

## ASPECTEN VAN INFORMATIEVERWERKING

28

Een eerste kennismaking met

VAZAL

een programma-pakket voor het

- opschonen
- verbeteren
- verwerken
- opslaan
- presenteren

van gegevens

ing. J.B.H.M. van Gils

Nota's van het Instituut zijn in principe interne communicatiemiddelen, dus geen officiële publikateis.

Hun inhoud vaireert sterk en kan zowel betrekking hebben op een eenvoudige weergave van cijferreeksen, als op een concluderende discussie van onderzoeksresultaten. In de meeste gevallen zullen de conclusies echter van voorlopige aard zijn omdat het onderzoek nog niet is afgesloten.

Bepaalde nota's komen niet voor verspreiding buiten het Instituut in aanmerking

14 JUNI 1982

ISW 159 196 - 02

## ASPECTEN VAN INFORMATIEVERWERKING

28

De nota's handelend over Aspecten van Informatieverwerking bevatten inlichtingen over de ontwikkeling van informatieverwerking binnen het instituut. Naast meer concluderende en toelichtende beschouwingen zal aandacht worden besteed aan het gebruik van programma's en programma-pakketten en zullen zakelijke inlichtingen over praktijkervaring met en toepassing van de informatieverwerking worden gegeven.

+018219+016420+017621+193422+0023/23+011924+009525-069626-0  
 +017919+016120+017321+193422+0027/23+011924+009325-067426-0  
 +017619+015920+016821+193422+0047/23+009524+007825-064426-0  
 +017219+015720+016321+193422+0075/23+007524+006125-064226-0  
 +020219+017220+021021+193422-0503/23+024024+019025-081526-0  
 +019019+016520+019721+193422-0184/23+017924+013725-071426-0  
 +019019+016720+021121+193422-0376/23+025524+019125-072826-0  
 +020319+017420+021921+193422-0514/23+033524+025425-076026-0  
 +020019+017520+021321+193422-0364/23+027924+021525-075226-0

**Programma-pakket  
 VAZAL**

- opschonen
- verbeteren
- verwerken
- opslaan
- presenteren

van zesevens

\$	\$	\$	\$\$\$\$\$\$\$\$	\$	\$	79226-0
\$	\$	\$	\$	\$	\$	72026-0
\$	\$	\$	\$	\$	\$	65226-0
\$	\$	\$	\$	\$	\$	61726-0
\$	\$	\$	\$	\$	\$	69626-0
\$	\$	\$	\$	\$	\$	67226-0
\$	\$	\$	\$	\$	\$	63226-0
\$	\$	\$	\$	\$	\$	60826-0
\$	\$	\$	\$	\$	\$	61726-0
\$	\$	\$	\$	\$	\$	67226-0
\$	\$	\$	\$	\$	\$	70626-0

**een eerste  
 kennismaking**

J.B.H.M. van Gils

+016519+014020+015021+193422+0079/23+006224+004725-063026-0  
 +016019+014120+014921+193422+0063/23+004724+002825-064326-0  
 +015819+014020+014621+193422+0052/23+002824+000825-065626-0  
 +015519+013820+014521+193422+0038/23+001324+000625-065426-0  
 +015319+013820+014321+193522+0029/23+000524+000625-065626-0  
 +015219+013720+014221+193422+0030/23+000524+000625-066126-0  
 +015019+013720+014121+193522+0007/23+000524+000625-065926-0  
 +014919+013520+013921+193522+0006/23+000424+000625-065826-0  
 +014819+013420+013921+193522+0006/23+000524+000625-065826-0  
 +014719+013420+014021+193522+0006/23+000624+000625-065426-0  
 +014719+013420+013921+193522+0006/23+000524+000525-065326-0  
 +014619+013320+013821+193522+0006/23+000424+000525-065126-0  
 +014519+013220+013721+193522+0006/23+000424+000525-064426-0  
 +014419+013220+013721+193422+0061/23+004024+001725-064426-0  
 +014319+013020+013421+193422+0079/23+005524+003125-061026-0  
 +014119+012920+013221+193422+0071/23+004924+003125-061126-0  
 +014019+012820+013221+193422+0062/23+004724+002625-062226-0  
 +014119+012820+013321+193422+0051/23+004324+001825-063626-0  
 +014119+012920+013521+193422+0028/23+003924+001425-064326-0  
 +014219+013020+013721+193422+0015/23+004424+001825-064726-0  
 +014319+013220+013921+193422+0009/23+004824+002225-065326-0  
 +014619+013420+014321+193522-0093/23+004824+004225-066526-0  
 +014819+013620+014521+193522-0080/23+004224+003525-067226-0  
 +014819+013620+014421+193422-0058/23+003224+002825-067126-0  
 +014819+013620+014621+193522-0061/23+003324+002825-066826-0  
 +015419+014320+015721+193522-0229/23+010724+008925-071026-0  
 +016219+015220+017521+193522-0374/23+016924+014325-073926-0  
 +016319+015120+016321+193522-0105/23+005k24+004725-072026-0  
 +016319+015320+016821+193422-0189/23+010024+008525-072726-0  
 +017419+016420+019121+193422-0441/23+022224+018125-079126-0

