ministerie van verkeer en waterstaat

koninklijk nederlands meteorologisch instituut

ruwe schatting van de grootte van te detecteren trends in de chemische samenstelling van de neerslag door middel van regressie-analyse

TA Buishand

technische rapporten TR-83

KNMI

T.A. Buishand (1986): "Rough estimates of the magnitude of detectable trends in the chemical composition of precipitation by means of regression analysis."
Royal Netherlands Meteorological Institute (KNMI), Techn. Rep. TR.- 83

Abstract

Regression analysis can be used to detect changes in the chemical composition of precipitation. The most simple model for this purpose reads:

$$y_i = a + b\, i/N + e_i, \qquad i = 1,\dots, JN$$

with y_i: concentration (or logarithm of concentration) for period i,
e_i: a disturbance term with zero mean and variance σ_e^2,
N : the number of precipitation samples in a year,
J : number of years.

The regression coefficient b denotes the expected annual change in concentration (or the relative mean annual change when logarithms are taken). Testing the null hypothesis b = 0 against b ≠ 0 can be done by Student's t-test. The probability of rejecting the null hypothesis increases with $|b|$. The quantity which is considered in this report is the value of $|b|$ for which the null hypothesis is rejected with a probability of 80%. Estimates of this detectable trend b_{80} were obtained from published statistical data on precipitation chemistry in The Netherlands. For the significance level α a value of 5% was chosen.

For monthly data over a period of 5 years the detectable trend b_{80} ranges from about 10% per year (nitrate and sulfate) to slightly over 20% per year (magnesium, zinc, chloride and sodium). If instead of monthly data from a single station monthly averages from a national network of 12 stations are considered, then b_{80} is about 10% for all main components. A denser network does not result in a further reduction of b_{80} due to increased cross correlation between the data from neighbouring sites.

The magnitude of the detectable trend b_{80} strongly depends on the record length (is roughly inversely proportional with $J^{3/2}$). For fixed J a further reduction of b_{80} requires multiple regression models with meteorological information and a systematic annual cycle as extra explanatory variables.

RUWE SCHATTING VAN DE GROOTTE VAN TE DETECTEREN TRENDS IN DE CHEMISCHE SAMENSTELLING VAN DE NEERSLAG DOOR MIDDEL VAN REGRESSIE-ANALYSE

T.A. Buishand

Inhoud

1. Inleiding

In een eerder memorandum (Buishand, 1984) werd ingegaan op de statistische problematiek rond het vaststellen van trends in de chemische samenstelling van de neerslag. Aangegeven werd op welke dingen men bij de statistische analyse moet letten om een zo groot mogelijke detectiekans te krijgen.
Een vraag blijft nog welke trends in de concentratie (of depositie) van de verschillende componenten bij een bepaalde netwerkconfiguratie gedetecteerd kunnen worden met een goede statistische methode. Een antwoord op deze vraag wordt in dit rapport gegeven. Hierbij is uitgegaan van gegevens van recente onderzoeken in Nederland te weten:

- 5 jaar metingen van de chemische samenstelling van de neerslag (Ridder, 1983). Hierin zijn enige beschouwingen over trend opgenomen.
- Trend-analyse van de chemische samenstelling van de neerslag op basis van gegevens van het RID-VEWIN meetnet (Aldenberg, 1986).
- Optimalisatie landelijk meetnet regenwaterkwaliteit (Van Egmond et al., 1985).

In het navolgende zullen de gegevens uit deze drie onderzoeken afzonderlijk worden behandeld. Daarna zullen de resultaten onderling vergeleken worden. Hierbij zal tevens worden aangegeven in hoeverre het aantal meetstations en meetfouten van invloed zijn.

2. Onderzoek Ridder betreffende de chemische samenstelling van de neerslag op basis van 5 jaar metingen

Ridder (1983) ging uit van de jaargemiddelden van de concentraties (uit maandconcentraties, gewogen met de maandsom van de neerslag) van de waarnemingsstations van het KNMI-RIV netwerk voor het tijdvak 1978-82. Voor elk station afzonderlijk werd een toets op trend uitgevoerd voor de componenten F, NO_3 en SO_4. Bij deze toets werd gebruik gemaakt van het volgende regressiemodel:

$$y_i = a + b\, i + e_i\,, \qquad i = 1,\ldots, J \qquad (1)$$

met y_i: concentratie in jaar i,
e_i: een storingsterm met verwachting nul,
J : het aantal jaren.

