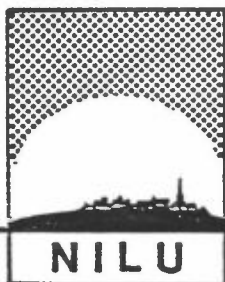


NILU  
OPPDRAGSRAPPORT NR 9/83  
REFERENCE: 22682  
DATO: AUGUST 1983

## *RESEPTORMODELLER*

Jan Schaug

Utført etter oppdrag  
fra Statens forurensningstilsyn



**NORSK INSTITUTT FOR LUFTFORSKNING**

POSTBOKS 130.- 2001 LILLESTRØM

NILU  
OPPDRAGSRAPPORT NR 9/83  
REFERENCE: 22682  
DATO: AUGUST 1983

***RESEPTORMODELLER***

Jan Schaug

Utført etter oppdrag  
fra Statens forurensningstilsyn

NORSK INSTITUTT FOR LUFTFORSKNING  
POSTBOKS 130, 2001 LILLESTRØM  
NORWAY

ISBN-82-7247-366-6

INNHALDSFORTEGNELSE

	Side
FORORD .....	5
SAMMENDRAG .....	7
1 INNLEDNING .....	9
2 BESKRIVELSE AV AKTUELLE MODELLER .....	11
2.1 Den kjemiske massebalanse metoden .....	14
2.2 Anrikningsfaktorer .....	17
2.3 Multivariable metoder .....	18
2.3.1 Faktoranalyse og hovedkomponentanalyse .....	18
2.3.2 Målrettet transformasjon av faktor- resultater ("Target transformation factor analysis") .....	22
2.3.3 Regresjonsanalyse .....	23
2.3.4 "Cluster" analyse .....	24
2.4 Datainnsamlingen .....	25
2.5 Prøvetaking .....	25
2.6 Kildesammensetningen .....	26
2.7 Komponenter og analytiske metoder .....	27
3 BRUK AV RESEPTORMODELLER I EKSISTERENDE UNDERSØKELSER .....	28
4 SUMMARY .....	30
5 REFERANSER .....	30
VEDLEGG 1 .....	35
VEDLEGG 2 .....	41
VEDLEGG 3 .....	49



## FORORD

Denne rapporten er utført av Norsk institutt for luftforskning etter oppdrag fra Statens forurensningstilsyn. Rapporten gir en beskrivelse av ulike typer reseptormodeller, krav til data-grunlaget og mulige anvendelser av disse modeller innenfor eksisterende og fremtidige luftforurensningsundersøkelser.



## SAMMENDRAG

Reseptormodeller er modeller som er basert på statistisk analyse av målte konsentrasjoner i et målepunkt og som resulterer i en identifisering og kvantifisering av kildenes bidrag til målepunktet.

Rapporten gir en oversikt over de mest aktuelle reseptormodellene og fordeler og ulemper ved disse. Vekten er lagt på de metodene som til nå har vært lite eller ikke benyttet ved vurdering av luftforurensninger i Norge. Dette er metoder som er viet stor oppmerksomhet i utlandet, særlig i USA. Metodene er blant andre:

- kjemisk massebalanse
- hovedkomponent analyse
- målrettet transformasjon av faktorresultater

Kjemisk massebalanse har til nå hovedsakelig blitt brukt for å analysere aerosoldata. Hovedkomponentanalysene kan brukes på alle typer data, og transformasjonsmetoden kan oppfattes som et bindeledd mellom de to andre metodene. I vedleggene er det gitt eksempler på bruken av modellene.

En god oppløsning av kildene krever bestemmelse av forholdsvis mange luftforurensningskomponenter. Aerosoler analyseres derfor ofte med metoder som f.eks. røntgenfluorescens (XRF), eller proton indusert røntgenemisjon (PIXE). Rene og ikke-reaktive filtermaterialer, helst teflon, må velges for prøvetaking av aerosoler. Ofte samles aerosoler i 2 fraksjoner med et skille omkring 2  $\mu\text{m}$  for å forhindre kjemiske reaksjoner på filteret, og for å få mer informasjon om forurensningene.



Modellene er aktuelle for flere prosjekter ved NILU. Modellene kan gi nyttig informasjon f.eks. i forbindelse med korrosjonsstudier, studier av langtransport av forurensninger og studier av aerosolforurensninger som resultat av utslipp fra flere kilder.

Metodene utnyttes best når en i aktuelle undersøkelser bruker flere i kombinasjon eller kombinerer dem med kildeorienterte spredningsmodeller. Reseptormodellene ses ikke som et alternativ til spredningsmodellene, men som et supplement til disse. Den kjemiske massebalansemodellen kan f.eks. identifisere og kvantifisere kildene for partikler og ikke reaktive gasser og derved forbedre eksisterende emisjonsoversikter som inngår i spredningsmodellene.

## RESEPTORMODELLER

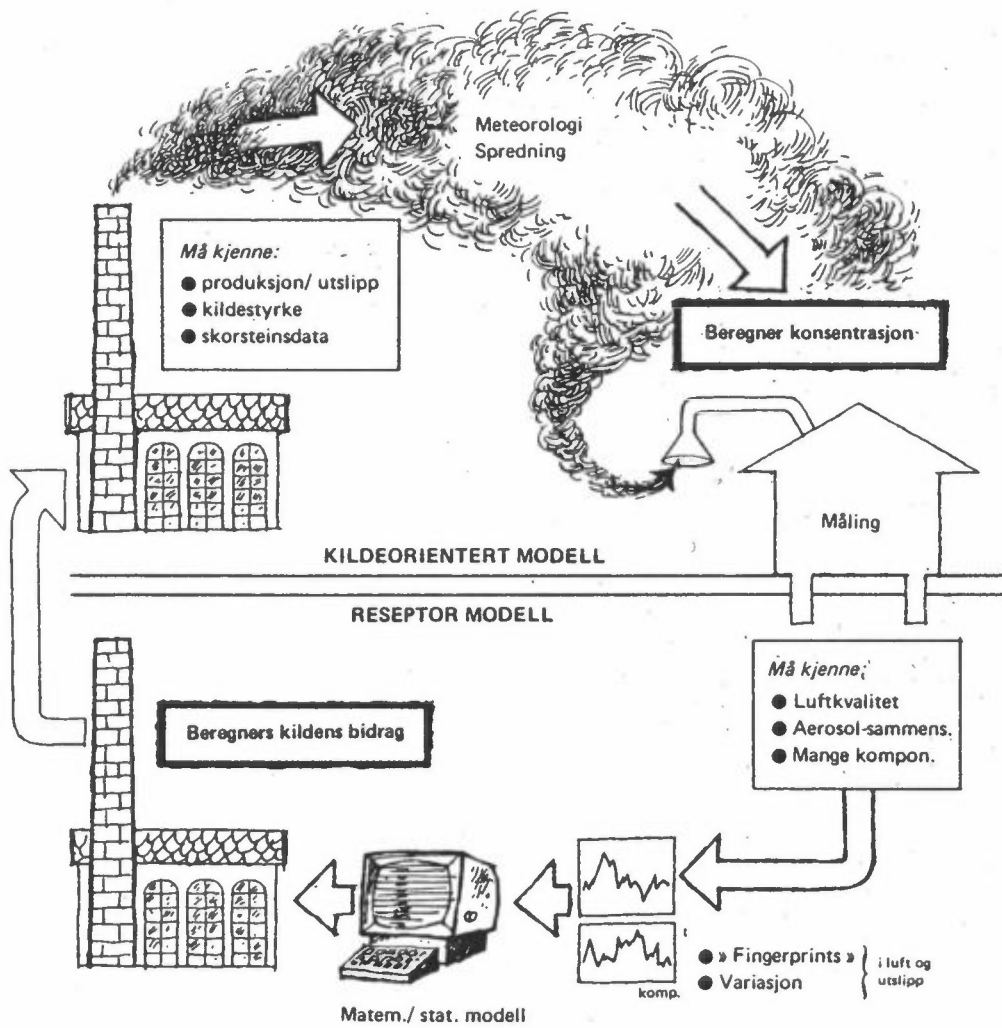
### 1. INNLEDNING

Spredningsmodeller har lenge vært det viktigste redskapet til vurdering av kildenes relative betydning for luftkvaliteten. Slike modeller krever ofte omfattende meteorologiske data og emisjonsoversikter for de aktuelle komponenter og områder. Enkelte modeller har med både reaktive og ikke-reaktive komponenter. Det nødvendige sett med kjemiske ligninger er derfor innebygget for å modellere transformasjonene mellom komponenter. En mengde undersøkelser av spredning og eksponering har vært utført med spredningsmodeller. For gassformige luftforurensninger har beregningene gitt gode resultater. For partikulære luftforurensninger har resultatet vært mindre tilfredsstillende. Dette kan skyldes flere forhold. Kildene er ofte ikke punktkilder og partikler med forskjellige størrelser spres og avsettes forskjellig. Kompliserte topografiske forhold kan gjøre spredningsberegningene vanskeligere å utføre.

Reseptormodeller er modeller som er basert på konsentrasjonsmålinger i et målepunkt (en reseptor). Beregningene resulterer i en identifisering av kildene og eventuelt en kvantifisering av kildenes bidrag til forurensningene i målepunktet. En av modellene beregner også nøyaktigheten i resultatet. I tillegg til emisjonsmålinger krever modellen et kjennskap til utslippenes sammensetning, men er ikke avhengig av omfattende utslippsoversikter. Topografien har ingen betydning for beregningene. Meteorologiske målinger er ofte nødvendige for å vurdere resultatene.

Reseptormodellene egner seg ikke til å (1):

- beregne virkningen av foreslåtte nye kilder
- beregne en forventet økning av belastningen i resipienten som følge av forandrede utslippsbetingelser (temperatur, mengde eller høyde).



Figur 1: Spredningsmodellen beregner luftkonsentrasjoner fra emisjonsoversikter. Reseptormodellen beregner kildebidragene fra luftkonsentrasjoner (2).

I tillegg til rene sprednings- og reseptormodeller har det vært laget modeller som prøver å kombinere det beste fra begge modelltypene til en blandet ("hybrid") modell. I USA ser US Environmental Protection Agency (EPA) dette som et meget lovende felt, og de vil satse på en videreutvikling på dette området (3). En kombinert bruk av kildeorienterte modeller og reseptormodeller vil forbedre beslutningsgrunnlaget idet de på en utfyllende måte kan kontrollere utgangsdata og konsentrasjonsfordelingene i rom og tid.

