

KONINKLIJK NEDERLANDS
METEOROLOGISCH INSTITUUT

De Bilt

Verslagen

V - 267

J.P. de Jongh

Analogenselectie m.b.v. de computer
voor de meerdaagse verwachting

De Bilt, 1975

Publikationsnummer: K.N.M.I. V-267 (M.O.)

U.D.C.: 551.509.315 :
551.509.318 :
551.509.32

Inhoud

	blz.
1. Inleiding	1
2. Vergelijking methodieken analogenselectie	2
2.1 Overzicht selectiemethoden analogen	3
2.2 Resultaten selectiemethoden analogen	6
2.3 Beoordeling operationele versie	7
3. Weervertaling m.b.v. analogen	10
3.1 Zonneschijn	11
3.2 Neerslag	13
3.3 Maximumtemperatuur	13
3.4 Minimumtemperatuur	14
3.5 Discussie	14

Analogenselectie m.b.v. de computer
voor de meerdaagse verwachting

J.P. de Jongh

1. Inleiding

Bij het opstellen van de meerdaagse verwachting (2 en 3 dagen vooruit) wordt hoofdzakelijk gebruikgemaakt van de 24, 48 en 72 uur prognoses van het 500 mbar vlak, zoals die door de Amerikaanse meteorologische dienst worden berekend. De meteorologen moeten via het verwachte stromingspatroon tot een weersverwachting komen. Eén van de hulpmiddelen hierbij is het gebruik van analogen. Hiervoor worden uit een bestand van 500 mbar kaarten van 1949 tot heden een aantal kaarten geselecteerd, welke een grote mate van analogie vertonen met de Amerikaanse 72 uur prognose. Indien ook de voorgeschiedenis van de geselecteerde kaarten redelijk overeenkomt met de voorgeschiedenis van de 72 uur prognose, dan wordt het "weer", dat bij de analoge situaties behoort, opgezocht.

Het selecteren van de analoge 500 mbar stromingspatronen is een tijdrovende bezigheid, omdat een groot aantal kaarten (\pm 750) moet worden bekeken.

In het W.R. 75-10* wordt een aantal methoden beschreven om met behulp van de computer uit een 500 mbar bestand analogen te selecteren. Het probleem is, een automatische selectiemethode te vinden, die zo goed mogelijk de selectie-criteria van de meteorologen simuleert.

In het W.R. 75-10 wordt de bruikbaarheid van een aantal selectiemethoden op een 4-tal prognoses uitgetoetst en onderling vergeleken. Dit vergelijkend onderzoek is nu op een veel groter aantal prognoses voortgezet. De methode waarvan de

* J.P. de Jongh en S. Kruizinga: Empirische orthogonale eigenvelden van het 500 mbar-vlak. W.R. 75-10, KNMI.

resultaten het beste overeenkwamen met de analogenselectie van de meteorologen, wordt momenteel dagelijks door de meerdaagse sectie gebruikt.

De analogenselectie met behulp van de computer moet in eerste instantie als een ruwe voorselectie worden beschouwd. De meteorologen vinden d.m.v. hun subjectieve selectiemethode gemiddeld 3 - 5 bruikbare analogen bij een 72 uur prognose uit een gegeven bestand van ongeveer 25 jaar. Het is de bedoeling dat zoveel mogelijk van deze gevallen in de computerlijst voorkomen, terwijl de lijst toch niet al te lang mag worden. In eerste instantie zijn computerlijsten vervaardigd waarop 20 analogen voorkwamen. In overleg met de meteorologen zijn deze lijsten later uitgebreid tot 30 analoge gevallen.

Een bijkomend voordeel van de analogenselectie met behulp van de computer is, dat op eenvoudige wijze een aantal weergrootheden in de computeruitvoer zijn op te nemen. Op deze manier wordt al direkt een overzicht van het bijbehorende weer verkregen.

2. Vergelijking methodieken analogenselectie

In W.R. 75-10 wordt een 7-tal selectiemethoden beschreven. Deze zullen in paragraaf 2.1 in het kort worden besproken. In paragraaf 2.2 zullen de resultaten van de verschillende methoden met elkaar worden vergeleken. In paragraaf 2.3 wordt nagegaan in hoeverre de methode, die in 2.2 het beste blijkt te voldoen, ook bruikbaar is in de operationele weerdienst.

Figuur 1 geeft een overzicht van de roosterpunten waarvan de 500 mbar hoogten worden gebruikt voor de analogenselectie.

2.1 Overzicht selectiemethoden analogen

Methode 1

Het 500 mbar patroon van de 72 uur prognose kan worden ontbonden in zgn. empirische orthogonale functies. (Voor een beschrijving van e.o.f. zie W.R. 75-10). Stel dat de scores op de eigenvectoren van de prog zijn:

$$a_1, a_2, \dots, a_n.$$

Stel de scores van de n^e dag uit het bestand zijn

$$\beta_{1n}, \beta_{2n}, \dots, \beta_{pn}.$$

De dag waarvoor de grootheid

$$D_n = \sum_{i=1}^p g_i (a_i - \beta_{in})^2$$

minimaal is, wordt als de beste analoog beschouwd. De dag met de op één na kleinste waarde van D_n wordt als de op één na beste analoog beschouwd, etc.

