDEN OM DE DETECTEERBAARHEID VAN LANGJARIGE TRENDS TE VERBETEREN DOOR SPLITSING IN GROEPEN

De detecteerbaarheid van langjarige trends wordt bepaald door de waarde van de residuele standaardafwijking σ_e . De grootte van de trend die met 80% kans gedetecteerd kan worden in een reeks van vijf jaren wordt gegeven door (Buishand e.a., 1988):

$$b_{80} = 0.25 \sigma_e \quad (5)$$

Dus een lage waarde van σ_e leidt tot een lage waarde van b_{80} , hetgeen wil zeggen dat een relatief kleine trend met 80% kans gedetecteerd kan worden.

Vergelijking (5) geldt voor een tweezijdige t-toets (onbetrouwbaarheidsdrempel 5%) bij een volledige reeks, d.w.z. een reeks waarin voor elke maand een waarde voor de concentratie of depositie is opgegeven. In de mariene groep komen de meeste ontbrekende maanden voor. In Appendix D wordt afgeleid dat voor deze groep de waarde van b_{80} ongeveer 15% hoger is dan uit vergelijking (5) volgt.

Een verbetering van de detecteerbaarheid van langjarige trends zou zonder meer bereikt zijn indien σ_e binnen de twee groepen kleiner zou zijn dan σ_e voor het gehele bestand. Een dergelijke situatie doet zich voor als de cluster-analyse resulteert in vrij homogene groepen, die onderling sterk verschillen. Bij de hier beschouwde onderverdeling komt de waarde van σ_e binnen de groepen echter ongeveer overeen met die zonder splitsing. Dit is bij alle vier componenten het geval. In een dergelijke situatie is alleen iets te verdienen als een trend zich voordoet in één van de groepen, terwijl hij in de andere groep verwaarloosbaar is of tegengesteld verloopt. Het volgende voorbeeld illustreert dit.

Bij de logaritmen van de ammoniumconcentraties is $\sigma_e = 0.25$ voor de mariene groep. Uit vergelijking (5) volgt dan dat $b_{80} = 0.063$. Vermenigvuldiging met een factor 1.15, vanwege het ontbreken van mariene regendagen in sommige maanden, levert een trend van 7.2% per jaar op. Een trend van deze grootte kan dus met 80% kans gedetecteerd worden in de regenwatermonsters van de mariene groep. Stel nu dat er geen trend is in de continentale groep. Daar de mariene groep een bijdrage levert van 27% in de totale depositie (tabel 5), komt een trend van 7.2% per jaar in deze groep neer op een trend van 1.9% per jaar in de totale depositie. Bij de toepassing van het regressiemodel op de concentraties van alle regenwatermonsters is σ_e gelijk aan 0.28. Voor deze waarde van σ_e kan uitgerekend worden dat een trend van 1.9% per jaar in een reeks van vijf jaren slechts met 10% kans gedetecteerd kan worden.

Voor de ammoniumconcentraties in de continentale groep ($\sigma_e = 0.27$) geldt dat een trend van 6.8% per jaar met 80% kans gedetecteerd kan worden in een reeks van vijf jaren. Bij afwezigheid van een trend in de mariene groep leidt deze trend tot een verandering van 5% per jaar in de gemiddelde concentratie van alle regenwatermonsters. Een dergelijke trend heeft een detectiekans van ongeveer 50% indien het model toegepast wordt op alle regenwatermonsters. De winst door het opdelen in groepen is hier dus kleiner dan bij een sterke trend in de mariene

Tabel 5. Totale neerslaghoeveelheid (mm) en depositie ($\mu\text{mol}/\text{m}^2$)

	Marien	Continentaal	Alle monsters	% marien
Neerslaghoeveelheid	821	906	1726	48
Hydroniumdepositie	13300	39800	53100	25
Ammoniumdepositie	35300	97900	133200	27
Nitraatdepositie	16300	52000	68300	24
Sulfaatdepositie	16200	45300	61500	26

groep. Het is zelfs zo dat in dit geval net zo goed met het landgemiddelde uit het LMR ($\sigma_e = 0.19$) gewerkt kan worden. Met dit landgemiddelde kan in een reeks van vijf jaren een trend van 5% per jaar met ruim 80% kans gedetecteerd worden.

