

K O N I N K L I J K N E D E R L A N D S
M E T E O R O L O G I S C H I N S T I T U U T

D e B i l t

WETENSCHAPPELIJK RAPPORT

W.R. 77-2

G.J. Cats

Berekening van de geowind

De Bilt, 1977

Publikationsnummer: K.N.M.I. W.R. 77-2 (MO)

U.D.C.: 551.55 :
551.511.32

Abstract

A method for calculating the geostrophic wind, that is, by definition, the vector $\left(\frac{1}{\rho f} \left(\frac{-\partial p}{\partial y}, \frac{\partial p}{\partial x}\right)\right)$ (so a multiple of the gradient of air pressure), from synoptic air pressure observations in the Netherlands, is presented. The method smoothes the observed air pressure values by means of a principal components analysis, and interpolates the principal components by means of a simple two-dimensional Taylor series.

The accuracy of the method can be estimated only in an indirect way: the pressure gradient is not measured directly. Probably the method leads to a root mean square deviation of 0.5 m/s in the length and of 5° in the direction of the geostrophic wind.

With the help of the principal components analysis a procedure was developed to complete missing pressure observations, with a root mean square deviation that reaches about 0.7 mbar in the case of four missing observations. The principal components analysis was further used to estimate the 10-meter wind from the air pressure observations.

This appeared, without taking instationarity effects into account yet, possible for the station "De Bilt" (06260) with a root mean square (r.m.s.) error of 1.5 m/s and for the station "Zestienhoven" (06344) with a r.m.s. error of 1.9 m/s in both horizontal wind components, corresponding to a r.m.s. error of 1.2 and 1.8 m/s in the length and 24° and 24° in the direction of the 10-meter wind in De Bilt and Zestienhoven respectively.

Samenvatting

Een methode om de geowind, die gedefinieerd wordt als de vector $\left(\frac{1}{\rho f} \left(-\frac{\partial p}{\partial y}, \frac{\partial p}{\partial x} \right) \right)$ (dus als een veelvoud van de luchtdrukgradiënt), te berekenen uit de synoptische luchtdrukwaarnemingen in Nederland, wordt gepresenteerd. De methode gebruikt ter afvlakking van de gemeten luchtdrukwaarden de hoofdcomponentenanalyse en ter interpolatie van de hoofdcomponenten een eenvoudige tweedimensionale Taylor-reeks.

De nauwkeurigheid van de methode kan alleen op indirecte wijze worden geschat: immers, de luchtdrukgradiënt wordt niet direct gemeten. Vermoedelijk leidt de methode tot een standaardafwijking van 0.5 m/s in de grootte en van 5° in de richting van de geowind.

Met behulp van de hoofdcomponentenanalyse werd een procedure ontwikkeld waarmee ontbrekende luchtdrukwaarnemingen kunnen worden ingevuld, met een standaardfout die oploopt tot ongeveer 0.7 mbar bij vier ontbrekende waarnemingen.

De hoofdcomponentenanalyse werd nog toegepast om uit de luchtdrukwaarnemingen de grondwind te schatten: dit bleek, nog zonder rekening te houden met instationariteitseffecten, mogelijk, met voor De Bilt een standaardfout van 1.5 m/s en voor Zestienhoven een standaardfout van 1.9 m/s in beide horizontale windcomponenten, overeenkomend met een standaardfout van 1.2 resp. 1.8 m/s in de lengte en 24° resp. 24° in de richting van de grondwind te De Bilt resp. Zestienhoven.

Inhoud

- I Inleiding

 - II Analysemethoden van het luchtdrukveld
 - 1. Inleiding
 - 2. Hoofdcomponentenanalyse
 - 3. "Objectieve" analyse conform analyse van het 500 mbar vlak
 - 4. Analyse op een beperkt aantal stations
 - 5. Interpolatiemethode van Gandin

 - III Hoofdcomponentenanalyse
 - 1. Inleiding
 - 2. Tweedimensionaal voorbeeld
 - 3. Hoofdcomponenten - aanvullingen en toepassing
 - 3.a Notatie
 - 3.b Berekening van de hoofdvelden binnen Nederland
 - 4. Interpolatie van de hoofdvelden
 - 4.a Doel
 - 4.b Methode

 - IV Verificatie van de methode
 - 1. Inleiding
 - 2. Statistische beschouwing
 - 3. Vergelijking geowind - gemeten wind
 - 3.a Vergelijking geowind - 200-meter (Cabau)wind
 - 3.b Vergelijking geowind - 10-meter wind

 - V Conclusies en aanbevelingen
- Dankbetuiging
- Appendix

I Inleiding

Onder geostrofische wind, kortweg "geowind", wordt verstaan de wind, die waait aan de grond, indien geen wrijving optreedt, bij een luchtdrukveld met rechte equidistante isobaren in stationaire omstandigheden, oftewel die wind, waarvan de componenten zijn $\left(-\frac{1}{f\rho} \frac{\partial p}{\partial y}, \frac{1}{f\rho} \frac{\partial p}{\partial x} \right)$. Hierin is, zoals gebruikelijk, p de luchtdruk, f de coriolisparameter, ρ de luchtdichtheid en x en y de west-oost resp. zuid-noord as.

De geowind is interessant, omdat hij vaak op eenvoudige wijze tot een benadering voor de werkelijke wind, b.v. op 10 meter hoogte, kan leiden en bovendien omdat hij, in de vorm van een luchtdrukgradiënt, optreedt in de bewegingsvergelijkingen van de atmosfeer, of de planetaire grenslaag.

Daarom is een poging in het werk gesteld de geowind te berekenen; hierbij staan eigenlijk alleen de synoptisch waargenomen luchtdrukken ter beschikking van de berekening, omdat nu eenmaal geen gradiënten worden waargenomen, terwijl de grondwind als te onbetrouwbaar, onder andere wegens locale effecten, moet worden aangemerkt.

In hoofdstuk II zullen enkele methoden om de geowind te bepalen worden besproken; het zal blijken dat van deze methoden de hoofdcomponentenanalyse het best voldoet. In hoofdstuk III wordt deze nader bekeken en in hoofdstuk IV worden enkele verificatiemogelijkheden besproken. Tenslotte geeft hoofdstuk V conclusies en aanbevelingen.

Steeds zal worden uitgegaan van de Nederlandse situatie: een dicht net van meetpunten met kleine verschillen. "Dicht" betekent hierbij dat de onderlinge afstand van de stations veel kleiner is dan de lengteschaal van de overheersende druksystemen, "kleine verschillen" dat de verschillen tussen de stations vaak niet veel groter zijn dan de meetfouten, die op de stations ontstaan door ijk- en afleesfouten en een standaardwaarde van, zeg, 0.3 mbar hebben.

II Analysemethoden van het luchtdrukveld

1. Inleiding

Om uit de waargenomen luchtdrukken de luchtdrukgradiënt te kunnen bepalen, is het het meest voor de hand liggend eerst het luchtdrukveld zelf te analyseren. Onder een analyse wordt verstaan een interpolatie door middel van een gladde functie, waarbij rekening moet worden gehouden met fouten in de waarnemingen.

In het volgende zullen de hoofdcomponentenanalyse, de "objectieve analyse" (conform die van het 500 mbar vlak) en de door de weerkamer toegepaste analyse ter bepaling van de geowind door middel van interpolatie op drie punten ter sprake komen.

2. Hoofdcomponentenanalyse

Bij een hoofdcomponentenanalyse wordt het luchtdrukveld geschreven als een lineaire combinatie van "hoofdvelden", oftewel "hoofdcomponenten". De ervaring leert dat slechts enkele hoofdvelden "significant"*) optreden en die zijn dan tamelijk glad, terwijl alle andere tamelijk grillige structuren hebben. Helaas is het niet mogelijk onderscheid te maken tussen meetfouten en locale effecten.

Bij een voorlopig onderzoek bleek ook in de Nederlandse situatie bovenstaande splitsing in "significante" en overige hoofdvelden uitvoerbaar. Bij deze experimenten en bij onderzoek naar de toepasbaarheid van de hoofdcomponentenanalyse in combinatie met een n^{de} graads polynoombenadering voor de hoofdvelden werd gebruik gemaakt van nog niet gepubliceerde resultaten, die dhr. W.J.A. Kuipers al had verkregen door soortgelijk onderzoek op een luchtdrukbestand van Nederland uit de jaren 1970 t.e.m. 1973.

De resultaten van deze experimenten waren zódanig, dat besloten werd deze methode te prefereren boven de andere, die hieronder

*) In afwijking van het gebruik in de statistiek wordt het woord "significant" hier gebruikt om aan te duiden dat slechts een klein aantal hoofdvelden een grote bijdrage geeft aan de beschrijving van het luchtdrukveld.

worden beschreven.

Opgemerkt dient te worden, dat de methode volkomen objectief is.

3. "Objectieve" analyse conform analyse van het 500 mbar vlak

Voor toepassing van numerieke verwachtingsmodellen dient het grootschalige luchtdrukveld te worden geanalyseerd. Momenteel is hiertoe een objectieve-analyse programma op het KNMI operationeel; voor een beschrijving zie [1].

Het bleek onmogelijk de parameters, die dit model enige subjectiviteit geven, zo te kiezen dat van het Nederlandse luchtdrukveld een betrouwbare analyse kon worden verkregen: de invloedssferen van de stations moesten, om een niet al te grillig patroon te verkrijgen, in de orde van 200 km gekozen worden, zodat het programma over grote gedeelten van Nederland slechts de gemiddelde luchtdruk kon reproduceren. Bij kleinere invloedssferen konden de geowinden lokaal oplopen tot 70 m/s, hetgeen bepaald onrealistisch was.

4. Aanlyses op een beperkt aantal stations

Gezien het feit dat de luchtdrukverschillen binnen Nederland vaak in grootte vergelijkbaar zijn met de standaardmeetfouten van de stations, verdient het aanbeveling zoveel mogelijk stations in de analyse te betrekken. Gebruik van een klein aantal stations moet alleen daarom al worden afgeraden. Bovendien is het a priori niet duidelijk welke stations gekozen zouden moeten worden. Om deze redenen zijn dergelijke analysemethoden niet nader onderzocht.

5. Interpolatiemethode van Gandin

De zogenaamde "optimale interpolatiemethode van Gandin" wordt in een appendix besproken; het blijkt daar dat onder enkele voor de hand liggende veronderstellingen deze methode equivalent is aan de hoofdcomponentenanalyse.

III Hoofdcomponentenanalyse

1. Inleiding

De theorie die ten grondslag ligt aan de hoofdcomponentenanalyse wordt elders uitvoerig beschreven, zie b.v. Lawley en Maxwell [2].

Hier zal aan de hand van een tweedimensionaal voorbeeld worden verduidelijkt hoe de methode van de hoofdcomponentenanalyse werkt; in paragraaf 3 volgen dan de aanvullingen voor en de toepassing op de Nederlandse situatie. Paragraaf 4 behandelt de interpolatie van de hoofdvelden.

2. Tweedimensionaal voorbeeld

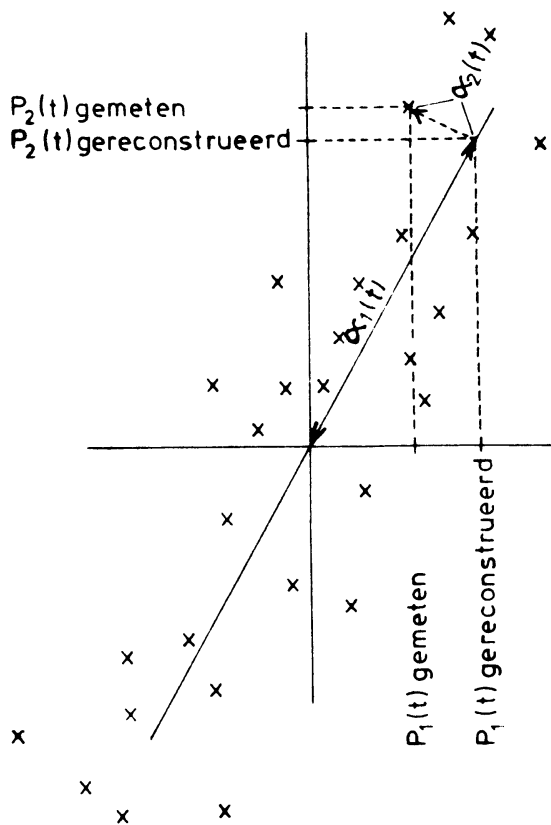
De hoofdcomponentenanalyse is in essentie niets anders dan een voorschrift om het te beschrijven luchtdrukveld te schrijven als een lineaire combinatie van orthonormale basisvectoren. De basisvectoren worden zo gekozen, dat de i^{de} basisvector zo veel mogelijk variantie van het luchtdrukveld beschrijft, onder de voorwaarde dat hij genormeerd is en loodrecht staat op de voorafgaande basisvectoren $1 \dots i-1$.

Tweedimensionaal wordt dit als volgt:

Veronderstel dat twee meetreeksen, $p_1(k)$ en $p_2(t)$ gegeven zijn, en dat men een correlatie tussen die verwacht. Beide meetreeksen zijn per veronderstelling behept met toevallige meetfouten, die ongecorrleerd optreden. Men denke hierbij b.v. aan de luchtdrukwaarnemingen van twee dicht bij elkaar gelegen stations.

Men mag veronderstellen dat de meetreeksen een gemiddelde 0 hebben; indien dit namelijk niet het geval is, kan men overgaan op een nieuwe meetreeks $p_i'(t) = p_i(t) - \frac{1}{N} \sum_{t=1}^N p_i(t)$.

De meting op tijdstip t kan in een grafiek worden weergegeven, waarbij $p_1(t)$ de abscis en $p_2(t)$ de ordinaat wordt van het punt dat behoort bij de meting. Er ontstaat dan onderstaande figuur.



Bij de analyse van de meetreeksen wil men de meetfout kunnen scheiden van de opgetreden luchtdruk. Onder de veronderstelling dat de meetfouten niet gecorreleerd optreden met de meetwaarde en met elkaar, geeft dan de correlatie tussen $p_1(t)$ en $p_2(t)$ een startpunt voor de analyse.

In de figuur is de lijn, waarop alle meetpunten zouden liggen indien er geen meetfout was, en de correlatie tussen $p_1(t)$ en $p_2(t)$ 100%, getekend. De afstand van een meetpunt tot die lijn, $\alpha_2(t)$ genoemd, of ook de "score van de tweede hoofdcomponent", is dan een maat voor de meetfout.

Dit kan men in formules als volgt behandelen:

Schrijf $p_i(t)$ als som van een werkelijke waarde $P_i(t)$ en een meetfout $\rho_i(t)$, en neem aan dat $P_1(t)$ en $P_2(t)$ 100% correleren:

$P_2(t) = \mu P_1(t)$. Dan geldt:

$$\overline{P_i \rho_j} = 0 \quad \text{en} \quad \overline{\rho_i \rho_j} = \rho P_1^2 \delta_{ij}$$

voor zekere ρ , onder de aanname dat $\overline{\rho_1^2} = \overline{\rho_2^2} (= \rho P_1^2)$.

Hierin is $\overline{\quad}$ de tijdmiddelingsoperator.

De covariantiematrix $\overline{p_i p_j} = \begin{pmatrix} 1+\rho & \mu \\ \mu & \mu^2+\rho \end{pmatrix} P_1^2$ heeft als eigenwaarden:

$1 + \mu^2 + \rho$ en ρ , en als eigenvectoren:

$$E_1 = \frac{1}{\sqrt{1+\mu^2}} (1, \mu) \quad \text{resp.} \quad E_2 = \frac{1}{\sqrt{1+\mu^2}} (-\mu, 1).$$

E_1 ligt dan langs de in de figuur getekende "lijnen van lineaire regressie", en $\alpha_1(t)$, de afstand van de projectie van $(p_1(t), p_2(t))$

op die lijn tot de oorsprong wordt gegeven door:

$$a_1(t) = (1+\mu^2)^{-\frac{1}{2}} (p_1(t) + \mu p_2(t)) = (1+\mu^2)^{\frac{1}{2}} P_1(t) + (1+\mu^2)^{-\frac{1}{2}} (\rho_1 + \mu \rho_2)$$

Verder is $a_2(t) = (1+\mu^2)^{-\frac{1}{2}} (-\mu p_1(t) + p_2(t)) = (1+\mu^2)^{-\frac{1}{2}} (-\mu \rho_1 + \rho_2)$

en a_2 bevat dus inderdaad alleen de meetfouten.