Voor de gewichten is $g_1 = g_2 = \dots g_{20} = 1$ genomen en $g_{21} = g_{22} \dots = g_{30} = 0$.

Voor deze methode is een veld van 30 roosterpunten gebruikt. Dit heeft tot resultaat dat er ook maar 30 scores en 30 gewichten zijn.

Methode 2

Deze methode is dezelfde als methode 1, behalve dat $g_1 = g_3 = 5$ (zijnde de gewichten van de eigenvectoren, die boven Nederland een grote bijdrage leveren) en $g_2 = g_4 = g_5 \dots g_{20} = 1$, g_{21} t/m $g_{30} = 0$.

Methode 3

In het voorafgaande is ernaar gestreefd analogen te selecteren door de afstand tussen de score-vector van de uitgangstoestand en de score-vector van de analoge dag minimaal te maken. Een andere methode om analogen te zoeken is dat de score-vector van de prognose zoveel mogelijk parallel is aan de score-vector van de analoge dag, d.w.z. het inproduct tussen de score-vectoren moet maximaal zijn. Dit komt overeen met de eis dat

$$I_n = \frac{\sum_{i=1}^P (\alpha_i - \beta_{in})}{\left(\sum_{i=1}^P \alpha_i^2 \right) \cdot \left(\sum_{i=1}^P \beta_{in}^2 \right)}$$

maximaal moet zijn.

Meteorologisch betekent dit, dat vooral op de posities van de druksystemen wordt gelet en minder op de intensiteit van de systemen.

Methode 4

Behalve het selecteren van analogen met behulp van scores kan men voor de selectie ook rechtstreeks uitgaan van de hoogten in de roosterpunten. Stel h_{oj} is de hoogte van het 500 mbar vlak in het j^e roosterpunt van de uitgangstoestand en h_{nj} is de hoogte van het 500 mbar vlak in het j^e roosterpunt op dag n , dan is

$$D_n = \sum_{j=1}^{63} (h_{oj} - h_{nj})^2$$

een maat voor de "afstand" tussen de velden.

De dag waarvoor D_n minimaal is, levert de beste analoge situatie op.

Deze methode heeft als voordeel boven de selectie met behulp van de scores dat het op eenvoudige wijze mogelijk is, door het toevoegen van gewichtsfactoren, de analogie in bepaalde delen van

het kaartgebied zwaarder te laten wegen dan in andere gedeelten. De "afstand" D_n wordt dan:

$$D_n = \sum_{j=1}^{63} g_j \cdot (h_{oj} - h_{nj})^2.$$

De gewichten van alle roosterpunten waarvan de hoogten over de gehele periode 1949 t/m 1970 aanwezig zijn, worden 1 gesteld; de gewichten van de overige roosterpunten zijn 0 genomen.

Methode 5

Deze methode is dezelfde als methode 4 maar de gewichten van 30 roosterpunten in de omgeving van Nederland zijn 3 genomen; buiten dit gebied zijn de gewichten als bij methode 4 genomen.

Methode 6

Bij het selecteren van analogen zijn de meteorologen vooral geïnteresseerd in de vorm van het stromingspatroon en ze zijn in mindere mate geïnteresseerd in de absolute hoogte van de 500 mbar isohypsen. Wel wordt de absolute waarde van de isohypse in De Bilt en naaste omgeving van belang geacht. Afwijkingen van meer dan 10 decameter zullen niet of nauwelijks worden getolereerd. Het is mogelijk dat voor een goede analoog de gemiddelde waarde van de 500 mbar hoogten in de 63 roosterpunten enige decameters afwijkt van de gemiddelde hoogte van de roosterpunten in de prognose. Om deze reden zijn ook analogen bepaald waarbij de standaardafwijking tussen de prognose en de analoog is berekend volgens de volgende formule:

$$D_n = \sum_j g_j (h_{nj} - h_{oj})^2 - \frac{\left(\sum_j g_j (h_{nj} - h_{oj}) \right)^2}{\sum_j g_j}$$

waarbij de gewichten van alle roosterpunten, die compleet in het bestand aanwezig zijn, 1 genomen worden.

Methode 7

Deze methode is dezelfde als methode 6, behalve dat de gewichten in een 30-tal roosterpunten in de omgeving van Nederland 3 zijn genomen.

2.2 Resultaten selectiemethoden analogen

Bij de Amerikaanse prognoses van de maanden oktober 1974 t/m december 1974 zijn volgens alle zeven boven omschreven methoden met behulp van de computer analogen geselecteerd en vergeleken met de resultaten van de meteorologen. Hierbij dient het volgende te worden opgemerkt:

1. Alleen die dagen uit het 500 mbar bestand waarvan het dagnummer in het jaar minder dan 15 verschilt met het dagnummer van de prognose, kunnen als analoog optreden.
2. Het 500 mbar materiaal bevat gegevens van de jaren 1949 t/m 1970. Er kunnen dus door de computer geen analogen worden gevonden uit de jaren 1971 tot heden, terwijl de meteoroloog dit wel kan.
3. Indien de datum van de analoog van de computerlijst slechts één dag verschilt met de datum die de meteoroloog vindt, wordt aangenomen dat de analoog toch is gevonden. Omdat de meteoroloog de kaarten vlak voor en na de analoog bekijkt, is het toestaan van één dag verschuiving wel reëel.

