

KONINKLIJK NEDERLANDS  
METEOROLOGISCH INSTITUUT

De Bilt

Verslagen

V - 265

Th. de Crook

Experimentele weersvoorspelling op lange  
termijn met de ADAM  
(ADaptive Arithmetical Method)

De Bilt, 1975

Publikationsnummer: K.N.M.I. V-265 (M.B.W.)

U.D.C.: 551.509.331 :  
551.509.323

Experimentele weersvoorspelling op lange termijn  
met de ADAM  
(ADaptive Arithmetical Method)  
deel 2<sup>\*</sup>)

Th.de Crook

INHOUD

0. Inleiding
1. Korte omschrijving van de methode
2. Experimenten met de ADAM
  - 2.1. Verschillende versies
  - 2.2. Indikatie, interpretatie en verifikatie van de voorspellingen
  - 2.3. Resultaten
  - 2.4. De significantie van de verschillende versies
  - 2.5. Konklusie.

\* ) Deel 1 is het verslag V-225 "Experimentele weersvoorspelling op lange termijn met de ADAM" van J. van Galen en Th. de Crook. (1969)

.....

## 0. Inleiding

In verslag V-225 "Experimentele weersvoorspelling op lange termijn met de ADAM (ADaptive Arithmetical Method), deel 1" is een methode ontwikkeld, waarmee voor de komende maand de gemiddelde etmaaltemperatuur voor de maand en de dekaden kan worden voorspeld.

In hoofdstuk 2 van deel 1 werden de versies O, A, E, T en U beschreven met de erbij behorende resultaten. Thans zijn echter meer resultaten ter beschikking en is bovendien de L-versie toegevoegd (zie punt 2.1.). In punt 2.3. zijn de resultaten in de tabellen 1 en 2 weergegeven. De interpretatie en de verifikatie van de voorspellingen zijn objectief gemaakt (zie punt 2.2.), zodat nu snel een objectieve uitspraak over de kwaliteit van de methode kan worden gedaan. In punt 2.4. wordt de significantie van de verschillende versies getoetst met de t-toets. Het verslag wordt afgesloten met enkele voorzichtige konklusies (zie punt 2.5.).

## 1. Korte omschrijving van de methode

De methode is ontwikkeld met het doel de maandgemiddelde etmaaltemperatuur te De Bilt te voorspellen. Teneinde tot de voorspelling te komen, wordt gebruik gemaakt van de volgende gegevens:

- a. De maandgemiddelde etmaaltemperatuur van de te voorspellen maand van 1881 t/m 1955.
- b. De Grosswetterlagen van iedere dag in een periode voorafgaande aan de te voorspellen maand van 1881 t/m 1955.
- c. De Grosswetterlagen van iedere dag in een periode voorafgaande aan de te voorspellen maand (van bv. 1969).

De ADAM bestaat in principe uit drie delen, t.w. de voorbereiding tot het leerproces, de éénmalige leerfase en de voorspelling.

### 1e. De voorbereiding tot het leerproces

Dit omvat de bepaling van de deviaties van de "klimatologische luchtdruk" (klimatologie) in drie nader te bepalen perioden (vanaf 1881). Informatie inzake de luchtdruk wordt ons verschaft door de "Grosswetterlagen" (zie b.).

### 2e. De éénmalige leerfase

Uit de deviaties (van drie perioden) en de gemiddelde etmaaltemperatuur van de te voorspellen maand (zie a.) van 1881 t/m 1955

worden volgens een bepaald procédé (zie hoofdstuk 1, deel 1) 6 x 20 prediktors met bijbehorende gewichten samengesteld.

3e. De voorspelling

Voor elk jaar wordt voor de betreffende te voorspellen maand een voorspelling gemaakt voor de maandgemiddelde etmaaltemperatuur met behulp van de deviaties en de 6 x 20 geselecteerde prediktors.

Op soortgelijke wijze kunnen ook de dekadegemiddelde etmaaltemperaturen voorspeld worden, alsmede de neerslaghoeveelheden en de gemiddelde zonnenschijnduur voor de maand en de dekaden. Voor een uitvoerige beschrijving van de ADAM wordt verwezen naar verslag V-225.

