

Databases Nederlands milieubeleid

Rob Aalbers*

Onderzoek Centrum voor Financieel Economisch Beleid (OCFEB)

en

Herman R.J. Vollebergh**

Vakgroep Algemene Economie

Onderzoek Centrum voor Financieel Economisch Beleid (OCFEB)

Erasmus Universiteit Rotterdam

OCFEB Research Memorandum 0102, 'Environmental Policy, Economic Reform and Endogenous Technology', Working Paper Series 6

* Corresponderend auteur: Rob Aalbers, OCFEB, Erasmus Universiteit Rotterdam, Postbus 1738, 3000 DR Rotterdam, e-mail: aalbers@few.eur.nl

** Vollebergh dankt de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) voor ondersteuning.

Research Memorandum 0102

OCFEB
Room H6-21
Erasmus University Rotterdam
P.O. Box 1738
3000 DR Rotterdam
The Netherlands
Telephone +31 10 408 2430
Telefax +31 10 408 9173
E-mail info@ocfeb.nl
Website www.ocfeb.nl

Inhoudsopgave

Samenvatting	5
1. Inleiding	7
2. Korte beschrijving van de emissieregistratie	8
3. De individuele emissieregistratie (ER-I)	10
3.1 Doelgroepen in de ER-I.....	10
3.2 Opbouw van een record in de ER-I.....	10
3.3 De gegevens binnen een record.....	11
3.4 De steekproefmethodiek.....	13
3.5 De longitudinale opbouw van de database.....	14
3.6 De wijze van emissievaststelling.....	15
3.7 De aanwezigheid van zogenaamde koppelingsvariabelen.....	16
4. Productie- en investeringsstatistieken (CBS)	17
4.1 Productiestatistieken voor bedrijven met 20 of meer werknemers.....	17
4.1.1 De gegevens binnen een record.....	17
4.1.2 Steekproefmethodiek.....	18
4.1.3 De longitudinale opbouw van de productiestatistieken.....	18
4.1.4 De wijze van gegevensverzameling.....	18
4.1.5 De aanwezigheid van zogenaamde koppelingsvariabelen.....	18
4.2 Productiestatistieken voor bedrijven met 1 tot 19 werknemers.....	18
4.2.1 De gegevens binnen een record.....	19
4.2.2 Steekproefmethodiek.....	19
4.2.3 De longitudinale opbouw van de productiestatistieken.....	19
4.2.4 De wijze van gegevensverzameling.....	19
4.2.5 De aanwezigheid van zogenaamde koppelingsvariabelen.....	19
5. De milieubalans	20
5.1 Inleiding.....	20
5.2 Bijschatting verbrandingsemissies.....	20
5.3 Bijschatting procesemissies.....	21
5.4 Bijschatting niet-energetisch gebruik van fossiele brandstoffen.....	22
5.5 Stortplaatsen.....	23
5.6 Overige procesemissies.....	23
5.7 Inzicht in de betrouwbaarheid van de schattingen voor de gebruiker.....	23
6. Koppeling van TNO- en CBS-bestanden op bedrijfsniveau	25

Samenvatting

In deze studie is nagegaan hoe zowel het proces van gegevensverzameling van emissies naar de lucht als de verdere verwerking van de gegevens voor de milieubalans tot stand komen. De verzameling van deze gegevens vindt plaats door de Nederlandse Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek (TNO). De zogenaamde Individuele Emissieregistratie geeft een zo nauwkeurig mogelijk beeld van de uitstoot van vervuilende stoffen naar de lucht door grote individuele vervuilers. Op basis van deze gegevens wordt de uitstoot (bij)geschat van bedrijven die niet in de steekproef zijn opgenomen. Dit gebeurt door de bestanden van TNO te koppelen aan de Productie- en Investeringsstatistieken van het Centraal Bureau voor de Statistiek (CBS). Dit rapport bespreekt de emissieregistratie in detail en gaat na hoe de bijschattingen ten behoeve van de Milieubalans precies tot stand komen. Uitgebreid wordt ingegaan op de gevolgde steekproefmethodiek en de longitudinale opbouw van beide databases en wordt gedetailleerd aangegeven hoe de bijschattingen voor de milieubalans precies verlopen. Tenslotte wordt uiteengezet hoe op basis van een alternatieve koppelingmethode tussen de CBS en TNO databases gedetailleerd inzicht kan worden verkregen in de effecten van het Nederlandse milieubeleid op de milieuproductiviteit van bedrijven.

1. Inleiding

In het kader van het Nationaal Milieu Beleidsplan (NMP) worden binnen Nederland gegevens verzameld over de uitstoot door individuele bedrijven van vervuilende stoffen. De verzameling van deze gegevens vindt plaats door twee instituten. De Nederlandse Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek (TNO) verzamelt gegevens over de uitstoot van vervuilende stoffen naar de lucht, terwijl het Rijksinstituut voor Integraal Zoetwaterbeheer en Afvalwaterbehandeling (RIZA) gegevens verzamelt over de uitstoot van vervuilende stoffen naar zowel water als bodem.

De door TNO en RIZA verzamelde gegevens vormen het uitgangspunt voor de Nederlandse milieubalans. Op basis van deze gegevens wordt de uitstoot (bij)geschat van bedrijven die niet in de steekproef zijn opgenomen. Dit gebeurt door de bestanden van TNO en RIZA te koppelen aan de productie- en investeringsstatistieken van het Centraal Bureau voor de Statistiek (CBS).

Op dit moment bestaat onvoldoende inzicht in zowel het proces van gegevensverzameling als de verdere verwerking van die gegevens voor de milieubalans. Ten aanzien van de gegevensverzameling is bijvoorbeeld onduidelijk uit welke variabelen de verschillende databases (TNO, RIZA, en CBS) bestaan en welke steekproefmethodiek is gehanteerd bij de samenstelling van de database. Ten aanzien van de verwerking van de gegevens is onduidelijk op welke wijze de databases worden gekoppeld en hoe voor niet geënquêteerde bedrijven de uitstoot wordt bijgeschat. Het Ministerie van Economische Zaken heeft OCFEB dan ook gevraagd om een onderzoek te verrichten naar de totstandkoming van de milieubalans. Dit onderzoek heeft zich alleen gericht op de koppeling tussen de databases van TNO en CBS. De koppeling tussen de databases van het RIZA en CBS vindt echter op een vergelijkbare wijze plaats.

In hoofdstuk 2 wordt een korte beschrijving gegeven van de emissieregistratie zoals die door CBS en TNO in opdracht van het Ministerie van VROM wordt uitgevoerd. In hoofdstuk 3 wordt de inhoud van de Individuele Emissieregistratie (ER-I) besproken. Daarnaast wordt onder andere aandacht besteed aan de steekproefmethodiek en de longitudinale opbouw van de database. In hoofdstuk 4 worden de Productie- en Investeringsstatistieken besproken zoals die door het CBS worden samengesteld. Ook hier wordt aandacht besteed aan de inhoud van de database, de steekproefmethodiek en de longitudinale opbouw. In hoofdstuk 5 worden de bijschattingen voor de milieubalans besproken. Het betreft hier de emissies van bedrijven die niet in de ER-I zijn opgenomen. Tot slot wordt in hoofdstuk 6 een alternatieve koppelingsmethode tussen de CBS en TNO databases toegelicht. Als deze koppeling in de toekomst inderdaad tot stand wordt gebracht, dan kan op basis daarvan een gedetailleerd inzicht worden verkregen in de effecten van het Nederlandse milieubeleid op de milieuproductiviteit van bedrijven.