Een voorbeeld van de computeruitvoer wordt in figuur 2 gegeven. Als eerste worden de datum en de 500 mbar hoogten van de 72 uur prognose in de 63 roosterpunten gegeven. Vervolgens worden de data van de 20 geselecteerde analogen vermeld. Achter iedere datum wordt het bijbehorende weer gegeven in de vorm van de maximumtemperatuur in de vijf hoofdstations en het gemiddelde over de vijf hoofdstations, de minimumtemperatuur in de vijf hoofdstations en het gemiddelde over de vijf hoofdstations, de windrichting en windsnelheid in Eelde om 12.00z, het aantal

uren zonneshijn te De Bilt, het aantal stations met neerslag ≥ 0.3 mm in de periode van 18.00z - 06.00z (maximaal tien stations) en het aantal millimeters neerslag te De Bilt van 18.00z - 06.00z. Verder wordt nog opgegeven het gemiddelde hoogteverschil tussen de analoog en de 72 uur prognose in de 63 roosterpunten (kolom "gem"). In de kolom "rang" staat het rangnummer van de analoog. De beste analoog heeft rang 1, etc.

Tabel I geeft een overzicht van het aantal analogen, dat zowel door de meteoroloog als met behulp van de computermethoden, is gevonden. Het blijkt dat dit aantal groot is voor de methoden 3, 6 en 7. Bij deze methoden wordt vooral gelet op de vorm van het stromingspatroon en minder op de absolute hoogte van het veld. Blijkbaar komt dit het beste overeen met de selectiecriteria van de meteoroloog. Methode 7 geeft het grootste aantal corresponderende analogen. Bij deze methode wordt vooral de nadruk gelegd op het stromingspatroon in de omgeving van Nederland. Dit gebeurt kennelijk ook door de meteorologen.

In het W.R. 75-10 worden voor een viertal Amerikaanse 72 uur prognoses de resultaten van de verschillende analogen-selectiemethoden vergeleken. Ook daar blijkt dat methode 7 de beste resultaten oplevert. Om deze redenen werd besloten per 1 januari 1975 de lijst met analogen, berekend volgens methode 7, op routinebasis te laten uitdraaien. Vanaf die datum is een onderlinge vergelijking van de verschillende methoden niet meer mogelijk. De meteorologen nemen namelijk regelmatig analogen uit de lijst over. Op deze manier zal methode 7 bevoordeeld worden ten opzichte van de andere methoden.

2.3 Beoordeling operationele versie

Hoewel in paragraaf 2.2 is gebleken, dat bij methode 7 de grootste overeenkomst bestaat tussen de analogen, die zijn geselecteerd door de meteorologen en die, welke zijn geselecteerd volgens de computermethode, is het nog niet duidelijk in hoeverre deze methode ook op zinvolle wijze bruikbaar is in de

weerdienst. Hoewel de criteria voor het al of niet bruikbaar zijn natuurlijk niet scherp vastliggen, moet toch wel van de computermethode worden geëist, dat voor ieder individueel geval een groot gedeelte van de bruikbare analogen wordt gevonden.

De beoordeling van de mate van analogie tussen twee kaarten is een subjectieve zaak. De beoordeling hangt af van het gewicht dat aan bepaalde verschillen tussen de kaarten wordt gehecht. Bijvoorbeeld, van een analoog kan de stroming in de omgeving van Nederland bijzonder goed overeenkomen met de stroming van de 72 uur prognose maar de absolute hoogte van de 500 mbar kaarten verschilt nogal; bij een andere analoog zal de hoogte weer beter overeenkomen, maar zal de stroming wat verschillen. Hoe men deze afwijkingen beoordeelt, verschilt van meteoroloog tot meteoroloog. Dit heeft tot gevolg dat de ene meteoroloog een bepaalde analoog wel zal gebruiken en de andere niet. Vooral op dagen dat geen kaarten met een uitgesproken analogie zijn te vinden, kunnen er zodoende gemakkelijk verschillen optreden tussen de analogen, die gevonden zijn door de meteorologen en die, welke op de computerlijst voorkomen.

In overleg met de meteorologen zijn in de operationele versie nog twee veranderingen aangebracht, nl.

1e het aantal analogen in de computerlijst is uitgebreid van 20 tot 30;

2e de analogen worden gezocht in die dagen van het bestand waarvan het dagnummer in het jaar minder dan 20 verschilt met het dagnummer van de prognose. Bij de experimenten in hoofdstuk 2 is een verschuiving van slechts 15 dagen toegelaten.

Voor de maanden mei en juni 1975 zijn alle analogen, die op de computerlijst voorkomen, beoordeeld door twee onafhankelijke personen. Deze beoordeling geschiedde in drie klassen, nl. goed (+), matig (0) en slecht (-). Uit deze beoordeling zijn de volgende punten naar voren gekomen:

- In de 41 lijsten met analogen, die voor de maanden mei en juni zijn berekend, komen in het totaal 162 analogen voor, die met 0, 0/+ of + zijn beoordeeld en derhalve bruikbaar

zijn. Dit wil zeggen, dat er gemiddeld vier bruikbare analogen op een computerlijst voorkomen. Dit is ook ongeveer het aantal dat de meteorologen vinden.