nota 1317

**Instituut voor  
 Cultuurtechniek en  
 Waterhuishouding**

afdeling wiskunde

versie:  
 Januari 1982

+018819+017520+020921+193422-0624/23+032024+025425-082526-0  
 +019019+017320+019121+193422-0236/23+019824+015925-074426-0  
 +017819+016120+017521+193422-0000/23+010324+007925-066026-0  
 +017319+015820+016921+193422+0043/23+012724+010325-062826-0  
 +017019+015920+017021+193422-0000/23+015924+013525-061226-0  
 +018819+018020+019621+193622-0535/23+023224+018625-084926-0  
 +018919+017920+019521+193422-0455/23+019924+015125-084126-0  
 +018819+017520+019021+193422-0304/23+013324+010525-082526-0  
 +018619+017420+018921+193422-0282/23+012324+009725-082726-0  
 +018519+017220+018621+193422-0286/23+012624+009925-082026-0

## I N H O U D

	blz.
1. INLEIDING	1
2. MEER INFORMATIE OVER VAZAL	1
3. CENTRALE OPSLAGVORM	2
4. REGISTRATIE-VERMINKINGEN EN MEETFOUTEN	2
5. OMZETTING NAAR GEWENSTE GROOTHEDEN	4
6. PRESENTATIE VAN HET (BEWERKTE) BESTAND	5
7. ANDERE OPSLAGVORM	5
8. SPECIFIEKE BEWERKINGEN	6
9. DE PROGRAMMA'S IN VAZAL	6
9.1. Overzicht van de indirect command files	7
9.2. Overzicht van de hoofdprogramma's	7
10. SCHEMA: AANSLUITMOGELIJKHEDEN TUSSEN PROGRAMMA'S EN FILES	10
LITERATUUR	12

## 1. INLEIDING

Het programma-pakket VAZAL kan worden beschouwd als gereedschap bij het opschonen, verbeteren, verwerken, opslaan en presenteren van gegevens. Hieronder wordt uiteengezet, wat er met het getallenmateriaal van de onderzoeker gebeurt en wat hij er mee kan doen met behulp van VAZAL.

Ieder programma van VAZAL kan conversationeel werken (de gebruiker achter de terminal antwoordt op geprogrammeerde vragen). De antwoorden, die de gebruiker via de terminal kan geven, kunnen ook vooraf worden samengesteld tot een file en in die vorm aan een hoofdprogramma worden aangeboden.

VAZAL biedt een reeks subroutines met handleidingen. Programma's buiten VAZAL kunnen van deze subroutines gebruik maken.

## 2. MEER INFORMATIE OVER VAZAL

In VAZAL is de ervaring verwerkt, die is opgedaan met oudere versies van het pakket (VAN GILS, 1976). Hierbij is dankbaar gebruik gemaakt van de opmerkingen van zelfrekenende onderzoekers, die met de oudere versies werkten.