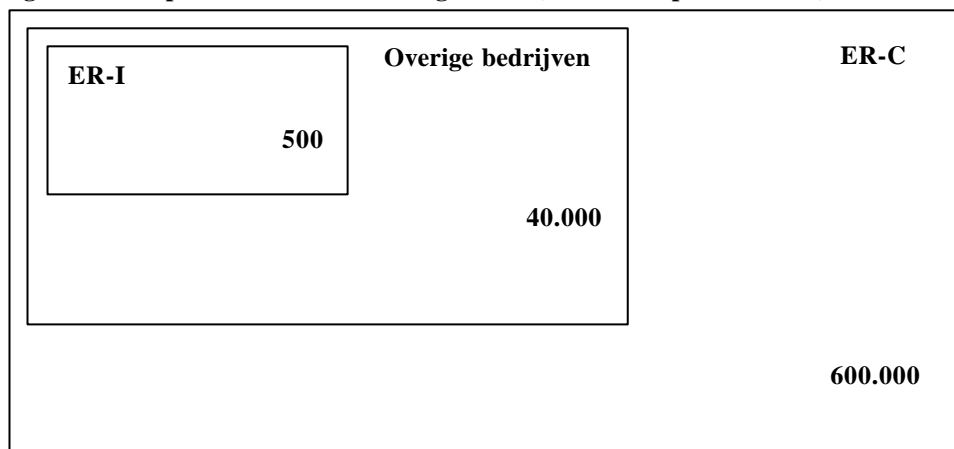
2. Korte beschrijving van de emissieregistratie

De emissieregistratie heeft tot doel inzicht te verkrijgen in het niveau en de verspreiding van emissies van milieuverontreinigende stoffen in Nederland. Zij omvat zowel de emissies van industriële als niet-industriële bronnen. De registratie van industriële bronnen geschiedt voor een beperkt aantal grote vervuilers op individuele basis, dat wil zeggen dat de emissies worden bepaald op het niveau van het individuele bedrijf. De gegevens over deze bedrijven worden opgeslagen in de Individuele EmissieRegistratie (ER-I). De ER-I bevat gegevens van ongeveer 500 bedrijven (meting 1998). Deze 500 bedrijven zijn verantwoordelijk voor 60-80% van de totale emissies naar de lucht en zijn dus verreweg de grootste vervuilers naar de lucht.

Mede op basis van de ER-I worden de emissies van de 40.000 kleinere industriële puntbronnen in kaart gebracht. Op deze wijze ontstaat inzicht in de omvang en samenstelling van de totale industriële emissies. Hierbij wordt onder andere gebruik gemaakt van de productie- en investeringsstatistieken (PS en IS) van het Centraal Bureau voor de Statistiek (CBS). Zowel de bedrijven uit de PS en IS van het CBS als de bedrijven uit de ER-I worden gerangschikt naar SBI-nummer. Per SBI-categorie (meestal 2-digitaal niveau) worden vervolgens de emissies van de kleinere industriële bedrijven bepaald aan de hand van een kengetal, bijvoorbeeld de fysieke productie (in tonnen gemeten) van alle Nederlandse bedrijven tezamen ten opzichte van de productie van de bedrijven uit de ER-I.

Tot slot worden de emissies van niet-industriële bronnen (huishoudens, verkeer, etc.) berekend. Dit gebeurt met behulp van statistische gegevens, zoals het aantal inwoners, woningen, arbeidsplaatsen, auto's, etc. De emissiegegevens van alle puntbronnen in Nederland - ongeveer 600.000 - worden vervolgens samen gebracht in de Collectieve EmissieRegistratie (ER-C).

Figuur 1: De opbouw van de emissieregistratie (met aantal puntbronnen)



In de hiernavolgende paragrafen zal nader worden ingegaan op de wijze waarop de gegevens door respectievelijk TNO en het CBS worden verzameld en verwerkt. Daarnaast zal aandacht worden besteed aan de longitudinale opbouw van de databases en de vraag of de bestanden op het niveau van de individuele bedrijven aan elkaar te koppelen zijn.

3. De individuele emissieregistratie (ER-I)

In dit hoofdstuk zal de opbouw en samenstelling van de individuele emissieregistratie worden besproken. Na een kort overzicht van de sectoren die in de ER-I worden bijgehouden zal de opbouw van een record in de ER-I worden besproken alsmede de gegevens die (in principe) voorhanden zijn binnen een record. Vervolgens zullen aan bod komen: de wijze waarop de gegevens worden verzameld, de betrouwbaarheid van de gegevens, de longitudinale opbouw van de database in combinatie met de gehanteerde steekproefmethodiek en tot slot de aanwezigheid van mogelijke koppelingvariabelen.

3.1 Doelgroepen in de ER-I

Voor het beleid is het van belang om te weten hoe de bedrijven uit de ER-I over de doelgroepen (zoals die door VROM worden gehanteerd) zijn verdeeld. Tabel 1 geeft de verdeling van de 709 individueel geregistreerde vestigingen weer over de doelgroepen in termen van SBI-codes alsmede het totaal aantal vestigingen dat in het ER-I systeem is opgenomen. Met een vestiging wordt hier een (onderdeel van een) bedrijf bedoeld dat specifiek verbonden is aan één locatie (zie ook paragraaf 3.2). Het betreft hier overigens gegevens uit 1993. Uit de tabel blijkt dat met name de grote procesindustrie zwaar vertegenwoordigd is.

Tabel 1: Verdeling van de ER-I vestigingen over de doelgroepen in 1993

	SBI-code	Aantal vestigingen	Aantal werknemers
Raffinaderijen	23201	9	4.619
Energiesector	11, overig 23, 40	57	6.841
Afvalverwijderingsbedrijven	9002-4, 37	15	1.393
Chemische Industrie	24	118	63.231
Overige Industrie		510	177.176
Totaal		709	253.260