- Over de maanden mei en juni 1975 hebben de meteorologen in het totaal 211 analogen bij het opstellen van de meerdaagse verwachting gebruikt; hiervan zijn er 21 uit de jaren 1971 tot heden. Omdat de 500 mbar gegevens van deze jaren op de magneetband ontbreken, kunnen deze analogen dus niet op de computerlijst staan. In het totaal zijn er dus 190 analogen, die ook op de computerlijsten kunnen voorkomen. Hiervan blijken er 125 inderdaad op de computerlijsten voor te komen, indien een verschuiving van twee dagen wordt toegestaan. D.w.z. indien de meteoroloog een analoog op 14 april vindt en op de computerlijst komt 16 april voor, wordt aangenomen dat de analoog wel is gevonden.

Van de 65 analogen, die niet op de lijst voorkomen, blijken er 11 met O/+ of met + beoordeeld te zijn. Dit zijn dus zeer bruikbare analogen, die ten onrechte gemist zijn. Anderzijds zijn er ook nog een aantal analogen (15) op de lijst gevonden, die met O/+ of met + zijn beoordeeld en niet door de meteoroloog zijn gebruikt. Het totale aantal goed bruikbare analogen, dat door de meteorologen is gevonden, blijkt dus nagenoeg gelijk te zijn aan het aantal dat op de computerlijst voorkomt.

- Op de computeruitvoer wordt ook nog een aanduiding gegeven omtrent de mate van analogie in de kolom "rang". De beste analoog krijgt een 1 toegewezen, etc. Nagegaan is in hoeverre deze rang ook enige betekenis heeft. Bij de analogen, die door de meteoroloog zijn gevonden, is het rangnummer op de computerlijst opgezocht. Ook zijn de rangnummers bekeken van de analogen, die twee dagen of minder verschillen met de datum, die de meteoroloog vond. Indien dan de dag met het laagste nummer wordt genomen, blijkt dat er 84 analogen zijn met rang < 11 ; dat er 24 analogen zijn met rang ≥ 11 en ≤ 20 en 21 analogen met rang > 20 . Onder de analogen met rangnummer 1 t/m 10 komen dus de meeste bruikbare analogen voor, mits men ook nog de

data in beschouwing neemt, die twee dagen of minder verschillen met de data uit de computerlijst. Het aangeven van een rangnummer heeft dus wel zin. Dit rangnummer wil echter niet zeggen dat de analoog met rangnummer 1 meteorologisch gezien beslist beter moet zijn dan de analoog met rangnummer 2, etc.

3. Weervertaling met behulp van analogen

Behalve de lijst met analogen, die dagelijks voor de weerkamer wordt gemaakt, wordt er nog een tweede lijst door de computer vervaardigd. Deze geeft behalve de data van de analogen ook nog een systematisch overzicht van het weer dat op de data van de analogen optrad. Figuur 3 geeft een overzicht van de gewijzigde computeruitvoer. Onder de lijst met data wordt een overzicht gegeven van het percentage zonnenschijn te De Bilt, verdeeld in vier klassen, zoals ook op de kritiekkarten gebruikelijk is. (0%, 1%-29%, 30%-59% en $\geq 60\%$). Het aantal analogen dat tot de verschillende klassen behoort staat opgegeven. Omdat de zonnenschijngegevens uit de jaren 1949 en 1950 niet aanwezig waren op het moment dat de magneetband werd vervaardigd, zal het totale aantal in de klassen kleiner dan 30 kunnen zijn.

Ook van de neerslag in De Bilt gedurende het etmaal van 18.00z - 18.00z is een frekwentieverdeling gegeven, waarbij de klassenindeling overeenkomt met die van de kritiekkarten. (< 0.3 mm, 0.3 mm-1.4 mm, 1.5 mm-4.4 mm en ≥ 4.5 mm).

Bij de landelijke neerslag is uitgegaan van tien waarnemingsstations. De indeling is volgens de kritiekkart in de klassen D, MD, V, PL, MR. De aantallen in de klassen hebben echter betrekking op de neerslag van 18.00z - 06.00z, terwijl de meteoroloog een kritiekkart invult, die betrekking heeft op de landelijke neerslag van 18.00z - 18.00z. Om deze reden zijn de gegevens van de landelijke neerslag niet verder uitgewerkt.

Op de onderste twee regels van de computeruitvoer staan de maximum- en de minimumtemperatuur, gemiddeld over de vijf hoofdstations en gerangschikt naar grootte.

Nagegaan is in hoeverre deze frekwentieverdelingen kunnen bijdragen tot het opstellen van de weersverwachting. Hiervoor zijn voor acht maanden dat door de computer analogen zijn geselecteerd (november 1974 t/m juni 1975) de kritiekkarten K 17-3 ingevuld. Dit zijn de kritiekkarten van de verwachting 3 dagen vooruit. De resultaten van deze kritiekkarten zijn vergeleken met de resultaten, die de weerkamer op de overeenkomstige dagen behaalde.

3.1 Zonneschijn

Om een kritiekkart op een eenduidige wijze te kunnen invullen, is het nodig een bepaalde "aanstreeptaktiek" vast te leggen. Dit is als volgt gedaan. Van de vier klassen waarin de zonneschijnduur is verdeeld, worden steeds twee naast elkaar liggende klassen aangestreept en wel die klassen waarin de klimatologische kans kleiner is dan de kans, zoals die door de analogenlijst wordt gegeven. Dit kan moeilijkheden geven, indien bijvoorbeeld klasse I en klasse III bij de analogen zeer sterk vertegenwoordigd zijn. Om na te gaan of dit effect veelvuldig aanleiding geeft tot verschillen bij het aanstrepen van de kritiekkart, zijn door twee personen onafhankelijk van elkaar de kritiekkarten voor de zonneschijnduur ingevuld. Verschillen kwamen echter zo zelden voor, dat mag worden aangenomen dat de kritiekkart op een genoeg eenduidige wijze kan worden ingevuld.