Reseptormodeller er behandlet i en rapport fra Environmental Research & Technology, Inc. (4), og her vil en finne en mengde referanser til arbeider på dette feltet. EPA sender ut en serie publikasjoner om reseptormodeller som er nyttig for den som vil sette seg bedre inn i dette feltet. Foreløpig er det utarbeidet en rapport som gir en oversikt over modellene (5), og en rapport spesielt om den kjemiske massebalanse-modellen (6).

## 2. BESKRIVELSE AV AKTUELLE MODELLER

Reseptormodellene er gjerne inndelt i tre hovedgrupper:

- kjemiske metoder
- fysiske metoder
- mikroskopiske metoder

De kjemiske metodene kan inndeles videre i undergrupper:

- kjemisk massebalanse-modell
- anrikningsfaktor-modell
- multivariabel metoder
  - faktoranalyse
  - hovedkomponent analyse ("principal component analysis")
  - målrettet transformasjon av faktorresultater ("target transformation factor analysis")

- regresjonsanalyse
- "cluster" analyse
- tidsserie analyse
- analyser av romlig datavariasjon

De fysiske og mikroskopiske metodene stilles gjerne opp sammen med de kjemiske i oversikter av denne typen. De to førstnevnte metodegruppene kan imidlertid betraktes som analytiske verktøy selv om enkelte av dem vil gi kvantitativ informasjon om kildepåvirkningen for en del kildekategorier. Radioisotopmetoden f.eks. er en fysisk metode som kan brukes til å skille forurensninger fra fossilt brennstoff og biomasse ved hjelp av karbonisotopen  $^{14}\text{C}$ . De kjemiske modellene egner seg best til å løse praktiske oppgaver i forbindelse med luftforurensninger, og de viktigste modellene vil bli behandlet nærmere. Enkelte modeller vil bli behandlet med eksempler i vedlegg.

Throgmorton og Axtell (7) har gitt en oversikt over krav til arbeidsmengde, erfaring, regnemaskin og datagrunnlaget for de forskjellige metodene og modellene. I tillegg til metoder som allerede er nevnt, finner en i tabell 1 fysiske og mikroskopiske metoder som f.eks. "scanning" elektron mikroskopi (SEM) og kvalitative estimater av virkningen av nærliggende punktkilder ("microinventory").

Tabell 1: Teknikker for beregning av kildebidrag  
(Throgmorton and Axtell (7)).

Capabilities (1) Application to								
Method	Technique	Coarse Mode	Fine Mode	Short Term Episodes	Resource Requirements (2)			
					Man-power	Skill	Com-puter	Data
1	Optical microscopy	SQ	-	x	4	5	1	2
2	SEM microscopy	SQ	SQ	x	4	5	1	2
3	Automated SEM-XRF	Qn	-	x	5	5	3	2
4	Enrichment Factor	SQ	SQ	x	1	2	2	3
5	Time Series	QL	QL		1	2	2	3
6	Spatial Series	QL	QL	x	1	2	2	3
7	Chemical Mass Balance	Qn	Qn	x	4	5	2	5
8	Multivariate methods	Qn	Qn		3	5	2	5
9	Dispersion modeling	Qn	Qn		2	4	3	5
10	X-ray diffraction	Qn	-	x	4	5	1	2
11	Microinventory	SQ	SQ		3	2	1	2
12	Radioisotope Analysis	Qn	Qn	x	5	5	1	2

(1) Source apportionment capabilities:

SQ = semi-quantitative

QL = qualitative

Qn = quantitative

(2) Judged on a scale of 1 to 5, 5 being the most resource intensive

I mars 1982 arrangerte EPA et "workshop" med tittelen "Status of Source Apportionment Methods: Quail Roost II" hvor eksperter fra de viktigste miljøene for slike studier i USA var samlet. De fikk seg forelagt et simulert datamateriale som var beregnet med spredningsmodeller, og et virkelig datasett samlet i Houston, Texas (8). Blant konklusjonene fra dette møtet var:

- Den kjemiske massebalanse-metoden og multivariabel metoder ga resultater som lå innenfor 10-15% fra de sanne verdiene i det simulerte datasettet.
- Anvendt på datasettet som besto av virkelige målinger, oppnådde deltakerne en overensstemmelse innenfor 20-50%.

Forskjellene i resultatene skyldes vesentlig forskjeller i kildematerialet, og viser betydningen av riktige utslipps-sammensetninger i de enkelte kildene.

## 2.1 Den kjemiske massebalanse metoden

Massebalansemodellen er det siste tilskudd til den samling beregnings-metoder som kalles reseptormodeller. Den er utviklet med tanke på partikulære luftforurensninger, men kan også brukes på gasser. Modellen kan kvantifisere den enkelte kildes bidrag til resipienten og beregne usikkerheten i bidraget. Begrensningene i bruken og/eller nøyaktigheten i resultatet er (9):

- usikkerhet i reseptormålingene (imisjonsmålingene)
- usikkerhet i kildesammensetningene
- systematiske feil i reseptormålingene og/eller kildesammensetningene
- kjemiske reaksjoner og avsetninger av enkelte komponenter etter utslippet
- forvrent bilde av kildenes relative betydning p.g.a. sterk avhengighet av vindstyrke og vindretning
- flere utslipp med lik sammensetning

De nødvendige kjente data er massekonsentrasjonen av hver komponent i målepunktet (enhet  $\mu\text{g}/\text{m}^3$ ). Denne størrelsen betegnes  $C_i$ . Videre må en kjenne den relative sammensetning av hver kilde som betegnes  $F_{ij}$ .  $F_{ij}$  er den brøkdelen komponent i utgjør av den totale massen i kilde  $j$  ved måling i utslippet. Målt i reseptoren betegnes den samme brøkdelen  $F'_{ij}$ . I de tilfellene hvor den relative sammensetningen av utslippet ikke forandres mellom kilde og målepunkt, er:

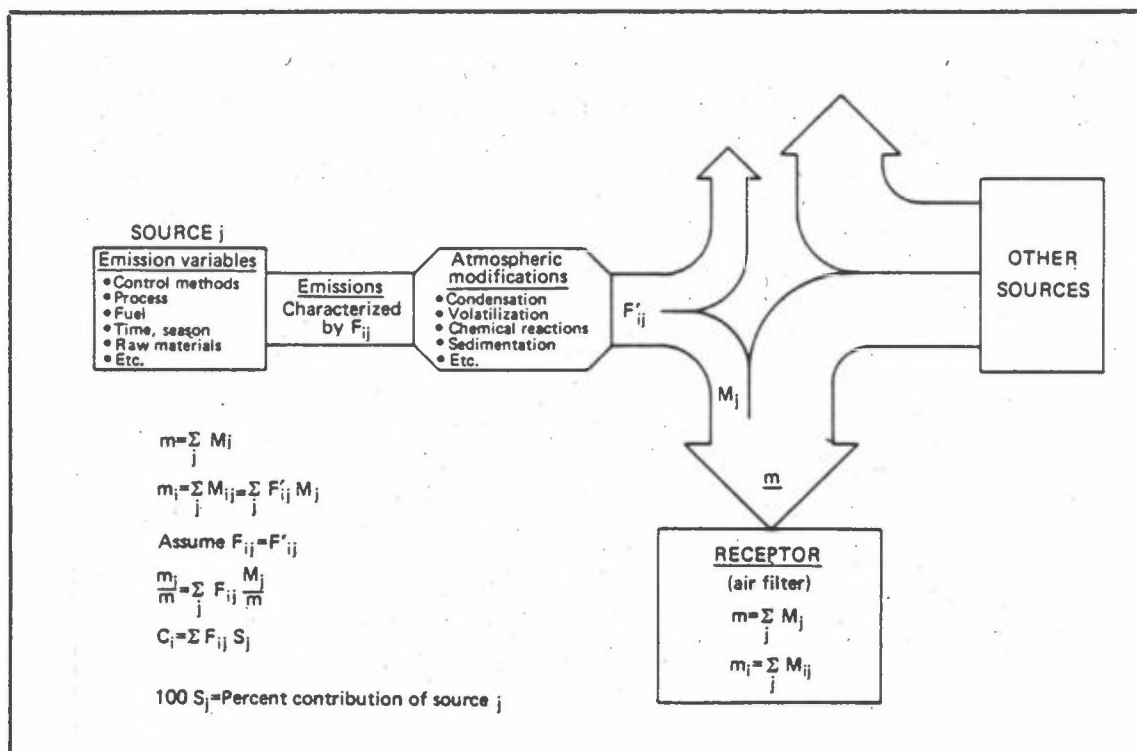
$$F'_{ij} = F_{ij}$$

Bidraget fra kilde  $j$  til resipienten, "kildebidraget", betegnes  $S_j$ . Enheten for kildebidragene kan være  $\mu\text{g}/\text{m}^3$ . Modellen som brukes til å bestemme kildebidragene er:

$$C_i = \sum_{j=1}^p F_{ij} S_j$$

Her er  $p$  antall kilder. Metoden kan ikke skille kilder med lik sammensetning, f.eks. veistøv og jordstøv. Kildene grupperes i slike tilfeller i kategorier med lik sammensetning, og  $p$  betegner i så fall antall kildekategorier.

Modellen er skjematisk fremstilt i figur 2. Her betegner  $M_j$  massen av forurensninger fra kilde  $j$ .  $M_{ij}$  er massen av komponent  $i$  fra kilde  $j$ .  $m$  er total masse og  $m_i$  massen av komponent  $i$  målt i reseptoren.



Figur 2: Skjematisk fremstilling av den kjemiske massebalansemodellen (Cooper, Watson (10)).

Ligningssystemet som modellen gir, kan løses når antall komponenter som bestemmes er lik eller større enn antall kilder. Beregningene kan gjøres på flere måter, men den vanligste metoden er å foreta en minste kvadraters tilpasning av kildebidragene.



Summerer en kildebidragene  $S_j$  vil en kunne få en differanse mellom summen og det en faktisk har målt i reseptorpunktet. Hvis summen er vesentlig større enn den målte mengden, må minst en kilde være overestimert. Er summen derimot vesentlig mindre enn den målte mengden forurensninger, vil dette være en indikasjon på at en eller flere viktige kilder mangler i beregningene.