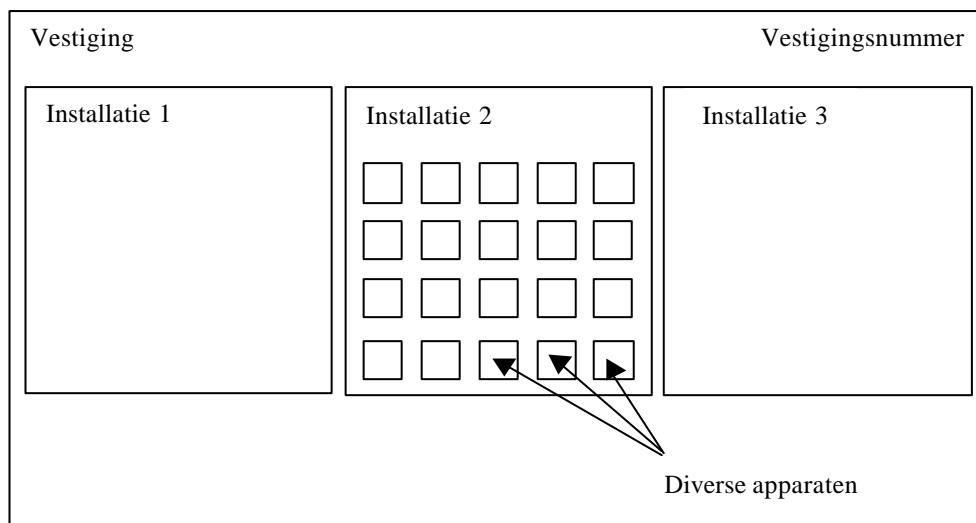
bron: VROM

3.2 Opbouw van een record in de ER-I

De gegevens in de ER-I zijn opgebouwd in records. Elk record geeft daarbij een vestiging weer (zie figuur 2). Als men binnen de ER-I dus spreekt over een 'bedrijf' dan bedoelt men daarmee altijd een vestiging. Dit kan, maar hoeft niet, samen te vallen met de meer gebruikelijke vorm van een bedrijf als juridische of economische eenheid. Een voorbeeld maakt dit duidelijk. Stel een onderneming (B.V. of N.V.) heeft binnen Nederland twee of meer vestigingen. Dan is deze onderneming één bedrijf in de juridische of economische zin. Binnen de ER-I is echter sprake van twee vestigingen, een op elke productielocatie. Als het bedrijf vervolgens zou besluiten om deze

productielocaties samen te voegen, dan neemt het aantal vestigingen binnen de ER-I af van twee naar één, terwijl het aantal bedrijven constant blijft.

Figuur 2: Opbouw van een record in de ER-I



Elke vestiging is opgebouwd uit een of meerdere installaties. Installaties worden onderscheiden naar gelang de (fysieke) producten die worden gemaakt. Zo worden in een elektriciteitscentrale zowel stoomketels als turbines als installaties aangemerkt. Ketels produceren daarbij stoom, turbines elektriciteit. Installaties worden slechts onderscheiden naar hun type, niet naar hun aantal. Zo heeft een elektriciteitscentrale met vier stoomketels en vier turbines, slechts twee installaties, namelijk de installatie 'stoomketel' en de installatie 'turbine'. Wel wordt bijgehouden hoeveel installaties er van elk type zijn, maar deze installaties zitten dus als één record in de database.

Elke installatie is op zijn beurt weer opgebouwd uit een of meerdere apparaten. De installatie 'stoomketel' bestaat dus uit verschillende apparaten, waaronder een brander. Ook hier geldt dat apparaten slechts worden onderscheiden naar hun type en niet naar hun aantal. Als stoomketels bijvoorbeeld twee branders hebben, dan heeft de installatie 'stoomketel' slechts één apparaat 'brander'.

3.3 De gegevens binnen een record

De gegevens binnen een record zijn gekoppeld aan een van drie niveaus binnen een record, te weten vestiging, installatie of apparaat. Naarmate gegevens op een lager aggregatieniveau bekend zijn, bestaat meer inzicht in de wijze waarop de vervuiling ontstaat. De eerste categorie gegevens zijn de vestigingsgegevens, zoals de naam en het adres van de vestiging. De tweede categorie gegevens heeft betrekking op installaties

binnen de vestiging. Als van een bedrijf bekend is welke installaties er zijn, dan zal over het algemeen ook bekend zijn welke emissies deze installaties veroorzaken. Zijn voor een bepaald bedrijf geen installaties bekend, dan zijn de emissies alleen op bedrijfsniveau bekend. De derde categorie gegevens heeft betrekking op apparaten binnen een installatie.

De volgende gegevens zijn - in principe - op het niveau van de vestiging, een installatie en een apparaat bekend:

1) Op het niveau van een vestiging:

- De officiële naam van de vestiging;
- Het adres en de plaats van de vestiging;
- Het bedrijfsregistratienummer;
- De naam van het bedrijf waarheen de correspondentie moet worden gezonden;
- De instantie die toezicht uitoefent op het bedrijf ten aanzien van de luchtmissies;
- Het totaal aantal werknemers in vaste dienst;
- Het jaar waarover de gegevens zijn verzameld.

2) Op het niveau van de installatie:

- Installatienummer;
- Het aantal samengenomen gelijksoortige installaties;
- De naamgeving van de fabricage van een bepaald product, bijv. 'bereiding van';
- De ontwerpcapaciteit: voor elektriciteitscentrales bijv. het elektrisch vermogen in kW;
- De werkelijk productie in het betreffende jaar;
- Het aantal bedrijfsuren gedurende welke de productie heeft plaatsgevonden, waarbij stand-by als productie wordt aangemerkt;

3) Op het niveau van het apparaat:

- Het apparaatnummer;
- Het aantal samengenomen gelijksoortige apparaten;
- De naam van het soort apparaat dat de emissie veroorzaakt;
- De ontwerpcapaciteit: voor een brander bijv. het thermisch vermogen.

Afhankelijk van het niveau van waarneming worden emissiepunten en emissies waargenomen per bedrijf, installatie of apparaat. Binnen de ER-I worden drie soorten emissiepunten onderscheiden, namelijk een schoorsteen, een gebouw en de buitenlucht.

Van elk emissiepunt is - in principe - bekend:

- Het emissiepuntnummer;
- Het aantal samengenomen emissiepunten van dezelfde soort en die gesitueerd zijn binnen eenzelfde kaart-vierkant (een oppervlakte van 5 bij 5 km);
- De naam waaronder het emissiepunt bij het bedrijf bekend is;
- De periode dat een emissie plaats vindt;

- De totale tijdsduur van de emissie in uren per jaar;
- De totale omvang van de emissie gedurende de periode dat de emissie plaats heeft.

Van elke emissie is - in principe - bekend:

- Het soort emissie waarom het gaat (de chemische samenstelling);
- De hoeveelheid van de stof die wordt geëmitteerd;
- De emissievorm (vast, vloeibaar, gas);
- De wijze van meting (zie paragraaf 3.6);
- De oorzaak van eventuele veranderingen in het niveau van de emissie;
- Eventueel doorlopen reinigingen voordat de emissie plaatsvindt en de bestemming van de afgevangen stoffen (afval, recycling of eindproduct).

In principe worden alle emissies naar de lucht die vrijkomen binnen het productieproces geregistreerd. Enkele voorbeelden daarvan zijn: kooldioxide (CO₂), koolmonoxide (CO), stikstofoxiden (als NO₂), VOS, NMVOS, formaldehyde, toluen, etheen, methaan, ammoniak en dioxines.