Daar ook voor de drie andere componenten (hydronium, sulfaat, nitraat) het aandeel van de mariene groep ongeveer 25% bedraagt (zie tabel 5) en σ_e ongeveer dezelfde waarde heeft als bij toepassing van het regressiemodel op alle regenwatermonsters, geldt ook voor deze componenten dat de huidige opdeling in groepen het meeste oplevert indien zich een sterk afwijkende trend in de mariene groep voordoet. Het is echter de vraag of dergelijke trends te verwachten zijn.

Een cluster-analyse op basis van de concentraties heeft als bezwaar dat de neerslag binnen een groep uit verschillende gebieden afkomstig kan zijn. In de continentale groep werden bijvoorbeeld dagen met aanvoer van sterk verontreinigde mariene luchtmassa's aangetroffen (par. 2). Voor een trendonderzoek is dit niet wenselijk, daar de grootte van een trend meestal aan de oorsprong van de neerslag gekoppeld zal zijn. Bij een onderzoek over de chemie van de neerslag in het oostelijk deel van Noord-Amerika voerden Fernau en Samson (1990) een cluster-analyse uit op gegevens van de bovenlucht. De clusters bestonden uit groepen van dagen met ongeveer dezelfde stromingspatronen. In de meeste gevallen was de spreiding in de concentraties en deposities binnen een cluster groot.

Een onderverdeling op basis van bovenluchtgegevens leidt in principe tot grotere ver-

schillen in de chemische samenstelling binnen de groepen dan een cluster-analyse op basis van de concentraties, omdat de laatste de verschillen in chemische samenstelling binnen een groep zo klein mogelijk tracht te maken. Men mag dan ook niet verwachten dat een nadere studie naar de onderverdeling in groepen tot een aanzienlijk kleinere spreiding van de concentraties binnen de groepen zal leiden. De mogelijkheden om met een splitsing van het databestand in clusters de detecteerbaarheid van langjarige trends te verbeteren blijven daarom beperkt.

6 CONCLUSIES

Door middel van cluster-analyse is het mogelijk de dagwaarden van de chemische samenstelling van de neerslag in groepen onder te verdelen. Bij de ECN-gegevens van Lelystad waren twee duidelijk verschillende groepen te onderscheiden: een groep met neerslagmonsters van vrijwel uitsluitend mariene oorsprong en een groep die nagenoeg alle neerslagmonsters bij continentale luchtaanvoer omvatte. Voor de in dit onderzoek bestudeerde componenten (hydronium, ammonium, nitraat en sulfaat) was de gemiddelde concentratie in de continentale groep een factor 2 à 3 maal zo hoog als die in de mariene groep.

Alleen bij hydronium werd een significant stijgende trend gevonden. Deze trend werd echter voor een belangrijk deel veroorzaakt door één sterk afwijkende maand. Zowel voor de regenwatermonsters in de mariene als die in de continentale groep was voor alle beschouwde componenten de amplitude a_1 van de cosinusterm in de modellen (3) en (4) aanzienlijk kleiner dan bij de maandwaarden van alle regenwatermonsters. De relatief grote amplitude in dit laatste geval kan voor een belangrijk deel verklaard worden door het seizoenverloop van de neerslaghoeveelheid binnen de continentale en de mariene groep. Zo levert de relatief hoge neerslaghoeveelheid in de continentale groep gedurende het voorjaar en de zomer een belangrijke bijdrage tot de hoge concentraties in deze seizoenen.

Bij de ECN-gegevens verschilt het seizoenverloop van de concentraties van die bij de meetgegevens van het Landelijk Meetnet Regenwatersamenstelling (LMR). Dit moet niet alleen worden toegeschreven aan het gebruik van wet-only vangers op het ECN-meetpunt, maar ook aan het feit dat bij deze vangers de regenwatermonsters niet geanalyseerd werden bij kleine neerslaghoeveelheden en de bemonstering gestaakt werd bij vorst en sneeuwval. Door het gebruik van wet-

only vangers en een stringente controle op verdachte waarnemingen kan het tijdsverloop van de concentraties van de hydroniumionen veel beter verklaard worden dan bij de meetgegevens van het LMR.

Daar de spreiding van de concentraties binnen de groepen groot is, zijn er slechts geringe mogelijkheden om door bemonstering op dagbasis de detecterebaarheid van langjarige trends te verbeteren. Er zijn perspectieven in situaties dat een systematische verandering in de ene groep sterk verschilt van die in de andere groep.