De hier beschreven versie van VAZAL staat geïnstalleerd bij de Landinrichtingsdienst te Utrecht op de PDP 11/70 van TWR.

De hier behandelde versie van VAZAL is die van januari 1982. Alleen de meest recente versie van VAZAL wordt onder UIC [50,60] beschikbaar gehouden. Deze "eerste kennismaking" staat in file VAZAL.TXT. Gebruikers vinden een meer gedetailleerde inleiding in file INLEIDING.TXT. Deze is ook in nota-vorm uitgebracht (VAN GILS, 1982).

Enige mappen met de meest recente listings van de VAZAL-files worden op de afdeling Wiskunde van het instituut beschikbaar gehouden.

### 3. CENTRALE OPSLAGVORM

De programma's van VAZAL gebruiken voor de gegevens een nader omschreven specifieke standaardvorm van opslag, waarin de informatie steeds kan worden geraadpleegd en naar believen kan worden gewijzigd. Deze opslagvorm wordt datafile genoemd. Programma's buiten VAZAL kunnen van deze datafile gebruik maken.

In een datafile staan de gegevens als gehele getallen in een matrix opgeslagen. Het adres van een getal wordt gevormd door het volgnummer van de rij (record) en het volgnummer van de kolom. Een uitverkoren waarde voor een getal, de onbekend-code, duidt aan dat het getal als onbekend moet worden beschouwd. De gehele getallen hebben een bereik van -32767 tot +32767.

De programma's van VAZAL zijn ingericht met de gedachte, dat in een kolom van de datafile de reeks getallen van een lopende variabele (als een tijdreeks) wordt opgeslagen en in een record de getallen behorend bij eenzelfde parameterwaarde, bijvoorbeeld tijd, plaats, of enquetennummer. Zo krijgen bijvoorbeeld de grondwaterstands-metingen, op een plaats verricht, een kolomnummer en bevat ieder record in de betreffende kolom de meting behorende bij eenzelfde datum. Alle getallen van een scan van een datalossier zijn geregistreerd met een kloktijd en worden opgeslagen in 1 record van een datafile.

### 4. REGISTRATIE-VERMINKINGEN EN MEETFOUTEN

De ervaring heeft geleerd, dat in reeksen getallen (gegevens) fouten kunnen voorkomen. Dit kunnen menselijke, maar ook instrumentele fouten zijn. Door een type- of ponsfout ontstaat dikwijls een enkel fout getal, een fout van een automatische meteropnemer veroorzaakt een reeks foute getallen in een kolom, een registratie-verminking van een datalogger veroorzaakt nogal eens een reeks getallen in

foute adressen in een of twee opeenvolgende records.

Registratie-verminkingen en type- en ponsfouten kunnen het beste worden verbeterd in de onbewerkte registratie, omdat tengevolge van conversies in programma's informatie verloren kan gaan, en, omdat een fout in de registratie een reeks foute getallen kan veroorzaken. De characters, waaruit de registratie is opgebouwd, kan men verbeteren met behulp van een editor. Onleesbare getallen en onjuiste adresseringen worden opgespoord tijdens de omzetting van de geregistreerde characters tot getallen in binaire code in core. Na verbetering van deze fouten in de registratie moet de conversie worden herhaald.

De omzetting wordt uitgevoerd door programma INPUT. Uit de registraties worden allereerst de scans (de verzameling getallen die in een record wordt opgeslagen) samengesteld en vervolgens worden de characters in de scans geconverteerd tot getallen en onderwijl worden de adressen van de getallen in de datafile bepaald. INPUT toetst hierbij op:

- de lengte van de scan (facultatief)
- het voorkomen van niet-numerieke characters in de scan (facultatief)
- het voorkomen van niet te decoderen characters in de scan (onjuiste volgorde van de characters)
- het optreden van overschreden adressen in de datafile als gevolg van verminkte en verloren scheidingstekens tussen de getallen en tussen de scans (ingeval scheidingstekens in de registratie zijn opgegeven).

Programma INPUT kan hierbij de adressen vermelden waarin onbekend-codes zijn geschreven tengevolge van registratie-verminkingen en tengevolge van lege plaatsen in de registraties in plaats van getallen.