De kildebidragene som finnes, representerer den kombinasjonen av kilder som passer best til de målte data. Det er derfor viktig å bestemme et stort antall komponenter både i utslippene og i reseptoren for å oppnå et godt resultat. Aerosoler samles ofte i to fraksjoner. Massebalansemetoden kan erfaringsmessig oppløse 6-8 kildekategorier.

Tidsvariasjoner i sammensetningen eller i utslippsmengden til en kilde kan gjøre at prøver samlet ved forskjellige tidspunkt gir forskjellige resultater. Det kan derfor være fornuftig å gjøre beregningene på ett sett av prøver og foreta en midling av resultatene til slutt. Erfaringsmessig vil en midling av de endelige resultatene gi et bedre resultat enn midling av datamaterialet med en påfølgende beregning med massebalansemodellen. Prøvetakingstiden som brukes er typisk 24 timer eller kortere.

Den kjemiske massebalanse-modellen kan omformes og utvides til også å bestemme usikkerheten i resultatet. Modellen er da mer komplisert.

Vedlegg 1 viser et eksempel på bruk av modellen.

En vesentlig mangel ved massebalanse-modellen er muligheten til å få med reaktive komponenter i beregningene. Dette er spesielt aktuelt for sulfat som dannes ved oksydasjon av svoveldioksyd, og som ofte utgjør en betydelig del av den totale aerosolmassen i en prøve. Det arbeides derfor med å få transformasjoner av kjemiske komponenter med i modellen.

Hvis en kilde har en komponent som er typisk for denne kilden, kan denne kildens bidrag lett beregnes fordi en ikke behøver å ta hensyn til andre kilder.

$$C_t = F_{tj} S_j$$

Her betegner  $t$  en komponent som er unik for kilde  $j$  og  $C_t$  er konsentrasjonen av denne komponenten i resipienten. Den brøkdelen komponent  $t$  utgjør av kilde  $j$ 's totale utslipp,  $F_{tj}$  må også være kjent. Kildebidraget  $S_j$  kan dermed beregnes fra ovenstående ligning.

## 2.2 Anrikningsfaktorer

Anrikningsfaktoren kan defineres slik:

$$E_{\text{aerosol}} = \frac{\left(\frac{C_i}{C_r}\right)_{\text{aerosol}}}{\left(\frac{C_i}{C_r}\right)_{\text{referanse}}}$$

Her skal "referanse" betegne en referansekilde og  $r$  et referanseelement. Som referansekilde velges ofte jordstøv fordi avblåsningen av slikt støv ofte gir et vesentlig bidrag til den totale mengde støv på et filter. Denne kilden omgir også ofte målestedet og er derfor uavhengig av vindretningen.

$C_i/C_r$  blir konsentrasjonsforholdet mellom element  $i$  og et referanseelement  $r$ . Eksempler på referanseelementer er aluminium, jern og silisium. Anrikningsforholdet blir derved forholdet mellom  $C_i/C_r$  i aerosolen og i f.eks. jordstøv. Hvis forholdet er nær 1, er jordstøv den vesentlige kilden til elementet  $i$ , men er forholdet større enn 1, bidrar en annen stor kilde til forurensningen av element  $i$  i resipienten. Metoden er enkelt å bruke, men brukt som over, vil den ikke kvantifisere kildebidraget. I de tilfeller der en kilde har en typisk komponent, kan metoden utvikles videre og vil gi et semikvantitativt estimat for denne kildens bidrag (4). Metoden blir da svært lik den kjemiske massebalanse metoden.

### 2.3 Multivariable metoder

Disse metodene er basert på tidsvariasjonen i et data-materiale. Foruten de kjemiske komponentene vil ofte fysiske variable som vindretning, temperatur osv. bli brukt som data i modellene. Viktige fordeler ved disse metodene er:

- de kan inkludere fysiske variable (vind, temperatur, etc).
- en behøver ikke kjenne antall kilder. (Dette blir ofte et resultat av beregningene.)
- de kan behandle gassformige komponenter og sekundære partikler.

Ulempen med metodene er at de er basert på tidsvariasjonen i datamaterialet og store utslipp med liten variasjon kan derfor bli maskert av svakere utslipp med stor tidsvariasjon.

#### 2.3.1 Faktoranalyse og hovedkomponent-analyse

Formålet med en klassisk faktoranalyse er å forklare variasjonen i en rekke målte variable ved hjelp av et lite antall virkende faktorer. Disse faktorene er av to typer, felles faktorer som påvirker flere variable, og unike faktorer som er spesifikke for en variabel. Antall faktorer kan ideelt sett bestemmes og er mindre enn antall målte variable. Alle faktorene er uavhengige av hverandre (ortogonale).

I det følgende betegner  $Z_{jk}$  observasjon nr. k av variabel nr j i standardisert form. På denne formen hvor  $z_{jk}$  er dimensjonsløs, har middelerdi 0 og varians 1, blir kovariansmatrisen og korrelasjonsmatrisen identiske. Ulempen med standardiserte variable er at de mister sin umiddelbart lesbare form.

M er antall felles faktorer.

$P_{ik}$  er beregnet verdi av felles faktor i for observasjon nr k av de variable.  $P_{ik}$  kalles for faktor "scores" på engelsk og inneholder tidsvariasjonen til de variable, når en regner på en tidsrekke. "Scores" brukes også på norsk.

$A_{ji}$  kalles faktor ladning ("factor loadings") og inneholder sammenhengen mellom faktorer og variable.

$U_{jk}$  er verdien av den unike faktor for variabel j og observasjon nr k.  $d_j$  som på engelsk kalles "uniqueness", er et uttrykk for den del av variansen felles-faktorene ikke kan forklare. Bidragene til  $d_j$  kommer fra variabel j's spesifikke egenskaper og fra usikkerheten i målingene av variabel j.

Faktormodellen er:

$$Z_{jk} = \sum_{i=1}^M A_{ji} P_{ik} + d_j U_{jk}$$

I en hovedkomponent-analyse ("principal component analysis") gjøres ingen antagelser om faktorene (=hovedkomponentene) bak datamaterialet. I stedet blir første faktor definert som den faktor som gir den beste lineære kombinasjon av de variable. Med best menes i denne sammenheng den lineære kombinasjonen som forklarer mer av variansen i datamaterialet enn en hvilken som helst annen liniær kombinasjon som kan lages. Når effekten av den første faktoren er fjernet fra datamaterialet, defineres annen faktor som den faktor som kan forklare mest av den resterende variansen. Faktorene funnet på denne måten blir også uavhengige, og en vil få like mange faktorer (hovedkomponenter) som variable.

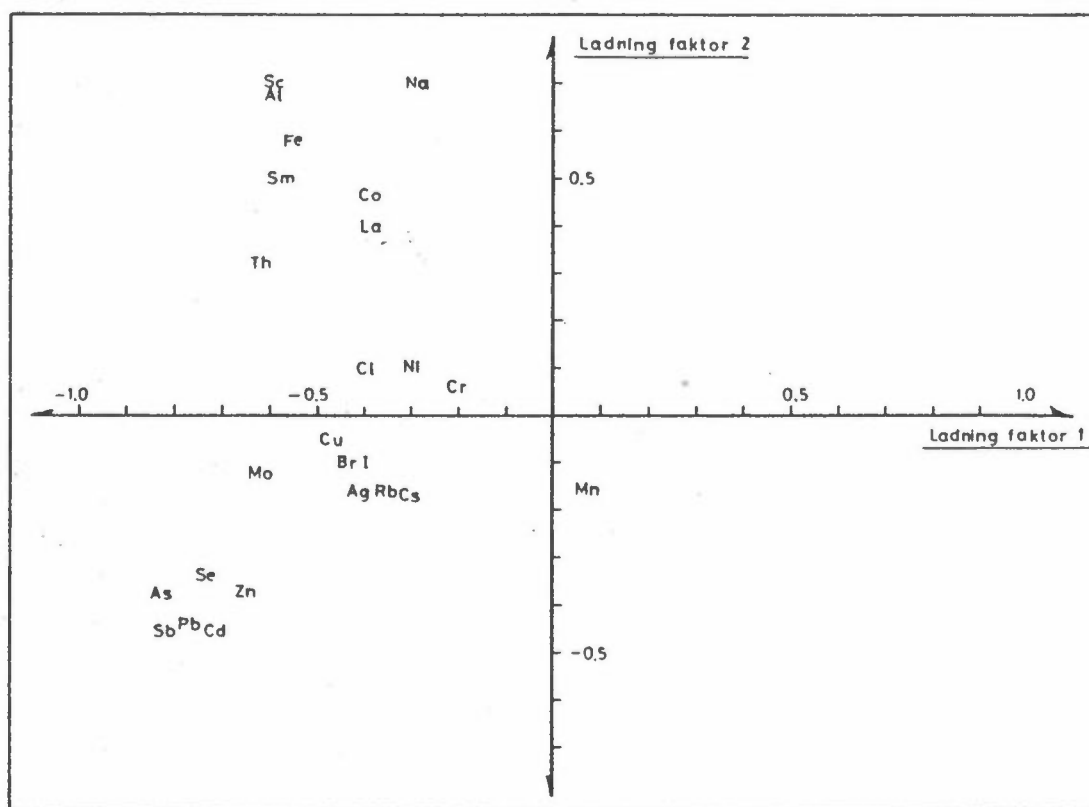
Modellen for en hovedkomponent-analyse blir:

$$Z_{jk} = \sum_{i=1}^N A_{ji} P_{ik}$$

$N$  er antall hovedkomponenter. Det kan vises at ladningene  $A_{ji}$  blir korrelasjonen mellom variabel  $j$  og faktor  $i$  når variable og faktorer er i standard tilstand.

Faktor ladningene er tidsuavhengige og vil fortelle hvilke faktorer som har betydning for de enkelte variable. Tidsvariasjonen i datamaterialet er samlet i faktorene  $P$  for tids-serier. Ved siden av en tabell over ladningene vil en tabell som forteller hvor stor del av den totale variansen hver faktor forklarer, være viktig. Når resultatet skal brukes på luftforurensningsdata, vil en ta hensyn til de faktorene som tilsammen forklarer 70-90% av variansen i materialet. Et eksempel er gitt i Vedlegg 2.