LITERATUUR

- Asman, W.A.H., J. Slanina and J.H. Baard, 1981. Meteorological interpretation of the chemical composition of rain-water at one measuring site. *Water, Air and Soil Pollution*, 16, 159-175.
- Atkinson, A.C., 1985. *Plots, Transformations and Regression - An introduction to graphical methods of diagnostic regression analysis*. Clarendon Press, Oxford.
- Berge, E., 1988. Time-trends of sulfate and nitrate in precipitation in Norway (1972-1982). *Atmospheric Environment*, 22, 333-338.
- Buishand, T.A., 1984. Detectie van trends in de chemische samenstelling van de neerslag. KNMI memorandum FM 84-42 (intern document).
- Buishand, T.A., 1986. Ruwe schatting van de grootte van te detecteren trends in de chemische samenstelling van de neerslag door middel van regressie-analyse. KNMI Technisch Rapport TR-83.
- Buishand, T.A., G.T. Kempen, A.J. Frantzen, H.F.R. Reijnders and A.J. van den Eshof, 1988. Trend and seasonal variation of precipitation chemistry data in the Netherlands. *Atmospheric Environment*, 22, 339-348.
- Buishand, T.A. and M.A.J. van Montfort, 1989. Changes in the chemical composition of atmospheric precipitation in the Netherlands during the period 1978-1987. Vakgroep Wiskunde Landbouwniversiteit, Technical Note 89-02.
- Davies, T.D., G. Farmer and R.J. Barthelmie, 1990. Use of simple daily atmospheric circulation types for the interpretation of precipitation composition at a site (Eskdalemuir) in Scotland, 1978-1984. *Atmospheric Environment*, 24A, 63-72.

- Fernau, M.E. and P.J. Samson, 1990. Use of cluster analysis to define periods of similar meteorology and precipitation chemistry in eastern North America. Part II: Precipitation patterns and pollutant deposition. *Journal of Applied Meteorology*, 29, 751-761.
- Janssen, A.J., 1988. Enkele aspecten van jaarlijkse variaties in meteorologie en luchtverontreiniging. ECN rapport 88-129.
- Kempen, G.T., T.A. Buishand, H.F.R. Reijnders, A.J. Frantzen en A.J. van den Eshof, 1986. Tijdreeksanalyse van de meetgegevens van het Landelijk Meetnet Regenwatersamenstelling. KNMI publikatie 170, RIVM rapport nr. 218203002.
- Martin, A. and F.R. Barber, 1984. Acid gases and acid in rain monitored for over 5 years in rural east-central England. *Atmospheric Environment*, 18, 1715-1724.
- Rohde, H. and L. Granat, 1984. An evaluation of sulfate in European precipitation. *Atmospheric Environment*, 18, 2627-2639.
- Skeffington, R.A., 1984. The chemistry of bulk precipitation at a site in southeast England - I. Small-scale spatial variation, frequency distributions and variation with time. *Atmospheric Environment*, 18, 1683-1693.
- Slanina, J., J.H. Baard, W.L. Zijp and W.A.H. Asman, 1983. Tracing the sources of the chemical composition of precipitation by cluster analysis. *Water, Air and Soil Pollution*, 20, 41-45.
- Slanina, J., J.J. Möls and J.H. Baard, 1990. The influence of outliers on results of wet deposition measurements as a function of measurement strategy. *Atmospheric Environment*, 24 A, 1843-1860.
- Weisberg, S., 1985. *Applied Linear Regression*, 2nd edition. John Wiley, New York.

APPENDIX A

Resultaten van de regressie-analyse

In tabel A1 zijn de resultaten voor de logaritmen van de concentraties samengevat, terwijl in tabel A2 dit voor de logaritmen van de deposities is gedaan. De gemiddelden $\overline{\ln C}$ en $\overline{\ln D}$ hebben steeds betrekking op het rekenkundige gemiddelde. Een sterretje bij een geschatte parameter geeft aan dat de waarde significant is bij een onbetrouwbaarheidsdrempel van 5%; twee sterretjes geven aan dat de waarde tevens significant is bij een onbetrouwbaarheidsdrempel van 1%. De significantie van de parameters a_1 en b werd getoetst met een F-toets, de hypothese $c = 0$ werd getoetst met een eenzijdige t-toets. Bij deze laatste toets luidt het alternatief $c > 0$ (negatieve correlatie tussen concentratie en neerslaghoeveelheid; positieve correlatie tussen depositie en neerslaghoeveelheid).