Als er een trend is dan geeft b de verandering aan in de verwachtingswaarde van de y_i's na één jaar. Immers

$$E(y_{i+1}) - E(y_i) = b. \tag{2}$$

Met behulp van Student's t-toets werd de hypothese b = 0 (geen trend) getoetst met als alternatief b ≠ 0 (een stijging of daling in de gemiddelde concentratie). Voor de onbetrouwbaarheid α werd een waarde van 5% gekozen.

Indien b enigszins van nul afwijkt zal dit meestal geen aanleiding geven om de nulhypothese te verwerpen ten gunste van de alternatieve hypothese. De kans dat de nulhypothese verworpen wordt neemt toe naarmate b sterker van 0 verschilt. De waarde van $|b|$ die met p% kans aanleiding geeft tot het verwerpen van de nulhypothese zullen we aanduiden als b_p.
Bij de t-toets neemt b_p monotoon toe met p; voor $p = \alpha$ is $b_p = 0$ en als $p = 100$ dan is $b_p = \infty$. We kunnen b_p schrijven als

$$b_p = \delta_p \, \sigma_{\hat{b}} \tag{3}$$

met δ_p: de niet-centraliteitsparameter als functie van p.
$\sigma_{\hat{b}}$: de standaardafwijking van de kleinste kwadraten schatting $\hat{b}$ van b.
De relatie tussen p en δ_p is weergegeven in fig. 1 van Buishand (1984). Voor J = 5 jaar (3 vrijheidsgraden) en p = 80% geldt:

$$b_{80} = 4{,}26 \, \sigma_{\hat{b}}. \tag{4}$$

Een jaarlijkse toename van 4,26 $\sigma_{\hat{b}}$ in de gemiddelde concentratie leidt dus in 80% van de gevallen tot het verwerpen van de nulhypothese in een reeks van 5 opeenvolgende jaarwaarden. Bij overgang van jaarwaarden naar maandwaarden kan een kleinere trend met 80% kans reeds gedetecteerd worden doordat δ_p dan kleiner is terwijl $\sigma_{\hat{b}}$ nagenoeg niet verandert (Buishand, 1984). Voor maandwaarden geldt ruwweg bij een reeks van 5 jaren (58 vrijheidsgraden):

$$b_{80} \approx 2{,}85 \, \sigma_{\hat{b}} \, . \tag{5}$$

Tabel 1 geeft voor F, NO_3 en SO_4 de grootte van b_{80} volgens deze vergelijking. De waarde van $\sigma_{\hat{b}}$ in deze tabel is de mediaan van de geschatte

standaardafwijking van $\hat{b}$ van 12 stations op basis van jaarwaarden voor het tijdvak 1978-82.
In de laatste kolom van tabel 1 is b_{80} uitgedrukt als percentage van het landgemiddelde. Dit getal blijkt voor alle drie componenten ongeveer 7% te bedragen.

3. Trend-analyse van de gegevens van het RID-VEWIN meetnet

De trend-analyse van het voormalige RID heeft betrekking op 14-daagse meetgegevens van 27 stations voor het tijdvak januari 1979 - december 1982 (4 jaren). Een toets op trend werd uitgevoerd voor de componenten H, NH_4, NO_3 en SO_4. Daarbij werd gebruik gemaakt van het regressiemodel 1 met J = 4. In plaats van het jaargemiddelde werd echter uitgegaan van de jaarmediaan (mediaanwaarde van alle meetgegevens in het betreffende kalenderjaar). De regressiecoëfficiënten a en b werden geschat door middel van een gewogen kleinste kwadratenmethode.
Voor de bepaling van de trend die met 80% kans gedetecteerd kan worden in een reeks maandgegevens over een tijdvak van 5 jaren (b_{80}) mogen we in principe weer uitgaan van verg. 5. In deze vergelijking heeft $\sigma_{\hat{b}}$ betrekking op de standaardafwijking van de geschatte regressiecoëfficiënt $\hat{b}$ op basis van een reeks van 5 jaren. Daar het RID-onderzoek betrekking heeft op meetgegevens over een tijdvak van 4 jaren en daar var $\hat{b}$ ruwweg omgekeerd evenredig is met J^3 (Buishand, 1984, verg. 8) zijn de geschatte standaardafwijkingen van $\hat{b}$ uit dit onderzoek gecorrigeerd met een factor $\sqrt{4^3/5^3} = 0,72$. De mediaanwaarde (over 27 stations) van deze gecorrigeerde standaardafwijkingen is gegeven in tabel 2, tezamen met de grootte van b_{80}. Voor NO_3 en SO_4 zijn de waarden voor b_{80} ruim een factor 1,5 hoger dan de getallen uit het onderzoek van Ridder (1983) in tabel 1.
Naast een trend-analyse van gemeten concentraties werd ook een trend-toets uitgevoerd op berekende deposities. De waarden voor b_{80} voor deposities zijn vermeld in tabel 3. De procentuele jaarlijkse veranderingen, die met 80% kans gedetecteerd kunnen worden zijn van dezelfde orde van grootte als bij de gemeten concentraties.