Sammenhengen mellom de variable og faktorene kan anskueliggjøres grafisk ved at faktorladningen (korrelasjonen mellom faktorer og variable) plote inn i to-dimensjonale ortogonale (rettvinklede) aksekors. Det kan lages en grafisk fremstilling for hvert par av faktorer. Figur 3 er et eksempel på en slik fremstilling.



Figur 3: Sporelementer i bladmose (*Hylocomium splendens*).  
 Faktorladninger for faktor 1 og 2.  
 (Rambæk, Semb, Schaug, Steinnes (11)).

Variable som skyldes samme faktor forekommer ofte som en gruppe i slike grafiske fremstillinger. For å lette den fysiske tolkningen av resultatene ser en gjerne at korrelasjonen mellom en gruppe variable og en enkelt faktor er så sterk som mulig. Samtidig bør denne gruppes korrelasjon med alle andre faktorer være minimal. I figur 3 kan korrelasjonen mellom gruppen Se, As, Zn, Sb, Pb og Cd og faktor 1 f.eks. forbedres ved at aksekorset roteres i papirplanet om origo,  $45^{\circ}$ . Korrelasjonen med faktor 1 blir nær  $-1$  og med faktor 2 nær  $0$ . Rotasjonen er et eksempel på en enkel transformasjon av faktoreresultatene. I praksis må en velge mer kompliserte metoder for

å ta hensyn til alle variable samtidig. Transformasjonsmetoder er beskrevet av f.eks. Malinowski og Howery (12) og Harman (13). Erfaringsmessig synes det ikke å være nødvendig å bruke transformasjonsmetoder på luftforurensningsdata.

For å forklare faktorene som kildekategorier må en ha kjennskap til de aktuelle kildenes sammensetning. Disse kilde-sammensetningene må sammenlignes med gruppene av variable som er godt korrelert med de enkelte faktorer.

Metodene som er omtalt her er basert på tidsvariasjonen i datamaterialet. Eventuelle store kilder med liten variasjon kan bli maskert av svakere kilder med stor variasjon i prøveperioden. Det er derfor viktig å foreta målinger over en tilstrekkelig lang tidsperiode. Resultatet vil gi en oversikt over kildetyper som forårsaker variasjonen i datagrunnlaget og ikke nødvendigvis betydningen av den enkelte kilde i resipienten. Metoden kan derfor med fordel brukes til å identifisere kildene, for deretter å bli etterfulgt av beregninger med den kjemiske massebalanse-metoden i de tilfeller hvor dette er mulig. Videre beregninger kan også gjøres med målrettet transformasjon av faktoranalyse.

### 2.3.2 Målrettet transformasjon av faktorresultater ("Target transformation factor analysis")

Som det fremgår av metodens navn, er dette en transformasjonsmetode som brukes på faktorresultatene, og hensikten er å fastslå hvilke utslipp som ligger bak faktorresultatene. For å bruke metoden må de aktuelle utslippenes kvantitative sammensetning være godt kjent.

Metoden kan ses på som et bindeledd mellom den kjemiske massebalanse-metoden og faktoranalyse (eventuelt hovedkomponentanalyse) og gir mulighet til å sammenligne potensielle utslipp, en etter en, med faktorresultatene. Beregningene gjennomføres ved å bruke faktorresultatene og et av utslippenes sammensetning til å beregne en trans-

formasjonsmatrise. Denne matrisen brukes igjen sammen med faktorresultatene til å reprodusere det aktuelle utslippets sammensetning. Hvis det utslippet en har valgt å prøve påvirker resipienten, dvs. passer til faktorresultatene, vil den reproduserte sammensetning være tilnærmet lik den opprinnelige.

Foruten å idenfisere utslippene som har betydning for resipienten, kan metoden forbedre kjennskapet til de enkelte utslipps sammensetning.

I de tilfellene hvor en kan identifisere kildene, og sammensetningen av disse er godt kjent, kan et også foreta en beregning som gir kildenes relative betydning for resipienten. Begrensningene i denne beregningen blir som for den kjemiske massebalanse-metoden, og er vesentlig avhengig av hvor godt en kjenner eller kan beregne kildesammensetningene.

En beskrivelse av metoden og et eksempel på bruken presenteres i Vedlegg 3.

### 2.3.3 Regresjonsanalyse

Denne metoden er en teknikk som brukes til å analysere sammenhengen mellom en avhengig variabel og et sett uavhengige variable. Metoden kalles multippel regressjonsanalyse, når minst to uavhengige variable er med i beregningene. Den avhengige variabel betegne ofte  $Y$ , og de uavhengige  $X_1, X_2, \dots, X_N$ . Modellen er med disse betegnelsene:

$$Y = A + \sum_{i=1}^N B_i X_i$$

Her er  $A$  og  $B_1, B_2, \dots, B_N$  konstanter.  $B_i$  kalles for regressjonskoeffisientene.  $N$  er antall uavhengige variable.



Modellen er lineær og forutsetter at den avhengige variable kan uttrykkes ved hjelp av en sum. Multippel regresjon kan imidlertid anvendes også i mer kompliserte sammenhenger ved å ta i bruk transformasjonsmetoder og andre hjelpemidler.

Metoden som er kvantitativ, er mye brukt i forbindelse med luftforurensninger. Den kan f.eks. brukes til å bestemme kildebidragene i tilfeller hvor kildenes sammensetning er kjent og kildene inneholder karakteristiske komponenter. Hvis  $X_i$  er kilde  $i$ 's karakteristiske komponent, representerer  $B_i X_i$  denne kildens bidrag til konsentrasjonen av komponent  $Y$  i målepunktet.  $A$  blir summen av bidragene fra alle kilder som ikke er spesifisert i modellen. Daisey f.eks. brukte modellen til å kvantifisere kildebidragene til partikulære PAH-forurensninger (14) i New York City. Bestemmelse av forskjellige materialers korrosjonshastighet uttrykt ved konsentrasjonene av luftforurensninger og klimavariabeler er et annet eksempel på bruken av denne modellen (15).

#### 2.3.4 "Cluster" analyse

Denne metoden er en samling metoder som brukes for å oppdage grupperinger i et datasett. Resultatet av en slik analyse er et sett med grupper av variable hvor hver variabel innen en gruppen har en naturlig assosiasjon til de andre elementene i samme gruppe. Det er derfor nødvendig å definere et kriterium for likhet. I vanlig "cluster" analyse vil de variable bare være medlem av en gruppe, men i "fuzzy cluster" analyse vil en variabel kunne være medlem av flere grupper. "Cluster" analyser er behandlet f.eks. av Anderberg (12). I figur 3 utgjør elementene Se, As, Zn, Sb, Pb og Cd f.eks. en gruppe eller et "cluster". Thrane (17) har brukt "cluster" analyse og regresjonsanalyse til å estimere PAH-bidraget fra aluminiumverk i prosent av total mengde PAH en rekke målesteder i Sundsvall.

## 2.4 Datainnsamlingen

Ved bruk av reseptormodeller må prosjekter planlegges fra starten med dette for øyet, fordi disse teknikkene stiller spesielle krav til prøvetakingen og valg av komponenter. Vanligvis må mange komponenter bestemmes for å kunne gi en god karakteristikk av alle kildene, og valget av analysemetode faller derfor på de metoder som kan gi mange komponenter samtidig. Røntgenfluorescens er f.eks. en mye brukt metode for analyse av aerosolkonsentrasjoner.

Lokaliseringen av prøvetakingsstedene er bestemt av formålet med undesøkelsen. Foreliggende meteorologiske data og spredningsberegninger er nyttige i denne forbindelse. Det er fornuftig å velge enkelte prøvetakingssteder slik at de blir utsatt for de høyeste konsentrasjonene fra de viktigste kildene. Spredningsmodellene kan beregne beliggenheten av slike områder. Hvis en kan finne meteorologiske forhold hvor et slik prøvetakingssted kan anses påvirket av bare en kilde, kan en ved målinger få et bedre kjennskap til forurensningene fra denne kilden, og finne eventuelle karakteristiske komponenter.

Det er all grunn til å skaffe til veie mest mulig informasjon om målestedene og kildene på forhånd, og legge vekt på en god planlegging. Det resultatet en oppnår blir ikke under noen omstendighet bedre enn de data en legger inn i modellen. I det følgende vil en se litt nærmere på prøvetaking, kildekarakterisering og analytiske metoder.

## 2.5 Prøvetaking

Hvis en skal bruke den kjemiske massebalanse-modellen på aerosoler, vil det være en fordel å fraksjonere luftens aerosoler i en grov og en fin fraksjon. Til dette brukes ofte en to-trinns prøvetaker som skiller mellom aerosoler større og mindre enn 2  $\mu\text{m}$ . Også for multivariable metoder vil det være en fordel å bruke to fraksjoner for støvet for å samle mer informasjon. Foruten å gi mer informasjon, vil en fraksjo-

nering av støvprøver forhindre kjemiske reaksjoner mellom den basiske grov-fraksjonen og den sure fin-fraksjonen. Metodene kan imidlertid også brukes på data samlet i en fraksjon, på gasskonsentrasjoner, på meteorologiske data m.m.

Det foretrukne filtermaterialet for massebalanse-metoden og for faktoranalyse er teflon, fordi dette filtermaterialet er ikke-reaktivt og rent. Også micropore og nucleopore filter kan anbefales. Kvartsfiltere er urene, gir systematiske feil i målingene og bør derfor ikke brukes.

Prøvetakingstiden er som oftest 24 timer. For å studere belastningen under spesielle vindforhold, kan det være aktuelt å bruke kortere prøvetakingstid.

## 2.6 Kildesammensetningen

Det er i utgangspunktet fire forskjellige måter å skaffe tilveie dette datamaterialet. Disse er:

- målinger i utslippet fra de aktuelle kilder
- målinger i skyen på nedvindsiden av et aktuelt utslipp. Målingene kan kontrolleres ved å slippe et sporstoff (f.eks.  $SF_6$ ) ut i kilden
- fra litteraturen
- fra bibliotek opprettet med reseptormodellering for øyet

Målinger i utslippet vil gi sammensetningen, men denne er ikke nødvendigvis representativ for nedfallet i resipienten. Hvis målingene utføres i piper med høy temperatur kan flyktige elementer kondensere og store partikler være sedimentert før forurensningene når fram til målestedet. Såvel målinger fra fly som fra ballonger har vært brukt og bør vurderes brukt for å utføre målinger i røykfanen et stykke fra kilden.