Tabel II laat van de maand februari de frekwentieverdelingen voor de geselecteerde analogen zien. Tevens geeft de tabel een overzicht van de opgetreden en aangestreepte klassen.

Tabel III geeft een overzicht van de PI van de zonneschijnduur voor dag 3. Het blijkt dat voor 184 Amerikaanse 500 mbar prognoses met de computer analogen zijn bepaald. De gemiddelde PI, die de weerkamer over deze 184 gevallen haalde, is 0.03. Deze waarde is representatief voor de resultaten van de meteorologen. Uit het overzicht van gescoorde PI's, zoals dit door de OD maandelijks wordt uitgegeven, blijkt dat over de periode juni 1974 t/m

juni 1975 ook een PI van 0.03 is gescoord. Met de computerlijst is een PI van 0.11 te behalen. Dit is duidelijk méér dan de weerkamer. De PI, die met de analogenmethode wordt gescoord, blijkt in zeven van de acht maanden die zijn onderzocht hoger te zijn dan die, welke door de meteorologen is gescoord. Het verschil in PI is dus systematisch en wordt niet door een enkele uitschieter bepaald.

Uit tabel III blijkt ook dat het aantal treffers, dat door de meteorologen is behaald, iets groter is dan volgens de computerlijst (trefferpercentage resp. 68 en 66%). De inzet van de meteorologen is echter veel groter dan bij de analogenmethode.

3.2 Neerslag

Op de kritiekaart is behalve de zonneschijnduur ook de neerslag te De Bilt ingevuld. De volgende aanstreeptaktiek is gevolgd. Aangestreept is "droog" (klasse I) of "neerslag" (klasse II, III en IV). Alleen indien zowel voor klasse I als voor klasse II de frekwentie bij de analogen veel groter is dan de klimatologische frekwentie, zijn deze beide klassen aangestreept.

Tabel IV geeft een overzicht van de gevonden resultaten. Het blijkt dat de PI, die door de weerkamer wordt bereikt (0.08), nagenoeg gelijk is aan de PI, die met de analogenmethode wordt gescoord (0.09). Het verschil van 0.01 in de PI is beslist niet significant. Het trefferpercentage is zowel bij de meteorologen als bij de analogenmethode 63.

3.3 Maximumtemperatuur

In de computerlijst is ook de maximumtemperatuur, gemiddeld over de vijf hoofdstations, opgenomen. Nagegaan is in hoeverre deze grootte, die door de meteorologen wordt voorspeld, ook uit de computerlijst kan worden bepaald. Hiervoor is het verschil tussen de prikwaarde die door de meteorologen wordt opgegeven en de opgetreden maximumtemperatuur vergeleken met het verschil tussen de prikwaarde die uit de analogenlijst volgt en de opgetreden maximumtemperatuur. De prikwaarde volgens de analogenlijst

is als volgt bepaald. Onder aan de analogenlijst wordt een overzicht gegeven van de maximumtemperaturen, naar grootte gerangschikt. Hierin wordt een interval ter lengte van 5° C opgezocht, wat overblijft indien de extreme waarden (zowel de extreem hoge als de extreem lage waarden) worden weggelaten. Als prikwaarde wordt de middelste waarde van het gevonden interval van 5° genomen.

Uit tabel V blijkt dat de weerkamer de maximumtemperatuur gemiddeld iets nauwkeuriger voorspelt dan met de analogenlijst gebeurt. De gemiddelde afwijking

$$\frac{1}{N} \sum |R_{\text{obs}} - T_{\text{for}}|$$

is voor de weerkamer 1.73° en bij de analogenmethode 1.78. Het is echter nog een zeer redelijk resultaat, zeker wanneer we bedenken dat de winter 1974-1975 bijzonder zacht was. Het voorspellen van extreme waarden zal vooral bij de analogenmethode zelden voorkomen, omdat het aantal extreem zachte wintermaanden in het computerbestand gering is.

Er zijn echter nog twee oorzaken waarom de voorspelling van de maximumtemperatuur met behulp van de analogenmethode moeilijkheden zal geven:

1. Bij het voorspellen van de maximumtemperatuur is de voorgeschiedenis van belang. Bijvoorbeeld het al of niet aanwezig zijn van sneeuw op het continent zal in de winter van invloed zijn op de maximumtemperatuur. Dergelijke gegevens zal de meteoroloog bij zijn voorspelling gebruiken. De analogenmethode in de huidige vorm heeft echter geen enkel "geheugen"; iedere dag worden 30 analogen geselecteerd louter en alleen op de vorm van het stromingspatroon.
2. Uit de 72-uur prognoses worden in de weerkamer RVA-gebieden berekend. Deze worden gebruikt bij het tekenen van een 72-uur prognose van de grondkaart. Op deze grondkaart worden fronten getekend, die de scheiding tussen luchtmassa's met verschillende temperatuur aangeven. De positie van de fronten en de eigenschappen van de aanwezige luchtmassa's zijn een hulpmiddel

bij het opstellen van de verwachting voor de maximumtemperatuur. Het is maar zeer de vraag of de analogenmethode, die rechtstreeks met 500 mbar gegevens werkt, ook deze synoptische informatie gebruikt.