Met behulp van $a_1(t)$ kan men nieuwe waarden voor $p_1(t)$ en $p_2(t)$ construeren:

$$p_1^{\text{reconstr}}(t) = a_1 E_{1,1} = P_1(t) + \frac{\rho_1 + \mu \rho_2}{1+\mu^2}$$

$$p_2^{\text{reconstr}}(t) = a_1 E_{1,2} = \mu p_1^{\text{reconstr}}(t)$$

Onvermijdelijk is uiteraard dat deze nog een term, die de meetfout bevat, hebben, maar bewezen kan worden dat de som van de varianties van deze twee termen: $(1+\mu^2)^{-1} \frac{(\rho_1 + \mu \rho_2)^2}{\rho^2 P_1^2}$ minimaal is, of nauwkeuriger geformuleerd:

$p_1^{\text{reconstr}}(t)$ en $p_2^{\text{reconstr}}(t)$ zijn dusdanig genormaliseerde lineaire combinaties van de oorspronkelijke waarnemingen $p_1(t)$ en $p_2(t)$, dat

i) het verwachte verband $p_2(t) = \mu p_1(t)$ wordt ingebouwd:

$$p_2^{\text{reconstr}}(t) = \mu p_1^{\text{reconstr}}(t)$$

ii) de variantie van de afstand van $(p_1(t), p_2(t))$ tot die ideale lijn minimaal is.

Men breidt dit alles gemakkelijk uit tot n dimensies, indien men beschikt over n meetreeksen. Men gaat in het algemeen dan niet uit van een 100% correlatie tussen twee waarnemingen, maar tussen een k -tal, $k < n$. Weer dient de covariantiematrix gediagonaliseerd te worden, en met de eigenvectoren bij de eigenwaarden

$\lambda_1 \dots \lambda_k$ ($\lambda_1 > \lambda_2 > \dots > \lambda_k > \dots > \lambda_n$) stelt men k scores $\alpha_1 \dots \alpha_k$ samen, waarmee weer nieuwe waarden voor de waarnemingen worden gereconstrueerd. Bij gegeven k is dan de som der varianties van de bijdragen van de meetfouten in deze nieuwe waarden weer de minimaal haalbare.

In dit verslag worden, zoals in de literatuur gebruikelijk is, de eigenvectoren van de covariantiematrix de hoofdvelden genoemd (merk op dat zij dimensieloos zijn); echter, in afwijking van het in de statistiek gebruikelijke, worden de eerste k hoofdvelden hier "significant optredend" genoemd.

3. Hoofdc componenten - aanvullingen en toepassing

3.a Notatie

Er zal gebruik gemaakt worden van de volgende notaties:

- $p_i(t)$ = luchtdruk van station i op tijdstip t
 i = stationsnummer $i = 1 \dots n$, n = aantal stations
 t = tijd, $t = 1 \dots T$, T = aantal metingen per station
 $\langle \rangle$ = plaatsmiddelingoperator
 $\bar{}$ = tijdsmiddelingoperator
 $\pi_i(t)$ = $p_i(t) - \langle p \rangle(t)$
 E_{ij} = j^e component van i^e hoofdveld
oftewel: waarde van i^e hoofdveld in station j
 $\alpha_i(t)$ = score van i^e hoofdveld op tijdstip t
 $\alpha_i(t) = E_{ij}(\pi_j - \bar{\pi}_j)$
 \vec{r} = (x, y) plaatsvector
 k = aantal "significante" hoofdvelden

Tenzij anders vermeld, zal worden gesommeerd over een index, die in twee factoren van één produkt voorkomt; Griekse indices duiden hierbij op een sommatie over de eerste k , en Latijnse op één over de eerste n natuurlijke getallen.

3.b Berekening van de hoofdvelden binnen Nederland

3.b.1 Gebruikelijk is de hoofdvelden te berekenen als de eigenvectoren van de covariantiematrix Σ van de luchtdrukken:

$$\Sigma_{ij} = \overline{(p_i(t) - \bar{p}_i)(p_j(t) - \bar{p}_j)}$$

Omdat Nederland kleiner is dan de schaal van de overheersende luchtdrukssystemen zullen de waarnemingen op alle stations in het algemeen tegelijk omhoog of omlaag gaan. Daarom werd besloten van de luchtdrukwaarden eerst het landelijk gemiddelde af te trekken, zodat de hoofdvelden de eigenvectoren werden van de matrix

$$\Sigma_{ij} = \overline{(\pi_i(t) - \bar{\pi}_i)(\pi_j(t) - \bar{\pi}_j)}$$

Als k het aantal "significante" hoofdvelden is, wordt $\pi_i(t) - \bar{\pi}_i$ benaderd door:

$$\pi_i'(t) = \alpha_{\lambda_i}(t) E_{\lambda_i} = E_{\lambda_i} E_{\lambda_j} (\pi_j(t) - \bar{\pi}_j), \text{ oftewel:}$$

$$p_i(t) \text{ wordt benaderd door } p_i'(t) = \pi_j'(t) + (\bar{p}_i - \langle \bar{p} \rangle) + \langle p \rangle(t)$$

k kan nu worden bepaald door te eisen dat $p_i(t) - p_i'(t)$ een variantie heeft, die kleiner is dan het kwadraat van de standaardmeetfout van de stations.

3.b.2 Ter beschikking stonden de uurlijkse waarnemingen over de periode 01-01-71 tot en met 31-12-74 van de stations

008, 135, 147, 200, 202, 210, 220, 225, 229, 230/235, 240, 245, 250, 260, 265, 270, 272, 275, 280, 290, 300, 308, 310, 320, 325, 330, 344, 350, 370, 375, 380, 601, 603, 604 en 605.

230 en 235 werden als hetzelfde station beschouwd.

De 20 onderstreepte stations hadden gedurende deze periode 1459 simultane luchtdrukwaarnemingen om 0000 gmt en werden daarom gebruikt voor de berekening van de hoofdcomponenten.

Bovendien werden nog hoofdcomponenten bepaald uit waarnemingen om 1200 gmt van dezelfde stations, aangevuld met 325 (1454 simultane waarnemingen). De verschillen tussen deze twee stellen hoofdvelden waren zo klein, dat werd besloten alleen de 0000 gmt analyses verder te gebruiken.

3.b.3 Bij de benadering van het luchtdrukveld gebruikt de hoofcomponentenanalyse alle n stations. Er werd daarom een procedure ontwikkeld om ontbrekende waarnemingen in te vullen.

De procedure is gebaseerd op de volgende gedachtengang: Na invulling van de ontbrekende waarden ontstaat een complete verzameling gegevens, waarop de hoofdcomponentenanalyse kan worden toegepast. Deze geeft dan weer een waarde in de ontbrekende stations, en er ontstaat een stelsel lineaire vergelijkingen in de ontbrekende waarden als men die nieuw berekende waarden gelijk stelt aan de ingevulde.

In formule wordt dit:

Veronderstel dat de waarden van stations $i=1\dots r$ ontbreken.

We eisen $\pi_i(t) - \bar{\pi}_i = \pi_i'(t)$ ($i=1\dots r$), waarin $\pi_i(t)$ de ingevulde, dus onbekende waarde is, en $\pi_i'(t)$ m.b.v. de onbekende $\pi_i(t)$ ($i=\dots r$) en de gemeten $\pi_i(t)$ ($i=r+1\dots n$) berekend wordt. $\bar{\pi}_i$ is bekend uit de andere waarnemingen op station i .

We krijgen achtereenvolgens:

$$\pi_i(t) - \bar{\pi}_i = E_{\lambda i} E_{\lambda j} (\pi_j(t) - \bar{\pi}_j) \quad i=1\dots r$$

met $\pi_i(t) = p_i(t) - \left(\frac{1}{n} \sum_{j=1}^r p_j(t) + \frac{1}{n} \sum_{j=r+1}^n p_j(t)\right)$ en $\bar{\pi}_i = \bar{p}_2 - \langle p \rangle$

zodat

$$\sum_{j=1}^r A_{ij} p_j = - \sum_{j=r+1}^n A_{ij} p_j + \sum_{j=1}^n A_{ij} \bar{\pi}_j \quad (i=1\dots r) \quad (\text{III.3.b.3;1})$$

waarin $A_{ij} = T_{ij} - \frac{1}{n} \sum_{k=1}^n T_{ik}$ met $T_{ij} = \delta_{ij} - E_{\lambda i} E_{\lambda j}$.

Vergelijkingen III.3.b.3;1 vormen het stelsel van r lineaire vergelijkingen waaruit p_i ($i=1\dots r$) opgelost kunnen worden.

3.b.4 Bovenstaande procedure werkte bevredigend, behalve bij weglating van station 300.

Wanneer men de waarnemingen van zeven stations wegliet, konden zij met een standaardafwijking van 0.3 mbar gereproduceerd worden. (Echter: dit werd steekproefsgewijze voor slechts enkele combinaties van die zeven stations uitgevoerd). Station 300 daarentegen bleek slechts tot op 10 mbar benaderbaar te zijn.

Besloten werd dan ook dit station niet te gebruiken in de hoofdcomponentenanalyse, zodat als gebruikte stations blijven de stations die zijn opgenomen in tabel 1, waarin ook de coördinaten en de waarden van $\bar{\pi}_i$ zijn opgenomen. (De x-as is de west-oostlijn door Cabauw, en de y-as de zuid-noordlijn door Cabauw; de eenheid is 20 km). Verder is dus $\bar{\pi}_i$ de mate waarin het tijdsgemiddelde van het station i afwijkt van het landelijk gemiddelde, dit laatste ook gemiddeld over de tijd; dit landelijk tijdsgemiddelde over de periode 1971 t/m 1974, om 0000 gmt, bedraagt 1015.774 mbar.

Een analyse van het gemiddelde luchtdrukveld ($\bar{\pi}_i$) treft men aan in figuur 1.

station	iii	i	x (20 km)	y (20 km)	$\bar{\pi}_i$ (mbar)
Valkenburg	210	1	-1.8	1.1	-0.170
Texel	220	2	-2.1	5.9	-0.367
IJmuiden	225	3	-1.1	2.8	-0.006
Den Helder/De Kooij	230/235	4	-0.7	5.2	-0.360
Schiphol	240	5	-0.4	1.8	-0.105
Terschelling	250	6	1.0	7.8	-0.422
De Bilt	260	7	0.9	0.7	0.145
Soesterberg	265	8	1.3	1.0	0.041
Leeuwarden	270	9	2.7	7.0	-0.534
Deelen	275	10	3.4	0.5	0.146
Eelde	280	11	5.6	6.5	-0.468
Twente	290	12	6.9	1.8	-0.093
Vlissingen	310	13	-4.6	-2.7	0.282
Hoek van Holland	330	14	-2.8	0.1	0.311
Zestienhoven	344	15	-1.8	-0.1	0.025
Gilze-Rijen	350	16	0.0	-2.1	0.311
Eindhoven	370	17	1.6	-3.0	0.308
Volkel	375	18	2.8	-1.7	0.431
Beek	380	19	3.1	-5.8	0.525

Tabel 1. Lijst van gebruikte stations, hun coördinaten en hun gemiddelde luchtdruk t.o.v. het landelijk gemiddelde (1015.774 mbar).

Table 1. List of used stations, their coordinates and the deviation of their mean pressure from the area-mean (1015.774 mbar).

3.b.5 Uit de waarnemingen van de stations uit tabel 1 zijn de hoofdvelden berekend; een analyse van de eerste zeven hoofdvelden vindt men in de figuren 2 t/m 8.

3.b.6 Bepaling van het aantal significante hoofdvelden kan geschieden aan de hand van de volgende criteria:

- a) de significante hoofdvelden zijn niet al te grillig;
 - b) de door de overblijvende hoofdvelden gedekte variantie van het luchtdrukveld is vergelijkbaar met die van de meetfout.
- In tabel 2 staat de door elk hoofdveld "verklaarde" variantie. Deze variantie is gelijk aan de bijbehorende eigenwaarde van de covariantiematrix ($[2]$).

rangnr. hoofdveld	verklaarde variantie (mbar ²)
1	16.85
2	5.03
3	.30
4	.28
5	.18
6	.15
7 t/m 18 samen	.61
19	0
totaal	23.40

Tabel 2. De door de hoofdvelden "verklaarde" variantie.

Table 2. Variance described by the principal components.

De stations hebben een standaardmeetfout van ongeveer 0.2 mbar. Dit getal werd op de volgende wijze geschat:

- I Een systematische fout. Hoewel de systematische fouten uit de waarnemingen worden verwijderd door het aftrekken van het tijdsgemiddelde, zijn er aanwijzingen dat er toch nog een vorm van systematische fouten overblijft, die b.v. alleen optreden bij bepaalde windrichtingen. Een grootte-orde schatting van deze fouten verkrijgt men uit de analyse van het gemiddelde veld: Uit figuur 1 blijkt dat het grillige karakter ervan verdwijnt als men de waarde van de stations .10 à .15 mbar mag verschuiven.

II Een toevallige fout: Deze bestaat uit de afleeson nauwkeurigheid (ook weer .10 à .15 mbar). (Afleesfouten, die b.v. 1 of 10 mbar kunnen bedragen, of fouten in de toepassing van hoogte- en temperatuurcorrecties, zijn in het algemeen uit het bestand verwijderd door de klimatologische dienst van het KNMI. Deze vergroten normaliter echter de meetfout nog).

Het aantal mee te nemen hoofdvelden dient daarom zo groot te zijn, dat de niet-gedekte variantie ongeveer $19 \times (.2)^2 \text{ mbar}^2$ bedraagt. Dat aantal blijkt 5 te zijn.

Bovendien blijkt uit de analyses (fig. 2 t/m 8) dat dan ook aan criterium a) wordt voldaan.

3.b.7 Met behulp van de hoofdvelden op de 19 stations uit tabel 1 werd weer als in sectie 3 een procedure ter aanvulling van eventueel ontbrekende waarnemingen geschreven. (Deze procedure is in algol beschikbaar).

Nu werden bij weglating van twee waarnemingen die twee waarnemingen gereproduceerd met een standaardafwijking van 0.5 mbar; bij vier weggelaten waarnemingen bedroeg deze 0.7 mbar. Dat deze getallen groter zijn dan de vergelijkbare uit paragraaf 4 moet aan de zeer beperkte omvang van de steekproeven worden geweten.

Opgemerkt dient te worden, dat in alle gevallen slechts één station echt slecht werd benaderd; dit was niet steeds hetzelfde station en hierin werd dan ook geen aanleiding gevonden nog een station te verwerpen.

3.b.8 Interessant is nog de invloed van een meetfout op een station op de berekende luchtdrukwaarden op dat en andere stations.