Litteraturen er en mulig kilde for datainnsamlingen om typiske kildekategorier. I USA har EPA lagt vekt på å samle til-

gjengelige data og er i ferd med å bygge opp et bibliotek av tilgjengelige kildesammensetninger.

## 2.7 Komponenter og analytiske metoder

Valget av komponenter som skal analyseres, vil blant annet være bestemt av formålet med undersøkelsen. Såvel aerosoler, gasser og nedbørkomponenter kan være aktuelle, og en rekke analysemetoder kan derfor komme i bruk. Atomabsorpsjon og automatiserte spektrofotometriske metoder er eksempler på hyppig brukte analysemetoder for uorganiske luftforurensninger. For å få analysert mange komponenter i en analyse med relativt lave omkostninger blir røntgenfluorensens (XRF) mye brukt. Likeledes vil nøytronaktiveringsanalyse og proton induisert røntgenemisjon (PIXE) være aktuelle analysemetoder. Organiske luftforurensninger analyseres oftest med gasskromatografiske og massespektrometriske metoder. Den kjemiske massebalanse modellen krever imidlertid et bredt spektrum av komponenter og en vil kommentere kravene til aerosolundersøkelser med denne metoden litt nærmere. Det er viktig å bestemme de betydeligste forurensningskomponentene og komponenter som er spesifikke for enkelt kilder. Kildespesifikke elementer er f.eks. bly og bromid for biltrafikk, nikkel og vanadium for oljeforbrenning og natrium som marin komponent. Generelt må kildespesifikke elementer ses i sammenheng med industri og annen virksomhet i området hvor undersøkelsen pågår. Andre eksempler på elementer som kan være kildespesifikke er kalsium for kalkstein, silisium og aluminium for veistøv, selen og antimon for kullforbrenning. EPA (6) har laget en liste over komponentene som bør bestemmes i forbindelse med slike undersøkelser. Denne er gjengitt i tabell 2. Her er de viktigste komponentene understreket. Undersøkelser har imidlertid også vært gjort med godt resultat med færre komponenter. Hvis en vil forklare den totale støvmengden i resipienten, bør prøvene være samlet i to fraksjoner.

Tabell 2: Aerosoler. Anbefalt datagrunnlag for kjemisk massebalanse og andre reseptormodeller.

Total masse	Cl	Fe	Pb
<u>Na</u>	<u>K</u>	Ni	SO <sub>4</sub>
<u>Mg</u>	<u>Ca</u>	Cu	NO <sub>3</sub> <sup>4</sup>
<u>Al</u>	V	Se	Org.C
Si	Cr	Br	Elem.C
<u>S</u>	Mn	Ba	
As	Sr	Cd	
Sn	Sb		

### 3. BRUK AV RESEPTORMODELLER I EKSISTERENDE UNDERSØKELSER

Datamaterialet samlet på NILU består av kjemiske og meteorologiske variable fordelt over tid og rom. En rekke reseptormodeller eller -metoder som er blitt nevnt, såsom regresjonsanalyser, korrelasjonsanalyser og analyser av romlig variasjon i materialet, har vært brukt rutinemessig gjennom mange år. Faktoranalyser, hovedkomponent analyser og den kjemiske massebalansemodellen har til nå hovedsakelig vært brukt eksperimentelt for å utprøve metodikken ved NILU. Metodene er meget gode og vil trolig bli brukt i stadig større utstrekning. Enkelte modeller vil utvilsomt bli mer anvendelige i fremtiden på grunn av forbedringer i teorien.

Korrosjon og erosjon av forskjellige materialer påkoster samfunnet store beløp og uerstattelige gjenstander forringes på grunn av luftforurensninger. Et av hovedmålene med korrosjonsforskningen er å finne de sammenhenger som finnes mellom miljøvariable og korrosjonshastighet. Regresjonsanalyser har vært brukt i stor utstrekning til dette formål, og i tillegg vil det være naturlig å prøve å etablere sammenhengene i datamaterialet med hovedkomponent analyse eventuelt faktoranalyse.

Det foretas kontinuerlig måling av luftkvaliteten i tettbygde strøk. Slike undersøkelser kan ha til hensikt å fastslå nivået av forurensninger, de kan være ledd i kontrolltiltak for å vurdere påvirkningen fra spesifikke stasjonære kilder eller de kan settes i gang for å vurdere spredte kilder. Samles et

datamateriale med tilstrekkelig antall komponenter, kan hovedkomponentanalyser gi nyttig informasjon om kildene. I forbindelse med aerosolmålinger er massebalanse modellen vel egnet til å fastslå kildenes relative betydning. Mindre undersøkelser som analyseres for et stort antall komponenter, tilgjengelige meteorologiske data og kjennskapet til de forventete kilder, kan bestemme modellvalget og analyseprogrammet for større undersøkelser.

Tilpasning og bruk av reseptor-orienterte modeller vil bli forsøkt gjennomført i noen basisundersøkelser i det statlige program for forurensningsovervåkning. I forslaget til plan for basisundersøkelse i Mo i Rana inngår således såvel reseptor-orienterte modeller som kilde-orienterte spredningsmodeller i dataanalysen samt en kombinasjon av disse.

Vedfyringens bidrag til luftforurensninger i Elverum er til vurdering, og som sporstoff for vedfyringen har en  $^{14}\text{C}$ . En hovedkomponent analyse kan muligens få frem sammenhengen mellom temperatur,  $^{14}\text{C}$  og andre komponenter. Datamaterialet er idag for lite til å få frem en meningsfylt sammenheng.

Nordisk Ministerråd finansierer en rekke prosjekter for å kvantifisere miljøeffekter ved energiproduksjon (MIL-prosjektene). Reseptormodeller vil bli brukt i MIL-4prosjektet for å finne bidragene fra ulike kilder (17).

I Europa er det etablert et samarbeid om målinger av luftforurensninger ("European Monitoring and Evaluation Programme") med et måleprogram som omfatter svovelforbindelser og i mindre grad nitrogenkomponenter. I EMEP inngår syv bakgrunnsstasjoner i det statlige program for overvåkning av langtransportert forurenset luft og nedbør. Målestedene ligger i områder med liten eller ingen påvirkning fra lokale kilder. Her er det igang en undersøkelse med hovedkomponentanalyse for å kartlegge sammenhengen mellom samtlige kjemiske variable, beregnede konsentrasjoner, meteorologiske variable og kildekategorier. Videre vil en prøve å sammenligne et stort antall målesteder i inn- og utland for samme kjemiske komponent med

hovedkomponentanalyse for bedre å lære å kjenne dette data-materialet.

#### 4. SUMMARY

This report describes the most important receptor models based on literature studies. The chemical mass balance model, principal component analysis, and target transformation factor analysis have been emphasized as methods useful for several ongoing air pollution studies in Norway.

The requirements for the collection of data in the receptor and sources are discussed, and the annexes briefly give calculations or results from three selected problems.

The receptor models should be used in combination with each other in order to utilize the particular strength of each model.

The models do not seem to give an alternative to the source oriented dispersion models. Improved results may be obtained by combination of the two model categories to hybrid models, or by application of both model categories to the same problem.

#### 5 REFERANSER

- |     |   |  |
|-----|---|--|
| (1) | Core, J.E.<br>Cooper, J.A.<br>Hanrahan, P.L.<br>Cox, W.M. | Particulate dispersion model evaluation: A new approach using receptor models.<br><u>J. Air Poll. Contr. Ass.</u> , <u>32</u> ,<br>1142-1147 (1982). |
| (2) | NILU  | Årsberetning 1982. Bearbeidet fra referanse 5.   |

- (3) Pace, T.G. The role of receptor models for revised particulate matter standard.  
I: Speciality conference on receptor models applied to contemporary pollution problems, Danvers, Mass., 1982. Proceedings. Pittsb. PA., Air Pollution Control Ass., 1982. P. 18-28.
- (4) Environmental Research & Technology, Inc., The state of the art of receptor models relating ambient suspended particulate matter to sources. Los Angeles, March 1981. Document No. P-A422. Prepared for U.S. Environmental Protection Agency, Contract No. 68-02-2542 Task 8.
- (5) Receptor model technical series. Volume I. Overview of receptor model application to particulate source apportionment. Research Triangle Park, North Carolina, Environmental Protection Agency, 1981. (EPA-450/4-81-016a).
- (6) Receptor model technical series. Volume II. Chemical mass balance. Research Triangle Park, North Carolina, Environmental Protection Agency, 1981. (EPA-450/4-81-016b).
- (7) Throgmorton, J.A. Axtell, K. Digest of ambient particulate analysis and assessment methods. Research Triangle Park, North Carolina, Environmental Protection Agency. (EPA-450/3-78-113). P. 108-114.
- (8) Stevens, R.K. Pace, T.G. Status of source apportionment techniques - Quail Roost II.  
I: Speciality conference on receptor models applied to contemporary pollution problems. Danvers, Mass., 1982. Proceedings. Pittsb., PA., Air Pollution Control Ass., 1982. P. 46-59.



- (9) Watson, J.G. Overview of receptor model principles. I: Speciality conference on receptor models applied to contemporary pollution problems. Danvers, Mass., 1982. Proceedings. Pittsb., PA., Air Pollution Control Ass., 1982. P.6-17.
- (10) Cooper, J.A.  
Watson, J.B. Receptor oriented methods of air particulate source apportionment. J.Air Poll. Contr. Ass., 32, 1116-1125 (1980).
- (11) Rambæk, J.P.  
Semb, A.  
Schaug, J.  
Steinnes, E. Unpublished data.
- (12) Malinowski, E.R.  
Howery, D.G. Factor analysis in chemistry. New York, John Wiley & Sons, 1980.
- (13) Harman, H.H. Modern factor analysis. Chicago, The University of Chicago Press, 1964.
- (14) Daisey, J.M. Receptor source apportionment models for two polycyclic aromatic hydrocarbons. I: Speciality conference on receptor models applied to contemporary pollution problems. Danvers, Mass., 1982. Proceedings. Pittsb., PA., Air Pollution Control Ass., 1982. P. 348-357.
- (15) Haagenrud, S. NORDFORSK. Atmosfærisk korrosjon med spesielt hensyn til forholdene på bygninger. Delprosjekt 1.3: Bestemmelse av korrosjonen under korte forurensningsepisoder utendørs. Lillestrøm 1977. (NILU TN 14/77.)
- (16) Anderberg, M.R. Cluster analysis for applications. New York, Academic Press, 1973.
- (17) Thrane, K.E. Polycyclic aromatic hydrocarbons in ambient air in Sundsvall, Sweden. Lillestrøm 1982. (NILU OR 40/82.)