Eventuele overblijvende storende foute getallen tengevolge van type- en ponsfouten en registratie-verminkingen kunnen worden behandeld als meetfouten. Het visueel opsporen van verminkingen in de registratie lukt niet altijd volledig, het opsporen ervan met behulp van de programma's van VAZAL was tot nu toe (voor 100%) betrouwbaar.

De criteria voor het opsporen van meetfouten zullen per onderzoek moeten worden vastgesteld. De criteria kunnen bijvoorbeeld zijn:

- liggen de waarnemingen binnen een gesteld acceptatie-interval? (bijv. liggen de uren tussen 0 en 23)
- liggen de waarnemingen binnen het fysisch/technisch mogelijke bereik?
- is het verschil tussen twee opeenvolgende waarnemingen niet te groot?

Onder de voorwaarde dat aan zulke criteria wordt voldaan kunnen met getallen in records in de datafile de volgende bewerkingen worden uitgevoerd:

- |  |                   |
|--|-------------------|
| - tabelleren                               | (programma TABEL) |
| - plotten                                  | (programma PLOT)  |
| - op het beeldscherm brengen en verbeteren | (programma BETER) |
| - transformeren                            | (programma TRANS) |
| - kopiëren                                 | (programma COPY)  |

Met behoud van hun toetsingsmogelijkheden kunnen de programma's van VAZAL worden gewijzigd. Dan kan bij meer gespecificeerde toepassingen veelal meer informatie omtrent de adressering en de waarde van de geregistreerde getallen worden gebruikt.

Programma MERGE kan de scans samenstellen en schrijven als characters. Deze bewerking kan aan de omzetting van INPUT voorafgaan. De foutmeldingen van MERGE zijn minder omvangrijk.

## 5. OMZETTING NAAR GEWENSTE GROOTHEDEN

Uit het verbeterde bestand in de datafile zullen bruikbare getallenreeksen moeten worden gevormd. Afhankelijk van het verkregen materiaal zal de onderzoeker hiervoor een aantal bewerkingsstappen opgeven. Dit kunnen bijvoorbeeld zijn:

- sorteren van de waarnemingen (bijvoorbeeld per meteropnemer)



- herzien van de waarnemingen, die buiten het meetbereik liggen
- schalen van de meetwaarden tot gewenste eenheden
- bepalen van het aantal intervallen (scans) zonder registratie
- berekenen van hoeveelheden per eenheid (bijvoorbeeld mm. per dag) uit een variërend aantal scans.

TRANS is een algemeen werkend transformatie-programma dat getallen in de datafile voorwaardelijk kan transformeren. De instructies ervoor kunnen zonder fortrankennis worden opgesteld. TRANS kan uitgebreid worden met subroutines, die op het betreffende onderzoek gericht zijn, waardoor het aantal instructies sterk kan worden beperkt.

## 6. PRESENTATIE VAN HET (BEWERKTE) BESTAND

Programma TABEL voorziet in een mogelijkheid om de getallen in de datafile overzichtelijk in een tabel weer te geven. Het programma stelt de gebruiker in staat door voorwaardelijk te tabelleren een gecombineerde automatische en visuele controle van grote aantallen gegevens uit te voeren. Hiertoe worden tevens per kolom in de tabel het maximum, het minimum, het aantal bekende getallen en het aantal onbekende getallen opgegeven.

Programma PLOT kan voorwaardelijk per scan de getallen met een eigen symbool per lopende variabele plotten in een raster met vaste maas-afmetingen. PLOT wordt hoofdzakelijk gebruikt voor het weergeven van tijdreeksen. Met de grafieken is een visuele controle op het regelmatig verlopen van uitgevoerde metingen mogelijk.

## 7. ANDERE OPSLAGVORM

Vele bestaande programma's en programma-pakketten kunnen niet lezen van de hier gedefinieerde datafile, maar wel van een sequentiële ASCII-file. Programma OUTPUT schrijft voorwaardelijk de getallen in de datafile naar een dergelijke opslag.