Veronderstel dat de waarneming op station j vermeerderd wordt met een meetfout ε_j . Uit de formules in sectie 1 ziet men dat hierdoor de benadering van de luchtdruk, p'_i , verschuift met

$$\begin{aligned} \Delta p'_i &= E_{\lambda_i} E_{\lambda_j} \left(\varepsilon_j - \frac{1}{n} \sum_{k=1}^n \varepsilon_k \right) + \frac{1}{n} \sum_{k=1}^n \varepsilon_k \\ &= (\delta_{ij} - A_{ij}) \varepsilon_j, \text{ met } A_{ij} \text{ gedefinieerd als in sectie 3.} \end{aligned}$$

De matrix $\delta_{ij} - A_{ij}$ is in tabel 3 opgenomen.

station waarop afwijking invloed heeft/
influenced station

	station waar afwijking optreedt/ → influencing station																		
	210	220	225	230/5	240	250	260	265	270	275	280	290	310	330	344	350	370	375	380
210	11	3	10	7	10	2	8	8	2	3	-3	-1	11	14	11	7	2	2	-7
220	3	86	14	20	0	-7	-2	0	-6	-6	9	-3	-5	1	-1	-2	-9	-6	14
225	10	14	13	12	12	1	10	11	1	6	-4	1	-2	10	8	5	-1	3	-10
230/5	7	20	12	17	10	21	6	6	9	1	-2	-5	-6	6	4	1	-2	1	-5
240	10	0	12	10	13	6	12	12	3	7	-7	1	-2	11	9	6	3	6	-10
250	2	-7	1	21	6	71	-2	-3	34	-6	-3	-14	-5	-3	-2	-5	5	1	8
260	8	-2	10	6	12	-2	13	14	0	12	-4	8	-7	7	7	8	6	10	-6
265	8	0	11	6	12	-3	14	15	-1	13	-4	10	-11	6	5	8	6	12	-5
270	2	-6	1	9	3	34	0	-1	28	2	25	9	4	0	1	-3	0	-1	-4
275	3	-6	6	1	7	-6	12	13	2	16	8	19	-11	0	2	7	9	14	2
280	-3	9	-4	-2	-7	-3	-4	-4	25	8	73	34	13	-3	-1	-5	-6	-6	-5
290	-1	-3	-1	-5	1	-14	8	10	9	19	34	34	-9	-5	-1	4	6	11	2
310	11	-5	-2	-6	-2	-5	-7	-11	4	-11	13	-9	69	25	20	10	5	-10	11
330	14	1	10	6	11	-3	7	6	0	0	-3	-5	-25	20	15	8	1	-1	-10
344	11	-1	8	4	9	-2	7	5	1	2	-1	-1	20	15	13	8	4	2	-4
350	7	-2	5	1	6	-5	8	8	-3	7	-5	4	10	8	8	11	11	10	12
370	2	-9	-1	-2	3	5	6	6	0	9	-6	6	5	1	4	11	18	15	27
375	2	-6	3	1	6	1	10	12	-1	14	-6	11	-10	-1	2	10	15	18	19
380	-7	14	-10	-5	-10	8	-6	-5	-4	2	-5	2	11	-10	-4	12	27	19	68

Tabel 3. De matrix $\delta_{ij}-A_{ij}$ (i = rij-, j = kolomnummer).

De invloed van een waarnemingsfout ϵ_j (kolomvector), op station j gemaakt, verandert de door de hoofdcomponentenanalyse berekende luchtdruk op station i met een bedrag $(\delta_{ij}-A_{ij}) \epsilon_j$. Getallen in procenten.

Table 3. The matrix $\delta_{ij}-A_{ij}$ (i = number of row; j = number of column).

A change ϵ_j in the measurement at the "influencing" station j changes the by means of the principal component analysis calculated pressure with an amount of $(\delta_{ij}-A_{ij}) \epsilon_j$ at the "influenced" station i. Figures in per cents.

Hoewel de matrix symmetrisch lijkt, is hij dat niet; de vier getallen links boven bijvoorbeeld luiden to in de vierde decimaal:

.1081	.0317
.0316	.8620

Er zijn enkele opvallend grote getallen; met name bij station 220, en de "clusters" 250-270-280-290; 310-330 en 370-375-380.

Het grote diagonaalelement bij 220 betekent dat een verandering van de luchtdrukmeetwaarde op dit station vrijwel alleen dit station beïnvloedt. Men kan hieruit de conclusie trekken, dat het misschien zinvol is 220 te verwerpen bij de analyses; in dit onderzoek werd dit nog niet uitgevoerd.

3.b.9 Ten slotte werd ter verificatie van de gebruikte formules de standaardafwijking van het benaderde veld ten opzichte van het gemeten veld berekend; deze afwijking is gelijk aan

$$\left[\frac{1}{T} \sum_{t=1}^T \frac{1}{n} \sum_{i=1}^n [p_i(t) - p'_i(t)]^2 \right]^{\frac{1}{2}}$$

Met medeneming van de eerste vijf hoofdvelden zou bovenstaande uitdrukking conform sectie 6 voor de 1459 oorspronkelijke waarnemingen 0.20 mbar moeten bedragen.

Na invulling van enkele ontbrekende waarnemingen conform sectie 7 werd het aantal meetdagen 1461; de uitdrukking nam toen wederom een waarde aan van 0.20 mbar om 0000 gmt, terwijl de toepassing op de 1200 gmt metingen een resultaat van 0.31 mbar gaf.

Aangezien de 1200 gmt metingen min of meer als onafhankelijk van de 0000 gmt metingen mogen worden beschouwd, zijn beide standaardafwijkingen voldoende laag om vertrouwen in de toegepaste techniek te rechtvaardigen, hoewel het getal van 0.31 mbar zo ver boven de standaardmeetfout op de stations uitgaat, dat het niet geheel zinloos lijkt ook nog eens het gedrag van een stelsel hoofdcomponenten berekend op de 1200 gmt metingen te onderzoeken. Dit werd echter in dit onderzoek niet gedaan.

4. Interpolatie van de hoofdvelden

4.a Doel

De hoofdvelden worden door statistische technieken op objectieve wijze verkregen uit de meetreeksen van de stations maar zij worden slechts berekend in die stations.

Om nu een geowind of een luchtdrukgradiënt te kunnen bepalen, is het noodzakelijk de hoofdvelden te interpoleren.

Met de notatie uit paragraaf III.2 heeft men:

$p'_i(t) = E_{\lambda_i} E_{\lambda_j} (\pi_j(t) - \bar{\pi}_j) + \bar{p}_i - \langle \bar{p} \rangle + \langle p \rangle (t)$; hierin moet nu de "plaats-index" i worden vervangen door de plaatsvector \vec{r} :

$$p'(\vec{r}, t) = E_{\lambda}(\vec{r}) E_{\lambda_j} (\pi_j(t) - \bar{\pi}_j) + \bar{p}(\vec{r}) - \langle \bar{p} \rangle + \langle p \rangle (t) \quad (\text{III.4.a;1})$$

Men ziet dat zowel de gebruikte hoofdvelden als het gemiddelde veld, \bar{p}_i , moeten worden geïnterpoleerd.

Merk nog op dat de tijdsafhankelijke delen van het rechterlid van III.4.a;1 geen index i dragen ($\langle p \rangle (t) = \frac{1}{n} \sum_{j=1}^n p_j(t)$) en dus dat het rechterlid, als de interpolatie van E_{λ_i} en \bar{p}_i eenmaal is uitgevoerd, alleen wordt bepaald door de metingen óp de stations: het gemeten luchtdrukveld behoeft niet te worden geïnterpoleerd.

4.b Methode

4.b.1 Uit de figuren 2 en 3 blijkt dat de eerste twee hoofdvelden goed benaderbaar zijn door een lineaire functie, d.w.z.:

$$E_i(\vec{r}) = e_0^i + e_1^i x + e_2^i y \quad (i = 1, 2)$$

De derde, vierde en vijfde zijn nog steeds redelijk glad van structuur; omdat zij zeker niet lineair zijn, werd hiervoor een tweede-graads aanpassing gekozen:

$$E_i(\vec{r}) = e_0^i + e_1^i x + e_2^i y + e_3^i x^2 + e_4^i xy + e_5^i y^2 = e_k^i r^k \quad (i = 3, 4, 5) \quad (\text{III.4.b;1})$$

waarin dus de notatie:

$$r^0 = 1, r^1 = x, \dots, r^5 = y^2 \quad \text{wordt gebruikt.}$$

Deze formule geldt voor $i = 3, 4$ en 5 , maar om vergelijkbare nauwkeurigheid in de eerste twee hoofdcomponenten te bereiken, werden deze ook kwadratisch benaderd.

Ten slotte werd nog de index $i = 0$ in gebruik genomen voor het gemiddelde veld: $\bar{\pi}_i$, geïnterpoleerd tot $\bar{\pi}(\vec{r}) = E_0(\vec{r}) = e_k^0 r^k$.

Men vindt in de lineaire regressie tussen E_i ($i = 0 \dots 5$) en de r^k ($k = 0 \dots 5$) de in tabel 4 vermelde coëfficiënten e_k^i .

In deze tabel staat ook de onnauwkeurigheid in de coëfficiënten vermeld, dit als maat voor de significantie; als zodanig kan ook nog gebruikt worden de zogenaamde multipele correlatiecoëfficiënt, die ook nog is vermeld.

e_k^i		$i = 0$	1	2	3	4	5	(dimensie)
k	$r^k =$							
0	1	130 \pm 40	-92 \pm 4	48 \pm 6	-130 \pm 40	-160 \pm 50	100 \pm 60	1
1	x	5 \pm 12	-0.5 \pm 1.0	-79.3 \pm 1.7	-32 \pm 10	0 \pm 14	47 \pm 16	(20 km) ⁻¹
2	y	-78 \pm 11	62.6 \pm 0.9	10.9 \pm 1.6	-58 \pm 9	6 \pm 12	5 \pm 15	(20 km) ⁻¹
3	x ²	-2 \pm 3	0.0 \pm 0.3	0.6 \pm 0.5	0 \pm 3	9 \pm 4	-15 \pm 4	(20 km) ⁻²
4	xy	-1 \pm 3	-0.4 \pm 0.3	-2.2 \pm 0.5	10 \pm 3	-21 \pm 4	-5 \pm 5	(20 km) ⁻²
5	y ²	-1.0 \pm 1.8	0.33 \pm 0.16	-0.3 \pm 0.16	13.8 \pm 1.6	7 \pm 2	0 \pm 2	(20 km) ⁻²
(dimensie)		10 ⁻³ mbar	10 ⁻³	10 ⁻³	10 ⁻³	10 ⁻³	10 ⁻³	(20 km) ⁻²
mult.corr. coeff.		0.952	0.999	0.998	0.932	0.875	0.822	

Tabel 4. De coëfficiënten e_k^i . De getallen in het schema dienen te worden vermenigvuldigd met de dimensies, die staan vermeld in de laatste kolom en de voorlaatste rij van de tabel. De laatste rij bevat de multipele correlatiecoëfficiënten tussen E_{ij} en $e_k^i r_j^k$ (waar \vec{r}_j de plaatsvector is van het k^e station).

Table 4. The coefficients e_k^i as they appear in eq. III.4.b;1. The figures in the scheme have to be multiplied by the dimensions that are indicated in the last column and the last-but-one row in the table. The last row contains the multiple correlation coefficients between E_{ij} and $e_k^i r_j^k$ (where \vec{r}_j is the position vector of the station k).

4.b.2 De geïnterpoleerde velden zijn geanalyseerd in de figuren 9 t/m 14. Een nadeel van de gebruikte interpolatiemethode is dat het interpolatieschema de hoofdvelden, die zijn bepaald om er exact mee te kunnen rekenen, nogmaals afvlakkend bewerkt. Deze afvlakking heeft voor het gemiddelde veld nog wel betekenis, en het is zelfs mogelijk een schatting te maken van de systematische fout op de stations aan de hand van die afvlakking, maar voor de hoofdcomponenten $E_i(\vec{r})$ ($i = 1 \dots 5$) verdient het aanbeveling een interpolatieprocedure te zoeken, die de "meet"punten E_{ij} exact reproduceert. In afwachting van de vondst van zo'n schema werd de kwadratische Taylor-reeks voorlopig aangehouden.

Men kan zich dan afvragen of men niet evengoed of beter het luchtdrukveld instantaan kwadratisch kan interpoleren, dus zonder tussenkomst van de hoofdcomponentenanalyse. Het bezwaar echter is dat men dan niet precies weet wat men doet (er is geen a priori reden om aan te nemen dat het luchtdrukveld kwadratisch is), terwijl een hoofdcomponentenanalyse niet-systematische meetfouten kan detecteren, aan de hand van een lange reeks waarnemingen per station gewogen. Toevallig blijkt dan dat de eerste vijf hoofdcomponenten een min of meer kwadratische vorm hebben.

4.b.3 De standaardafwijking van de benadering door de met behulp van de Taylor-polynomen benaderde hoofdvelden ten opzichte van de werkelijk gemeten luchtdruk werd bepaald. Er werd gevonden ($p_i^a(t)$ = polynoombenadering voor $p_i(t)$):

$$\left[\frac{1}{T} \sum_{t=1}^T \frac{1}{n} \sum_{i=1}^n (p_i(t) - p_i^a(t))^2 \right]^{\frac{1}{2}} = 0.24 \text{ mbar, op het tot } T = 1461$$

aangevulde bestand om 0000 gmt, terwijl die standaardafwijking om 1200 gmt 0.34 mbar bedroeg.

Deze getallen liggen dusdanig dicht bij de overeenkomstige uit sectie III.3.b.9, dat geconcludeerd mag worden dat de enkelvoudige kwadratische polynoombenadering acceptabel is.

4.b.4 Het is mogelijk de restvariantie van de schattingsfout voor het luchtdrukveld te verminderen door de polynoombenadering voor de hoofdcomponenten als nieuwe, niet-orthonormale basis te kiezen. Men berekent dan de scores niet uitgaande van de hoofdcomponenten, maar uitgaande van de inverse van de matrix, die als i - j component heeft $E_{ij}^P = E_i^P(\bar{x}_j)$, de waarde van de polynoombenadering van het i^e hoofdveld in waarnemingsstation j , en wel als $\alpha_i^P = (E^P)^{-1}_{ij} p_j$, waarin α_i^P de i^e score is.

De verbetering in de schatting is echter niet significant, waar de getallen uit de vorige sectie al nauwelijks afwijken van die uit sectie III.3.b.9.

IV Verificatie van de methode

1. Inleiding

Om enige indruk te verkrijgen van de nauwkeurigheid en de betrouwbaarheid van de gebruikte hoofdcomponentenanalyse en de polynoombenadering staan enkele wegen open:

- a) de puur statistische: van de hoofdvelden kan de reproduceerbaarheid worden nagegaan (J.P. de Jongh en S. Kruizinga [3]), evenals van de polynoombenadering.
- b) men kan trachten de hoofdcomponentenanalyse te vergelijken met een handanalyse van het luchtdrukveld, of met andere analyses.
- c) men kan een vergelijking maken tussen geowind en gemeten wind.

2. Statistische beschouwing

Omdat voor de polynoombenadering reeds een indruk van de significantie d.m.v. multipele correlatiecoëfficiënten is gegeven in paragraaf III.4.b.1, zal hier nog uitsluitend iets gezegd worden over de hoofdvelden zelf.

De relevante formules zijn [3]:

$$\text{Variantie van } \lambda_k = \frac{2}{T} \lambda_k^2 \quad (\text{geen sommatie over } k)$$

$$\text{Variantie van } w_{ik} = \frac{w_{ik}^2}{2T} + \frac{\lambda_i^2}{T} \sum_{\substack{j=1 \\ j \neq i}}^n \frac{w_{jk}^2}{(\lambda_i - \lambda_j)^2}$$

waarin $\lambda_k = k^e$ eigenwaarde van de covariantiematrix,
en $w_{ik} = \sqrt{\lambda_i} E_{ik}$ (geen sommatie over i).

T is in principe het aantal onafhankelijke waarnemingen. Aangezien hier 1459 dagelijkse waarnemingen zijn gebruikt, terwijl het luchtdrukveld een "correlatietijd" heeft van ongeveer twee dagen, wordt uitgegaan van $T = 750$; diit leidt tot de conclusie dat de eigenwaarden tot op 5% zijn gegeven. De nauwkeurigheid van de i^e eigenvector wordt bepaald door de noemer $(\lambda_i - \lambda_j)^2$. Voor $i = 1$ en

$i = 2$ blijft de schade beperkt, en er wordt een onzekerheid van ongeveer 0.03 gevonden in de componenten van de eerste twee eigenvectoren. Voor $i = 3$ en $i = 4$ treedt "contaminatie" op, omdat λ_3 en λ_4 dicht bij elkaar liggen; men moet ook eigenlijk de 3e en de 4e eigenvector beschouwen als een basis voor de tweedimensionale deelruimte bij $\lambda_3 = \lambda_4$, zonder dus een voorkeur aan de hand van de eigenwaarde voor één van beide aan de dag te leggen. Hetzelfde geldt voor de eigenwaarden λ_5 en λ_6 , en vanuit dit standpunt gezien is het incorrect het aantal significante hoofdvelden gelijk aan 5 te kiezen. Elke keuze van $k > 4$ vertoont echter dit euvel.

Samenvattend kan worden gezegd, dat de eerste vier hoofdvelden en de polynoombenaderingen voldoende relevant zijn uit statistische overwegingen.