4. Onderzoek Van Egmond, Kesseboom en Onderdelinden betreffende de optimalisatie van het landelijk meetnet regenwatersamenstelling

Voor de bepaling van de nauwkeurigheid van ruimtelijke interpolatie en de nauwkeurigheid van gebiedsgemiddelden gaan Van Egmond et al. (1985) uit van de logaritmen van de maandwaarden van 12 stations van het KNMI-RIV net. Om uit de gegevens van dit onderzoek iets te zeggen over de grootte van de trend, die met 80% kans gedetecteerd kan worden gaan we uit van een iets ander regressiemodel dan verg. 1, namelijk:

$$\ln y_i = a + b\, i/N + e_i, \qquad i = 1,\ldots,JN \tag{6}$$

met y_i: concentratie op tijdstip i,

e_i: een storingsterm met verwachting 0 en variantie σ_e^2,

N : het aantal meetgegevens in één jaar (hier N = 12),

J : het aantal jaren.

Uit verg. 6 volgt:

$$y_i = e^{a + b\, i/N}\, e^{e_i}, \qquad i = 1,\ldots,JN \tag{7}$$

Voor de verwachtingswaarde van de concentratie op tijdstip i geldt derhalve:

$$E(y_i) = e^{a + b\, i/N}\, E(e^{e_i}) \tag{8a}$$

en voor de verwachtingswaarde van de concentratie op tijdstip i + N:

$$E(y_{i+N}) = e^{a + b\, i/N + b}\, E(e^{e_{i+N}}). \tag{8b}$$

Deling van verg. 8b door verg. 8a geeft:

$$E(y_{i+N})/E(y_i) = e^b \approx 1 + b \tag{9}$$

daar de storingstermen e_i en e_{i+N} dezelfde kansverdeling hebben.
De regressiecoëfficiënt b geeft dus bij benadering de <u>relatieve verandering</u> van de verwachte concentratie na een periode van 1 jaar. Deze coëfficiënt is vergelijkbaar met de laatste kolom in de tabellen 1 en 2.
Voor de variantie van de kleinste kwadratenschatter $\hat{b}$ van b geldt bij

benadering (Buishand, 1984):

$$\sigma_{\hat{b}}^2 \approx \frac{12\ \sigma_e^2}{J^3\ N} \ . \tag{10}$$

Voor N = 12 en J = 5 krijgen we:

$$\sigma_{\hat{b}} \approx 0{,}089\ \sigma_e . \tag{11}$$

Substitutie van verg. 11 in verg. 5 geeft voor de trend die met 80% kans gedetecteerd kan worden in een reeks maandgegevens over een tijdvak van 5 jaren:

$$b_{80} = 2{,}85\ \sigma_{\hat{b}} \approx 0{,}25\ \sigma_e \ . \tag{12}$$

De publikatie van Van Egmond et al. (1985) geeft echter niet σ_e, maar de volgende twee spreidingsmaten:

- de variantie ten opzichte van het tijd-ruimtegemiddelde,
- de variantie ten opzichte van het ruimtegemiddelde.

Alvorens nader in te gaan op de trend-analyse zullen we deze spreidingsmaten beknopt behandelen.