Van genoemde methode zijn een aantal versies ontwikkeld. Voor elke versie zijn een groot aantal voorspellingen gemaakt, waarvan een prestatie-index is berekend. Deze prestatie-index is een maat voor de kwaliteit van de verschillende versies en tevens een middel om de versies met elkaar te kunnen vergelijken.

2. Experimenten met de ADAM

2.1. Verschillende versies

De versie L.

De O, A, E, T en U versie zijn beschreven in deel 1 (verslag V-225). Bij de L-versie wordt de informatie inzake de luchtdruk ons niet verschaft door de "Grosswetterlagen", maar door de luchtdrukwaarden in de roosterpunten.

Voor het overige is de L-versie identiek aan de U-versie d.w.z. de klimatologie en de deviaties worden bepaald in drie aaneensluitende perioden van 15 dagen, de prediktors zijn drievoudig met als algemene vorm  $a_{\pm} b_{\pm} c_{\pm}$  ( $a_{\pm}$  is positieve of negatieve deviatie, zo ook  $b_{\pm}$  en  $c_{\pm}$  is positief of negatief deviatieverschil).

2.2. Indikatie, interpretatie en verificatie van de voorspellingen

In punt 1.4. van deel 1 zijn de drie indikaties  $U^{(B)}$ ,  $U^{(N)}$  en  $U^{(A)}$  gevormd. Zo is  $U^{(B)}$  een indikatie (getal) met betrekking tot het optreden van B(elow). De interpretatie van deze drie op objectieve wijze tot stand gekomen indikaties is ook objectief gemaakt, waarbij

ruimte is voor het aangeven van 1, 2 of 3 klassen. In het geval van drie klassen (B, N of A) is er in feite geen voorspelling gedaan. De interpretatie gaat als volgt.

Zet de drie indikaties op volgorde (de grootste eerst); de kleinste (dus laatste) indikatie wordt afgetrokken van alle drie indikaties (de laatste indikatie is nu dus nul); als de eerste en tweede indikatie nul zijn dan is de inzet = 1 en is de voorspelling B, N of A en dus juist (treffer = 1); als de tweede indikatie groter of gelijk is aan de helft van de eerste indikatie dan is de inzet =  $2/3$  en is de temperatuur die hoort bij de eerste of de tweede indikatie gelijk aan de werkelijke temperatuur dan is de voorspelling juist (treffer = 1), anders niet (treffer = 0); als de tweede indikatie kleiner is dan de helft van de eerste indikatie dan is de inzet =  $1/3$  en is de temperatuur die hoort bij de eerste indikatie gelijk aan de werkelijke temperatuur dan is de voorspelling juist, anders niet.

Voorbeeld:

Stel:  $U^{(B)} = -2$  ,  $U^{(N)} = +1$  en  $U^{(A)} = +18$  en de werkelijke temp. is A(bove).

Zet nu de drie indikaties op volgorde, dan krijgen we:

$$U^{(A)} = +18 , U^{(N)} = +1 \text{ en } U^{(B)} = -2$$

Trek de kleinste indikatie af, dat geeft:

$$U^{(A)} = +20 , U^{(N)} = +3 \text{ en } U^{(B)} = +0$$

De tweede indikatie is kleiner dan de helft van de eerste indikatie. Volgens het bovenstaande is de inzet =  $1/3$  en is de voorspelling A en is juist (treffer = 1).

Voorbeeld:

Stel:  $U^{(B)} = +13$  ,  $U^{(N)} = +15$  en  $U^{(A)} = +3$  en de werkelijke temp. is B(elow).

We krijgen dan:

$$\begin{aligned} U^{(N)} &= +15 , U^{(B)} = +13 \text{ en } U^{(A)} = +3 \\ U^{(N)} &= +12 , U^{(B)} = +10 \text{ en } U^{(A)} = +0. \end{aligned}$$

De tweede indikatie is groter dan de helft van de eerste.