3. Vergelijking geowind - gemeten wind

3.a Vergelijking geowind - 200m (Cabauw)-wind

Voor een aantal ad hoc gevallen werd de geowind vergeleken met de 200m-wind te Cabauw. Op één geval na bleek steeds overdag de geowind ongeveer 10% sterker te zijn dan de 200m-wind, en iets geruimd (10° à 20°). Een voorbeeld vindt men in fig. 16.

Aangezien één en ander is zoals verwacht moet worden, bevestigen deze resultaten de betrouwbaarheid van de berekende geowind; men kreeg de indruk dat de geowind tot op ongeveer 0.5 m/s en 5° nauwkeurig geacht mag worden. Men zij echter gewaarschuwd dat deze getallen, bij gebrek aan een werkelijk gemeten geowind, maar schattingen zijn!

3.b Vergelijking geowind - 10 meter wind

Voor de 1461 1200 gmt gevallen in de jaren 1971 tot en met 1974 werd de grondwind, zoals die gemeten was in De Bilt (260) en Zestienhoven (344) vergeleken met de berekende geowind.

Hierbij werd de volgende procedure toegepast.

- 1) Situaties met zwak veranderlijke wind ($dd = 99$) werden verworpen.
- 2) De wind te De Bilt werd gecorrigeerd voor de ruwheid van het terrein door vermenigvuldiging met de beschuttingsfactoren van Wieringa [4], die luiden:

bij windrichting tussen en is de beschuttingsfactor:

345°	15°	1.17
15°	45°	1.18
45°	75°	1.19
75°	105°	1.19
105°	135°	1.15
135°	165°	1.20
165°	195°	1.21
195°	225°	1.23
225°	255°	1.33
255°	285°	1.39
285°	315°	1.38
315°	345°	1.27

- 3) Vervolgens werd een lineaire regressie uitgevoerd tussen de west-oost respectievelijk de zuid-noord component van de wind enerzijds en de score van de hoofdvelden anderzijds (de score is de sterkte waarmee het hoofdveld optreedt; in $\pi'_i(t) = E_{\lambda i} E_{\lambda j} (\pi_j(t) - \bar{\pi}_j)$ is dus $\alpha_\lambda = E_{\lambda j} (\pi_j(t) - \bar{\pi}_j)$ de score van hoofdveld λ).

Er geldt, dankzij de lineaire regressie:

$\vec{u} = \vec{u}_\lambda \alpha_\lambda + \vec{u}_0$, waarin \vec{u} de 10-meter-windvector is, en \vec{u}_λ de regressiecoëfficiënten zijn. We schrijven $\alpha_0 = 1$, dus $\vec{u} = \vec{u}_\lambda \alpha_\lambda$ ($\lambda = 0 \dots 5$).

Verder geldt:

$$p'(\vec{r}, t) = \alpha_\lambda E_\lambda(\vec{r}) + \langle p \rangle (t) \quad (\lambda = 0 \dots 5),$$

waarin weer $\alpha_0 = 1$ en $E_0(\vec{r}) = \pi(\vec{r})$; bedenken dat de hoofdvelden geïnterpoleerd zijn: $E_\lambda(\vec{r}) = e_i^\lambda r^i$ (λ en $i = 0 \dots 5$), vindt men:

$$p'(\vec{r}, t) = \alpha_\lambda e_i^\lambda r^i + \langle p \rangle (t), \quad \text{en}$$

$$\frac{\partial p'(\vec{r}, t)}{\partial \vec{r}} = \alpha_\lambda e_i^\lambda \frac{\partial r^i}{\partial \vec{r}}, \quad \text{waaruit voor de gewind volgt:}$$

$$\vec{u}_g = \frac{1}{f\rho} \vec{\varepsilon} \alpha_\lambda \bullet_i^\lambda \frac{\partial r^i}{\partial \vec{r}}, \quad \text{waarin } \vec{\varepsilon} \text{ de tensor } \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \text{ voorstelt.}$$

Daarnaast kan men het tensorverband:

$$\vec{u} = \vec{D} \vec{u}_g \quad \text{opschrijven.}$$

\vec{D} zal afhangen van vele factoren, waarvan terreinomstandigheden en stabiliteit vermoedelijk het belangrijkste zijn. Terreinomstandigheden worden ingecalculeerd door de correctiefactoren van Wieringa, en de invloed van de stabiliteit wordt hopelijk gereduceerd door de 1200 gmt situatie te beschouwen. Onder deze veronderstellingen wordt \vec{D} dus constant, en representeert zowel draaiing als remming door wrijving.

Er ontstaat

$$\vec{u}_\lambda \alpha_\lambda = \vec{u} = \vec{D} \vec{u}_g = \vec{D} \vec{E} \frac{1}{f\rho} \alpha_\lambda e^{\lambda} e^i \frac{\partial r^i}{\partial r}$$

De tijdsafhankelijkheid treedt alleen op in α_λ ; na uitdeling ontstaat:

$$\vec{u}_\lambda = \vec{D} \vec{E} \frac{1}{f\rho} e^{\lambda} e^i \frac{\partial r^i}{\partial r}$$

Hier staan voor $\lambda = 0 \dots 5$ steeds twee vergelijkingen voor de tensor \vec{D} . \vec{D} wordt uitgedrukt als produkt van remming en draaiing:

$$\vec{D} = \gamma \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$$

en in tabel 5 worden de berekende γ 's en φ 's gegeven voor $\lambda = 0 \dots 4$.

	De Bilt		Zestienhoven	
$\lambda =$	γ	φ	γ	φ
0	.55	+3°	.71	+18°
1	.41	+23°	.49	+21°
2	.43	+21°	.52	+25°
3	.05	+51°	.30	-14°
4	2.3	+50°	.31	-74°

Tabel 5. Wrijvingsfactoren en draaihoeken van de grondwind te De Bilt en Zestienhoven ten opzichte van de geowind voor de eerste vier hoofdvelden en het gemiddelde veld ($\lambda = 0$).

Table 5. Frictional factors and angles of rotation of the 10 m wind at De Bilt and at Zestienhoven with respect to the geostrophic wind for the first four principal components and the average field ($\lambda = 0$).

Men kan uit de gegevens van tabel 5 concluderen dat het veronderstelde verband zeker bestaat voor de eerste en tweede eigenvector; de draaiing en afremming ten gevolge van wrijving (resp. 20° en 55%) lijken redelijk. De derde en vierde eigenvector geven ook hier weer geen blijk van significantie. Dit heeft enkele verklaringen:

- a) het derde en het vierde hoofdveld zijn niet significant te scheiden;
- b) de gradiëntbenadering door de polynoomafgeleide $\frac{\partial}{\partial \vec{r}} e_i^\lambda r^i$ is niet goed. Dit laatste zou bijvoorbeeld te De Bilt inderdaad voor de derde eigenvector een te lage en voor de vierde een te hoge γ opleveren, zoals blijkt uit een vergelijking van fig. 4 met fig. 12 respectievelijk fig. 5 met fig. 13.

Dit laatste kan ook nog als verklaring dienen voor de afwijkingen bij $\lambda = 0$, dus bij het gemiddelde veld.

Een nevenresultaat van dit onderzoek is de mogelijkheid de grondwind uit de parameters van de geowind te schatten; er geldt:

$$\begin{aligned} u_{260} &= u_\lambda a_\lambda \pm 1.7 \text{ m/s} & u_{344} &= u_\lambda a_\lambda \pm 2 \text{ m/s} \\ v_{260} &= v_\lambda a_\lambda \pm 1.6 \text{ m/s} & v_{344} &= v_\lambda a_\lambda \pm 1.8 \text{ m/s} \end{aligned}$$

waarin u_{260} en v_{260} de met de beschuttingsfactoren vermengvuldigde west-oost respectievelijk zuid-noord component van de wind te De Bilt zijn, en u_{344} en v_{344} die te Zestienhoven (zonder beschuttingsfactoren).

De multipele correlatiecoëfficiënt is in alle vier de gevallen 0.90.

Men kan overigens algemeen schrijven:

$$\begin{aligned} u_{260} &= (1 \pm 4) \text{ m/s} & u_{344} &= (3 \pm 9) \text{ m/s} \\ v_{260} &= (1 \pm 7) \text{ m/s} & v_{344} &= (1 \pm 9) \text{ m/s} \end{aligned}$$

waaruit men ziet dat de windcomponenten door de benadering $u_\lambda a_\lambda$ ongeveer een factor 4 beter worden beschreven dan door hun gemiddelden.

Men kan deze componentsgewijze gegevens nog terugrekenen naar ware lengte en richting:

De lengte wordt dan in De Bilt gereproduceerd met een standaardafwijking van 1.2 m/s, en de richting met een standaardafwijking van 24° . De analoge getallen in Zestienhoven zijn respectievelijk 1.8 m/s en 24° .

Dit alles heeft uiteraard betrekking op de metingen in de jaren 1971-1974.

V. Conclusies en aanbevelingen

Uit het voorafgaande mag worden geconcludeerd dat de methode van de hoofdcomponentenanalyse voor het luchtdrukveld boven Nederland goed voldoet in statistische zin, maar ook in individuele gevallen. Het aantal mee te nemen hoofdvelden dient vier of vijf te zijn.

Met deze hoofdvelden en een eenvoudige interpolatieprocedure ervoor kan redelijk betrouwbaar een geostrofische wind aan de grond worden berekend, zoals door de ad hoc toetsing aan windgegevens van 200 m hoogte geconcludeerd mag worden.

Een methode om ontbrekende waarnemingen in te vullen, uitgaande van de overige waarnemingen, bleek succesvol.

De statistische relevantie van de berekende hoofdvelden is niet zo groot, dat een splitsing naar dag-nacht of land-zee effecten nu al gemotiveerd kan worden. Bij een uitgebreider gegevensbestand kan dit echter aanbeveling verdienen.

Verdere aanbevelingen voor onderzoek zijn, zoals reeds genoemd in de voorafgaande hoofdstukken

- a) een onderzoek naar het noodzakelijke aantal stations: nu zijn er 19 gebruikt en na middeling blijven er 18 onafhankelijke over; met deze 18 worden de vijf parameters (de scores) bepaald. Het is dus niet onmogelijk dat een aantal stations in alle omstandigheden overbodig is.
- b) in de analyse van de gemiddelde luchtdruk (fig. 1) kan men een motivatie vinden tot onderzoek naar de oorzaak van de systematische meetfout op vele stations.
- c) ten slotte is al gewezen op het mogelijke nut van een beter interpolatieschema voor de hoofdvelden.

Dankbetuiging

Hierbij wil ik de heer W.J.A. Kuipers danken voor enkele stimulerende aanwijzingen, en de heer S. Kruizinga voor het nauwkeurig doorlezen van het manuscript, dat mede tot stand kon komen dankzij het feit dat ik over een aantal computerprogramma's van zijn hand kon beschikken.

Referenties

- [1] "Beschrijving van programma's voor de EL-X8".
KNMI Wetenschappelijk Rapport 69-3, 1969.
- [2] D.N. Lawley en A.E. Maxwell,
"Factor analysis as a statistical method",
London; Butterworths, 1971.
- [3] J.P. de Jongh en S. Kruizinga,
"Empirische orthogonale eigenvelden van het 500 mbar vlak",
KNMI Wetenschappelijk Rapport 75-10, 1975.
- [4] J. Wieringa en P.J.M. van der Veer,
"Nederlandse windstations 1971-1974",
KNMI Verslag V-278, 1976.

Appendix

De interpolatiemethode van Gandin

De notatie in het volgende is conform die in hoofdstuk III van dit verslag.

Bij een aantal meetreeksen $p_i(t)$ kan men de correlatiecoëfficiënten $\overline{p_i p_j}$ uitzetten tegen de ruimtelijke afstand tussen de punten i en j (die plaatsvectoren \vec{r}_i en \vec{r}_j hebben).

Als blijkt dat $\overline{p_i p_j} = f(\vec{r}_i; \vec{r}_j)$ voor een zekere functie f , dan kan men voor de luchtdruk op een willekeurige plaats \vec{r} eisen:

$$\overline{p_i p(\vec{r})} = f(\vec{r}_i, \vec{r}) \quad (A1)$$

met dezelfde functie f , en voor alle i .

De interpolatiemethode van Gandin bepaalt nu de zogenaamde gewichtsfuncties $g_i(\vec{r})$, zodanig dat (A1) geldt als men schrijft: $p(\vec{r}, t) = g_i(\vec{r}) p_i(t)$.

We zullen bovenstaande formuleren in " π -taal" door de vervanging $p \rightarrow \pi - \bar{\pi}$.

In fig. 15 is een aantal van de $\frac{(\pi_i - \bar{\pi}_i)(\pi_j - \bar{\pi}_j)}{(\pi_i - \bar{\pi}_i)^2 (\pi_j - \bar{\pi}_j)^2}$ uitzet

als functie van $|\vec{r}_i - \vec{r}_j|$. Er is duidelijk weinig verband.

Men kan echter het volgende opschrijven:

$$\begin{aligned} \overline{(\pi_i - \bar{\pi}_i)(\pi_j - \bar{\pi}_j)} &= E_{ki} E_{kl} \overline{(\pi_l - \bar{\pi}_l)(\pi_m - \bar{\pi}_m)} E_{nj} E_{nm} \\ &= E_{ki} E_{kl} \sum_{lm} E_{mn}^T E_{nj} = E_{ki} \lambda_k \delta_{kn} E_{nj} = \\ &= \sum_k \lambda_k E_{ki} E_{kj} \end{aligned}$$

Om dus de covariantie $\overline{(\pi_i - \bar{\pi}_i)(\pi_j - \bar{\pi}_j)}$ als functie van \vec{r}_i en \vec{r}_j te schrijven, kunnen we gebruik maken van het feit dat E_{ki} en E_{kj} te schrijven zijn als functie van \vec{r}_i en \vec{r}_j respectievelijk (zie

hoofdstuk III, par. 2 van dit verslag): $E_{ki} = E_k(\vec{r}_i)$.

Bij meenemen van slechts de eerste k ($=5$) hoofdvelden $E_{1i} \dots E_{ki}$, welke een duidelijk verband vertonen met de geografische ligging van de meetpunten, ontstaat:

$$\overline{(\pi_i - \bar{\pi}_j)(\pi_j - \bar{\pi}_i)} = \sum_{\lambda} \lambda_{\lambda} E_{\lambda i} E_{\lambda j} \quad (A2)$$

Men bewijst nu gemakkelijk dat de gewichtsfuncties $g_i(\vec{r})$ gegeven worden door de uitdrukking $E_{\lambda i} E_{\lambda}(\vec{r})$, waarin $E_{\lambda}(\vec{r})$ de geïnterpoleerde waarde van het λ^e eigenveld ter plaatse \vec{r} is.

Immers:

$$\pi(\vec{r}) - \bar{\pi}(\vec{r}) = g_i(\pi_i - \bar{\pi}_i) \quad \text{leidt tot}$$

$$\overline{(\pi(\vec{r}) - \bar{\pi}(\vec{r}))(\pi_j - \bar{\pi}_j)} = \overline{g_i(\pi_i - \bar{\pi}_i)(\pi_j - \bar{\pi}_j)}$$

Het linkerlid is, volgens (A1), gelijk aan: $\sum_{\lambda} \lambda_{\lambda} E_{\lambda}(\vec{r}) E_{\lambda j}$, voor elke j en het rechterlid, volgens (A2), $g_i \sum_{\lambda} \lambda_{\lambda} E_{\lambda i} E_{\lambda j}$, ook voor elke j . Na vermenigvuldigen links en rechts met $E_{\mu j}$ ontstaat:

$\lambda_{\mu} E_{\mu}(\vec{r}) = g_i \lambda_{\mu} E_{\mu i}$ waarin niet sommatie over μ wordt geïmpliceerd.

Deling door λ_{μ} geeft dan het te bewijzen als resultaat.

Met deze uitdrukking voor $g_i(\vec{r})$ vinden we:

$$\pi(\vec{r}) - \bar{\pi}(\vec{r}) = E_{\lambda i}(\pi_i - \bar{\pi}_i) E_{\lambda}(\vec{r}) = a_{\lambda} E_{\lambda}(\vec{r})$$

en dit is precies gelijk aan de door middel van de hoofdcomponenten-analyse verkregen uitdrukking.