Tabel A1. Geschatte parameters bij het regressiemodel (3) met concentraties in $\mu\text{mol/l}$ en neerslaghoeveelheden in mm

(M) Marien (24 maanden)

Component	$\overline{\ln C}$	R ²	\hat{a}_1	$\hat{\phi}_1(^{\circ})$	\hat{b}	\hat{c}	$\hat{\sigma}_e$
Hydronium	2.75	0.02	0.08	130	0.001	-0.20	0.42
Ammonium	3.85	0.40	0.08	238	0.114	0.96*	0.25
Nitraat	3.04	0.10	0.04	178	0.073	0.24	0.27
Sulfaat	2.99	0.04	0.05	8	-0.004	0.29	0.23

(C) Continentaal (31 maanden; bij de resultaten van hydronium in de tweede regel is februari 1983 weggelaten)

Component	$\overline{\ln C}$	R ²	\hat{a}_1	$\hat{\phi}_1(^{\circ})$	\hat{b}	\hat{c}	$\hat{\sigma}_e$
Hydronium	3.54	0.29	0.59*	202	0.191	1.33	0.63
	3.63	0.23	0.42*	197	0.055	1.39	0.48
Ammonium	4.70	0.25	0.02	177	-0.106	0.96	0.27
Nitraat	4.43	0.35	0.33*	201	-0.095	1.17	0.33
Sulfaat	3.89	0.28	0.15	190	-0.102	1.18*	0.29

(M&C) Alle regendagen (33 maanden; bij de resultaten van hydronium in de tweede regel is februari 1983 weggelaten)

Component	$\overline{\ln C}$	R ²	\hat{a}_1	$\hat{\phi}_1(^{\circ})$	\hat{b}	\hat{c}	$\hat{\sigma}_e$
Hydronium	3.27	0.48	0.60**	184	0.222*	-0.28	0.51
	3.35	0.52	0.52**	180	0.124	0.62	0.38
Ammonium	4.40	0.53	0.28**	161	0.032	2.21**	0.28
Nitraat	3.67	0.62	0.50**	178	0.036	1.38*	0.31
Sulfaat	3.61	0.59	0.31**	169	-0.007	2.14**	0.26

Tabel A2. Geschatte parameters bij het regressiemodel (4) met deposities in $\mu\text{mol}/\text{m}^2$ en neerslaghoeveelheden in mm

(M) Marien (24 maanden)

Component	$\overline{\ln D}$	R ²	\hat{a}_1	$\hat{\phi}_1(^{\circ})$	\hat{b}	\hat{c}	$\hat{\sigma}_e$
Hydronium	5.88	0.80	0.04	56	0.032	0.39**	0.53
Ammonium	6.98	0.88	0.09	277	0.129	0.34**	0.34
Nitraat	6.17	0.85	0.06	317	0.083	0.36**	0.40
Sulfaat	6.12	0.89	0.11	353	0.023	0.37**	0.37

(C) Continentaal (31 maanden; bij de resultaten van hydronium in de tweede regel is februari 1983 weggelaten)

Component	$\overline{\ln D}$	R ²	\hat{a}_1	$\hat{\phi}_1(^{\circ})$	\hat{b}	\hat{c}	$\hat{\sigma}_e$
Hydronium	6.69	0.75	0.48	211	0.173	0.44**	0.61
	6.81	0.79	0.34	206	0.049	0.42**	0.46
Ammonium	7.85	0.84	0.02	283	-0.088	0.39**	0.30
Nitraat	7.14	0.84	0.31*	206	-0.079	0.38**	0.35
Sulfaat	7.04	0.87	0.09	209	-0.098	0.40**	0.29

(M&C) Alle regendagen (33 maanden; bij de resultaten van hydronium in de tweede regel is februari 1983 weggelaten)

Component	$\overline{\ln D}$	R ²	\hat{a}_1	$\hat{\phi}_1(^{\circ})$	\hat{b}	\hat{c}	$\hat{\sigma}_e$
Hydronium	7.01	0.77	0.65**	186	0.294*	0.36**	0.56
	7.12	0.77	0.55**	182	0.178	0.31**	0.44
Ammonium	8.14	0.79	0.29**	167	0.068	0.25**	0.32
Nitraat	7.40	0.80	0.53**	183	0.081	0.27**	0.37
Sulfaat	7.34	0.83	0.33**	172	0.032	0.26**	0.30