3.4 Minimumtemperatuur (T_N)

Op dezelfde wijze als voor de maximumtemperatuur is ook een verwachting voor de minimumtemperatuur, gemiddeld over de vijf hoofdstations, opgesteld. Dit bleek echter weinig zinvol. De resultaten zijn duidelijk slechter dan de resultaten van de weerkamer, zoals uit tabel VI blijkt. Waarschijnlijk is bij het opstellen van een verwachting voor de minimumtemperatuur meer detailinformatie nodig dan bij de analogenmethode wordt gebruikt.

3.5 Discussie

1. Bij het opstellen van statistische voorspelregels is het zinvol het materiaal op te splitsen in twee gedeelten, nl. een gedeelte waarop de voorspelregel wordt afgeleid (afhankelijk materiaal) en een gedeelte waarop de regel kan worden toegepast (onafhankelijk materiaal). Bij de voorspelling van de weergrootheden, zoals die met de analogenmethode is uitgevoerd, kan nauwelijks van afhankelijk of onafhankelijk materiaal worden gesproken. Er zijn een aantal analogenmethoden naast elkaar uitgetoetst gedurende de maanden oktober t/m december, om na te gaan welke methode het beste voldoet. Er is echter niet nagegaan welke van deze methoden de hoogste PI geeft. Het komt er dus op neer, dat het materiaal van alle acht maanden november '74 t/m juli '75 als onafhankelijk kan worden beschouwd.
2. Voor het voorspellen van de zonneschijnduur op dag 3 is de analogenlijst bijzonder zinvol. Volgens het memorandum van H. Daan (CWD-75-3) is de PI van 0.03, die door de sectie meerdaagse wordt behaald, niet significant verschillend van 0, terwijl de PI van 0.11, die uit de analogenmethode volgt, dit wel is.

3. Bij het voorspellen van de hoeveelheid neerslag te De Bilt en de maximumtemperatuur, gemiddeld over de vijf hoofdstations, kan de analogenmethode een zinvol hulpmiddel zijn. Bij het opstellen van een verwachting van de minimumtemperatuur, gemiddeld over de vijf hoofdstations, zal de analogenmethode waarschijnlijk geen bijdrage kunnen leveren.

Tabel I. Overzicht van het aantal analogen, dat zowel door de meteorologen als bij de computerselectie is gevonden. N = aantal selecties, dat door de computer is uitgevoerd.

	N	m e t h o d e						
		1	2	3	4	5	6	7
oktober	16	25	23	31	15	18	19	21
november	25	26	22	27	29	32	35	36
december	24	35	29	33	41	39	38	44
totaal	65	86	74	91	85	89	92	101

Tabel II. Frekwentieverdeling in de vier klassen van de zonneshijnduur, zoals die wordt gevonden uit de analogen van februari '75. De voorspelde klassen zijn onderstreept. Onder aan de tabel is de klimatologische frekwentieverdeling gegeven. Omdat zonneshijngegevens uit de jaren 1949 en 1950 ontbreken, is het totaal in de vier klassen in een aantal gevallen kleiner dan 30. In kolom 5 is aangegeven welke klasse is opgetreden.

k l a s s e				opgetreden
I	II	III	IV	
<u>17</u>	<u>10</u>	1	2	III
<u>14</u>	<u>4</u>	5	4	III
<u>12</u>	<u>3</u>	5	6	II
7	0	<u>4</u>	<u>13</u>	III
9	0	<u>7</u>	<u>10</u>	IV
11	4	<u>7</u>	<u>6</u>	IV
9	4	<u>9</u>	<u>7</u>	IV
9	4	<u>8</u>	<u>9</u>	IV
<u>15</u>	<u>2</u>	5	6	III
<u>17</u>	<u>5</u>	4	4	I
<u>18</u>	<u>4</u>	6	2	I
<u>18</u>	<u>6</u>	3	1	II
12	4	<u>6</u>	<u>6</u>	II
<u>13</u>	<u>8</u>	4	2	IV
<u>16</u>	<u>5</u>	2	5	I
<u>8</u>	<u>9</u>	3	2	III
<u>9</u>	<u>12</u>	3	0	IV
11	0	<u>6</u>	<u>13</u>	IV
10	0	<u>4</u>	<u>15</u>	IV
12	0	<u>4</u>	<u>12</u>	IV
5	2	<u>4</u>	<u>15</u>	III
8	2	<u>3</u>	<u>15</u>	IV
5	1	<u>6</u>	<u>18</u>	IV
8	3	<u>5</u>	<u>12</u>	IV
11	9	5	5	

Tabel III. Vergelijking van de resultaten van de zonnenschijnduurverwachting op dag 3, zoals die door de weerkamer en volgens de analogenmethode zijn behaald.