Figuuronderschriften

- Fig. 1 Een analyse van het gemiddelde luchtdrukveld boven Nederland. De getallen bij de stations geven de afwijking van het op dat station gemeten gemiddelde van het landelijk gemiddelde, dit laatste ook gemiddeld over de tijd, in μb .
- Fig. 2 Een analyse van de eerste hoofdcomponent, de getallen bij de stations zijn in 10^{-3} .
- Fig. 3 Als fig. 2, nu voor de tweede hoofdcomponent.
- Fig. 4 Als fig. 2, nu voor de derde hoofdcomponent.
- Fig. 5 Als fig. 2, nu voor de vierde hoofdcomponent.
- Fig. 6 Als fig. 2, nu voor de vijfde hoofdcomponent.
- Fig. 7 Als fig. 2, nu voor de zesde hoofdcomponent.
- Fig. 8 Als fig. 2, nu voor de zevende hoofdcomponent.
- Fig. 9 Een analyse van het gemiddelde luchtdrukveld boven Nederland, na interpolatie door een enkelvoudige tweedimensionale Taylor-reeks. De getallen bij de stations geven de afwijking van de door de interpolatiemethode geleverde waarde van het landelijk gemiddelde in μb .
- Fig. 10 Een analyse van de eerste hoofdcomponent na interpolatie d.m.v. een enkelvoudige tweedimensionale Taylor-reeks; de getallen bij de stations zijn in 10^{-3} .
- Fig. 11 Als fig. 10, nu voor de tweede hoofdcomponent.
- Fig. 12 Als fig. 10, nu voor de derde hoofdcomponent.
- Fig. 13 Als fig. 10, nu voor de vierde hoofdcomponent.
- Fig. 14 Als fig. 10, nu voor de vijfde hoofdcomponent.
- Fig. 15 De correlatiecoëfficiënten van π_i en π_j voor verscheidene waarden van i en j als functie van de afstand tussen de stations i en j .

Fig. 16 Metingen te Cabauw, d.d. 07-10-75 en berekende geowind;
de metingen zijn gemiddeld over het aangegeven halfuurs-
interval (tijd in gmt). In de figuur staat de windsnelheid
als functie van de hoogte en de grootte van de geowind;
onder de horizontale as staan de gemeten windrichting op
200 m hoogte en de richting van de geowind in graden ten
opzichte van het noorden.

-o-o-o-

Legends to Figures

- Fig. 1 An analysis of the mean pressure field above the Netherlands. The figures near the stations give the deviation of the measured mean from the mean that is averaged over the stations, in μb .
- Fig. 2 An analysis of the first principal component, the figures near the stations are in 10^{-3} .
- Fig. 3 As fig. 2, for the second principal component.
- Fig. 4 As fig. 2, for the third principal component.
- Fig. 5 As fig. 2, for the fourth principal component.
- Fig. 6 As fig. 2, for the fifth principal component.
- Fig. 7 As fig. 2, for the sixth principal component.
- Fig. 8 As fig. 2, for the seventh principal component.
- Fig. 9 An analysis of the mean pressure field above the Netherlands, after interpolation by means of a simple two-dimensional Taylor series. The figures near the stations give the deviation of the value found by the interpolation procedure from the mean that is averaged over the stations, in μb .
- Fig. 10 An analysis of the first principal component after interpolation by means of a simple two-dimensional Taylor series. The figures near the stations are in 10^{-3} .
- Fig. 11 As fig. 10, for the second principal component.
- Fig. 12 As fig. 10, for the third principal component. In

- Fig. 13 As fig. 10, for the fourth principal component.
- Fig. 14 As fig. 10, for the fifth principal component.
- Fig. 15 Correlation coefficients of π_i and π_j for several values of i and j as a function of the distance between the stations i and j .
- Fig. 16 Measurements at Cabauw, d.d. 07-10-75, and calculated geostrophic wind; the measurements have been averaged over the indicated half hour (time in gmt).
In the figure the windspeed is shown as a function of height, and the length of the geostrophic wind; under the horizontal axis the measured direction of the wind at a height of 200 m and the direction of the geostrophic wind in degrees with respect to the north are shown.