3.4 De steekproefmethodiek

De steekproefmethodiek die is gehanteerd bij de ER-I heeft van jaar tot jaar kleine of grote aanpassingen ondergaan. In de eerste ronde (1973 - 1978) werden alle vestigingen benaderd die uit milieuoogpunt relevant zouden kunnen zijn (zie ook Tabel 2). In totaal waren dit 6300 vestigingen. Omdat slechts een klein gedeelte van die bedrijven substantieel bijdroeg aan de industriële luchtverontreiniging, werd besloten een efficiëntere selectie van vestigingen te hanteren. Deze selectie is gemaakt op basis van de verwachte bijdrage van vestigingen aan de nationale emissie en op de verwachte invloed van de geëmitteerde stoffen op het milieu. Grote vuurhaarden, zoals elektriciteitscentrales en grote WKK centrales, zijn overigens in elke ronde geënkquêteerd. Dit leidde tot een gestage daling van het aantal geënkquêteerde vestigingen van 6300 in de eerste ronde tot ca. 500 vanaf de tiende ronde. Deze vestigingen kunnen worden beschouwd als de 500 grootste vervuilers naar de lucht binnen Nederland. Vestigingen die niet langer zijn meegenomen in de ER-I zijn toegevoegd aan de (veel minder detail bevattende) ER-C.

Tabel 2: Omvang van de ronden in de Individuele Emissieregistratie

Ronde	Aantal vestigingen	Aantal installaties ^a	Tijdvak emissies
1	6.300	20.000	1973 - 1978
2	1.100	2.700	1981 - 1984
3	1.500	2.900	1985 - 1987
4	700	2.370	1988 - 1989
5	709	2.540	1990 – 1991 ^b
6	799	3.203	1992
7	ca. 700	2.956	1993
8	ca. 700	3.000	1994
9	ca. 700	n.b.	1995
10	ca. 500	n.b.	1996
11	ca. 500	n.b.	1997
12	ca. 500	n.b.	1998

bron: VROM

^a n.b. = niet bekend.

^b In 1991 is slechts een halve ronde gemaakt

Naast selectie op basis van bovengenoemde criteria werden een aantal individuele bedrijven geselecteerd op verzoek van gemeenten en provincies. De achtergrond van deze verzoeken lag vaak in de sfeer van (controle van uitgegeven) milieuvergunningen. Samenvattend zijn dus vier criteria aan te duiden waarop bedrijven zijn geselecteerd:

- 1) De totale bijdrage van de emissies van een bepaalde vestiging aan de nationale emissie;
- 2) De verwachte invloed van de geëmitteerde stoffen op het milieu;
- 3) Grote vuurhaarden;
- 4) Op verzoek van gemeenten en provincies.

Hierbij is het van belang om in de gaten te houden dat deze criteria in de loop van de tijd meerdere malen en niet noodzakelijkerwijs in deze volgorde zijn toegepast.

3.5 De longitudinale opbouw van de database

De longitudinale opbouw van de ER-I database is zeer complex en per bedrijf verschillend. Om toch een redelijk inzicht te kunnen krijgen zullen hier een aantal eigenschappen van de ER-I kort worden toegelicht. Daarna zal worden bekeken voor hoeveel bedrijven er een complete steekproef voor de periode 1988/89 tot en met 1998 beschikbaar is.

In de 27 jaar dat de ER-I bestaat zijn er twaalf enquêteringsronden geweest.¹ In elke ronde wordt een bedrijf slechts één keer geënquêteerd.² Voor een individueel bedrijf zijn dus nooit meer dan twaalf waarnemingen beschikbaar. Deze waarnemingen beslaan de periode 1973 tot en met 1998. Sinds 1992 is de enquête jaarlijks uitgevoerd. Voor een bedrijf dat in alle twaalf rondes is geënquêteerd, is dus een waarneming beschikbaar in elk van de twaalf tijdvakken (zie Tabel 2). A priori is niet bekend welk bedrijf in welk tijdvak is geënquêteerd. Voor de jaren 1979 en 1980 zijn geen waarnemingen beschikbaar.

Omdat de steekproefmethodiek in de loop van de tijd steeds is aangepast aan de op dat moment geldende inzichten is het interessant om te kijken naar de longitudinale opbouw van de database. Een volledig overzicht van de longitudinale opbouw van de ER-I is vanwege de structuur van de database niet beschikbaar. Daarom hebben we aan TNO te Apeldoorn gevraagd voor hoeveel bedrijven er een continue reeks waarnemingen is vanaf de vierde ronde (1988/1989) tot en met de tiende ronde (1998). Uit raadpleging van de database bleek dat dit voor 253 tot 296 bedrijven het geval was.

Een uitzondering op de enquêteringsronden zoals die in tabel 2 is weergegeven vormen de grote vuurhaarden (bijv. elektriciteits- en WKK-centrales). Zij zijn verplicht om hun emissies jaarlijks te rapporteren aan de Europese Unie. Deze gegevens zijn opgenomen in de ER-I. Voor deze bedrijven is dus een complete reeks waarnemingen beschikbaar vanaf 1973. In totaal gaat het hier om 57 vestigingen in 1993 (zie ook tabel 1).

3.6 De wijze van emissievaststelling

De emissies van een bedrijf kunnen op verschillende manieren worden vastgesteld. Deze mogelijkheden zijn in aflopende mate van betrouwbaarheid: continue metingen, periodieke metingen, éénmalige metingen, berekeningen, emissiefactoren en schatting. In de rest van dit hoofdstuk zal kort op elk van deze methoden worden ingegaan.

- Continue metingen: geven de meest betrouwbare uitkomsten. Concentraties, debieten, bedrijfsuren en de jaarlijkse emissie zijn dan allemaal bekend. Dit type meting wordt voornamelijk toegepast bij vuurhaarden.
- Periodieke metingen: gedurende enkele keren per jaar worden metingen verricht. Voor een goede betrouwbaarheid is het van belang dat (i) de variatie over de tijd van concentraties en debieten gering is³; (ii) het aantal bedrijfsuren bekend is.

¹ De dertiende ronde, die betrekking heeft op 1999, wordt op dit moment verwerkt.

² Een uitzondering op deze regel is de vierde ronde waarin in 1989 slechts een kleine update is gemaakt en de vijfde ronde waarin in 1991 slechts een halve ronde is gemaakt.

³ Concentraties en debieten willen nog wel eens variëren als sprake is van batchgewijze productie. Bij het opstarten van installaties is in veel gevallen namelijk sprake van aanzienlijk hogere debieten en concentraties.

- Éénmalige metingen: concentraties en debieten worden slechts een keer gemeten. De jaarlijkse emissie wordt berekend aan de hand van het aantal bedrijfuren.
- Berekeningen: deze worden gebruikt als men de beschikking heeft over gegevens voor het opstellen van een massabalans en individuele metingen - om wat voor reden dan ook - niet worden verricht. Een ander voorbeeld is de emissiebepaling bij tankopslag. In dit geval wordt de jaarlijkse emissie berekend met behulp van bepaalde - uit de literatuur bekende - formules.
- Emissiefactoren: worden gebruikt als metingen totaal niet beschikbaar zijn. De hoogte van de emissiefactoren wordt bepaald aan de hand van een literatuuronderzoek. Dit type metingen wordt voornamelijk toegepast bij emissies van koolwaterstoffen uit verbranding, zware metalen en dioxines.
- Schatting: de minst betrouwbare mogelijkheid. De betrouwbaarheid is sterk afhankelijk van de persoon die de schatting maakt.