## 8. SPECIFIEKE BEWERKINGEN

In specifieke op het onderzoek gerichte bewerkingen op en met getallen in de datafile hoeft door VAZAL niet voorzien te zijn. De gebruiker kan eigen programma's laten aansluiten aan de datafile. Eigen programma's kunnen gewijzigde programma's van VAZAL zijn. De bewerkingsonderdelen in de programma's van VAZAL zijn veelal als subroutines geprogrammeerd en beschreven. Deze subroutines kunnen desgewenst worden overgenomen in specifieke programma's.

Enige technische hulp wordt de gebruiker geboden bij het aanmaken en gebruiken van zijn specifieke programma's en files door de indirect command files en door programma PRINT.

## 9. DE PROGRAMMA'S IN VAZAL

Alleen veelvuldig gebruikte taken zijn in VAZAL ondergebracht. Ze worden uitgevoerd door systeemprogramma's en hoofdprogramma's. De systeemprogramma's worden gestuurd door "indirect command files". Een indirect command file is een programma, dat commands aan systeemprogramma's geeft.

De handleiding van de indirect command files en de hoofdprogramma's beginnen met een korte beschrijving van wat het programma doet. De gebruiksaanwijzing in een handleiding geeft een formele beschrijving van de conservatieve instructie-input. Alleen omvangrijke taakonderdelen, die formeel beschreven zijn in de handleidingen van de subroutines, worden in het aanroepende hoofdprogramma toegelicht aan de hand van een voorbeeld.

De beginnende gebruiker zal met de toepassing van een programma weinig moeite hebben als hij de conversatie achter de terminal eerst een keer probeert met de handleiding ernaast.

## 9.1. Overzicht van de indirect command files

De indirect command files in VAZAL werken conversationeel. Ze kunnen ook met programma's buiten VAZAL werken. De taken van de indirect command files kunnen als volgt kort worden aangeduid:

- BATCH.CMD** starten van een hoofdprogramma in de BATCH QUEUE.  
Toepassing: - het uitvoeren van veel tijd vragende bewerkingen met hoofdprogramma's zonder eraan de terminal te binden.  
- het verzamelen van de conversatie met een hoofdprogramma op een te printen file.
- FOR.CMD** compileren en linken van een hoofdprogramma tot een "executable program".
- LOGIN.CMD** instellen van de terminal-karakteristieken van een DT80/1-terminal bij de login-procedure.
- PIPLI.CMD** verzamelen van filenamen in een opgegeven volgorde in directories.
- PRINT.CMD** afdrukken van files via de printer gekoppeld aan een DT80/1-terminal.
- PRUGE.CMD** verwijderen van alle vorige versies van files en de laatste versie op ;1 stellen.
- RUN.CMD** starten van een "executable program" eventueel met de programma-instructies verzameld in een file.

## 9.2. Overzicht van de hoofdprogramma's

De hoofdprogramma's kunnen conversationeel werken of kunnen de instructies van een file lezen. De taken van de hoofdprogramma's kunnen als volgt kort worden aangeduid:

- BETER.FIN** verbeteren van getallen in de datafile  
Toepassing: - het interactief opsporen van fouten en het verbeteren van getallen  
- kijken naar getallen.

**COPY.FIN**      kopiëren van getallen in een datafile  
 Toepassing: - het toevoegen van een deelbestand aan een  
                  hoofdbestand  
                  - het her-indelen en/of vergroten van een da-  
                  tafile  
                  - het overnemen van reeksen getallen uit een  
                  andere datafile.

**INPUT.FIN**    inlezen van getallen in een datafile  
 Toepassing: - het lezen en opslaan van getallen, waarin  
                  fouten kunnen voorkomen, uit vaste kolommen  
                  of kolommen tussen scheidingstekens  
                  - het opsporen van fouten in registraties.

**MERGE.FIN**    samenvoegen van registratie-records tot een reocrd per  
 scan in een ASCII-file  
 Toepassing: - het opsporen van fouten in registraties  
                  - het rechtstreeks toegankelijk maken van  
                  records (scans) met fouten met behulp van  
                  een nummering in de records  
                  - het voor het oog overzichtelijk maken van  
                  registraties bij gebruik van beeldscherm  
                  of listing in combinatie met programma  
                  PRINT.FIN.