<u>De variantie ten opzichte van het tijd-ruimtegemiddelde</u>

Stel dat x_{ij} de logaritme van de concentratie van station j is voor maand i. Het tijd-ruimtegemiddelde van M stations over een tijdvak van J jaren (NJ maanden) is dan:

$$\bar{x}_{..} = \frac{1}{MNJ} \sum_{i=1}^{NJ} \sum_{j=1}^{M} x_{ij} \ . \tag{13}$$

De variantie ten opzichte van $\bar{x}_{..}$ (of beter: het tweede moment ten opzichte van $\bar{x}_{..}$) van het j-de station wordt nu gedefinieerd als:

$$s_j^2(\bar{x}_{..}) = \frac{1}{NJ} \sum_{i=1}^{NJ} (x_{ij} - \bar{x}_{..})^2 \ . \tag{14}$$

Het gemiddelde van $s_j^2(\bar{x}_{..})$ over alle M stations duiden we aan als $\overline{s}^2(\bar{x}_{..})$. Voor de verwachtingswaarde van $\overline{s}^2(\bar{x}_{..})$ kunnen we aantonen (zie appendix):

$$E\{\overline{s}^2(\overline{x}_{..})\} \approx \frac{1}{M}\sum_{j=1}^{M}\sigma_j^2 + \frac{1}{M}\sum_{j=1}^{M}(\mu_j - \mu)^2 \tag{15}$$

met $\mu_j = E(x_{ij})$ en $\sigma_j^2 = \text{var}\ x_{ij}$; μ is het gemiddelde van de μ_j's (dit is de verwachtingswaarde van $\overline{x}_{..}$).

De variantie ten opzichte van het ruimtegemiddelde

Bij de variantie ten opzichte van het ruimtegemiddelde kijken we naar de afwijkingen van de x_{ij}'s ten opzichte van de maandgemiddelden $\overline{x}_{i.}$. Voor het j-de station wordt de variantie ten opzichte van $\overline{x}_{i.}$ gedefinieerd als:

$$s_j^2(\overline{x}_{i.}) = \frac{1}{NJ}\sum_{i=1}^{NJ}(x_{ij} - \overline{x}_{i.})^2 . \tag{16}$$

Het gemiddelde van $s_j^2(\overline{x}_{i.})$ over alle M stations duiden we aan als $\overline{s}^2(\overline{x}_{i.})$. De waarde van $\overline{s}^2(\overline{x}_{i.})$ hangt af van het aantal stations M. Als M = 1 dan is $\overline{s}^2(\overline{x}_{i.}) = 0$. In het algemeen neemt $\overline{s}^2(\overline{x}_{i.})$ monotoon toe met M. Al vrij gauw wordt echter een limietwaarde bereikt.

Het toetsen op trend bij meetgegevens van een enkel station

Indien in de historische reeks geen trend aanwezig is dan geldt $\sigma_e = \sigma_j$. Voor de meest belangrijke componenten (H, NH_4, NO_3 en SO_4) lopen de langjarige gemiddelden van de verschillende stations weinig uiteen, waardoor men de tweede term in het rechterlid van verg. 15 mag verwaarlozen. Een redelijke maat voor σ_e in verg. 12 is daarom $\overline{s}(\overline{x}_{..})$. Bij een aantal componenten leidt deze benadering echter tot een overschatting van b_{80}. Hierop zullen we aan het einde van deze paragraaf terugkomen.
Tabel 4 geeft voor de neerslaghoeveelheid en de concentraties van een groot aantal chemische componenten de grootte van de jaarlijkse verandering aan die met 80% kans gedetecteerd kan worden (b_{80}) in een reeks van maandgegevens van 5 jaren van een enkel station. De waarde van b_{80} loopt uiteen van ongeveer 10% (NO_3 en SO_4) tot ruim 20% (Mg, Zn, Cl en Na).

Het toetsen op trend bij landgemiddelden

Tot nu toe hebben we ons beziggehouden met het opsporen van een trend in een reeks van een enkel station. Een verbetering van het onderscheidingsvermogen is mogelijk door de trend-toets toe te passen op de maandgemiddelden van een aantal stations (Buishand, 1984). Een verbetering van het onderscheidingsvermogen houdt in dat de waarde van b, waarvoor in 80% van de gevallen de nulhypothese verworpen wordt, afneemt.