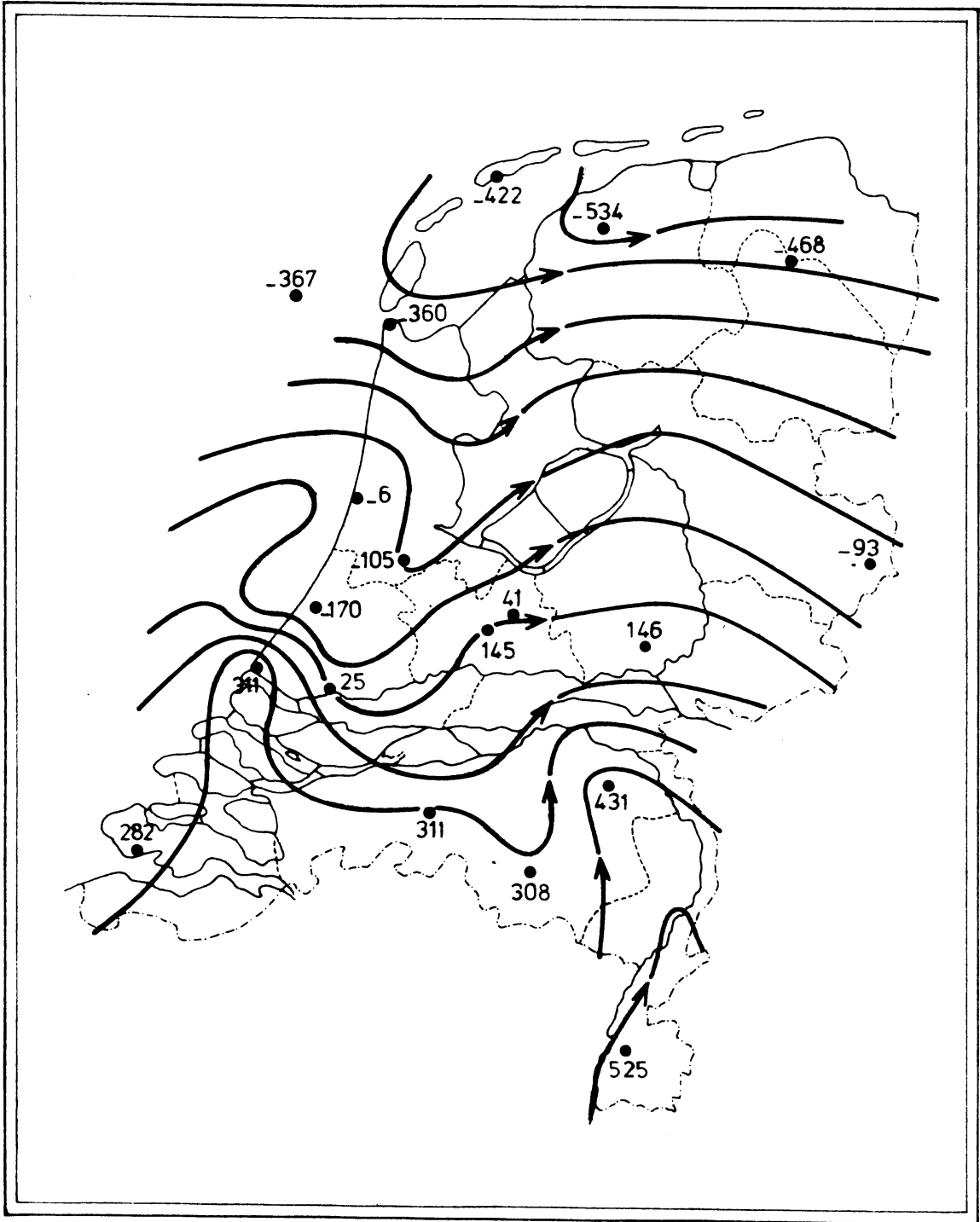


Fig.1

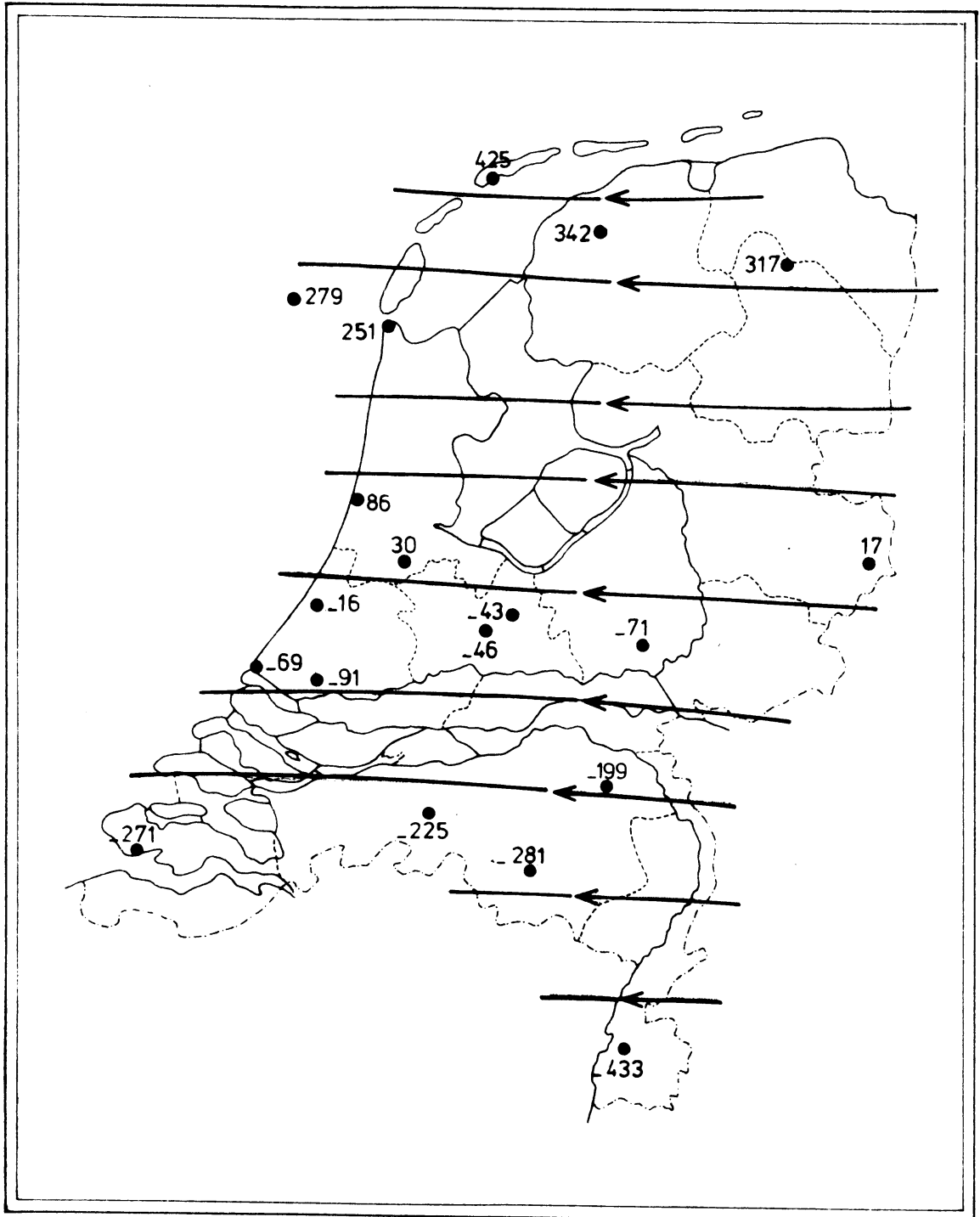


Fig.2

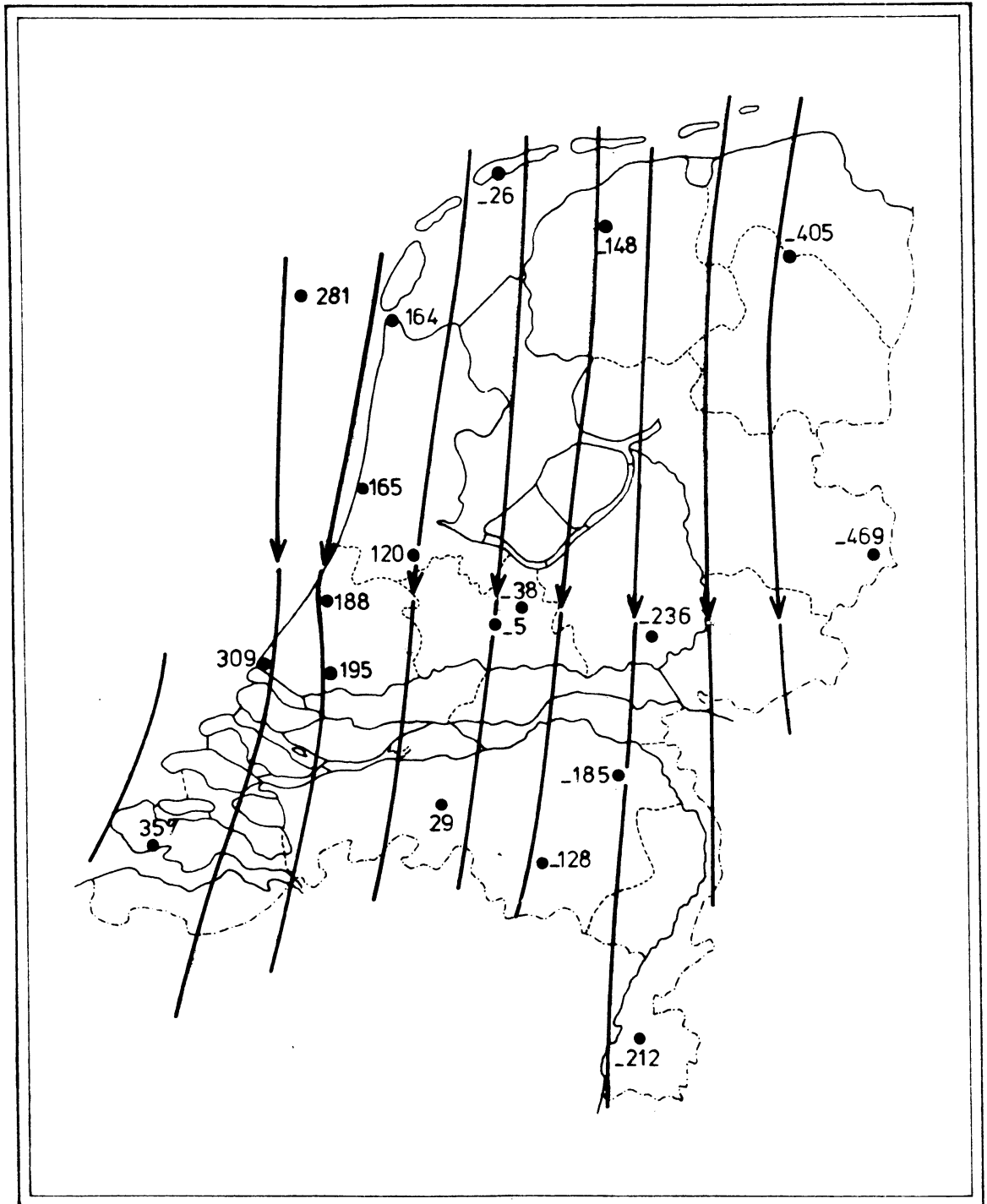


Fig. 3

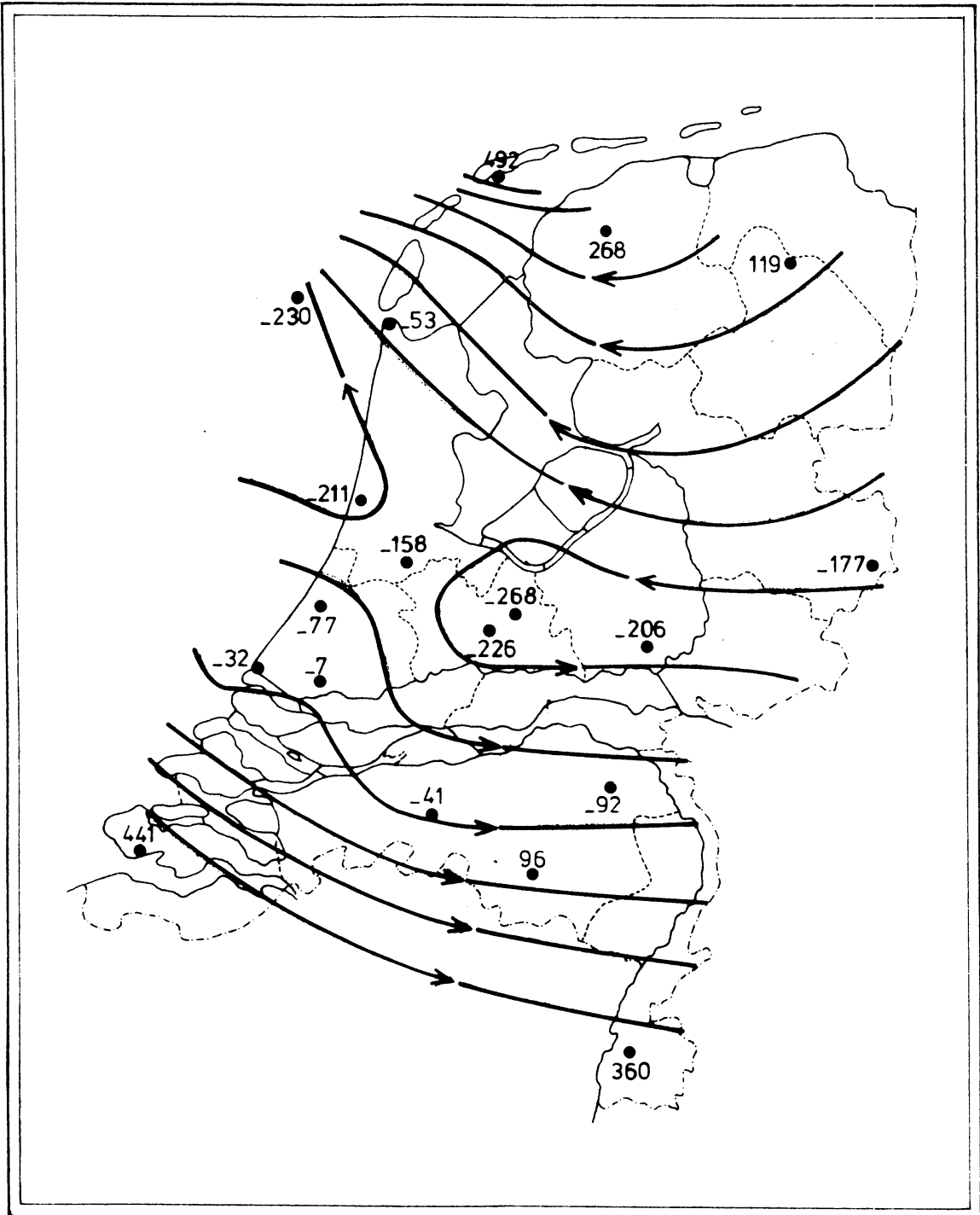


Fig. 4

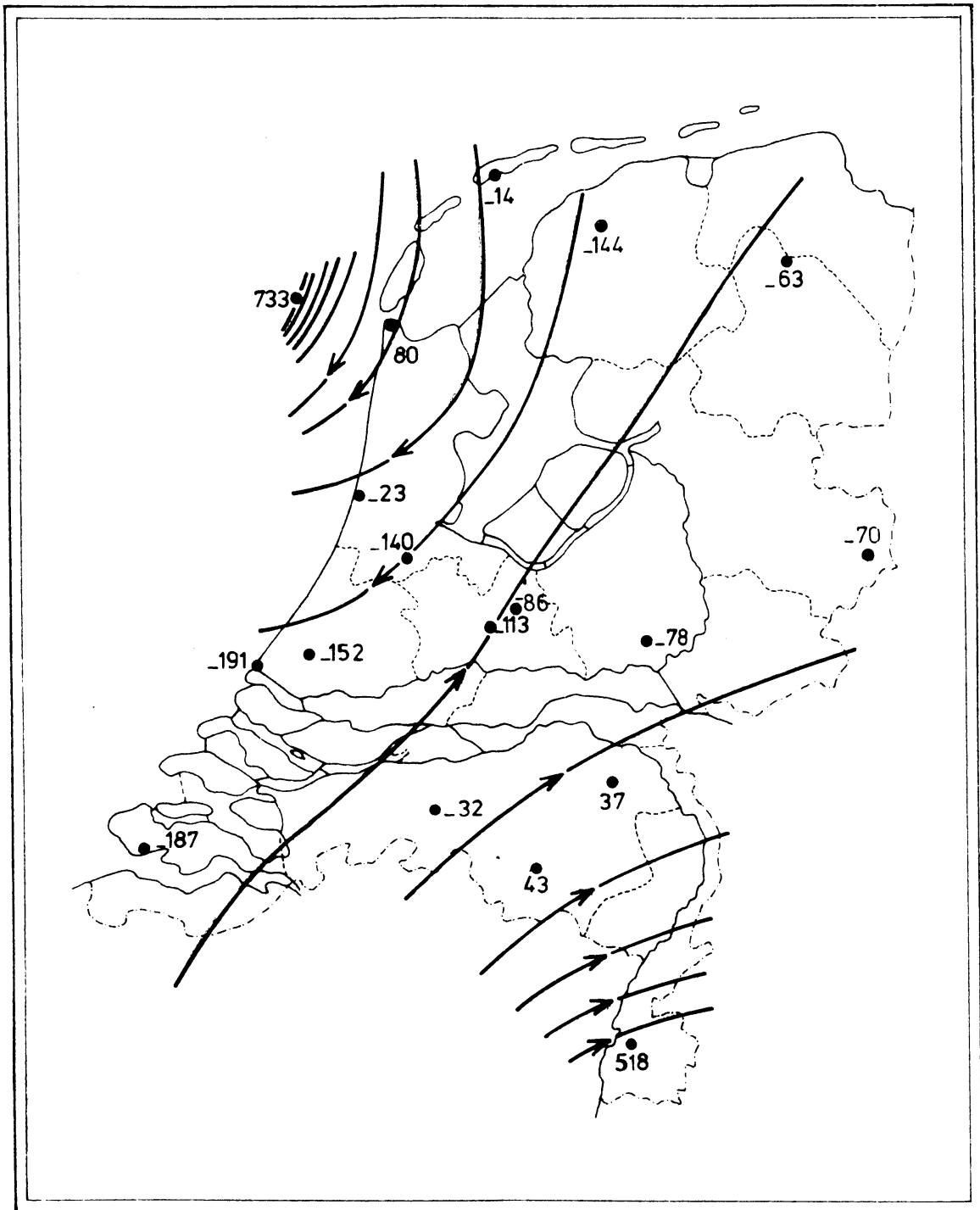


Fig.5