	aantal	PI		aantal treffers	
		weerkamer	analogen	weerkamer	analogen
november '74	26	0.0	0.10	18	15
december	23	0.14	0.05	18	16
januari '75	22	-0.08	-0.02	12	12
februari	24	0.09	0.20	17	17
maart	23	-0.11	0.06	13	13
april	21	-0.08	0.11	13	13
mei	23	0.14	0.24	19	19
juni	22	0.12	0.14	16	16
totaal	184	$\overline{PI} = 0.03$	$\overline{PI} = 0.11$	126	121

Tabel IV. Vergelijking van de resultaten van de verwachting van de hoeveelheid neerslag te De Bilt op dag 3, zoals die door de weerkamer en volgens de analogenmethode zijn behaald.

	aantal	PI		aantal treffers	
		weerkamer	analogen	weerkamer	analogen
november '74	26	0.12	0.10	17	17
december	23	0.01	0.19	12	16
januari '75	22	0.20	0.10	15	12
februari	24	0.05	0.18	15	19
maart	23	0.00	-0.02	13	12
april	21	0.12	0.09	14	12
mei	23	0.06	-0.02	15	13
juni	22	0.09	0.10	15	15
totaal	184	$\overline{PI} = 0.08$	$\overline{PI} = 0.09$	116	116

Tabel V. Vergelijking tussen de nauwkeurigheid van de voorspelling van de maximumtemperatuur, geniddeld over de vijf hoofdstations, zoals die door de weerkamer is gedaan en zoals die uit de analogen volgt. Als maat is genomen de grootheid $\Sigma |T_{\text{obs}} - T_{\text{for}}|$. Hierin is T_{obs} de opgetreden maximumtemperatuur, T_{for} de voorspelde maximumtemperatuur.

	aantal	$\Sigma T_{\text{obs}} - T_{\text{for}} $	
		weerkamer	analogen
november '74	26	33	37
december	23	27	33
januari '75	22	36	38
februari	24	42	63
maart	23	42	37
april	21	28	29
mei	23	53	42
juni	22	58	50
totaal	184	319	327

Tabel VI. Vergelijking tussen de nauwkeurigheid van de voorspelling van de minimumtemperatuur, gemiddeld over de vijf hoofdstations, zoals die door de weerkamer is gedaan en zoals die uit de analogenmethode volgt. Als maat is genomen de grootte $\Sigma |T_{\text{obs}} - T_{\text{for}}|$. Hierin is T_{obs} de opgetreden minimumtemperatuur, T_{for} de voorspelde minimumtemperatuur.

	aantal	$\Sigma T_{\text{obs}} - T_{\text{for}} $	
		weerkamer	analogen
november '74	26	46	51
december	23	46	52
januari '75	22	42	47
februari	24	52	59
maart	23	31	38
april	21	47	38
mei	23	28	33
juni	22	47	55
totaal	184	339	373