Als maat voor σ_e in verg. 12 moeten we nu de standaardafwijking $\sigma_{i.}$ van de maandgemiddelden $\bar{x}_{i.}$ gebruiken. Een geschikte schatter voor $\sigma^2_{i.}$ is de steekproefvariantie van de maandgemiddelden:

$$\hat{\sigma}^2_{i.} = \frac{1}{NJ} \sum_{i=1}^{NJ} (\bar{x}_{i.} - \bar{x}_{..})^2 \qquad (17)$$

Tussen $\hat{\sigma}^2_{i.}$ en de varianties ten opzichte van $\bar{x}_{i.}$ en $\bar{x}_{..}$ bestaat de volgende relatie (zie appendix):

$$\hat{\sigma}^2_{i.} = \bar{s}^2(\bar{x}_{..}) - \bar{s}^2(\bar{x}_{i.}). \qquad (18)$$

Tabel 5 geeft de grootte van b_{80} indien uitgegaan wordt van de maandgemiddelden van 12 KNMI-RIV stations. De getallen voor b_{80} in deze tabel zijn steeds kleiner dan die in tabel 4 voor een trend-toets op een meetreeks van een enkel station. De kleinste reductie in b_{80} (ongeveer 10%) treft men aan bij componenten met een sterke ruimtelijke correlatie (de neerslaghoeveelheid, NO_3 en SO_4). Voor de meeste componenten wordt een reductie van b_{80} van 30 à 40% gevonden. Een reductie van ruim 50% wordt aangetroffen bij K en F. Dit zijn componenten waarvoor er binnen Nederland vrij grote relatieve verschillen zijn in de gemiddelde concentraties (KNMI-RIV, 1983). De tweede term in het rechterlid van verg. 15 is dan niet meer verwaarloosbaar. Het gevolg hiervan is dat $\bar{s}^2(\bar{x}_{..})$ een te hoge schatting geeft voor σ^2_e, wat tot gevolg heeft dat de waarden voor b_{80} in tabel 4 te hoog zijn. Bij de getallen in tabel 5 is dit niet het geval daar verg. 18 ook bij plaatselijke verschillen in de gemiddelde concentraties een goede schatter voor $\sigma^2_{i.}$ oplevert.

5. Discussie

In de tabellen 1, 2 en 4 werd de grootte van de verandering in de concentratie aangegeven die met 80% kans gedetecteerd kan worden in een reeks meetgegevens van 5 jaren. Vergelijking van de getallen in deze tabellen leert dat de gegevens uit verschillende onderzoeken niet tot eensluidende resultaten leiden. De getallen voor b_{80} uit het onderzoek van Ridder (1983) in tabel 1 zijn lager dan die uit de trend-analyse van de meetgegevens van het RID-VEWIN meetnet in tabel 2 en die uit het onderzoek van Van Egmond et al. (1985) in tabel 4. Dit wordt enerzijds veroorzaakt doordat niet van hetzelfde datamateriaal is uitgegaan (KNMI-RIV gegevens of RID-VEWIN gegevens) en anderzijds door verschillende methoden die zijn gevolgd bij de statistische verwerking van de gegevens.

Een nadeel van de getallen uit het onderzoek van Ridder (tabel 1) is dat de geschatte waarde van $\hat{\sigma_b}$ gebaseerd is op jaarwaarden over een 5-jarig tijdvak. Deze schatting is weinig nauwkeurig doordat men bij elke reeks slechts drie vrijheidsgraden heeft. Het nemen van medianen over de 12 stations brengt alleen maar een aanzienlijke variantiereductie teweeg indien er geen of slechts een zwakke ruimtelijke correlatie is. Van de drie componenten in tabel 1 is dit alleen bij F het geval.

Bij de gegevens uit het onderzoek van Van Egmond et al. (1985) wordt een aantal benaderingen en aannames gemaakt. De getallen in tabel 4 kunnen om de volgende drie redenen aan de <u>hoge</u> kant zijn:

(i) Door verschillen in de gemiddelde concentraties binnen Nederland (zie par. 4).

(ii) Door de aanwezigheid van een trend in de historische reeks. De residuele variantie σ_e^2 moet geschat worden na verwijdering van deze trend, wat bij $\overline{s}^2(\overline{x}_{\bullet\bullet})$ niet het geval is.

(iii) Door de aanwezigheid van een seizoenverloop in de historische reeks. Dit heeft hetzelfde effect als (ii).

Om met het effect van een seizoenverloop rekening te houden is een uitgebreider regressiemodel dan verg. 6 nodig, zie Buishand (1984).

Tot nu toe hebben we ons beperkt tot de trend die kan worden opgespoord in een meetreeks van 5 jaren. De grootte van de trend die men kan detecteren is sterk afhankelijk van het aantal jaren J. Voor J = 10 jaren bedraagt b_{80} bijvoorbeeld ongeveer 1/3 van de waarde van b_{80} bij J = 5 jaren.