De keuze van de meetmethode is onder andere afhankelijk van de technische mogelijkheden, de kosten die de verschillende typen meetmethoden met zich mee brengen en de urgentie van een nauwkeurige meting. Binnen deze beperkingen wordt zoveel mogelijk voor de betrouwbaarste meetmethode gekozen. Op dit moment is het voor gebruikers van de TNO-database moeilijk om inzicht te krijgen in het soort meetmethode dat door TNO is gebruikt voor een specifieke stof, bijvoorbeeld dioxine. Het is dus voor gebruikers ook niet mogelijk om een schatting te krijgen van de betrouwbaarheid van een specifieke meting. Daarom is TNO van plan om in de nabije toekomst aan elk cijfer een betrouwbaarheidsindicator toe te voegen.

Sinds 1994 is het detail in de database sterk afgenomen. De reden daarvoor is dat de metingen namelijk niet meer door TNO, maar door de bedrijven zelf worden verricht en gerapporteerd. Dit gebeurt voornamelijk via de bedrijfsmilieuplannen. In concreto betekent dit dat voor de jaren na 1994 steeds minder gegevens op apparaats- en installatieniveau bekend zijn. Op dit moment is het - voor buitenstaanders - niet mogelijk om inzicht te krijgen in de betrouwbaarheid van de individuele metingen. Op termijn is TNO van plan om aan elke meting een betrouwbaarheidsindicatie toe te voegen, zodat dit probleem dan verholpen is.

3.7 De aanwezigheid van zogenaamde koppelingsvariabelen

Als koppelingsvariabelen kunnen worden gebruikt: de naam van de vestiging, het adres en plaats van de vestiging en de naam van het bedrijf waarheen de correspondentie moet worden gezonden. Alhoewel binnen de ER-I een bedrijfsregistratienummer bekend is, kan dit niet worden gebruikt om de ER-I aan andere databases te koppelen, omdat dit nummer alleen door TNO binnen de ER-I wordt gebruikt. Andere dataverzamelande instanties als CBS gebruiken andere bedrijfsregistratienummers voor hetzelfde bedrijf.

4. Productie- en investeringsstatistieken (CBS)

In dit hoofdstuk worden de productie- en investeringsstatistieken besproken, zoals die worden samengesteld door het Centraal Bureau voor de Statistiek. Daarbij moet in het oog worden gehouden dat niet zozeer de productiestatistieken zelf, maar de achterliggende microbestanden worden besproken. Deze bestanden bevatten gegevens op bedrijfsniveau en zijn - met uitzondering van namen en dergelijke - binnen de muren van het CBS voor iedereen toegankelijk. De productiestatistieken zijn oorspronkelijk opgezet voor alle bedrijven in de sector Industrie. Daarna is het CBS begonnen met de uitbreiding naar andere sectoren van de economie, zoals handel, transport en dienstverlening.

In dit rapport zal alleen uitgebreid worden stilgestaan bij de productiestatistieken voor de sector Industrie. De reden hiervoor is dat alleen deze bedrijven voor de milieubalans gekoppeld worden aan het bestand van TNO. De productiestatistieken voor bedrijven met 20 of meer werknemers worden behandeld in paragraaf 4.1, waarna in paragraaf 4.2 de productiestatistieken voor bedrijven met 1 tot 19 werknemers worden behandeld.

4.1 Productiestatistieken voor bedrijven met 20 of meer werknemers

Achtereenvolgens zullen aan bod komen de gegevens binnen een record, de steekproefmethodiek, de longitudinale opbouw van de database, de wijze van gegevensverzameling en de aanwezigheid van eventuele koppelingsvariabelen.

4.1.1 De gegevens binnen een record

Elke bedrijfseenheid heeft binnen de productie- en investeringsstatistieken zijn eigen record. Een bedrijfseenheid is daarbij gedefinieerd als de kleinste eenheid die economisch gezien zelfstandig is. Een onderdeel van een bedrijf wordt aangemerkt als een bedrijfseenheid als het zelfstandig zijn inkoop-, verkoop- en investeringsbeslissingen neemt en als het een boekhouding voert. Binnen de productiestatistieken heeft men - in principe - voor elk(e) bedrijf(seenheid) de beschikking over de volgende gegevens:

- Het algemeen bedrijfsregistratienummer (ABR);
- De naam van het bedrijf;
- Het aantal werknemers op de eigen loonlijst;
- Industriële verkopen;
- Verkopen buitenland;
- Voorraad mutaties producten, onderhanden werk en grond- en hulpstoffen;
- Inkoop materialen;
- Zelfvervaardigde investeringsgoederen;
- Energieverbruik in gulden (elektriciteit, gas en overig)
- Fysiek energieverbruik (elektriciteit in kWh en gas in m³);
- Lonen en uitgaven sociale voorzieningen;

- Milieuheffingen en zuiveringslasten
- Accijnzen;
- Restituties en subsidies;
- Afschrijvingen op vaste activa;
- Interest;
- Saldo winst c.q. verlies.

Een voorbeeld van de verzonden enquêteformulieren bevindt zich in bijlage 1.

4.1.2 Steekproefmethodiek

De productiestatistiek voor bedrijven met 20 of meer werknemers (PS20+) worden samengesteld voor alle bedrijfspgroepen in de bedrijfstak Industrie, met uitzondering van de Sociale Werkvoorziening. Hieronder vallen alle bekende industrietakken zoals bijvoorbeeld de chemie, de basismetale, de voedingsmiddelenindustrie en de papierindustrie. Alle bedrijven met 20 of meer werknemers worden aangeschreven met het verzoek om gegevens ter beschikking te stellen. Het aantal bedrijven dat jaarlijks wordt aangeschreven wisselt al naar gelang het aantal nieuwkomers en faillissementen. In totaal worden tussen de 6000 en 7000 bedrijven aangeschreven. Er is dus eigenlijk geen sprake meer van een steekproef, maar van een meting van alle bedrijven (met meer dan 20 werknemers) in de Nederlandse samenleving. De non-response is zeer laag (over het algemeen minder dan 5%).

4.1.3 De longitudinale opbouw van de productiestatistiek

De PS20+ worden bijgehouden sinds 1978. Jaarlijks worden alle bedrijven in de bedrijfstak Industrie aangeschreven om gegevens ter beschikking te stellen. Over het algemeen zal voor elk bedrijf binnen de sector Industrie een tijdreeks beschikbaar zijn vanaf 1978 tot heden.

4.1.4 De wijze van gegevensverzameling

Het CBS schrijft de desbetreffende bedrijven aan en vraagt hen gegevens ter beschikking te stellen. Deze worden bij aanlevering gecontroleerd en eventueel aangepast. Vervolgens worden de data samengebracht in een elektronische database.

4.1.5 De aanwezigheid van zogenaamde koppelingsvariabelen

Als koppelingsvariabelen kunnen worden gebruikt: de naam, locatie en postcode van het bedrijf.

4.2 Productiestatistiek voor bedrijven met 1 tot 19 werknemers

De productiestatistiek voor bedrijven met 1 tot 19 werknemers (PS19-) worden op vrijwel dezelfde wijze samengesteld dan de PS20+. Daarom zal in de volgende subparagrafen alleen aandacht worden besteed aan de verschillen van de PS19- ten opzichte van de PS20+.