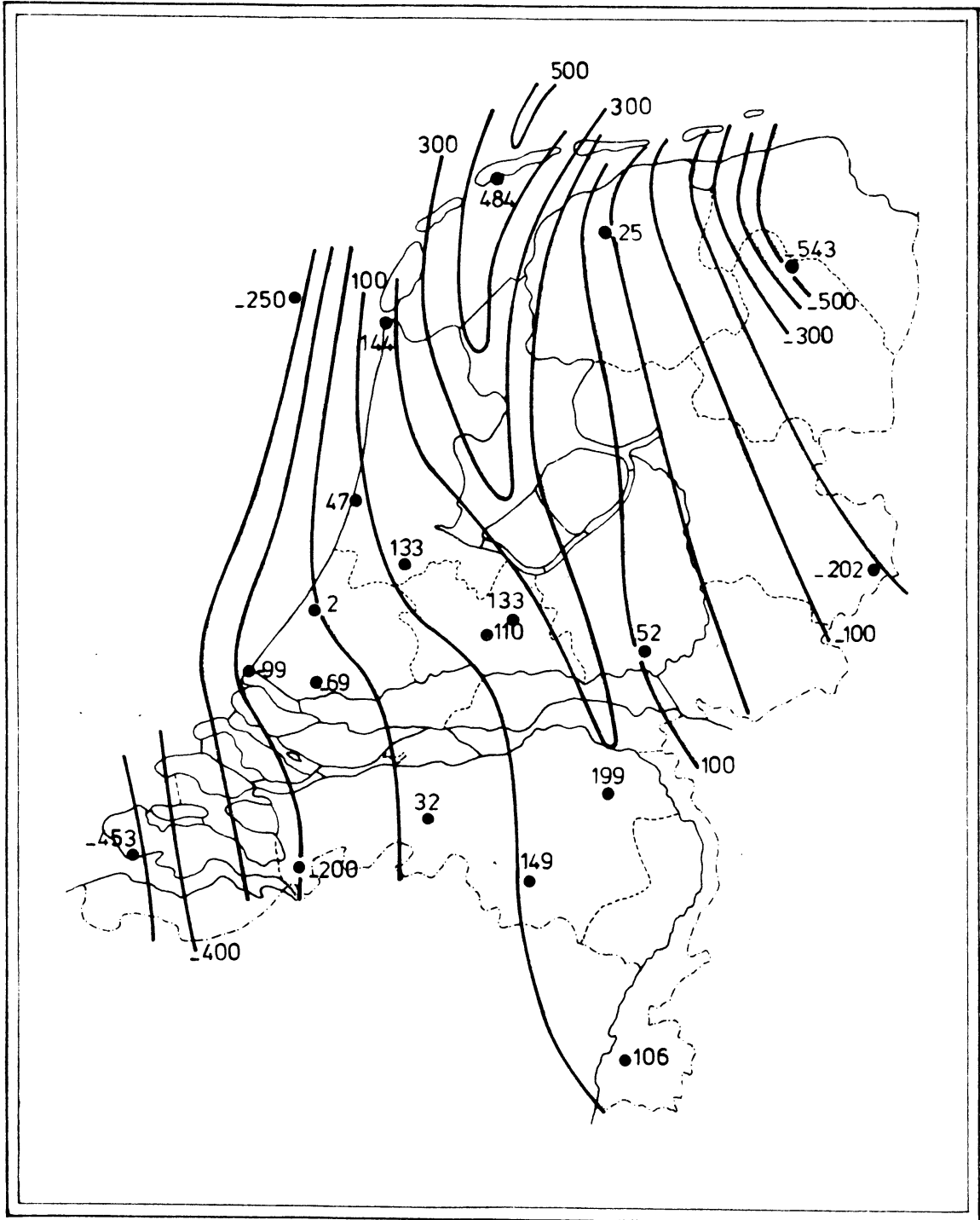


Fig.6

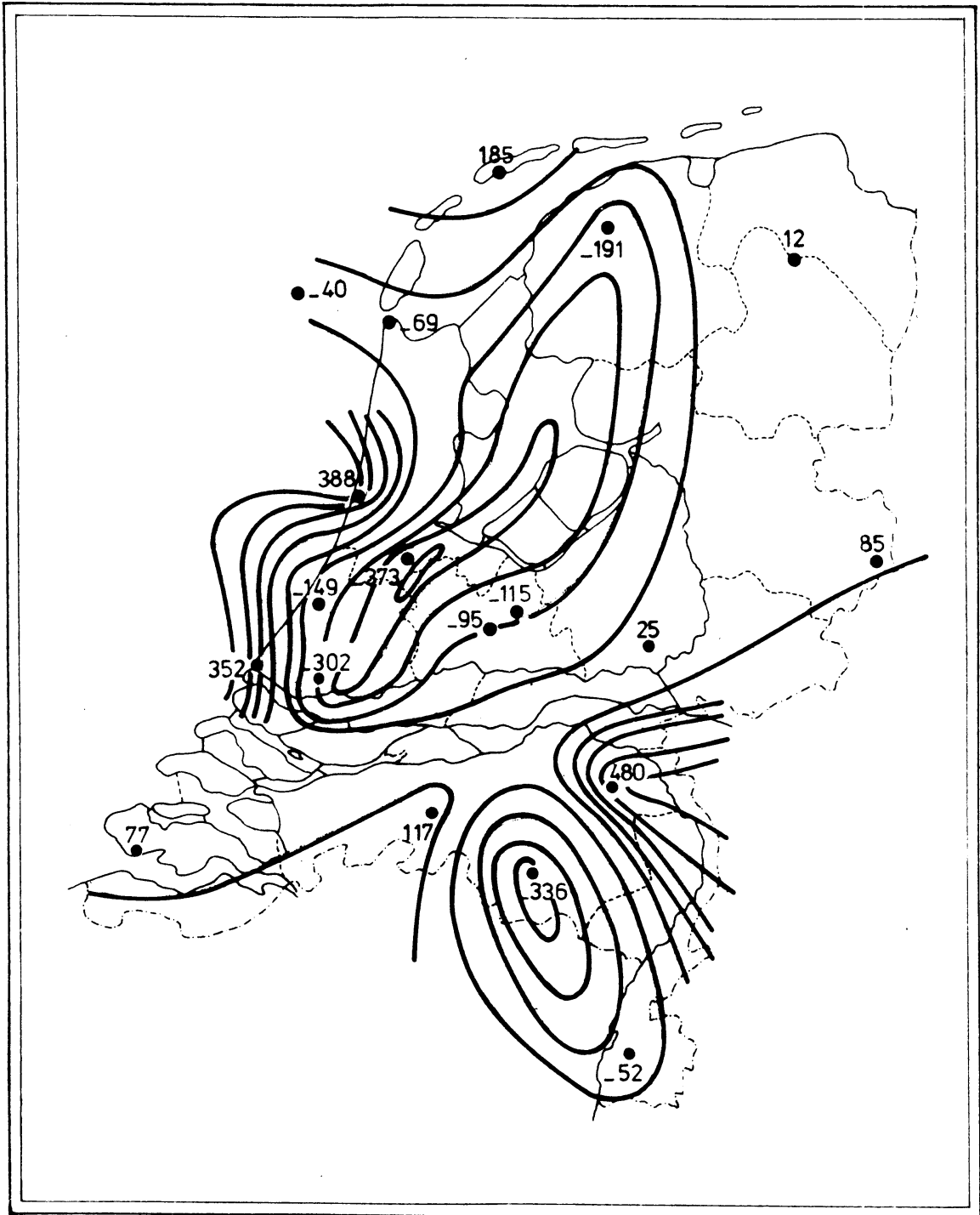


Fig 7

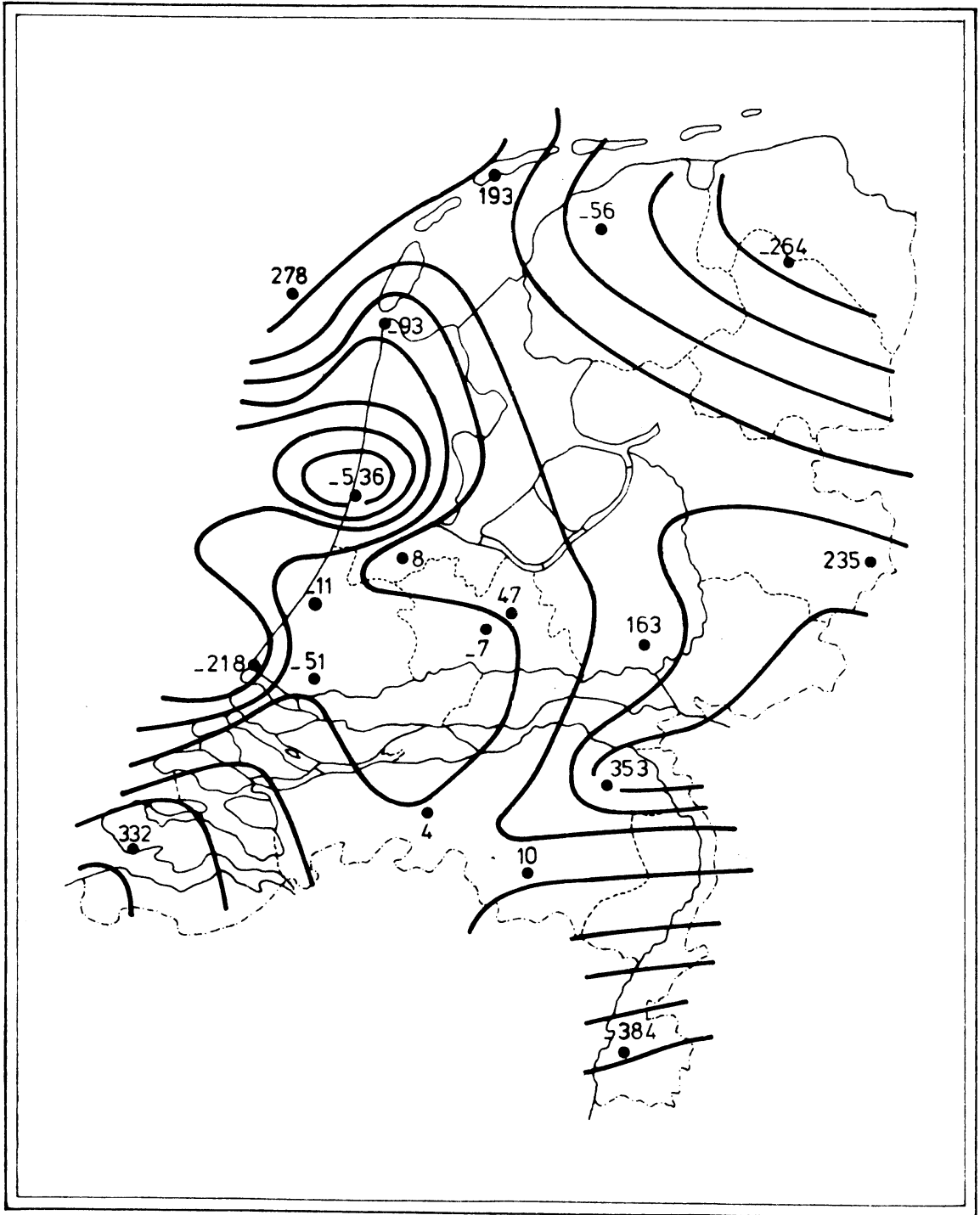


Fig. 8

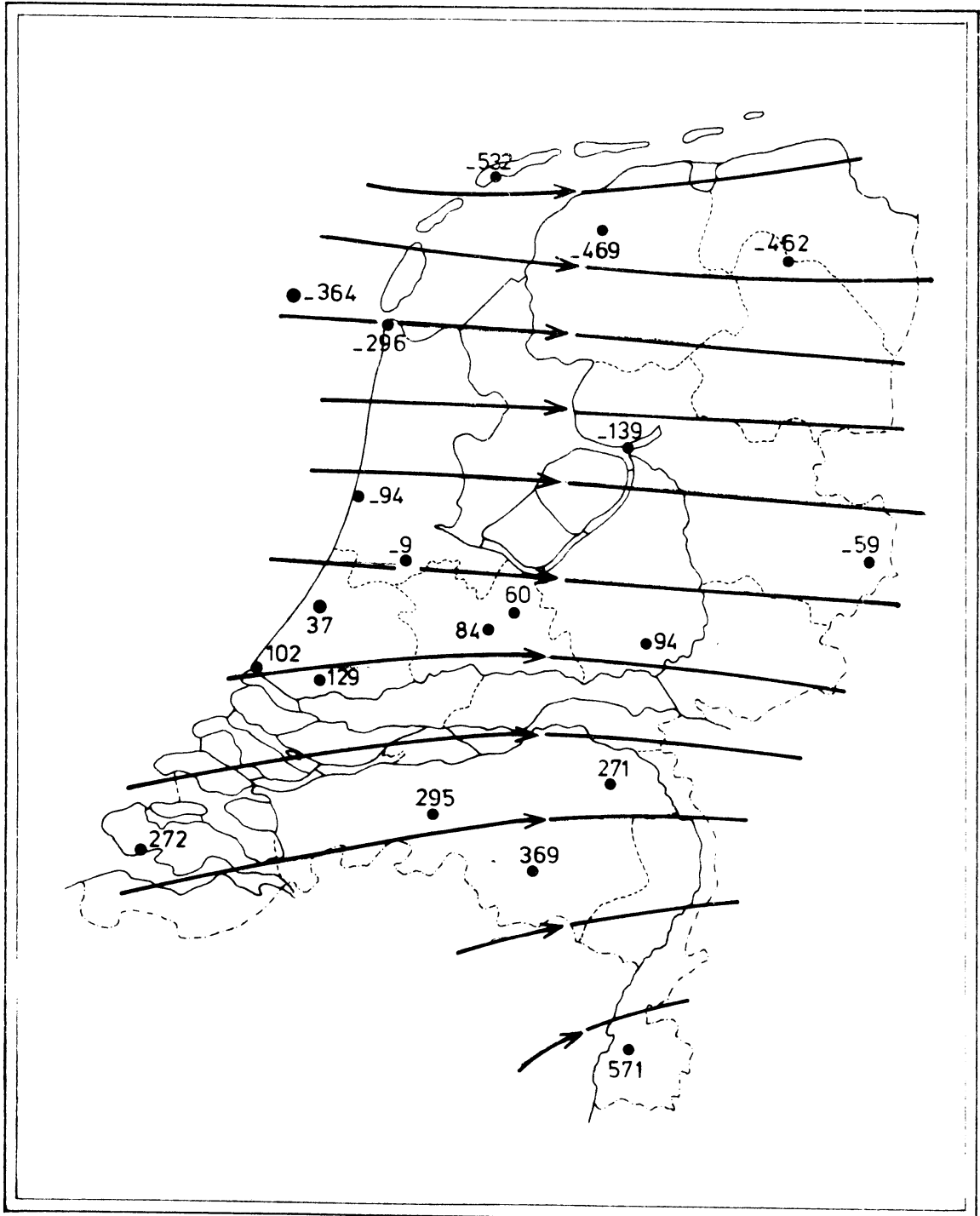


Fig. 9

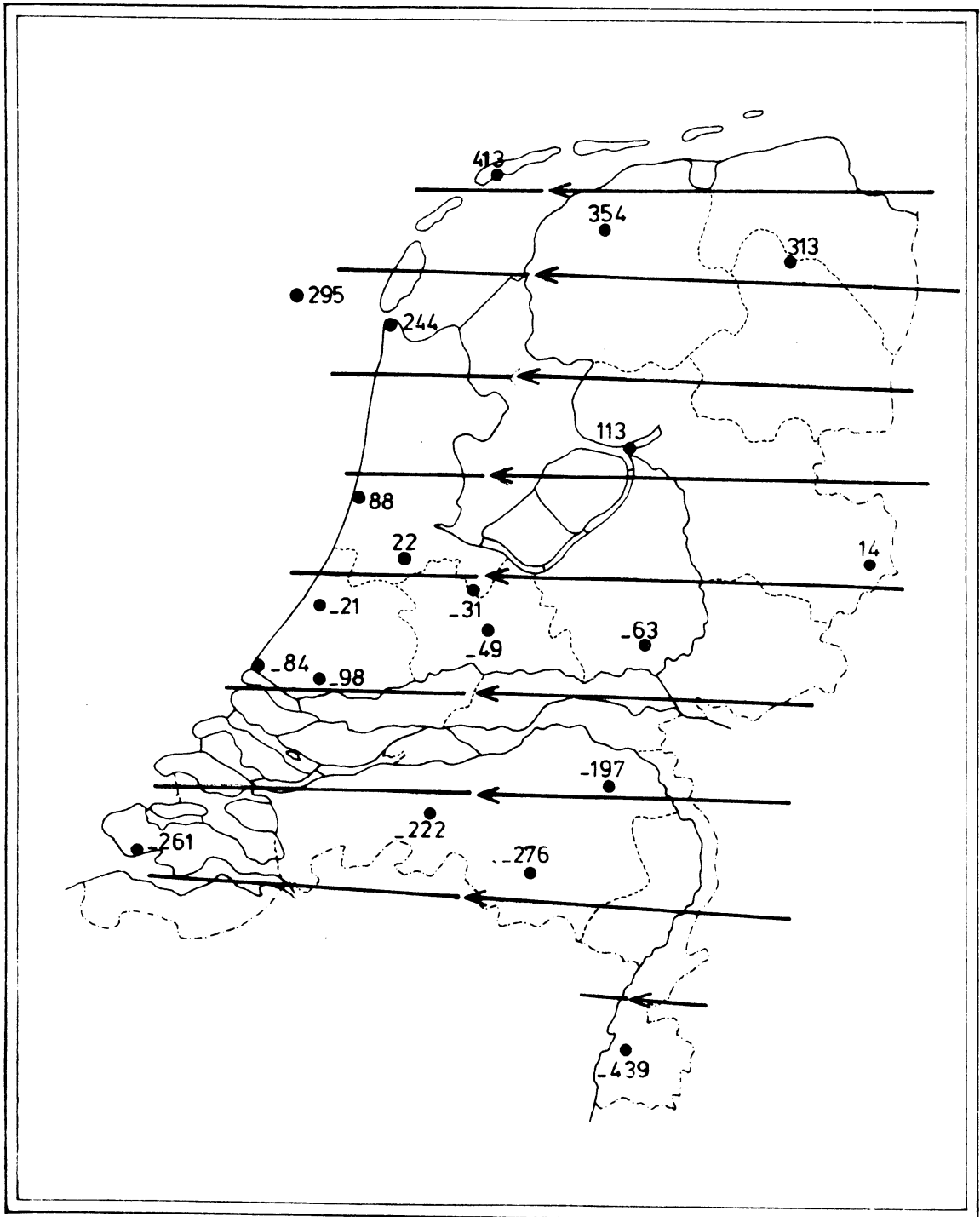


Fig.10

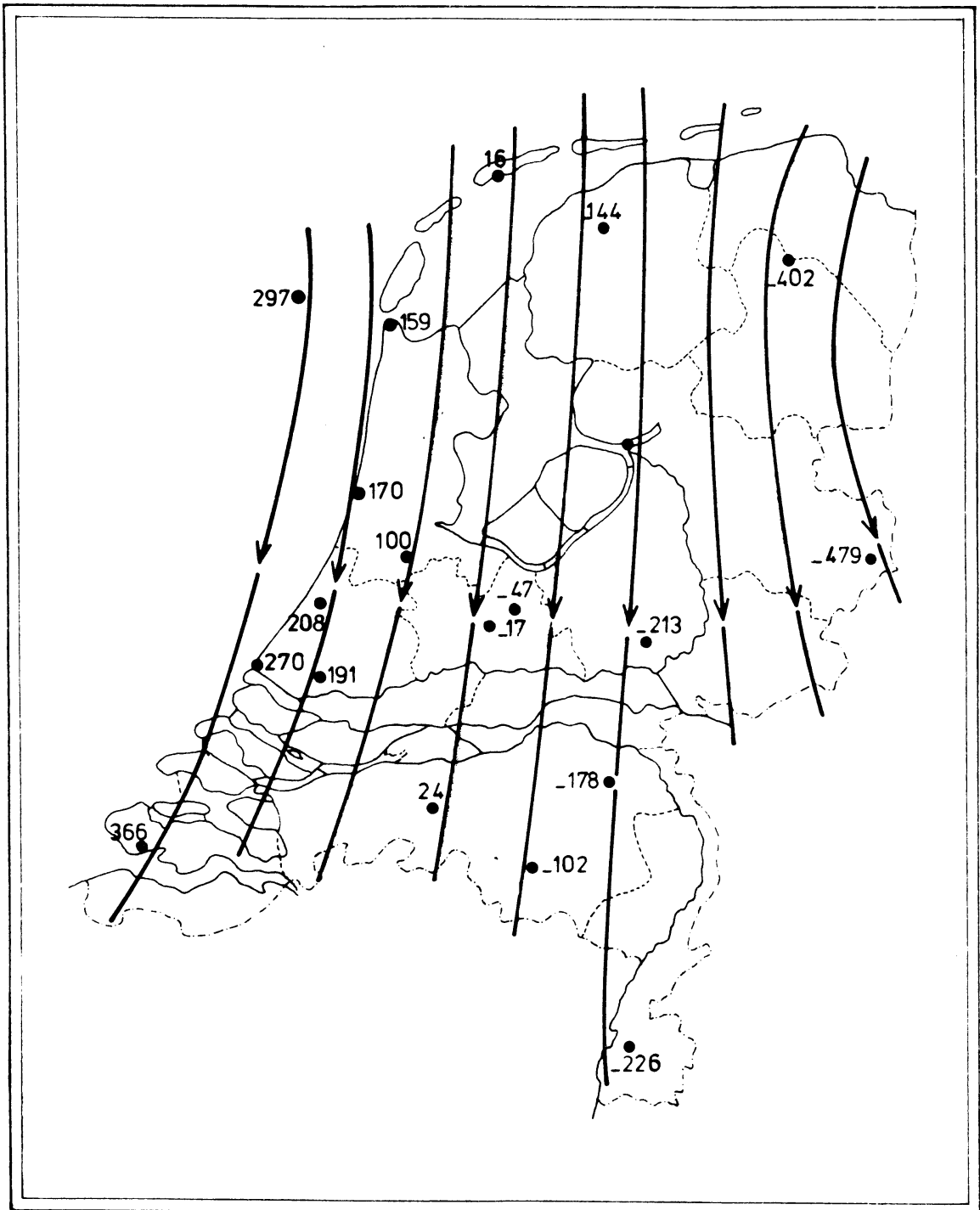