Bij een reeks van gemiddelden over 12 meetstations kan een kleinere trend gedetecteerd worden dan bij een reeks van een enkel station. Men kan zich afvragen of bij het middelen over meer dan 12 stations een nog kleinere trend

opgespoord kan worden. Dit is echter slechts het geval indien $\sigma^2_{i.}$ voor M > 12 nog afneemt met M.
In het rechterlid van verg. 18 is het vooral de variantie $\bar{s}^2(\bar{x}_{i.})$ ten opzichte van het ruimtegemiddelde $\bar{x}_{i.}$ die met M varieert. Deze bereikt echter al vrij gauw een limietwaarde met als gevolg dat $\hat{\sigma}^2_{i.}$ ook niet meer met M afneemt. Uit de gegevens van het onderzoek van Van Egmond et al. (1985) blijkt dat voor de meeste componenten $\hat{\sigma}_{i.}$ slechts met een paar procent afneemt als we van 6 naar 12 stations gaan. De hoogste afname (ongeveer 20%) wordt aangetroffen bij Zn. Men mag dan ook niet verwachten dat voor M > 12 de waarde van $\sigma_{i.}$ nog aanzienlijk zal afnemen. Uitbreiding van het aantal stations om de detecteerbaarheid van trends te verbeteren is daarom niet zinvol.

Tot nu toe zijn meetfouten buiten beschouwing gebleven. Een deel van de variantie van $\bar{x}_{i.}$ kan verklaard worden door meetfouten. Uit de gegevens van Van Egmond et al. (1985) volgt dat het aandeel van meetfouten in $\sigma^2_{i.}$ gering is. Deze is het grootste bij K en Zn (ongeveer 15% bij M = 12 stations). Hierbij moet worden opgemerkt dat de getallen van Van Egmond et al. nog aan de hoge kant zijn doordat microschaal variaties ook als meetfouten zijn opgevat. Verlaging van meetfouten heeft daarom nauwelijks invloed op $\sigma^2_{i.}$ en het onderscheidingsvermogen van een trend-toets.

De getallen van b_{80} in de tabellen 1 t/m 5 hebben steeds betrekking op maandwaarden van concentraties of deposities. In par. 2 werd uiteengezet dat overgang van jaarwaarden naar maandwaarden een winst in onderscheidingsvermogen oplevert. Een hogere bemonsteringsfrequentie dan één maal per maand levert echter nauwelijks een verdere winst in het onderscheidingsvermogen op (Buishand, 1984).

Recent onderzoek wijst erop dat het gebruik van meteorologische voorinformatie goede perspectieven biedt om de detecteerbaarheid van trends te verbeteren. Hierbij moet men denken aan de neerslaghoeveelheid die men weer kan opdelen in verschillende categorieën. Het is momenteel nog onduidelijk of men hiertoe moet overgaan op 24-uurs monstername.

6. Conclusies

Bij een meetnet van 12 stations kan voor de meeste componenten met behulp van een eenvoudig regressiemodel een systematische verandering in het langjarig gemiddelde van ongeveer 10% per jaar met 80% kans gedetecteerd worden bij meetgegevens over een tijdvak van 5 jaren (tabel 5). Bij metingen over een tijdvak van 10 jaren is de grootte van de te detecteren trend ongeveer een factor 3 lager. De hier gegeven getallen zijn slechts ruwe indicaties. Betere schattingen vereisen een tijdreeksen-analyse met een iets uitgebreider regressiemodel waarin o.a. rekening wordt gehouden met mogelijke seizoen-invloeden.

Het is niet mogelijk de detecteerbaarheid van trends te verbeteren door het meetnet verder uit te breiden of door de waarnemingsfrequentie te verhogen. Ook heeft verkleining van de meetfouten weinig effect. Slechts het gebruik van meteorologische voorinformatie biedt perspectieven om de detecteerbaarheid van trends te verbeteren.

Literatuur

Aldenberg, T. (1986). Exploratory Data Analysis en robuuste statistiek toegepast op regenwatermeetnetgegevens. Publikatie RIVM in voorbereiding.

Buishand, T.A. (1984). Detectie van trends in de chemische samenstelling van de neerslag. KNMI Memo FM-84-42. [Unpublished]

Egmond, N.D. van, Kesseboom, H. en Onderdelinden, D. (1985). Statistische Optimalisatie van het Landelijk Meetnet voor de Regenwaterkwaliteit. RIVM-rapport 218203001.