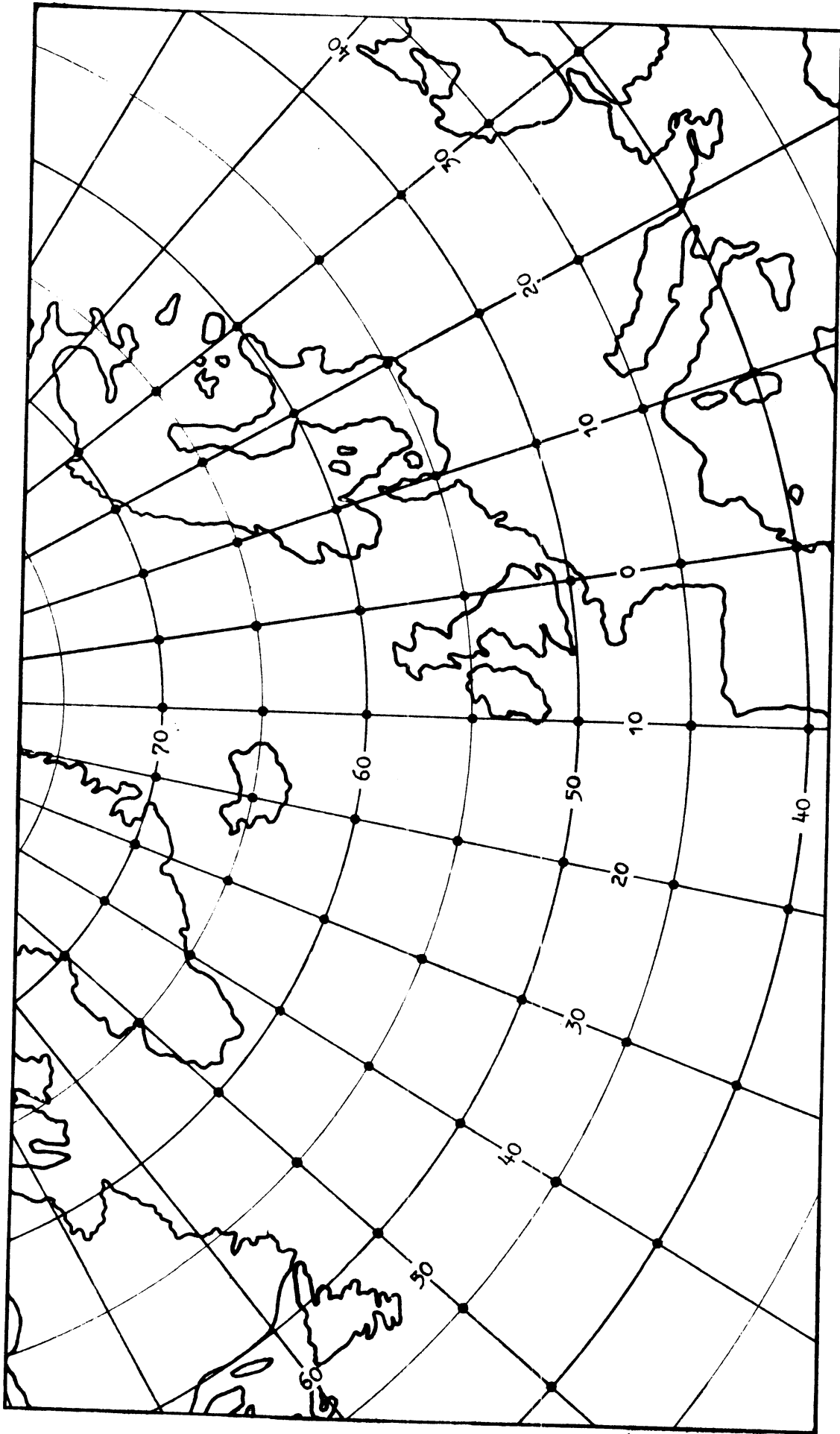


Fig. 1. ● = roosterpunten, die bij de analogeselectie worden gebruikt.

ANALOOG VAN DE PROG VAN 750404

HET BETREFFENDE HOOGTE VELD IS	542	555	564	565	560	548	538	531	526
538	557	570	570	570	558	543	530	523	521
534	555	572	571	571	557	534	519	517	524
531	550	570	571	571	557	531	513	518	530
531	548	566	571	571	558	537	521	526	540
542	550	562	568	568	560	540	534	540	553
558	555	560	567	567	563	556	548	548	564

Fig. 2.

DE DATA EN DE GEGEVENS VAN DE ANALOGEN ZIJN

DATUM	- MAXIMUM TEMPERAATUUR -				- MINIMUM TEMPERAATUUR -				DDFF	SS	RR	RR	RR	RR	GEM	RANG	
	260	280	310	380	230	380	310	280									260
550317	+5	+4	+4	+4	+4	-0	-1	+1	-1	+1	-0	3121	5	9	.8	-5.1	18
550318	+4	+2	+5	+4	+4	+0	-0	+1	-0	+1	+0	3412	4	7	1.7	-6.1	4
550319	+2	-0	+4	+2	+2	-1	-4	+0	-1	-1	-1	1109	3	10	3.8	-11.5	2
550320	+4	+3	+5	+4	+4	+4	-4	-1	-3	-1	-3	2707	6	4	.2	-16.6	11
560405	+6	+6	+5	+3	+6	+5	-1	-2	+1	-1	-0	3309	1	8	3.0	-2.5	12
560406	+6	+5	+5	+5	+5	+5	-1	-1	+0	-1	+1	114	9	8	.3	-5.4	14
570412	+7	+7	+8	+8	+8	+8	-1	-0	+2	+0	+2	2712	9	3	2.2	-.4	10
580408	+8	+6	+8	+7	+7	+7	+2	+1	+2	+1	+2	2808	3	6	.2	+2.2	15
580409	+6	+5	+8	+6	+6	+6	-2	-2	-0	-1	+2	3410	1	0	1.6	+2.9	8
580410	+5	+4	+6	+6	+4	+5	-1	+0	+0	-2	+1	3611	7	6	2.3	+2.6	7
620315	+4	+1	+4	+3	+4	+3	-4	-5	-0	-2	+0	408	8	7	.6	-11.8	9
620322	+6	+7	+5	+5	+4	+6	-2	-5	+0	-0	+1	9903	7	0	.0	-1.8	17
620323	+7	+7	+5	+6	+5	+6	+1	-1	+2	-3	+1	1306	3	1	.0	-4.0	6
620324	+5	+5	+4	+4	+5	+5	-0	-2	+1	-3	+1	310	9	0	.0	-1.8	16
700328	+7	+6	+5	+4	+5	+6	-0	-1	+1	-0	+1	3321	6	10	2.0	-4.4	19
700404	+7	+5	+7	+6	+6	+6	-0	-2	+2	-0	+1	1302	3	8	1.1	-7.0	20
700405	+6	+6	+5	+5	+6	+6	-1	-1	+1	+0	+1	3018	2	9	.4	-5.5	5
700406	+5	+2	+6	+6	+3	+5	+0	-0	+3	+1	+3	1102	0	9	9.0	-6.1	1
700407	+6	+5	+5	+6	+6	+6	-1	-1	+1	-1	+1	2508	6	9	.5	-4.4	3
700408	+6	+5	+6	+7	+6	+6	-2	-3	+1	+0	+0	1909	3	6	.6	-4.0	13

TMAX EN TMIN: GRADEN CELSIUS

DDFF: 1000 MAAL WINDRICHTING IN TIJNTALLEN GRADEN PLUS DE WINDSPELHED IN KILOPEU

SS: AANTAL UREN ZONNESCHYNSGEGEVENS ONTBREKEN

RR LAND: AANTAL STATIONS MET NEERSLAG GROTER DAN .3MM, MAXIMAAL 10 STATIONS

RR 260: NEERSLAG TE DE BILT IN MM. AFGETAP-OM 18 GIT