Fig.11

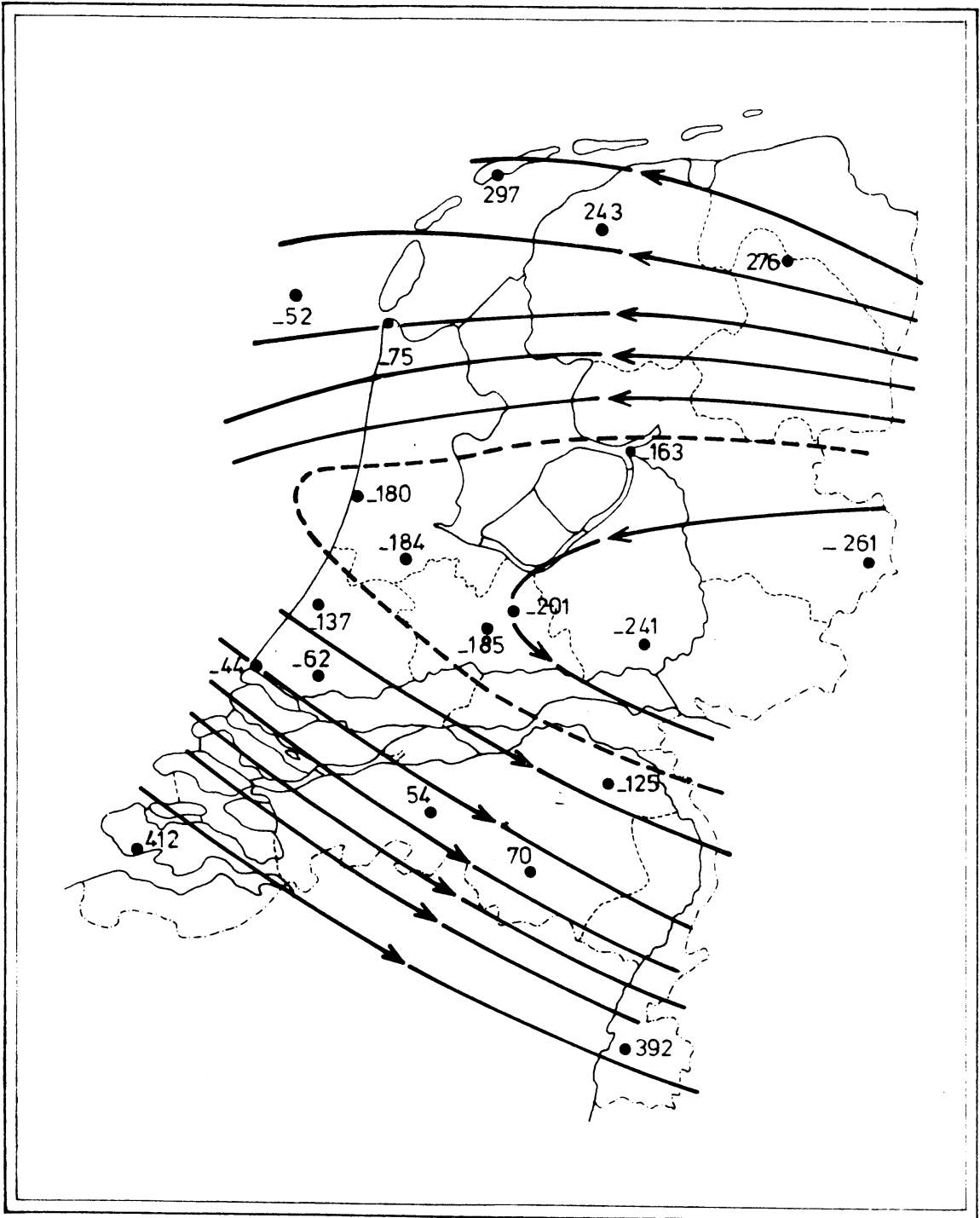


Fig.12

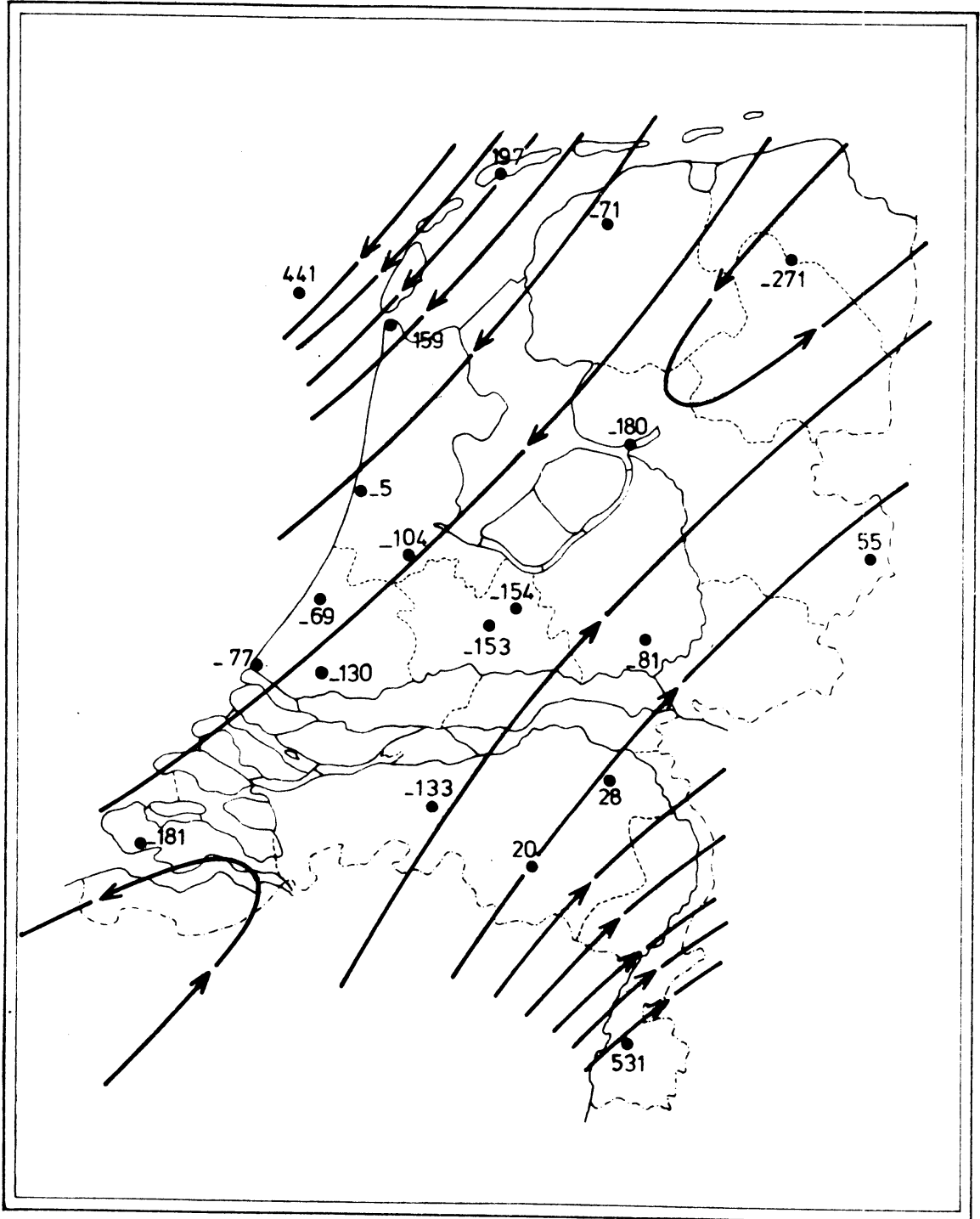


Fig.13

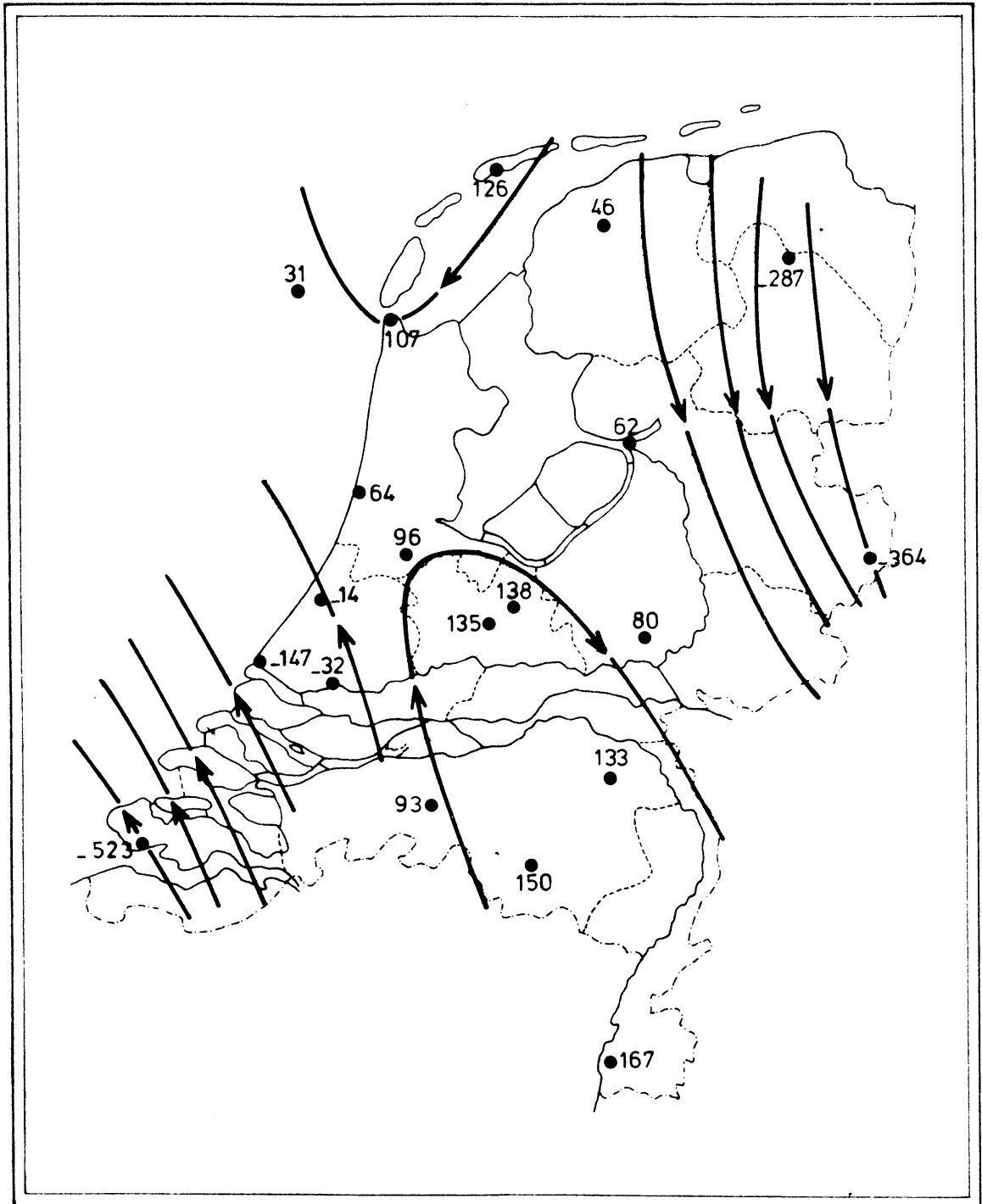


Fig.14

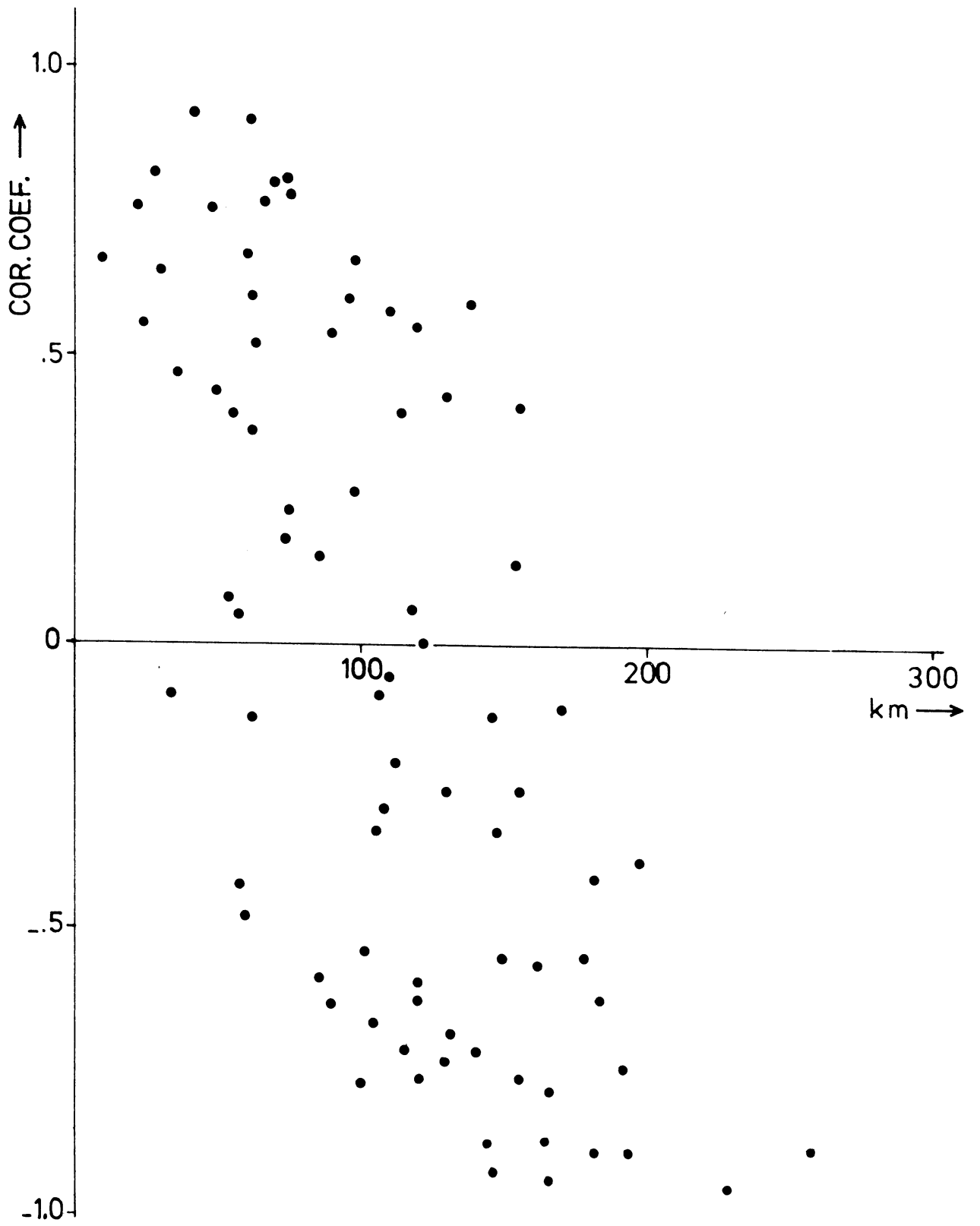


Fig.15

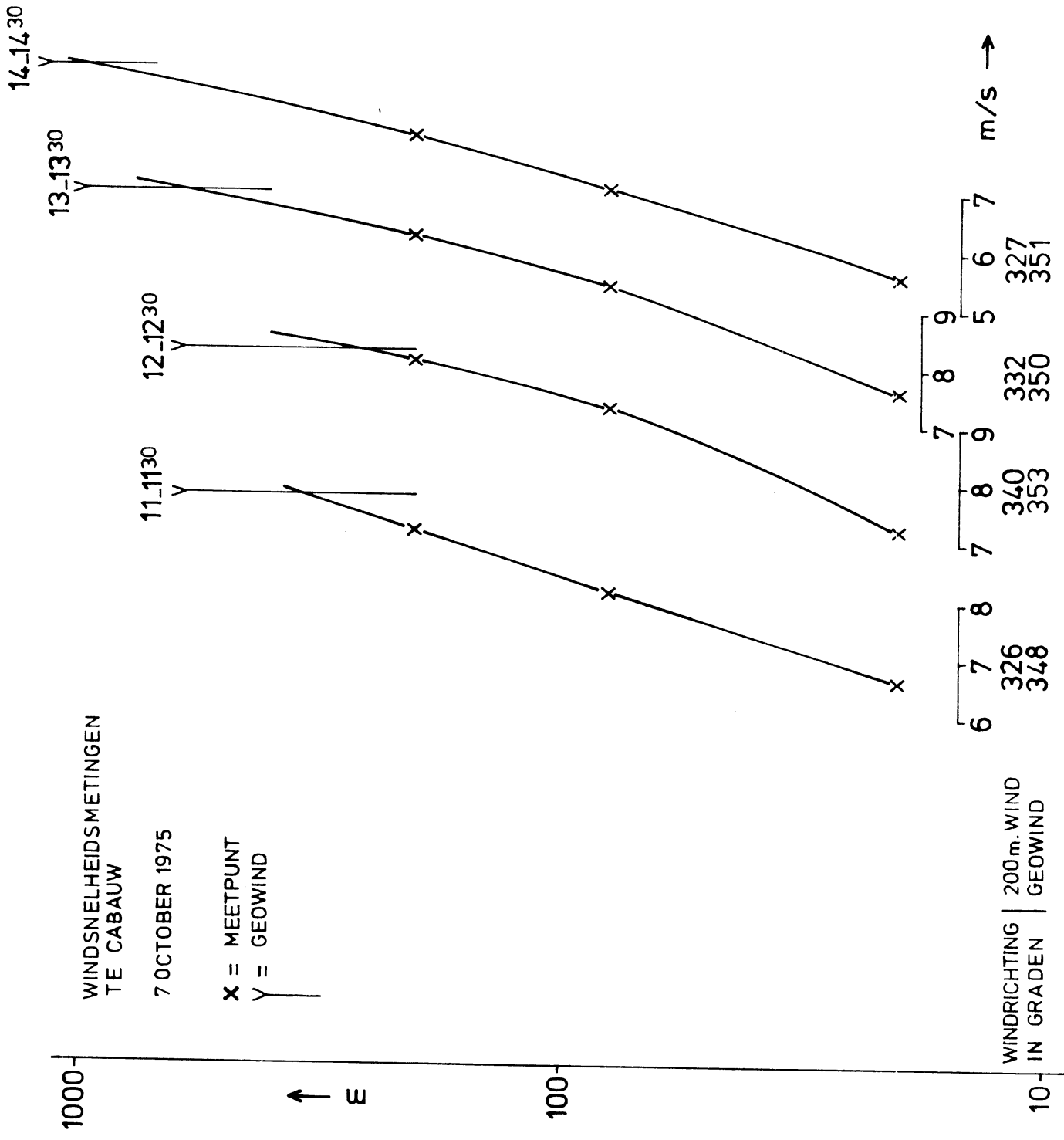


Fig.16