APPENDIX B

Gestudentiseerde residuen

Voor februari 1983 werd in par. 3 bij de concentratie en depositie van de hydroniumionen de waarde van het gestudentiseerde residu gegeven. Het begrip gestudentiseerd residu wordt in deze appendix nader toegelicht aan de hand van het volgende eenvoudige regressiemodel:

$$y_i = a + bx_i + e_i, \quad i = 1, \dots, n. \quad (\text{B1})$$

Evenals in de modellen (3) en (4) wordt verondersteld dat de residuen e_1, e_2, \dots, e_n onafhankelijke, normaal verdeelde kansvariabelen zijn met verwachting nul en alle met dezelfde standaardafwijking σ_e .

De onbekende regressiecoëfficiënten a en b kunnen geschat worden met behulp van de kleinste kwadratenmethode:

$$\hat{b} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}; \quad \hat{a} = \bar{y} - \hat{b} \bar{x}, \quad (\text{B2})$$

met \bar{x} en \bar{y} de gemiddelden van de x_i 's en de y_i 's. Met behulp van deze geschatte waarden kan bij elke y_i een residu \hat{e}_i berekend worden:

$$\hat{e}_i = y_i - \hat{a} - \hat{b} x_i. \quad (\text{B3})$$

De variantie van deze berekende residuen is kleiner dan de variantie σ_e^2 van de onbekende e_i 's. Er geldt [Weisberg (1985), vergn. (5.6) en (5.14)]:

$$\text{var}(\hat{e}_i) = \sigma^2(\hat{e}_i) = \sigma_e^2 \left[1 - 1/n - (x_i - \bar{x})^2 / S_{xx} \right], \quad (\text{B4})$$

$$\text{met } S_{xx} = \sum_{j=1}^n (x_j - \bar{x})^2.$$

Indien x_i sterk van \bar{x} verschilt, dan is de term $(x_i - \bar{x})^2 / S_{xx}$ niet verwaarloosbaar en zal $\sigma(\hat{e}_i)$ aanzienlijk kleiner zijn dan σ_e . Een schatting van $\sigma(\hat{e}_i)$ kan worden verkregen uit:

$$\hat{\sigma}^2(\hat{e}_i) = \hat{\sigma}_e^2 \left[1 - 1/n - (x_i - \bar{x})^2 / S_{xx} \right], \quad \text{met} \quad \hat{\sigma}_e^2 = \sum_{i=1}^n \hat{e}_i^2 / (n-2). \quad (\text{B5})$$

Om te beoordelen of y_i een uitschieter is, wordt \hat{e}_i gedeeld door $\hat{\sigma}(\hat{e}_i)$:

$$r_i = \hat{e}_i / \hat{\sigma}(\hat{e}_i). \quad (\text{B6})$$

De grootte r_i wordt vrij algemeen als het gestudentiseerde residu ("studentized residual") aangeduid. Er zijn echter auteurs (Atkinson, 1985) die r_i een gestandaardiseerd residu noemen.

Indien de i -de waarneming een uitschieter is, dan zal niet alleen \hat{e}_i hoog uitvallen, maar ook $\hat{\sigma}(\hat{e}_i)$ vanwege de invloed van \hat{e}_i op $\hat{\sigma}_e^2$. Er is een ander gestudentiseerd residu, het "studentized deletion residual" of "jackknife residual" waarbij de i -de waarneming bij het schatten van de regressiecoëfficiënten en de residuele standaardafwijking σ_e buiten beschouwing gelaten wordt. In plaats van \hat{e}_i in vergelijking (B3) wordt uitgegaan van:

$$\hat{e}_i^* = y_i - \hat{a}_{(i)} - \hat{b}_{(i)} x_i, \quad (\text{B7})$$

waarbij de index (i) aangeeft dat de regressiecoëfficiënten zonder de i -de waarneming geschat zijn. De variantie van \hat{e}_i^* wordt gegeven door:

$$\sigma^2[\hat{e}_i^*] = \sigma_e^2 \left\{ 1 + 1/(n-1) + \left[n(x_i - \bar{x}) / (n-1) \right]^2 / S_{xx(i)} \right\}, \quad (\text{B8})$$