- (18) Sivertsen, B. MIL4. Relative bidrag av forurensninger til menneske og miljø fra forskjellige kilder. Statusrapport mars 1983. Lillestrøm 1983. (NILU OR 16/83.)
- (19) Henry, R. Receptor models of airborne particulate sources. I: Seminar NILU, 16 October 1981.
- (20) Henry, R. Stability analysis of receptor models that use least squares fitting.  
I: Specialty conference on receptor models applied to contemporary pollution problems. Danvers, Mass., 1982.  
Proceedings. Pittsb., PA., Air Pollution Control Ass., 1982. P. 141-157.
- (21) Henry R. Mathematical foundations of principal component and factor analysis. I: Seminar NILU, 2 October 1981.
- (22) Lawson, C.L.  
Hanson, R.L. Solving least squares problems. Englewood Cliffs, N.J., Prentice-Hall Inc., 1974.
- (23) Henry, R. Principal component analysis of PAH data from Sundsvall. Lillestrøm, 1982. (NILU OR 32/1982.)



Vedlegg 1

***DEN KJEMISKE MASSEBALANSE MODELLEN***



## Vedlegg 1

**DEN KJEMISKE MASSEBALANSE MODELLEN**

Fremstillingen og eksemplet som er brukt i det følgende er basert på en seminarserie Ron Henry holdt på NILU vinteren 1981-1982 (19).

Den kjemiske massebalanse modellen har til nå i hovedsak vært anvendt på partikulære forurensninger. Modellen krever en prøve av luftkvaliteten i reseptoren som er analysert på en lang rekke komponenter, samt et godt kjennskap til utslippskildenes sammensetning. Som et enkelt eksempel brukes en støvprøve på 100  $\mu\text{g}$  støv pr  $\text{m}^3$  samlet i et målepunkt og analysert med følgende resultat:

Si	14.75 $\mu\text{g}/\text{m}^3$
Fe	5.75 "
Mn	0.83 "
Pb	1.65 "

Denne tallkolonnen er konsentrasjonene  $C$  av komponentene, f.eks. er  $C_3$  konsentrasjonen av Mn ( $0.83 \mu\text{g}/\text{m}^3$ ).

Likeledes må kildesammensetningen være kjent. Vi skal bruke følgende:

	Jordstøv	Stålverk	Bilavgasser
Si	0.28	0.05	0.0
Fe	0.05	0.20	0.05
Mn	0.001	0.05	0.01
Pb	0.0	0.01	0.30

Alle tall er her angitt som brøkdelen av total mengde støv fra den respektive kilde (dvs. multiplisert med 100 fås prosent).

Kildesammensetningen kalles  $F_{ij}$ .  $F_{32}$  f.eks. er brøkdelen Mn i det totale utslippet fra stålverk (0.05). Det forutsettes nå

at sammensetningen i de respektive kildene ikke forandres under transporten frem til målepunkt. Kildesammensetningene målt i reseptoren kalles  $F'_{ij}$ .  $F'_{ij}$  blir etter forutsetningen over lik  $F_{ij}$ , og kildebidragene  $S_j$  fra de respektive kildene kan beregnes.

For hver komponent  $i$  ( $i = Si, Fe, Mn, Pb$ )

$$C_i = \sum_{j=1}^3 F'_{ij} S_j$$

eller:

konsentrasjonen av silisium:

$$0.28 S_1 + 0.05 S_2 + 0.0 S_3 = 14.75 \mu\text{g}/\text{m}^3$$

konsentrasjonen av jern:

$$0.05 S_1 + 0.20 S_2 + 0.05 S_3 = 5.75 \mu\text{g}/\text{m}^3$$

konsentrasjonen av mangan:

$$0.001 S_1 + 0.05 S_2 + 0.01 S_3 = 0.85 \mu\text{g}/\text{m}^3$$

konsentrasjonen av bly:

$$0.0 S_1 + 0.01 S_2 + 0.30 S_3 = 1.65 \mu\text{g}/\text{m}^3$$

Kildebidragene,  $S_j$ , er da definert slik:

$S_1$  : total mengde jordstøv målt i reseptoren

$S_2$  : total mengde støv fra stålverk målt i reseptoren

$S_3$  : total mengde støv fra bilavgasser målt i reseptoren

I det generelle tilfellet kan ligningssystemet løses når det er bestemt minst like mange komponenter som kilder. Resultatet av beregningene blir kildebidragene. I vårt tilfelle har vi fire ligninger til å finne tre ukjente og vi benytter oss av minste kvadraters metode for å beregne bidragene.

Beregninger utført på vårt eksempel gir følgende kildebidrag:

Jordstøv, $S_1$	:	50 $\mu\text{g}$ støv/ $\text{m}^3$
Stålverk, $S_2$	:	15 " "
Bilavgasser, $S_3$	:	5 " "

Disse tre kildene gir tilsammen  $70 \mu\text{g st\ddot{o}v/m}^3$  i m\dd{a}lepunktet og vi m\dd{a}lte  $100 \mu\text{g/m}^3$ . Dette m\dd{a} bety at flere kilder enn v\dd{a}re tre bidrar til forurensningen i m\dd{a}lepunktet.

F\dd{o}lgende forhold vil generelt komplisere beregningene:

- 1) Usikkerheten i reseptorm\dd{a}linger
- 2) Usikkerhet i kildesammensetningen
- 3) Systematisk feil i reseptor og/eller kildesammensetning
- 4) Kjemiske reaksjoner og avsetninger av enkelte komponenter etter utslippet
- 5) Forvrenget bilde av kildenes relative betydning pga. sterk avhengighet av vindstyrke og vindretning
- 6) Kilder med lignende sammensetninger

Usikkerhetene i m\dd{a}lingene og systematiske feil er innarbeidet i en modifisert utgave av modellen, og det p\dd{a}g\dd{a}r arbeid for \dd{a} f\dd{a} med kjemiske reaksjoner i modellen. For \dd{a} f\dd{a} et riktig bilde av kildenes relative betydning b\dd{o}r en ta flere pr\dd{o}ver og foreta en midling av de endelige resultatene. Hvis en har kilder med n\dd{a}r samme kjemiske sammensetning, resulterer dette i en ustabilitet i beregningene som kan gi urealistiske kildebidrag. Massebalansemodellen gj\dd{o}r bruk av egenverdiberegninger, og \dd{a}rsaken til ustabiliteten er sm\dd{a} egenverdier. Disse problemene er beskrevet av Henry (20).

I eksemplet blir resultatet n\dd{a}r en tar med 10% usikkerhet i kilde- og reseptorm\dd{a}lingene, og ser bort fra systematiske feil, som f\dd{o}lger:

Jordst\dd{o}v, $S_1$	:	$50 \pm 5.5 \mu\text{g/m}^3$	(11%)
St\dd{a}lverk, $S_2$	:	$15 \pm 3.1 \text{ "}$	(20%)
Bilavgasser, $S_3$	:	$5 \pm 0.55 \text{ "}$	(11%)





Vedlegg 2

***HOVEDKOMPONENTANALYSE***



## Vedlegg 2

**HOVEDKOMPONENTANALYSE**

I det følgende vil en se litt nærmere på denne metoden illustrert med et enkelt eksempel. Datamaterialet er hentet fra bakgrunnsstasjonene Narbuvoll og Hummelfjell i Hedmark fylke. Dataene fra Narbuvoll er nedbørdata, og på Hummelfjell er det målt svoveldioksyd og sulfat i partikler. Alle data er døgnobservasjoner fra 1980. I tillegg til de variable som er tatt med nå, vil det være interessant å få med meteorologiske variable og konsentrasjoner beregnet med modeller. Dette vil trolig bli gjort i løpet av 1983. Fremstillingen som følger er stort sett basert på en seminarserie av Henry (21).

Utgangspunktet for beregningene er datamatriksen

$$C = \begin{pmatrix} C_{11} & C_{12} & \dots & C_{1N} \\ C_{21} & C_{22} & \dots & C_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ C_{M1} & C_{M2} & \dots & C_{MN} \end{pmatrix} \begin{array}{l} \text{variabel nr.} \\ \downarrow \\ \text{tid (prøvenr.)} \\ \rightarrow \end{array}$$

Her er  $C_{ji}$  den målte verdi (konsentrasjonen) av variabel  $j$  i prøvenummer  $i$ . Det er  $M$  variable, og i vårt eksempel er  $M=15$ . Antall prøver ( $N$ ) er 104. Datasettet må være fullstendig, slik at det ikke forekommer manglende enkeltobservasjoner i datamatriksen.

Middelverdiene og variansen beregnes, og alle enkeltobservasjoner overføres til standardform. Den standardiserte datamatrix betegnes  $Z$ .

Matriseproduktet  $Z^T Z$  vil bli en matrise hvor hvert element blir korrelasjonskoeffisient mellom to variable  $j$  og  $k$ . Matrisen er symmetrisk fordi korrelasjonen mellom variablene  $j$  og  $k$  må være lik korrelasjonen mellom  $k$  og  $j$ . Videre må elementene langs hoveddiagonalene være lik 1. Korrelasjons-

matrisen blir identisk med kovariansen for standardiserte variable.

Enhver matrise kan skrives som et produkt av 3 matriser. Denne oppsplittingen som på engelsk kalles "singular value decomposition", er beskrevet f.eks. av Lawson og Hanson (22). Datamatriksen  $Z$  kan derfor skrives som et produkt av to matriser hvis to av matrisene fra oppsplittingen slås sammen til en matrise:

$$Z = A P$$

Dette er modellen for hovedkomponent analyse på matriseform.  $P$  kalles faktor (eller komponent) "scores" på engelsk og inneholder tidsvariasjonen av faktorene. "Scores" brukes også på norsk.  $A$  kalles faktor (eller komponent) ladning, og inneholder sammenhengen mellom variable og faktorer. Det kan vises at ladningsmatrisen er korrelasjonen mellom faktorer og variable i standardtilstand.