KNMI-RIV (1983). Chemical composition of precipitation over the Netherlands. Annual Report 1982.

Ridder, T.B. (1983). 5 jaar metingen van de samenstelling van de neerslag. In: ZURE REGEN oorzaken, effecten en beleid; Proceedings symposium 's-Hertogenbosch, pp. 55-58.

Appendix

In deze appendix zullen we nader ingaan op de afleiding van de vergelijkingen 15 en 18. Beide vergelijkingen worden verkregen door $\overline{s}^2(\overline{x}_{..})$ te splitsen in twee componenten volgens een bekende methode uit de variantie-analyse.

Om verg. 15 af te leiden schrijven we $\overline{s}^2(\overline{x}_{..})$ als

$$\overline{s}^2(\overline{x}_{..}) = \frac{1}{M}\sum_{j=1}^{M} s_j^2(\overline{x}_{..}) \quad = \frac{1}{MNJ}\sum_{j=1}^{M}\sum_{i=1}^{NJ}(x_{ij} - \overline{x}_{.j} + \overline{x}_{.j} - \overline{x}_{..})^2$$

$$= \frac{1}{MNJ}\sum_{j=1}^{M}\sum_{i=1}^{NJ}(x_{ij} - \overline{x}_{.j})^2 + \frac{1}{M}\sum_{j=1}^{M}(\overline{x}_{.j} - \overline{x}_{..})^2$$

$$= \frac{1}{M}\sum_{j=1}^{M} s_j^2 + \frac{1}{M}\sum_{j=1}^{M}(\overline{x}_{.j} - \overline{x}_{..})^2 \tag{A1}$$

waarbij s_j^2 de steekproefvariantie is van de meetreeks van het j-de station. Voor de verwachtingswaarde van $\overline{s}^2(\overline{x}_{..})$ krijgen we:

$$E\{\overline{s}^2(\overline{x}_{..})\} = \frac{1}{M}\,\frac{NJ-1}{NJ}\sum_{j=1}^{M}\sigma_j^2 + \frac{1}{M}\sum_{j=1}^{M}(\mu_j - \mu)^2$$

$$+ \frac{1}{M}\sum_{j=1}^{M}\mathrm{var}(\overline{x}_{.j} - \overline{x}_{..}). \tag{A2}$$

Als NJ groot is dan is $(NJ-1)/(NJ) \approx 1$ en mag men de derde term in het rechterlid van verg. A2 verwaarlozen, hetgeen de volgende benadering oplevert:

$$E\{\overline{s}^2(\overline{x}_{..})\} \sim \frac{1}{M}\sum_{j=1}^{M}\sigma_j^2 + \frac{1}{M}\sum_{j=1}^{M}(\mu_j - \mu)^2. \tag{A3}$$

Om verg. 18 af te leiden schrijven we $\overline{s}^2(\overline{x}_{..})$ als

$$s^2(\overline{x}_{..}) = \frac{1}{MNJ}\sum_{j=1}^{M}\sum_{i=1}^{NJ}(x_{ij} - \overline{x}_{i.} + \overline{x}_{i.} - \overline{x}_{..})^2$$

$$= \frac{1}{MNJ}\sum_{j=1}^{M}\sum_{i=1}^{NJ}(x_{ij} - \overline{x}_{i.})^2 + \frac{1}{NJ}\sum_{i=1}^{NJ}(\overline{x}_{i.} - \overline{x}_{..})^2$$

$$= \frac{1}{M}\sum_{j=1}^{M} s_j^2(\overline{x}_{i.}) + \hat{\sigma}_{i.}^2 \quad = \overline{s}^2(\overline{x}_{i.}) + \hat{\sigma}_{i.}^2 \,. \tag{A4}$$

of

$$\hat{\sigma}_{i.}^2 = \overline{s}^2(\overline{x}_{..}) - \overline{s}^2(\overline{x}_{i.}). \tag{A5}$$

Tabellen

Tabel 1. Schatting van de grootte van de jaarlijkse verandering in de concentraties van de componenten F, NO_3 en SO_4 die in een reeks van 5 jaren van één station met 80% kans opgespoord kan worden indien uitgegaan wordt van maandwaarden, verg. 5. De geschatte waarden van $\hat{\sigma_b}$ zijn gebaseerd op jaargemiddelden van gemeten concentraties uit een onderzoek van Ridder (1983).