4.2.1 De gegevens binnen een record

Bij de PS19- is de vraagstelling beknopter dan bij de PS20+. De belangrijkste verschillen zijn dat slechts een totaalbedrag is opgenomen voor de energiekosten. Dat wil zeggen dat de onderverdeling naar elektriciteit, gas en overige brandstoffen ontbreekt. Ook is niet bekend hoeveel kWh en m³ het bedrijf heeft gebruikt. Daarnaast is voor veel posten slechts een totaalbedrag - zonder verdere uitsplitsing - beschikbaar. Voor de jaren 1996 en 1997 zijn de gegevens voor de PS19- op hetzelfde detailniveau beschikbaar als voor de PS20+. Het gaat hier wel om schattingen in plaats van waarnemingen. Een voorbeeld van de verzonden enquêteformulieren vindt u in bijlage 2.

4.2.2 Steekproefmethodiek

De PS19- worden samengesteld voor alle bedrijfstakken in de bedrijfstak Industrie, met uitzondering van de sociale Werkvoorziening. Hieronder vallen alle bekende industrietakken, zoals de chemie, de basismetalen, de voedingsmiddelenindustrie en de papierindustrie. Van de bedrijven met 19 of minder werknemers wordt slechts een beperkt aantal aangeschreven met het verzoek om gegevens ter beschikking te stellen. Het CBS gebruikt een gestratificeerde steekproef om bedrijven te enquêteren. Een algemeen beeld van deze steekproef valt niet te geven, omdat de daadwerkelijke enquête (wie wordt aangeschreven in een bepaald jaar) kan verschillen per sector en per type gegevens (handel, productie, fiscaal, etc.). In zijn algemeenheid geldt echter dat van de middelgrote bedrijven de helft jaarlijks wordt geënquêteerd. Voor kleine bedrijven is dit minder dan één op de tien bedrijven. De respons van de enquêtes ligt op ruim 80%.

4.2.3 De longitudinale opbouw van de productiestatistieken

De PS19- worden bijgehouden sinds 1990. Er is geen longitudinaal bestand c.q. panel van bedrijven uit de PS19- beschikbaar. Dit komt omdat het CBS gebruik maakt van het zogenaamde enquêtedruksysteem. Om de kleine bedrijven zoveel mogelijk te ontzien bij het invullen van de enquêtes worden bedrijven slechts één keer per zoveel jaar aangeschreven. Zo geldt voor middelgrote bedrijven die in 1998 zijn aangeschreven, dat zij niet aangeschreven worden in 1999 en vice versa.

4.2.4 De wijze van gegevensverzameling

Het CBS schrijft de desbetreffende bedrijven aan en vraagt hen gegevens ter beschikking te stellen. Deze worden bij aanlevering gecontroleerd en eventueel aangepast. Vervolgens worden de data samengebracht in een elektronische database.

4.2.5 De aanwezigheid van zogenaamde koppelingsvariabelen

Als koppelingsvariabele kan worden gebruikt: de naam, locatie en postcode van het bedrijf.

5. De milieubalans

5.1 Inleiding

Omdat de individuele emissieregistratie slechts inzicht geeft in de emissies van de 500 grootste vervuilers, vindt een bijschatting plaats van de overige industriële emissies. Dit zijn dus de emissies van de kleinere bedrijven. Bijschatting van de emissies van de bedrijven die niet in de ER-I zijn opgenomen vindt plaats door de productie- en investeringsstatistieken van het CBS - op SBI-niveau - te koppelen aan de ER-I. Daarbij wordt ook gebruik gemaakt van aanvullende informatie uit ander bronnen, zoals de Nederlandse Energie Huishouding (NEH)⁴ en de KoolWaterStoffen 2000 (KWS2000).⁵ Bijschatting van deze emissies vindt plaats in de volgende categorieën: verbrandingsemissies, procesemissies, niet-energetisch gebruik van fossiele brandstoffen, stortplaatsen en overige procesemissies. Elk van deze categorieën zal in de navolgende paragrafen worden besproken. Per categorie zal worden aangegeven wat de betrouwbaarheid is. Hiervoor wordt de volgende classificatie gebruikt:

- A: Een getal is gebaseerd op een groot aantal metingen aan representatieve installaties.
- B: Een getal is gebaseerd op een aantal metingen aan een deel van de voor de sector representatieve installaties.
- C: Een getal is gebaseerd op een beperkt aantal metingen, aangevuld met schattingen op basis van technische kennis over het proces.
- D: Een getal is gebaseerd op een gering aantal metingen, aangevuld met schattingen op basis van aannames.
- E: Een getal is gebaseerd op een technische berekening op basis van een aantal aannames.

5.2 Bijschatting verbrandingsemissies

Het betreft hier emissies door het verbranden van fossiele brandstoffen (olie, gas kolen en hout) in vuurhaarden, die niet individueel zijn geregistreerd. Dit zijn onder andere CO₂, SO₂ en NO_x. De bijschatting vindt plaats per bedrijfsgroep. Dit is meestal een sector, maar soms een onderdeel van een sector. De emissies worden berekend door het brandstofverbruik - afkomstig uit de Nederlandse Energie Huishouding (NEH) - te vermenigvuldigen met de voor die bedrijfsgroep geldende emissiefactoren. Deze emissiefactoren worden bepaald uit de emissiegegevens van de ER-I. Een voorbeeld maakt dit duidelijk. Stel dat het brandstofverbruik in een bepaalde bedrijfsgroep 2000 kton is. Van de bedrijven die tot deze bedrijfsgroep behoren zitten er een aantal in de ER-I. Van deze bedrijven weten we dat het brandstofverbruik 1500 kton en de uitstoot van NO_x 60 kton bedraagt. De totale NO_x emissies in deze bedrijfsgroep zijn dan: 60*

⁴ De jaarlijks door het CBS opgestelde NEH geeft inzicht in de energieproductie en het energieverbruik in Nederland door sectoren en naar energiedragers.

⁵ Afkomstig van het Informatiecentrum Milieuvergunningen (Infomill).

2000/1500 = 80 kton. Voor bedrijfstakken en brandstoffen die niet in de ER-I voorkomen wordt gewerkt met gemiddelde emissiefactoren.⁶ Het basisjaar van de gegevens uit de ER-I en de NEH is 1995. De kwaliteitsclassificatie van deze bijschattingen is B.

5.3 Bijschatting procesemissies

Het betreft hier procesemissies naar de lucht van bedrijven die niet individueel zijn geregistreerd. De bijschattingen voor de voedings- en genotmiddelenindustrie, grafische industrie, de rubber- en kunststofverwerkende industrie en de metalectro worden verzorgd door het RIVM. De bijschattingen voor de overige bedrijfstakken worden verzorgd door het CBS. Eerst zullen de CBS-bijschattingen worden behandeld, vervolgens de RIVM-bijschattingen.