met $S_{xx(i)}$ analoog aan S_{xx} , echter zonder de i -de waarneming. Een schatting van $\sigma^2[\hat{e}_i^*]$ wordt verkregen door $\hat{\sigma}_e^2$ te schatten op basis van de kwadraatsom van de $(n-1)$ residuen $\hat{e}_j^* = y_j - \hat{a}_{(i)} - \hat{b}_{(i)} x_j$, $j \neq i$. Het "studentized deletion residual" wordt dan gegeven door:

$$r_i^* = \hat{e}_i^* / \hat{\sigma}[\hat{e}_i^*]. \quad (\text{B9})$$

Er bestaat een monotone relatie tussen de r_i 's en de r_i^* 's [Atkinson (1985), verg. (3.13); Weisberg (1985), verg (5.18)]. Voor een multipel regressiemodel met p onbekende regressie-coëfficiënten luidt deze relatie:

$$r_i^* = r_i \left(\frac{n - p - 1}{n - p - r_i^2} \right)^{1/2} . \quad (B10)$$

Voor de regressiemodellen (3) en (4) in par. 3 is p gelijk aan 5. Bij hydronium neemt r_i^* in de gevallen (C) en (M&C) voor februari 1983 waarden aan tussen -4.7 en -4.1. Uit tabel E in Weisberg (1985) blijkt dat deze waarden net significant zijn bij een onbetrouwbaarheidsdrempel van 1%. Februari 1983 is derhalve voor hydronium een situatie apart.

APPENDIX C

Een benadering voor het seizoenverloop van de depositie

In figuur 3 is het seizoenverloop van de wortel van de neerslaghoeveelheid benaderd door de sinusoïde:

$$\sqrt{h_i} \approx \mu^* + a_1^* \cos(2\pi i/N - \phi_1^*) . \quad (C1)$$

Substitutie van deze benadering van $\sqrt{h_i}$ in vergelijking (4) leidt tot de volgende uitdrukking voor de gemiddelde depositie μ_m in maand m:

$$\mu_m \approx \exp \left[\mu_{\ln D} + 1/2 \sigma_e^2 + a_1 \cos(2\pi m/N - \phi_1) + c a_1^* \cos(2\pi m/N - \phi_1^*) \right],$$

$$m = 1, \dots, 12 \quad (C2)$$

waarbij $\mu_{\ln D} = a_0 + c \mu^*$. De afleiding van deze uitdrukking gaat analoog aan die van vergelijking (C18) in Kempen e.a. (1986). Een eventuele langjarige trend is in de uitdrukking voor μ_m buiten beschouwing gelaten.

Voor sulfaat zijn in figuur C1 de waarden van μ_m vergeleken met de werkelijke gemiddelde deposities. Het seizoenverloop in de depositie van deze component kan in de gevallen (M) en (C) vrijwel geheel worden toegeschreven aan het seizoenverloop van de neerslaghoeveelheid. Voor het geval (M&C) bereikt μ_m zijn maximum in de zomer, terwijl de werkelijke depositie maximaal is in het voorjaar. Vergelijking (C2) leidt in dit geval tot een onderschatting van de gemiddelde depositie van ongeveer 30% in het voorjaar, terwijl in de zomer de gemiddelde depositie met eenzelfde percentage overschat wordt. Dit wordt hoofdzakelijk veroorzaakt door het feit dat de benadering (C1) voor het seizoenverloop van de neerslaghoeveelheid in het geval (M&C) slecht voldoet, zoals uit figuur 3 blijkt. De coëfficiënt a_1^* in vergelijking (C1) verschilt in dit geval niet significant van nul (onbetrouwbaarheidsdrempel 5%). Eenzelfde verschijnsel treedt op bij de depositie van de andere componenten. De waarde van de parameter c is immers in het geval (M&C) voor deze componenten van dezelfde orde van grootte als die voor sulfaat, zie