Dus inzet =  $2/3$  en de voorspelling is N of B en is juist (treffer = 1).

De verifikatie kan geschieden met de prestatie-index:

$$P.I. = \frac{T - I_1}{N - I_2} \times 100$$

(zie resultaten).

De bovengenoemde manier van interpretatie en verifikatie zijn in het programma opgenomen, zodat nu voor iedere versie snel een groot aantal voorspellingen kan worden gemaakt en met de prestatie-index een uitspraak over de kwaliteit van de methode kan worden gedaan.

### 2.3. Resultaten

Om de kwaliteit van de verschillende versies te kunnen beoordelen en de versies onderling met elkaar te kunnen vergelijken wordt voor elke versie een prestatie-index (P.I.) berekend.

De prestatie-index (in %) wordt bepaald met de formule:

$$P.I. = \frac{T - I_1}{N - I_2} \times 100$$

waarbij T = het aantal opgetreden treffers,

N = het totaal aantal mogelijke treffers

$I_1$  = het aantal treffers dat op grond van het toeval optreedt (inzet), horende bij T

en  $I_2$  = het aantal treffers dat op grond van het toeval optreedt, horende bij N.

In tabel 1 zijn voor de maanden januari t/m december voor elke maand en voor elke dekade P.I.'s berekend van de versies E, T, U en L. De P.I.'s zijn berekend uit telkens 75 jaren bij het afhankelijke materiaal en uit gemiddeld 13 jaren bij het onafhankelijke materiaal. Zo is voor november de P.I. van de maand (voor het onafhankelijke materiaal) met versie E gelijke aan +6 (deze P.I. is berekend uit 13 voorspellingen). Bij het afhankelijke materiaal is de spreiding in de P.I.'s klein; voor het onafhankelijke materiaal is de spreiding aanmerkelijk groter.

Aan het einde van tabel 1 zijn gemiddelde P.I.'s berekend voor de maand en de dekade uit de maanden januari t/m december. Hierbij

zijn de P.I.'s van het afhankelijke materiaal bepaald uit  $12 \times 75 = 900$  voorspellingen. De P.I.'s van het onafhankelijke materiaal zijn bij de E- en T-versie berekend uit  $3 \times 12 + 9 \times 13 = 153$  voorspellingen, bij de U-versie uit  $6 \times 14 + 6 \times 15 = 174$  voorspellingen en bij de L-versie uit  $11 \times 12 + 11 = 143$  voorspellingen.

Zo is bijvoorbeeld voor de 1e dekade met de L-versie de berekende P.I. gelijk aan +4 (onafh. materiaal). Ook hier is de spreiding van de P.I.'s voor het onafhankelijke materiaal het grootst.

Bovendien is voor elke versie (zie onderaan tabel 1) een gemiddelde P.I. berekend voor zowel het afhankelijke als het onafhankelijke materiaal. Deze gemiddelde P.I.'s zijn berekend uit de P.I.'s voor de maand en de dekaden, dus uit  $4 \times 12 \times 75 = 3600$  voorspellingen voor het afhankelijke materiaal en voor het onafhankelijke materiaal bij de E- en T-versie uit  $4 \times 153 = 612$  voorspellingen, bij de U-versie uit  $4 \times 174 = 696$  voorspellingen en bij de L-versie uit  $4 \times 143 = 572$  voorspellingen.

Tabel 1 (T 24)