CBS-bijschattingen

De CBS-bijschattingen vinden plaats per bedrijfstak. De bijschattingen worden bepaald door de totale emissies van de individueel geregistreerde bedrijven te vermenigvuldigen met een zogenaamde ophoogfactor. Deze ophoogfactor is gebaseerd op een variabele uit de productiestatistieken die - naar inzicht van de desbetreffende onderzoekers - de hoogte van de emissies het beste verklaart. Een overzicht van deze variabelen voor een aantal bedrijfstakken kan gevonden worden in Tabel 3. De meest gebruikte variabelen voor het verklaren van de emissies en de bepaling van de ophoogfactor zijn (in die volgorde): de geproduceerde output in tonnen, stuks of liters, het aantal werkzame personen, de hoeveelheid gebruikte energie in joule en de omzet.

Voor bedrijfstak 17210 wordt de ophoogverklarende variabele gegeven door de hoeveelheid verschillende weefsels die worden geproduceerd. De ophoogfactor wordt dan gegeven door:

$$\text{Ophoogfactor} = (\text{ton weefsels uit bedrijfstak}) / (\text{ton weefsels van bedrijven uit ER-I}).$$

De bijschatting voor de emissies voor bedrijfstak 17210 is dan gelijk aan:

$$\text{Bijschatting emissies 17210} = (\text{Emissies van bedrijven uit de ER-I}) * (\text{ophoogfactor} - 1).$$

Stel bijvoorbeeld dat in bedrijfstak 17210 1000 ton weefsels produceert. 600 ton hiervan wordt door bedrijven geproduceerd die in de ER-I zitten. De uitstoot van VOS van deze bedrijven is gelijk aan 30 ton. De ophoogfactor is dan gelijk aan $1000/600 = 1,67$. De bijschatting van VOS voor deze bedrijfstak is dan $30 * (1,67 - 1) = 20$ ton. De totale emissie van VOS binnen bedrijfstak 17210 is dan $30 + 20 = 50$ ton.

⁶ Voor de emissie van distikstofoxiden worden literatuurgegevens gebruikt.

Tabel 3: Ophoogverklarende variabelen voor het compartiment lucht

SBI	Ophoogverklarende variabele	Eenheid
17210	hoeveelheid verschillende weefsels	ton
17240	werkzame personen	aantal
17300	hoeveelheid verbruikte vezels	ton
17400	opbrengst	gulden
17510	hoeveelheid vloerbedekking uit textiel	ton
17530	hoeveelheid gebonden textielvlies	ton
17540	opbrengst	gulden
19100	hoeveelheid huiden	m ²
19300	hoeveelheid geproduceerde schoenen	ton
20102	hoeveelheid verduurzaamd hout + SPIN document	ton
20200	fineer en ander plaatmateriaal	ton
20301	deuren (inclusief kozijnen daarvoor en drempels)	ton
21121	verschillende soorten papier en karton	ton
23202	energie	joule

bron: VROM

Alleen voor koolwaterstoffen zijn gegevens over de totale emissies bekend uit de KWS2000. De totale emissie van koolwaterstoffen in Nederland wordt in principe bepaald door de gegevens uit de KWS2000, behalve in het geval dat emissie uit de TNO-enquête hoger is. In dat geval worden de gegevens uit de TNO-enquête aangehouden. De kwaliteitsclassificatie van deze bijschattingen is C.

RIVM-bijschattingen

Voor vier sectoren worden de procesemissies door het RIVM bijgeschat. Het betreft de voedings- en genotmiddelenindustrie, de grafische industrie, de rubber- en kunststofverwerkende industrie en de metalectro-industrie. De bijschatting van emissies in deze bedrijfsgroepen vindt op diverse manieren plaats. Hierbij wordt onder andere gebruik gemaakt van gegevens uit de KWS2000. De kwaliteitsclassificatie van deze bijschattingen is D.

5.4 Bijschatting niet-energetisch gebruik van fossiele brandstoffen

Het betreft hier bijschattingen van emissies die optreden in processen waarbij energiedragers als grondstof voor de productie worden gebruikt en waarbij de aanwezige koolstof naar de lucht wordt geëmitteerd. Daarbij wordt verondersteld dat alle koolstof die niet in de eindproducten wordt opgenomen uiteindelijk CO₂ wordt. De bijschattingen vinden plaats door de gebruikscijfers van de diverse energiedragers (aardgas, olieproducten en kolenderivaten) te vermenigvuldigen met het koolstofgehalte van de diverse energiedragers. Daarbij wordt gecorrigeerd voor het percentage koolstof dat in de eindproducten wordt vastgelegd.

5.5 Stortplaatsen

Het betreft hier emissies (CH₄, NMVOS, benzeen, toluen, CFK & halonen, NO_x, H₂S, CO₂, CO, dioxine en fijn stof) uit stortplaatsen. De karakteristieken van individuele stortplaatsen worden gecombineerd met algemene emissiefactoren voor stortplaatsemisies. De gegevens over de individuele stortplaatsen zijn ontleend aan de jaarlijkse enquête door het Afvaloverlegorgaan (AOO), het RIVM en de Vereniging van Afvalverwerkers (VAVV). De algemene emissiefactoren zijn gebaseerd op een studie van de Grontmij. De kwaliteitsclassificatie is C.

5.6 Overige procesemissies

Voor een aantal processen wordt een afwijkend traject gevolgd voor de bijschatting. De reden hiervoor is dat bepaalde data al dan niet aanwezig zijn of uit een andere bron afkomstig zijn. Het betreft procesemissies die vrijkomen bij de delfstoffenwinning (kwaliteitsclassificatie: B), het bakken van brood (kwaliteitsclassificatie: B) en bij scheepswerven (kwaliteitsclassificatie: D).

5.7 Inzicht in de betrouwbaarheid van de schattingen voor de gebruiker

Naast de betrouwbaarheid van de bijschattingen is vooral de betrouwbaarheid van de cijfers aangaande de totale emissie in Nederland van belang. De schatting voor de totale emissie van een bepaalde stof is opgebouwd uit som van de uitstoot van een bepaalde stof uit de ER-I vermeerderd met de som van de bijschattingen van alle sectoren. Drie factoren zijn dan bepalend voor de betrouwbaarheid van de schatting van de totale emissie: de betrouwbaarheid van de metingen in de ER-I, de betrouwbaarheid van de bijschattingen voor een bepaalde stof in de diverse bedrijfspgroepen en de verhouding tussen de omvang van de metingen in de ER-I en de bijschattingen. Dit laatste element is van belang, omdat uit de betrouwbaarheid van de twee componenten (ER-I en bijschattingen) niet valt af te leiden hoe groot de betrouwbaarheid is van de schatting van de totale emissie. Stel bijvoorbeeld dat de cijfers uit de ER-I geen meetfouten bevatten, maar dat de bijschattingen de kwaliteitsclassificatie D bezitten. Als de cijfers uit de ER-I dan 10% van de totale emissie vormen en de bijschattingen dus 90%, dan is de schatting voor de totale emissie niet heel erg betrouwbaar. Als daarentegen de cijfers uit de ER-I dan 90% van de totale emissie vormen en de bijschattingen dus slechts 10%, dan is de schatting voor de totale emissie zeer betrouwbaar.