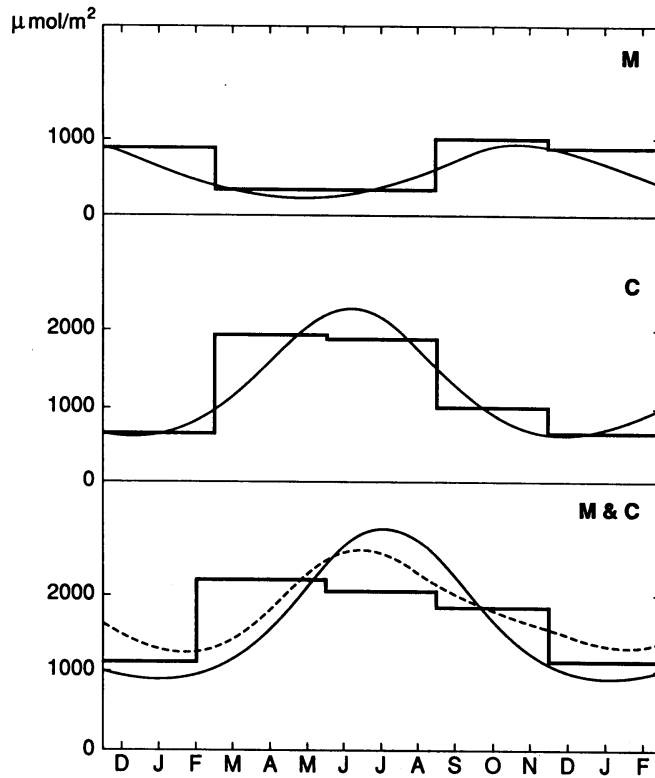


Fig. C1 Gemiddelde maandelijkse depositie van sulfaat. De rechte lijnen geven voor elk kwartaal de werkelijke maandgemiddelde depositie weer. De volle, vloeiende lijnen hebben betrekking op de benadering (C2) van μ_m ; de streeplijn in het geval (M&C) is verkregen door de benaderingen van μ_m voor de mariene en continentale regendagen bij elkaar op te tellen

tabel A2. Een betere benadering voor het seizoenverloop van de depositie wordt verkregen door de waarden van μ_m in de gevallen (M) en (C) bij elkaar op te tellen.

APPENDIX D

Het effect van ontbrekende maanden op de detecteerbaarheid van langjarige trends

Bij de modellen (3) en (4) wordt de grootte van de detecteerbare trend hoofdzakelijk bepaald door de variantie van de regressiecoëfficiënt \hat{b} . Voor deze variantie geldt bij benadering:

$$\text{var } \hat{b} \approx \frac{N^2 \sigma_e^2}{\Sigma(i-\bar{i})^2}. \quad (\text{D1})$$

De som in de noemer van deze uitdrukking loopt over alle maandindices van de concentratie (of depositie); \bar{i} is het gemiddelde van deze maandindices.

Beschouw nu een tijdvak met een lengte van K maanden, waarbij $K = k * n$. Indien er geen ontbrekende maanden in dit tijdvak zijn, dan is $\Sigma(i-\bar{i})^2$ gelijk aan:

$$S = \frac{1}{12} K(K^2-1) \approx \frac{1}{12} K^3 = \frac{1}{12} k^3 n^3. \quad (\text{D2})$$

Stel nu dat n oneven is en dat in elk van de k deeltijdvakken van n maanden de waarde voor de middelste maand ontbreekt. De som in de noemer van (D1) is dan gelijk aan:

$$S^* = \frac{1}{12} \left[kn(k^2 n^2 - 1) - kn^2(k^2 - 1) \right] \approx \frac{1}{12} \left[k^3 n^3 - k^3 n^2 \right]. \quad (\text{D3})$$

Het ontbreken van deze maanden leidt derhalve tot een toename van $\text{var}(\hat{b})$ met een factor:

$$S/S^* \approx k^3 n^3 / (k^3 n^3 - k^3 n^2) = 1 / (1 - 1/n). \quad (\text{D4})$$

Nu is $1/n$ gelijk aan de fractie f van de maanden waarvoor geen waarneming aanwezig is. Een redelijke benadering voor de variantie van \hat{b} in het geval van ontbrekende maanden lijkt derhalve:

$$\text{var}(\hat{b}) \approx 12 N^2 \sigma_e^2 / \left[K^3 (1-f) \right]. \quad (\text{D5})$$

Voor de regenwatermonsters in de mariene groep blijkt deze benadering goed te voldoen. In deze groep is voor ongeveer 25% van de maanden geen regenwatermonster aanwezig. De standaardafwijking van \hat{b} is daardoor een factor $\sqrt{1/(1-f)} = 1.15$ hoger dan die bij een volledige reeks. Daar de waarde van b_{80} vrijwel evenredig is aan de standaardafwijking van \hat{b} , moet de uitkomst van vergelijking (5) voor de mariene groep met eenzelfde bedrag vermenigvuldigd worden.