Ver- M. sie en dek	Onafh. materiaal				Afh. materiaal				
	E	T	U	L	E	T	U	L	
1,0	-25	-19	+ 0	-46	59	67	81	82	jan, maand jan. 1e dekade jan. 2e dekade jan. 3e dekade
1,1	+25	+25	+18	+13	58	63	81	79	
1,2	-13	- 6	+ 4	-33	51	58	84	79	
1,3	+ 0	+25	-21	+ 4	55	52	79	87	
2,0	+25	+25	+17	+41	61	62	82	77	februari
2,1	-27	-38	-20	-18	63	73	79	73	
2,2	+19	+19	+40	+18	57	66	85	83	
2,3	- 4	- 4	- 7	-36	55	66	88	86	
3,0	+19	+36	+17	+50	57	62	83	81	maart
3,1	+13	+42	+27	+25	62	69	77	74	
3,2	+25	+31	+43	+ 0	62	63	87	81	
3,3	+13	+19	- 3	+21	59	77	80	70	
4,0	+13	+ 8	+10	-21	51	64	83	79	april
4,1	+25	+ 2	+10	+13	56	70	82	81	
4,2	+ 2	+13	+30	+17	58	62	82	73	
4,3	+42	+25	+47	-21	64	67	87	73	
5,0	-10	-15	-10	+25	64	73	84	84	mei
5,1	+ 2	+ 2	+20	- 8	54	72	83	82	
5,2	+ 2	-10	-27	+17	49	49	88	81	
5,3	-10	- 4	- 3	- 8	57	61	85	83	
6,0	-10	-21	-13	+ 4	66	70	88	79	juni
6,1	-27	+ 2	- 0	-21	47	57	85	82	
6,2	-21	- 4	-33	-33	60	68	88	77	
6,3	-10	+ 2	+ 0	-25	59	61	86	83	

Vervolg tabel 1

7,0	-15	-21	-37	+29	59	66	79	83	juli
7,1	+13	+13	-10	+ 4	51	62	83	87	
7,2	+13	- 4	+ 0	+13	56	60	81	85	
7,3	- 4	+13	+ 3	+50	51	64	81	81	
8,0	- 4	- 4	+ 0	+29	61	66	87	87	augustus
8,1	- 4	+ 8	-14	+17	52	59	85	84	
8,2	-10	+ 2	- 4	- 8	51	55	80	87	
8,3	+ 8	+13	+ 0	-21	59	70	82	79	
9,0	+ 2	+13	+ 4	+ 0	64	70	87	81	september
9,1	+ 2	+ 2	-18	+29	51	65	86	85	
9,2	+25	+19	+25	+17	57	65	83	80	
9,3	+ 8	+13	- 7	+13	57	65	85	87	
10,0	- 4	- 4	-36	+ 0	67	67	83	87	oktober
10,1	- 4	+ 2	-21	- 4	52	69	85	83	
10,2	+13	+13	- 4	+33	50	64	84	85	
10,3	-15	- 4	+ 0	+ 0	57	63	85	83	
11,0	+ 6	+13	-39	-13	62	64	77	79	november
11,1	- 6	+ 0	+ 0	- 4	58	59	87	87	
11,2	+ 0	+ 6	+18	+ 4	57	62	74	80	
11,3	+ 6	+13	+ 7	- 4	51	59	77	82	
12,0	+ 0	+ 0	- 7	+ 8	53	61	83	73	december
12,1	+ 6	+13	-32	- 4	55	67	85	81	
12,2	+ 0	- 6	+29	+29	59	64	81	83	
12,3	+38	+38	+25	- 8	63	67	33	80	
maand	+ 0	+ 1	- 6	+ 9	60	66	83	81	
1e dek.	+ 2	+ 6	- 3	+ 4	55	65	83	82	
2e dek.	+ 5	+ 6	+10	+ 6	55	61	83	81	
3e dek.	+ 6	+12	+ 4	- 3	57	64	83	82	
gem.	+ 3	+ 6	+ 1	+ 4	57	64	83	82	

We zien dat voor het onafhankelijke materiaal de gemiddelde P.I. van de E, T, U en L-versie resp. +3, +6, +1 en +4 is. Voor de U-versie is de gemiddelde P.I. gelijk aan +1, d.w.z. dat de voorspelling nauwelijks beter is dan op grond van het toeval verkregen zou zijn.

Voor het afhankelijke materiaal zijn de gemiddelde P.I.'s voor de versies E, T, U en L resp. +57, +64, +83 en +82.