**OUTPUT.FIN**    schrijven van getallen van een datafile naar een ASCII-file  
 Toepassing: - het leesbaar maken van de gegevens in een  
                  datafile voor andere bestaande programma's,  
                  programma-pakketten en computers.

**PLOT.FIN**      plotten van getallen in een datafile  
 Toepassing: - het maken van grafieken met tijdsduurlijnen  
                  op een regeldrukker van gegevens in de da-  
                  tafile  
                  - het mogelijk maken van een visuele controle  
                  op het regelmatig verlopen van de uitvoer-  
                  de metingen.

**PRINT.FIN**    samenstellen tot een listfile  
 Toepassing: - overzichtelijke presentatie van output van  
                  programma MERGE

- outputprogramma gebruikt door indirect command file PRINT.CMD.

**TABEL.FIN** tabelleren van getallen in een datafile

Toepassing: - het overzichtelijk presenteren van de gegevens in de datafile

- het mogelijk maken van een gecombineerde automatische en visuele controle van grote aantallen gegevens.

**TRANS.FIN** transformeren van getallen in een datafile

Toepassing: - het sorteren van waarnemingen per meteropnemer

- het herzien van waarnemingen, die buiten het meetbereik liggen
- het schalen van meetwaarden tot gewenste eenheden
- het uitvoeren van eenvoudige berekeningen zoals bijvoorbeeld cumulaties tot hoeveelheden per tijdvak.

## 10. SCHEMA: AANSLUITMOGELIJKHEDEN TUSSEN PROGRAMMA'S EN FILES

De gebruikelijke procedure op de PDP-11/70 is, dat de gegevens worden aangeboden op magnetische tape (9-track tape) of als een pakket ponskaarten of worden ingetypt via de terminal.

De gegevens worden veelal eerst als een file op disk gezet. Dit is in het navolgende schema de "sequential ASCII input" (zie "INGANG"). De tape-driver of de card reader kunnen ook direct de "sequential ASCII input" geven, maar in een bestand op tape of ponskaart kan men moeilijker verbeteringen aanbrengen.

Zoals het schema laat zien worden de gegevens overgedragen van "sequential ASCII input" naar een datafile, van een datafile naar dezelfde of een andere datafile en van een datafile naar "sequential ASCII output". Sequential ASCII input en output zijn meestal disk-files. Sequential ASCII output is weer leesbaar voor andere programma's en kan overgezonden worden naar andere computers.

In elke vorm, waarin de gegevens staan, kunnen ze worden geraadpleegd en kunnen ze worden gewijzigd. Van sequential ASCII-files kunnen listings worden gemaakt, de files kunnen op het beeldscherm worden vertoont en de characters in de files kunnen met een editor worden veranderd. De getallen in de datafiles kunnen in tabellen en grafieken worden weergegeven op het beeldscherm en op papier en ze kunnen worden veranderd met diverse programma's van VAZAL en met programma's buiten VAZAL, die op een datafile aansluiten. De keuze van het programma, waarmee veranderingen in een datafile worden aangebracht hangt af van de gewenste bewerking.

In het schema worden de volgende begrippen gebruikt:

- sequential : input/output geschiedt van den beginne af van voor naar achter
- ASCII : American Standard Code of Information Interchange code van de characters in een PDP-computer
- formatted : input/output volgens een in het programma opgegeven formaat.



## LITERATUUR

- GILS, J.B.H.M. VAN, 1976. Aspecten van Informatie-verwerking 1. Computerverwerking van lange reeksen getallen. ICW-nota 935: pp. 15.
- GILS, J.B.H.M. VAN, 1982. Aspecten van Informatie-verwerking, 29. Mogelijkheden, bijbehorende denktrant en opbouw van VAZAL, een programma-pakket voor het opschonen, verbeteren, verwerken, opslaan en verbeteren van gegevens. ICW-nota 1318: pp. 31+5.