For å finne "scores" og ladning benytter en seg av kovariansmatrisen  $Z^T Z$ . Beregningene gjennomføres som en egenverdianalyse hvor egenverdiene til kovariansmatrisen bestemmes. Disse egenverdiene er et direkte uttrykk for variasjonen i datamaterialet.

Tabell 2.1: Middelerverdi, standard avvik og varians for 104 fullstendige datasett fra Narbuvoll og Hummelfjell 1980. Symbolene er definert etter tabellen.

VARIABEL	MIDDEL	STD.AVVIK	VARIANSE
WARA	4.21	4.85	23.5
MM	5.37	4.68	21.9
SO <sub>4</sub>	0.498	0.534	0.285
SO <sub>4</sub> C	0.460	0.526	0.277
NH <sub>4</sub>	0.182	0.287	0.0823
NO <sub>3</sub>	0.187	0.204	0.0414
NA <sub>3</sub>	0.370	0.507	0.257
MG	0.0550	0.0608	0.00370
CL	0.500	0.834	0.695
CA	0.278	0.184	0.0339
H	300.6	2892.	8365000.
K	0.164	0.169	0.0287
KOND	12.39	10.39	196.1
SUPA	0.396	0.538	0.290
SO <sub>2</sub>	1.00	0.765	0.586

De variable er:

WARA: utvaskingsforholdet:  $SO_4$  (nedbør)/ $SO_4$  (partikler)

MM : nedbørmengde

SO<sub>4</sub> : sulfatkonsentrasjoner i nedbør, mg S/l

SO<sub>4</sub>C: sulfatkonsentrasjoner i nedbør, korrigert for sjøsalt mg S/l

NH : ammoniumkonsentrasjonen i nedbør, mg N/l

NO : nitratkonsentrasjonen i nedbør, mg N/l

Na : natriumkonsentrasjonen i nedbør, mg/l

Mg : magnesiumkonsentrasjonen i nedbør, mg/l

Cl : kloridkonsentrasjonen i nedbør, mg/l

Ca : kalsiumkonsentrasjonen i nedbør, mg/l

H : hydroniumkonsentrasjonen i nedbør,  $\mu$ e/l (omregnet fra pH)

K : kaliumkonsentrasjonen i nedbør, mg/l

Cond: ledningsevnen i nedbør,  $\mu$ S/cm

SUPA: sulfatkonsentrasjonen i partikler,  $\mu$ g S/m<sup>3</sup>

SO<sub>2</sub> : svoveldioksydkonsentrasjonen i luften,  $\mu$ g S/m<sup>3</sup>

Den første variabelen er avledet av to andre. Variansen og middelkonsentrasjonen for hydroniumkonsentrasjonen er urealistisk høy. Det må derfor være minst en gal verdi i datamaterialet.

Variansen i datamaterialet som skyldes de enkelte faktorene er gitt ved egenverdiene til  $Z^T Z$ .

Tabell 2.2: Egenverdianalyse for 14 faktorer. Tabellen viser for hver faktor egenverdi, prosent forklart varians og kumulativ forklart varians.

	E	%VAR	K. %VAR
1	5.9728	42.66	42.66
2	3.0161	21.54	64.21
3	1.1762	8.40	72.61
4	1.0559	7.54	80.15
5	.9709	6.94	87.09
6	.7563	5.40	92.49
7	.3817	2.73	95.21
8	.2976	2.13	97.34
9	.1276	.81	98.25
10	.1061	.76	99.01
11	.0785	.56	99.01
12	.0401	.29	99.86
13	.0200	.14	100.00
14	.0000	.00	100.00

Fem faktorer forklarer over 87% av variasjonen i datamaterialet (tabell 2.2). For å identifisere kildene, ser en på faktorladningene i tabell 2.3).

Tabell 2.3: Faktor ladning for 14 faktorer.

MM	-.33	.37	.34	-.03	-.42	.66	.07	.13	.01	-.01	-.01	-.01	.00	.00
SO <sub>4</sub>	.93	.20	.26	-.02	-.07	-.07	-.06	-.05	.07	-.09	.03	.01	-.05	.00
SO <sub>4</sub> C	.91	.27	.26	-.01	-.06	-.08	-.07	-.05	.06	-.10	.04	.00	-.05	.00
NH <sub>4</sub>	.91	.17	.18	-.08	-.04	-.05	.14	-.01	-.24	-.10	-.05	.04	.05	.00
NO <sub>3</sub>	.87	.19	-.27	.03	.24	.05	-.14	-.26	-.10	.16	.05	.05	-.04	.00
NA	.22	-.91	.05	.16	-.21	.03	-.01	-.06	-.06	-.01	.21	-.03	-.03	.00
MG	.41	-.86	.04	-.09	-.14	.10	.07	-.03	.12	.03	-.05	.14	.01	.00
CL	.29	-.86	-.14	-.15	-.31	.03	-.03	-.00	-.08	.00	-.13	-.08	-.05	.00
CA	.79	-.18	.04	.07	.30	.01	.44	.19	.09	.03	.00	-.06	.00	.00
H	.02	-.18	.23	.94	.12	.07	-.08	.00	-.02	.00	-.07	.01	.00	.00
K	.29	-.35	.16	-.24	.70	.43	-.17	-.10	-.03	-.01	.00	-.01	.00	.00
COND	.93	-.05	.00	-.01	-.15	-.02	-.28	.12	.11	.02	-.05	-.05	.09	.00
SUPA	.77	.41	-.08	-.08	-.15	.12	.11	-.38	.00	.17	-.01	-.02	.01	.00
SO <sub>2</sub>	.35	.16	-.84	.19	-.01	.29	.03	-.04	.03	-.15	.01	.01	.00	.00

Den første faktoren er høyt korrelert med de tre viktige nedbørkomponenter sulfat, ammonium og nitrat, og dessuten med kalsium og ledningsevne i nedbøren, og sulfatkonsentrasjonen i aerosoler i luften. Verken hydroniumkonsentrasjonen i nedbøren eller svoveldioksydkonsentrasjonen er korrelert med denne faktor som er langtransporterte luftforurensninger.

Den andre faktoren er godt korrelert med de marine komponentene, klorid, natrium og magnesium og representerer derfor sjøsaltbidraget.

Liksom hydroniumionet er svoveldioksyd dårlig korrelert med første faktor, langtransporterte luftforurensninger. Svoveldioksyd er derimot som eneste variabel godt korrelert med tredje faktor. Det kan tenkes flere årsaker til dette. Outliere i datamaterialet er som over en mulig årsak. Dette resultatet behøver imidlertid ikke skyldes feil. Årsaken til lav korrelasjon kan også skyldes at datamaterialet bare omfatter dager med nedbør. Luftprøvesamlingen foregår ca 800 m høyere enn nedbørsamlingen, og prøvene kan derfor enkelte dager representere forskjellige luftmasser. Dette vil bidra til redusert korrelasjon for såvel svoveldioksyd som partikulært sulfat med første faktor.



Hydroniumkonsentrasjonen er godt korrelert som eneste variabel med fjerde faktor. Vi oppdaget feil i datamaterialet allerede ved gjennomsyn av middelveidier og spredning. Feil i datamaterialet, "outliers", fremkommer typisk ved høy ladning (god korrelasjon) mellom en eneste variabel og en faktor. Forøvrig ligger Narbuvoll tilstrekkelig langt inne i landet til at det vesentligste av den frie svovelsyren er omsatt til ammoniumsulfat. Selv uten feil ville derfor hydroniumkonsentrasjonen trolig være forholdsvis dårlig korrelert med faktoren for langtransporterte forurensninger.

Den femte faktoren er godt korrelert med kalium, men ikke med andre variable. Insekter i nedbørsamleren er en mulig kilde til kalium. Faktor fem kan også være påvirkningen av jordstøv som inneholder kalium.

En oppsummering av de fem viktigste faktorene i dette tilfellet blir:

- Faktor 1: Langtransporterte forurensninger.
- Faktor 2: Sjøsalt.
- Faktor 3: Svoveldioksyd
- Faktor 4: Hydroniumkonsentrasjonen (feil i datamaterialet).
- Faktor 5: Insekter i nedbørsamleren og/eller jordstøv.

## Vedlegg 3

**MÅLRETTET TRANSFORMASJON AV FAKTORRESULTATENE.**



## Vedlegg 3

**MÅLRETTET TRANSFORMASJON AV FAKTORRESULTATENE.**

Målrettet transformasjon av faktorresultatene ("Target transformation factor analysis", TTFA) er en transformasjonsmetode som kan brukes til å fastslå hvilke utslipp som ligger bak faktorresultatene og til å forbedre kjennskapet til de enkelte utslipp. Faktorresultatene angir hvor mye hvert utslipp (kilde eller kildekategori) forklarer av den totale variansen i datamaterialet. De direkte kildebidragene kan beregnes med massebalanse modellen etter en målrettet transformasjon, når utslippenes sammensetning er bestemt.

Gjennomføringen av en målrettet transformasjon er kort gitt i hovedteksten og metoden er grundig beskrevet av Malinowski and Howery (12). Metoden anskueliggjøres i dette vedlegget ved å sammenligne hovedkomponent metoden og den kjemiske massebalanse modellen. Videre vil en se på et eksempel.

Den kjemiske massebalansemodellen er:

$$(1) \quad C_i = \sum_{j=1}^P F_{ij} S_j$$

Dette uttrykket kan skrives på matriseform:

$$(2) \quad C = F S$$

I uttrykket (2) består C av en kolonne hvor hvert element i C representerer konsentrasjonen av en komponent som er bestemt i en prøve. F er en matrise hvor sammensetningen av hver kilde er representert ved en kolonne. Rekkefølgen av komponentene i

F's kolonner skal være samme som i C. S består av en kolonne hvor hvert element representerer et kildebidrag. p betegner antall kilder.

Uttrykket (2) kan brukes også når en har resultatene fra fler enn en prøve. C vil da inneholde en kolonne pr. prøve. F blir uforandret fordi den representerer utslippenes sammensetning. S derimot vil ha en kolonne pr. prøve liksom C og gi kildebidragene til hver prøve.