We zien dat de U- en L-versie bij het afhankelijke materiaal een duidelijk hogere gemiddelde P.I. hebben dan bij de E- en T-versie. Bij het onafhankelijke materiaal is dat duidelijk niet het geval.

De P.I.'s in tabel 2 hebben betrekking op de voorspellingen van de gemiddelde etmaaltemperatuur (T24) voor de maand en de dekaden. De voorspellingen zijn gemaakt met behulp van 6 \* 20 geselecteerde prediktors (zowel positieve als negatieve), zie deel 1. Worden alleen positieve prediktors gebruikt dan is de voorspelling in het algemeen slechter. Worden er minder of meer prediktors gebruikt (bv. 6 \* 10 prediktors) dan is de voorspelling in het algemeen ook slechter.

Tabel 2 (U-versie)

Ver- sie M. en dek.	Onafn. materiaal		Afh. materiaal	
	RR	S	RR	S
maand	-3	-3	83	83
1e dek.	+4	+1	84	83
2e dek.	-4	+1	82	84
3e dek.	-2	-5	83	82
gem.	-1	-1	83	83

In tabel 2 zijn met de U-versie gemiddelde P.I.'s berekend, die betrekking hebben op de voorspellingen van de neerslaghoeveelheden (RR) en de zonschijnduur (S) voor de maand en de dekaden. Hierbij zijn de P.I.'s van het afhankelijke materiaal bepaald uit 900 voorspellingen en de P.I.'s van het onafhankelijke materiaal uit 174 voorspellingen.

Bovendien is een gemiddelde P.I. berekend uit de P.I.'s voor de

maand en de dekaden, dus uit  $4 \times 900 = 3600$  voorspellingen voor het afhankelijke materiaal en uit  $4 \times 174 = 696$  voorspellingen voor het onafhankelijke materiaal. We zien dat voor het onafhankelijke materiaal de gemiddelde P.I. voor zowel RR als S -1 is, d.w.z. dat de voorspelling niet beter is dan op grond van het toeval verkregen zou zijn.

Voor het afhankelijke materiaal is de gemiddelde P.I. voor zowel RR als S gelijk aan +83, dus vrij hoog. De voorspellingen zijn gemaakt met behulp van  $6 \times 20$  geselecteerde prediktors.

#### 2.4. De significantie van de verschillende versies

De significantie wordt getoetst met behulp van de t-toets (zie verslag V-225, blz. 17).

$$\text{Voor de t-toets geldt: } t = \frac{(\bar{x} - \mu) \sqrt{n-1}}{\sigma}$$

Hierbij is  $\sigma$  de standaardafwijking en  $\mu$  de verwachtingswaarde.

Nemen we voor  $x$  de 48 P.I.'s volgens tabel 1 - E-versie, onafhankelijk materiaal - dan vinden we voor  $\mu = 0$ :

$$n = 48$$

$$\bar{x} = 2.96$$

$$\sigma = 15.60$$

$$\text{dus } t = \frac{(2.96 - 0) \sqrt{47}}{15.60} = 1.30.$$

Volgens de tabel op blz. 251 van het boek "Standard Mathematical Tables" (1959) komt dit overeen met een grens van 19%.

Stellen we de significantie-grens op 5%, dan wil dat zeggen dat de gemiddelde P.I. van +3 niet significant van nul verschilt, met andere woorden de E-versie is niet significant.

Voor de T-versie - zie tabel 1, onafhankelijk materiaal - vinden we:

$$n = 48$$

$$\bar{x} = 6.38$$

$$\sigma = 15.92$$

$$\text{dus } t = \frac{(6.38 - 0) \sqrt{47}}{15.92} = 2.75.$$

Dit komt overeen met een grens van minder dan 1%, dus de gemiddelde P.I. van +6 verschilt significant van de verwachtingswaarde  $\mu = 0$ ,

m.a.w. de T-versie is significant verschillend van de methode, waarbij de voorspelling op grond van het toeval wordt gedaan.