Component	$\hat{\sigma_b}$ µmol/l	b_{80} µmol/l	Landgemiddelde 1978-82 µmol/l	Detecteerbare trend in % per jaar
F	0,079	0,225	3	7,5
NO_3	1,64	4,67	60	7,8
SO_4	1,61	4,59	70	6,6

Tabel 2. Schatting van de grootte van de jaarlijkse verandering in de concentraties van de componenten H, NH_4, NO_3 en SO_4 die in een reeks van 5 jaren van één station met 80% kans opgespoord kan worden indien uitgegaan wordt van maandwaarden, verg. 5. De geschatte waarden van $\hat{\sigma_b}$ zijn ontleend aan een trend-analyse van de gegevens van het RID-VEWIN meetnet (Aldenberg, 1986).

Component	$\hat{\sigma_b}$ µmol/l	b_{80} µmol/l	Landgemiddelde µmol/l	Detecteerbare trend in % per jaar
H	2,26	6,45	25	25,8
NH_4	5,96	17,00	90	18,8
NO_3	3,37	9,61	70	13,7
SO_4	2,81	8,02	80	10,0

Tabel 3. Schatting van de grootte van de jaarlijkse verandering in de deposities van de componenten H, NH_4, NO_3 en SO_4 die in een reeks van 5 jaren van één station met 80% kans opgespoord kan worden indien uitgegaan wordt van maandwaarden, verg. 5. De geschatte waarde van $\hat{\sigma_b}$ zijn ontleend aan een trend-analyse van de gegevens van het RID-VEWIN meetnet (Aldenberg, 1986).

Component	$\hat{\sigma}_b$ µmol/maand	b_{80} µmol/maand	Landgemiddelde µmol/maand	Detecteerbare trend in % per jaar
H	132	378	1800	20,9
NH_4	303	864	6400	13,5
NO_3	150	428	4000	10,7
SO_4	191	544	4800	11,3

Tabel 4. Schatting van de grootte van de relatieve jaarlijkse verandering in de neerslaghoeveelheid en de concentraties van een elftal chemische componenten die in een reeks van 5 jaren van één station met 80% kans opgespoord kan worden indien uitgegaan wordt van maandwaarden, verg. 12. De varianties $\bar{s}^2(\bar{x}_{..})$ ten opzichte van het tijdruimtegemiddelde $\bar{x}_{..}$ zijn ontleend aan een onderzoek van Van Egmond et al. (1985).

Component	$\sigma_e^2 = \bar{s}^2(\bar{x}_{..})$	$\hat{\sigma_b}$	Relatieve jaarlijkse verandering b_{80}
Neerslag	0,32	0,051	0,15
H	0,31	0,050	0,14
NH_4	0,39	0,055	0,16
Na	0,90	0,085	0,24
K	0,46	0,061	0,17
Ca	0,51	0,063	0,18
Mg*	0,62	0,070	0,20
Zn	0,80	0,079	0,23
F	0,36	0,053	0,15
Cl*	0,79	0,079	0,23
NO_3	0,23	0,043	0,12
SO_4*	0,18	0,037	0,11

* Voor deze componenten is op basis van de Na-concentratie een correctie voor zeezout uitgevoerd.

Tabel 5. Schatting van de grootte van de relatieve jaarlijkse verandering in de neerslaghoeveelheid en de concentraties van een elftal chemische componenten die in een tijdvak van 5 jaren met 80% kans opgespoord kan worden indien uitgegaan wordt van de maandgemiddelden van 12 stations, verg. 12. De geschatte variantie van de maandgemiddelden $\hat{\sigma}^2_{i.}$ is berekend met behulp van verg. 18.

Component	$\sigma^2_e = \hat{\sigma}^2_{i.}$	$\sigma_{\hat{b}}$	Relatieve jaarlijkse verandering b_{80}
Neerslag	0,25	0,045	0,13
H	0,15	0,035	0,10
NH_4	0,24	0,045	0,13
Na	0,37	0,054	0,15
K	0,09	0,027	0,08
Ca	0,23	0,043	0,12
Mg*	0,35	0,053	0,15
Zn	0,29	0,048	0,14
F	0,08	0,025	0,07
Cl*	0,30	0,049	0,14
NO_3	0,18	0,037	0,11
SO_4*	0,12	0,031	0,09

* Voor deze componenten is op basis van de Na-concentratie een correctie voor zeezout uitgevoerd.