Alleen voor de bijschattingen bestaat op dit moment een redelijk inzicht in de betrouwbaarheid (denk aan de bovengenoemde kwaliteitsclassificatie). In de andere twee componenten bestaat ofwel geen inzicht in de betrouwbaarheid (cijfers ER-I) ofwel een beperkt inzicht (verhouding cijfers ER-I en bijschattingen). De verhouding tussen de cijfers uit de ER-I en de bijschattingen kan namelijk wel achterhaald worden, maar dit vergt enig rekenwerk. Overigens loopt de verhouding tussen cijfers uit de ER-I en de bijschattingen op het niveau van stoffen binnen de bedrijfspgroepen sterk uiteen. Zo is de bijdrage van grote bedrijven uit de ER-I aan de totale emissie van

stikstofoxiden binnen de bedrijfspgroep 'vezelindustrie' 100%. Daarentegen is de bijdrage van grote bedrijven aan de emissie van trichlooretheen binnen de bedrijfspgroep 'textiel- en kledingindustrie' slechts 20%. Wel is bekend dat gemiddeld genomen - over alle stoffen en alle bedrijfspgroepen - 80% van de emissies voor rekening komt van de bedrijven uit de ER-I en 20% voor rekening komt van de bijschattingen voor de kleinere bedrijven.⁷ Per stof en per bedrijfspgroep kan dit overigens - zoals gezegd - sterk verschillen. Voor een voldoende inzicht in de betrouwbaarheid van de gegevens uit de milieubalans is op dit moment dus onvoldoende informatie beschikbaar.

⁷ Dit is bekend op basis van de gegevens uit de eerste ronde van de TNO-enquête die betrekking had op 1973-1978.

6. Koppeling van TNO- en CBS-bestanden op bedrijfsniveau

Voor de totstandkoming van de milieubalans wordt de individuele emissieregistratie (ER-I) van TNO aan de productie- en investeringsstatistieken van het CBS gekoppeld. Deze koppeling wordt uitgevoerd - zie hoofdstuk 5 - op het niveau van bedrijfsgroepen (SBI's). Op deze wijze wordt inzicht verkregen in de (collectieve) milieuvervuiling van bedrijven die niet geregistreerd worden binnen de ER-I. Een alternatieve manier om de ER-I te koppelen aan de PS en IS is een koppeling te maken op het niveau van het *individuele bedrijf*. Op deze wijze ontstaat een bestand waarin per bedrijf informatie beschikbaar is over enerzijds economische gegevens en anderzijds milieugegevens.

Een koppeling op individueel niveau biedt een aantal interessante voordelen boven de bestaande koppeling per bedrijfsgroep. Op deze wijze kan heel expliciet de effectiviteit en efficiëntie van het milieubeleid worden geëvalueerd. Zo kan een betrouwbare inschatting worden verkregen van de milieuproductiviteit van bedrijven. Ook is het mogelijk om na te gaan hoe milieubeleid de technologische ontwikkeling - eventueel per sector - stimuleert. Daarnaast kan het effect van beleidsmaatregelen worden vastgesteld op de kosten die bedrijven moeten maken om aan de gestelde eisen te voldoen.

Voor een koppeling op individueel niveau is het noodzakelijk dat de wijze waarop gegevens voor de verschillende bestanden worden verzameld en gecatalogiseerd overeenkomen. In dit hoofdstuk zal daarom aandacht worden besteed aan de wijze waarop de TNO en CBS statistieken worden samengesteld. Binnen de Productie- en Investeringsstatistieken van het CBS wordt met de term 'bedrijfseenheid' de kleinste eenheid bedoeld die economisch gezien zelfstandig is. Daarmee wordt onder andere bedoeld dat de bedrijfseenheid zelfstandig beslissingen neemt ten aanzien van investeringen en in- en verkoop. Daarnaast moet de bedrijfseenheid zelfstandig een boekhouding voeren. Grote bedrijven zoals Philips bestaan uit enkele honderden bedrijfseenheden. Een bedrijfseenheid kan binnen de CBS-statistieken samenvallen met een vestiging of met meerdere vestigingen. In het eerste geval is een bedrijfseenheid gebonden aan één enkele locatie, in het tweede geval aan meerdere locaties. Op een locatie kunnen overigens meerdere bedrijfseenheden aanwezig zijn.

Binnen de Individuele Emissieregistratie (ER-I) valt een bedrijfseenheid altijd samen met een vestiging. Van elke bedrijfseenheid is dus een unieke locatie bekend. Ook in dit geval kunnen meerdere bedrijfseenheden op een locatie aanwezig zijn.

Een koppeling van de TNO- aan de CBS-bestanden is mogelijk op basis van de adresgegevens van de bedrijven. Gezien de omvang van het CBS-bestand (6000 à 7000 bedrijven) is de verwachting dat voor vrijwel alle vestigingen uit het TNO-bestand (500 tot 700 vestigingen) het bijbehorende bedrijf kan worden gevonden. Bij een koppeling kunnen in principe drie typen gekoppelde paren ontstaan:

1. Een bedrijf uit het CBS-bestand bestaat uit één vestiging en deze vestiging is opgenomen in het TNO-bestand.
2. Een bedrijf uit het CBS-bestand bestaat uit meerdere vestigingen en elk van deze vestigingen is opgenomen in het TNO-bestand.
3. Een bedrijf uit het CBS-bestand bestaat uit meerdere vestigingen, waarvan er tenminste één is opgenomen in het TNO-bestand en waarvan er tenminste één vestiging niet is opgenomen in het TNO-bestand.

In de gevallen 1 en 2 is sprake van een volledige matching van bedrijven uit de twee bestanden. In geval 3 is sprake van een onvolledige matching. De economische gegevens uit het CBS-bestand hebben immers betrekking op meer vestigingen dan de milieugegevens uit het TNO-bestand. Hoe groot de omvang is van elk van deze drie groepen valt helaas a priori niet in te schatten. Bij geen van de instituten (TNO en CBS) is op dit specifieke punt voldoende inzicht aanwezig in de samenstelling van de databases. Slechts door de koppeling daadwerkelijk uit te voeren kan hierin inzicht worden verkregen.

Er zijn verschillende mogelijkheden denkbaar om een op individueel niveau gekoppelde dataset te analyseren. Zo kan het onderzoek beperkt worden tot de data waarvoor een volledige matching (geval 1 en 2) beschikbaar is. Hiervoor is wel vereist dat er voldoende waarnemingen beschikbaar zijn. Als dit niet het geval is dan kan men overwegen om ook de waarnemingen waarvoor geen volledige matching (geval 3) beschikbaar is bij de analyse te betrekken. Met behulp van econometrische technieken kan dan gecorrigeerd worden voor de onvolledigheid van de matching. Voor een succesvolle analyse van de milieuproductiviteit van bedrijven is het in ieder geval van belang dat er een dataset van voldoende omvang beschikbaar is. Gezien de omvang van de TNO-dataset zal het noodzakelijk zijn om de CBS- en TNO-datasets over meerdere jaren te koppelen.

Een koppeling van de TNO- aan de CBS-dataset kan alleen - om redenen van geheimhouding - worden uitgevoerd door het CBS. Na koppeling kunnen onderzoekers binnen de muren van het CBS werken met de - anoniem gemaakte - data. Voor het gebruiken van de data en de daarvoor benodigde software rekent het CBS fl. 130,- per dagdeel.