Voor de U-versie - zie tabel 1, onafhankelijk materiaal - vinden we:

$$n = 48$$

$$\bar{x} = 1.25$$

$$\sigma = 20.72$$

$$\text{dus } t = \frac{(1.25 - 0)\sqrt{47}}{20.72} = 0.41.$$

Dit komt overeen met een grens van 68%, dus de gemiddelde P.I. van +1 verschilt niet significant van de verwachtingswaarde  $\mu = 0$ , m.a.w. de U-versie is niet significant.

Voor de L-versie - zie tabel 1, onafhankelijk materiaal - vinden we:

$$n = 48$$

$$\bar{x} = 3.90$$

$$\sigma = 21.97$$

$$\text{dus } t = \frac{(3.90 - 0)\sqrt{47}}{21.97} = 1.22.$$

Dit komt overeen met een grens van 23%, dus de gemiddelde P.I. van +4 verschilt niet significant van nul, m.a.w. de L-versie is niet significant verschillend van de methode, die berust op het toeval.

## 2.5. Konklusie

Op grond van de resultaten (zie tabel 1 en 2) kan de ADAM geen sukses worden genoemd.

De gemiddelde P.I.'s van de verschillende versies voor het afhankelijke materiaal zijn hoog, hetgeen voor de hand ligt als we bedenken dat met dit materiaal de prediktors samengesteld zijn en de voorspelling in feite een terug-voorspelling is.

Bij het onafhankelijke materiaal worden de gevonden prediktors toegepast op nieuw materiaal, hetgeen dan een "echte" voorspelling oplevert. De gemiddelde P.I.'s van de verschillende versies zijn klein voor het onafhankelijke materiaal. Blijkbaar zit in de gevonden prediktors niet voldoende informatie om veel goede voorspellingen te maken.

Bij het afhankelijke materiaal zijn we er in geslaagd de gemiddelde P.I.'s steeds groter te maken (vergelijk de verschillende versies: de O-versie heeft een gemiddelde P.I. van +52; de gemiddelde P.I. van de U-versie is +33). Maar een hogere gemiddelde P.I. bij het afhankelijke materiaal is blijkbaar geen garantie voor een hogere gemiddelde P.I. bij het onafhankelijke materiaal (zie tabel 1, onafh. materiaal).

Volgens de t-toets levert alleen de T-versie een gemiddelde P.I. (+6) op die significant verschillend is van de methode, waarbij de voorspelling op grond van het toeval wordt gedaan (in dit geval gemiddelde P.I. is gelijk nul). Alle andere versies zijn niet significant. D.w.z. alleen bij T-versie boeken we winst. Deze uitspraak is natuurlijk aanvechtbaar. Hoewel de significantie-grens minder is dan 1%, blijft de mogelijkheid bestaan dat de gemiddelde P.I. niet significant is. Dit wordt nog versterkt door de gemiddelde P.I. voor de maand die bij de T-versie slechts +1 is en zeker niet significant is. De hoogste gemiddelde P.I. treedt op bij de 3e dekade (+12), hetgeen we zeker niet hadden verwacht (eerder hadden we de hoogste gemiddelde P.I. verwacht bij de 1e dekade). De gemiddelde P.I. voor de 3e dekade is echter wel significant. De beste opeenvolging van gemiddelde P.I.'s vinden we bij de L-versie (voor de maand +9, de 1e dekade +4, de 2e dekade +6 en de 3e dekade -3). De versie is echter volgens de t-toets duidelijk niet significant (zie 2.4.).

Ook de gemiddelde P.I. voor de maand (+9) is niet significant. Met de L-versie boeken we dus beslist geen winst. Ook de gemiddelde P.I.'s in tabel 2 zijn niet significant, d.w.z. dat we met de U-versie geen winst boeken bij het voorspellen van de neerslaghoeveelheid en de gemiddelde zonneshijnduur.

Resumerend kunnen we zeggen:

Alleen de T-versie levert enige winst op bij het voorspellen van de gemiddelde etmaaltemperatuur. Deze konklusie is echter aanvechtbaar, omdat de winst voor een belangrijk deel wordt geleverd door de 3e dekade.

.....