Den kjemiske massebalansemodellen (2) skal nå sammenlignes med hovedkomponentmodellen:

$$(3) \quad Z_{jk} = \sum_{i=1}^N A_{ji} P_{ik}$$

eller på matriseform

$$(4) \quad Z = A P$$

$Z_{jk}$  er konsentrasjonen av komponent  $j$  i prøve nr.k.  $A_{ji}$  er ladningen av faktor  $i$  på komponent  $j$ , og  $P_{ik}$  er faktor  $i$ 's "score" på prøve nr.k.

I TTFA er det vanlig å bruke skalerte data, men dataene er ikke på standard form med middelerdi 0. Heretter tenkes brukt skalerte data og såvel Z som C byttes ut med et skalert data-sett som betegnes D. Modellene blir da:

$$(5) \quad D = F S$$

$$(6) \quad D = A P$$

I uttrykket (5) vil skaleringen av datamatriksen forårsake en tilsvarende skalering av sammensetningene av utslippene  $F$ .

Modellene i (5) og (6) er nå svært like og transformasjonen skal gjøres slik at  $F = A T$ , hvor  $T$  er en transformasjonsmatrise. Transformasjonsmatrisen  $T$  har en invers matrise  $T^{-1}$  og sammenhengen mellom en matrise og dens inverse er:

$$(7) \quad T T^{-1} = I$$

$I$  er en matrise hvor hvert element er 0 unntatt elementene på diagonalen fra øverste venstre til nederste høyre hjørne som er 1. Ved multiplikasjon av matriser oppfører  $I$  seg som tallet 1 i vanlig multiplikasjon. Uttrykket (6) kan omformes:

$$(8) \quad D = A I P$$

eller

$$(9) \quad D = A T T^{-1} P$$

Sammenlignes dette omformede uttrykket for hovedkomponent modellen (9) med den kjemiske massebalanse modellen (5) ser en:

$$(10) \quad F = A T$$

$$(11) \quad S = T^{-1} P$$

Sammensetningen av første utslipp i  $F$  er representert ved første kolonne i  $F$  som betegnes  $F_1$ . Ifølge reglene for matrisemultiplikasjon er  $F_1$  bestemt av  $A$  og første kolonne i transformasjonsmatrisen  $T$ , som betegnes  $T_1$ .

Hvis  $F_1$  betegner kolonne nr.1 i  $F$  (utslipp nr.1), og kolonne 1 i  $T$  betegnes  $T_1$  gjelder generelt:

$$(13) \quad F_1 = A \cdot T_1$$

I dette uttrykket er bare A kjent.

Beregningene i en målrettet transformasjon av faktorresultatene utføres i følgende fire trinn:

1. Faktorladningene A og egenverdiene beregnes med faktoranalyse eller hovedkomponent metoden.
2. Det velges en kilde 1 med sammensetning  $F_1$ .
3. Kolonne 1 i transformasjonsmatrisen,  $T_1$ , beregnes fra uttrykket (13). I dette inngår beregninger (minste kvadraters metode) for å finne beste tilpasning.
4.  $T_1$  brukes sammen med A i uttrykket (13) til å reprodusere sammensetningen  $F'_1$ . Hvis utslippet med sammensetningen  $F_1$  ikke har påvirket resipienten, vil tilpasningen i trinn 3 bli dårlig. Den reproduserte sammensetningen  $F'_1$  vil da være ulik den opprinnelige  $F_1$  og en ny kilde må prøves.

Antall kilder som har betydning for resipienten vil være kjent fra faktoranalysen. Beregningene i de fire trinnene gjentas for alle de aktuelle utslippene.

Henry har brukt TTFA på resultatene av en undersøkelse av polyaromatiske hydrokarboner (PAH) og fluorider i Sundsvall (23) i Sverige. I Sundsvall ligger et aluminiumverk med utslipp av PAH, og i tillegg vil biltrafikken trolig være en betydelig kilde for PAH.

Hovedkomponent metoden viste at 3 kilder forklarte fra 68 til 78% av variansen på målestedene. I tillegg til disse kildene var det 2-5 mindre kilder.

Kilden til fluorid er aluminiumsverket og hovedkomponent metoden viste at de lettere PAH-forbindelsene hadde en tendens til å samle seg om denne forbindelsen. De tyngre PAH-forbindelsene derimot samlet seg omkring coronen som som en vesentlig skyldes biltrafikk. De middels tunge PAH-forbindelsene hadde derimot ingen tilsvarende klar tendens.

Fra tilgjengelige emisjonsdata prøvde en først sammensetningen av utslippet fra bedriften med TTFA. Resultatet er gjengitt i tabell 3.1. I første kolonne er hver forbindelse angitt i form av et identifikasjonsnummer. Den antatte kildesammensetning som er i prosent av total mengde PAH, tar hensyn til 84% av PAH-mengden i utslippet.

Tredje kolonne gir den beste reproduserte kildesammensetning, og fjerde kolonne gir forholdet mellom den beste reproduserte og den antatte kildesammensetningen. Ved en perfekt tilpasning må dette forholdet være 1. Tilpasningen er generelt tilfredsstillende, men enkelte komponenter passer ikke.



Tabell 3.1: TTFA resultater for utslippet fra aluminiumverket.

PAH NUMBERS	ASSUMED SOURCE COMPOSITION*	BEST-FIT SOURCE COMPOSITION	BEST-FIT ----- ASSUMED
1010	6.80	-35.55	-5.23
1040	4.30	0.24	0.06
1050	6.00	-4.84	-0.81
1060	12.80	9.68	0.77
1070	3.00	3.42	1.14
1080	16.90	24.00	1.42
1090	4.50	3.32	0.74
1120	3.99	3.56	0.91
1130	5.80	9.04	1.56
1140	6.80	5.88	0.86
1150	1.00	1.46	1.46
1160	0.80	0.90	1.12
1170	0.76	-0.01	-0.01
1180	1.00	0.74	0.74
1190	1.70	2.06	1.21
1210	0.80	1.36	1.70
1220	1.20	0.82	0.68
1240	1.00	1.44	1.44
1260	2.60	2.83	1.09
1280	2.40	2.06	0.86

\* As a percentage of total PAH

Tabell 3.2 viser resultatene for biltrafikk. Forholdet mellom tilpasset og antatt utslipp stort sett innenfor en faktor to som er tilfredsstillende.

Tabell 3.2: TTFA resultater for biltrafikk.

PAH NUMBER	ASSUMED SOURCE COMPOSITION*	BEST-FIT SOURCE COMPOSITION	BEST-FIT
			ASSUMED
1010	16.50	-71.38	-4.33
1040	8.10	-3.80	-0.47
1050	0.00	-7.31	
1060	20.30	11.21	0.55
1070	4.30	6.75	1.57
1080	25.40	37.30	1.47
1090	8.00	4.39	0.55
1120	0.00	4.37	
1130	3.90	15.73	4.03
1140	2.60	6.76	2.60
1150	1.70	2.17	1.28
1160	1.70	1.39	0.82
1170	1.0	0.00	0.00
1180	2.60	1.74	0.67
1190	0.00	1.50	
1210	1.00	0.77	0.77
1220	0.30	0.22	0.74
1240	0.20	1.07	5.33
1260	0.90	0.79	0.88
1280	0.00	0.97	

\*As a percentage of total PAH.

Den tredje viktige kilden viste seg å være langtransporterte luftforurensninger av PAH. Tabell 3.3 viser sammenhengen mellom den antatte og den best tilpassede sammensetning for denne kildegruppen. Igjen finner en generelt et godt samsvar.

Tabell 3.3: TTFA resultater for langtransporterte luftforurensninger.

PAH NUMBER	ASSUMED SOURCE COMPOSITION*	BEST-FIT SOURCE COMPOSITION	BEST-FIT ASSUMED
1010	3.90	-31.30	-8.08
1040	0.50	-2.02	-4.05
1050	10.90	5.41	0.50
1060	6.50	5.72	0.88
1070	4.00	2.53	0.63
1080	27.00	28.40	1.05
1090	3.80	2.01	0.53
1120	0.90	0.85	0.94
1130	11.70	16.25	1.39
1140	8.20	9.91	1.21
1150	1.00	2.45	2.45
1160	1.20	1.53	1.27
1170	2.40	1.45	0.61
1180	3.99	5.42	1.39
1190	4.50	4.15	0.92
1210	2.40	2.24	0.93
1220	1.70	1.20	0.71
1240	1.20	1.13	0.95
1260	1.50	2.12	1.41
1280	0.39	0.74	2.47

\* As a percentage of total PAH.

Konsentrasjonene av enkelte lette komponenter i de tilpassede sammensetningene er negative. Årsaken kan være at disse forbindelsene ofte forbrennes. Effektiviteten til prøvetakingsutstyret kan dessuten for disse mest flyktige komponentene tenkes å være avhengig av meteorologiske forhold.



# NORSK INSTITUTT FOR LUFTFORSKNING

**NILU**

(NORGES TEKNISK-NATURVITENSKAPELIGE FORSKNINGSRÅD)  
POSTBOKS 130, 2001 LILLESTRØM  
ELVEGT. 52.

TLF. (02) 71 41 70

RAPPORTTYPE Oppdragsrapport	RAPPORT NR. OR 9/83	ISBN--82-7247-366-6
DATO AUGUST 1983	ANSV.SIGN. O.F.Skogvold	ANT. SIDER 58
TITTEL Reseptormodeller	PROSJEKTLEDER J. Schaug	
	NILU PROSJEKT NR. O-2682	
FORFATTER(E) Jan Schaug	TILGJENGELIGHET** A	
	OPPDRAAGSGIVERS REF.	
OPPDRAAGSGIVER Statens forurensningstilsyn		
3 STIKKORD (å maks. 20 anslag) Reseptormodell      Luftforurensning		
REFERAT (maks. 300 anslag, 5-10 linjer) Rapporten gir en beskrivelse av forskjellige reseptormodeller som f.eks. den kjemiske massebalanse modellen, hovedkomponent analyse og målrettet transformasjon av faktorresultater. Videre beskriver rapporten krav til datagrunnlaget og peker på mulige anvendelsesområder. De tre modellene som er nevnt over, behandles nærmere med eksempler i vedlegg.		
TITLE Receptor models		
ABSTRACT (max. 300 characters, 5-10 lines.)		

\*\*Kategorier: Åpen - kan bestilles fra NILU      A  
Må bestilles gjennom oppdragsgiver      B  
Kan ikke utleveres      C